DataVic Access Policy Guidelines

for the Victorian public sector (version 2)

August 2015

VICTORIA
State
Government

Currency

This is Version 2 of the DataVic Access Policy Guidelines, published in August 2015. Subsequent versions may be published from time to time at http://www.dtf.vic.gov.au/Victorias-Economy/Victorian-Government-intellectual-property-and-data-policies/DataVic-Access-Policy.

Inquiries

For any inquiries about the *DataVic Access Policy Guidelines for the Victorian public sector (version 2)* please contact:

data.vic@dtf.vic.gov.au
Department of Treasury and Finance
1 Treasury Place
Melbourne Victoria 3002
Australia

Contents

1. Introduction

1.1 DataVic Access Policy

Making datasets freely available to the public is the State's default position and where possible agencies must make datasets available with minimum restrictions, including the proactive removal of cost barriers.

1.2 DataVic Access Policy background

The DataVic Access Policy (the Policy) provides direction on licensing, pricing and management of Victorian government data so that it can be used and reused by the community and businesses.

The Victorian government holds, creates and collects a vast amount of data, ranging from demographic and economic to geospatial data. These datasets have the potential to drive innovation, reveal new research findings, create new business opportunities, and enable new services.

Around the world, governments are unlocking the value of its data by releasing it for public reuse. Governments have realised that many services can be developed externally. Rather than engaging directly in smart phone, web or other software developments, governments are releasing raw data and allowing the market to develop new and innovative products and services. In other jurisdictions, the market has delivered these products and services quickly and at no cost to government.

This Policy provides greater public access to Victorian government generated or owned data through the publication of datasets on the Victorian government Data Directory website, www.data.vic.gov.au (Data Directory).

1.3 Benefits of the Policy

Benefits of the Policy include

- stimulating economic activity and driving innovation and new services to the community and business;
- increasing productivity and improving personal and business decision making based on improved access to data;
- improving research outcomes by enabling access to primary data to researchers in a range of disciplines; and
- improving the efficiency and effectiveness of government by encouraging better management practices and use of the data.

1.4 Ministerial responsibility for the Policy

The Minister for Finance is the Minister responsible for administering the Policy through the Department of Treasury and Finance (DTF).

1.5 Policy intent

The intent of the Policy is to:

- enable public access to government data to support research and education;
- promote innovation;
- support improvements in productivity and stimulate growth in the Victorian economy; and
- enhance sharing of, and access to, information-rich resources to support evidence-based decision making in the public sector.

1.6 Policy Principles

The Principles underpinning the Policy are:

Table 1: DataVic Access Policy principles

	Principle	Chapter
1	Government data will be made available unless access is restricted for reasons of privacy, public safety, security and law enforcement, public health, and compliance with the law.	2-5
2	Government data will be made available under flexible licences.	6
3	With limited exceptions, government data will be made available at no or minimal cost.[1]	9
4	Government data will be easy to find (discoverable) and accessible in formats that promote its reuse.	7
5	Government will follow standards and guidelines relating to making datasets available and agency accountability for those datasets.	All

1.7 What agencies are covered by the Policy

The Policy and these supporting Guidelines apply to all agencies (that is, all departments and public bodies) of the State. 'Department' and 'Public body' are defined in the *Financial Management Act 1994*.[2] Public bodies include State business corporations and statutory authorities.

Accordingly, departments and public bodies as defined under the Financial Management Act must implement the Policy and these supporting Guidelines. Each agency must determine whether it is subject to the Policy on this basis.[3]

Implementation of the Policy and these supporting Guidelines will necessarily vary according to a number of factors, including the size, sophistication, intellectual property, data and needs of the agency. Agencies are encouraged to consider the Guidelines and to contact DTF with any queries.

---

[1] An agency may commercialise, or apply the Cost Recovery Guidelines to, government data if:

    a. it has an explicit statutory function to do so; or

    b. it has been explicitly authorised to do so by the Minister for Finance, because of a clear net benefit to the Victorian community.

[2] www.legislation.vic.gov.au/Domino/Web_Notes/LDMS/LTObject_Store/LTObjSt2.nsf/DDE300B846EED9C7CA257616000A3571/1C96E59A79C11194CA257761001F0276/$FILE/94-18a056.pdf

[3] Guidance as to the applicability of the Financial Management Act can be sought by checking the Annual Financial Report for the State of Victoria. The Report usually includes a list of significant controlled entities. All of the agencies in the controlled entities list are subject to the Policy. However, this list does not include all agencies subject to the Policy, and should therefore only be used indicatively. If an agency is unsure whether it is a public body under the Financial Management Act, it should seek legal advice.

1.8   Key actions of the DataVic Policy

These Guidelines also specify a set of key actions, identified by text boxes at the start of relevant chapters, to ensure compliance with the Policy. The list of key actions are:

1.   The Government's default position is that departments and public bodies datasets must be made available to the public.
2.   Datasets must be made available unless access is restricted for reasons of privacy, public safety, security, law enforcement, public health and compliance with the law.
3.   Datasets must be released in a machine-readable, reusable and open format.
4.   Personal, health and/or confidential information must be de-identified and aggregated.
5.   Creative Commons Attribution (CC BY) is the default licence for datasets released under the Policy.
6.   Metadata must be created for datasets released under the Policy.
7.   All datasets made available under the Policy are to be linked to the data directory, www.data.vic.gov.au.
8.   Agencies must make a determination about how often published datasets must be updated.
9.   Agencies must consider the Policy when developing and procuring datasets and databases.
10. Datasets will not be commercialised unless an agency has a statutory function to do so, or Ministerial approval is granted.
11. The agency head has overall accountability for implementing the Policy within their agency.
12. Each dataset made available must have an assigned custodian to ensure the dataset is managed through its lifecycle.
13. The progress of agencies compliance with the Policy will be reported to the responsible Minister and to Cabinet.

The list of key actions is also reproduced at Appendix 1: Key actions.

1.9   Currency

The Policy and these supporting Guidelines replace all previous policies on making government datasets available. This is Version 2 of the Policy Guidelines, published in August 2015. DTF is responsible for the maintenance of the Guidelines, and will keep a catalogue of each version. DTF welcomes feedback on the Guidelines and will review them on a regular basis. Agencies may suggest changes or additions to the Guidelines to DTF at data.vic@dtf.vic.gov.au.

Subsequent versions of the Guidelines may be published from time to time on the DataVic[4] and DTF[5] websites. DTF will inform agencies upon the publication of new versions of the Guidelines. Agencies should ensure that they are working with the current version of the Guidelines.

---

[4] www.data.vic.gov.au

[5] www.dtf.vic.gov.au/Victorias-Economy/Victorian-Government-intellectual-property-and-data-policies/DataVic-Access-Policy

1.10   Implementation support

DTF and the Department of Premier and Cabinet (DPC) are jointly responsible for the whole of Victorian Government implementation of the Policy.

Policy implementation support

DTF provides policy support, training, preparation of resources and is the first point of contact for all queries.

Comments and questions may be directed as follows:

Email: data.vic@dtf.vic.gov.au
Phone: (03) 9651 1880

Please forward requests for changes or additions to the Guidelines to data.vic@dtf.vic.gov.au for consideration.

DTF reviews the Guidelines annually.

Technical support for the Victorian government Data Directory

The linking of datasets on the Data Directory is managed by DPC. Technical queries in relation to the linking of datasets to the Data Directory should be referred to DPC.

Comments and questions may be directed as follows:

Email: info@data.vic.gov.au
Contact Form: www.data.vic.gov.au/contact
Web Services Portal: https://dsdbi.desk.com/
Phone: (03) 9651 8009

1.11   The Whole of Victorian Government Intellectual Property Policy Intent and Principles and supporting Guidelines

The Policy and supporting Guidelines must be read in conjunction with the Whole of Victorian Government Intellectual Property Policy (IP Policy) and supporting Guidelines.[6] The IP Policy sets out the State's approach to the management of Intellectual Property (IP). Government data often attracts IP protection. Accordingly, the DataVic and IP Policy are closely related, particularly in areas such as commercialisation and licencing. Agencies should consider these guidelines and contact DTF with any queries about the intersection between the policies.

Other policies that intersect the Policy are listed in Appendix 4: Related policies and standards.

---

[6]

www.dtf.vic.gov.au/Victorias-Economy/Victorian-Government-Intellectual-Property-and-data-policies/Intellectual-Property-Policy

2.   Making data available

Departments and agencies must take a number of steps to ensure that its data is made available and to comply with the *DataVic Access Policy*. The main steps are:

Table 2: DataVic Access Policy. The main steps

#	Step	Issues to consider	Chapter
1	Identify a dataset	• Existing and new datasets should be identified.	3
		• High-value datasets should be prioritised.	3.2
		• Some datasets are restricted for reasons of privacy, public safety, security and law enforcement, public health or compliance with the law.	4
2	Prepare the dataset for publication	• The dataset should be in an open format (e.g. CSV or XML).	5.1
		• The dataset should be de-identified of personal and/or confidential information.	5.2
		• A metadata record for the dataset should be created.	5.3
		• A data quality statement is recommended.	7.3
3	Select a licence	• An appropriate copyright licence must be selected.	6.1
		• The default licence is a Creative Commons Attribution 4.0 licence (CC-BY 4.0).	6.4
4	Publish the dataset	• The dataset should be uploaded to the department or agency's website or web service.	7.1
5	List the dataset at data.vic.gov.au	• Further information provided in the *DataVic Access Policy Dataset Publishing manual*.[7]	7.2
6	Manage the dataset	• Ensure currency of the dataset.	7.4
		• Manage feedback received from users via data.vic.gov.au.	7.6

Further detail can be found on each topic in the relevant chapters.

---

[7] www.dtf.vic.gov.au/Publications/Victoria-Economy-publications/IP-and-DataVic/DataVic-Access-Policy-Dataset-Publishing-Manual

3. Identifying datasets to be made available

This chapter contains the following key action:

1. The Government's default position is that departments and public bodies datasets must be made available to the public.

3.1 Definition of datasets

The Policy has adopted a broad definition of 'data'. The definition for data refers to datasets and databases owned and held by Victorian departments and public bodies and stored in formats including hardcopy, electronic (digital), audio, video, image, graphical, cartographic, physical sample, textual, geospatial or numerical form.

The Policy mandates that datasets must be made available in machine-readable, reusable and open formats. The Policy also applies to data made available in the form of an Application Programming Interface (API), web service or data tool (as long as the tool has a machine readable output).

The definition of datasets is intentionally broad to ensure all agencies consider a broad range of datasets to be made available.

Agencies are encouraged to review existing and new datasets and determine whether it is appropriate that they be made available.

3.2 What must be made available under the Policy

The Government's default position is that public sector datasets must be made available to the public, unless access is restricted for reasons of privacy, public safety, security and law enforcement, public health, and compliance with the law.

Agencies will be required to implement the Policy in **all** business areas that generate, create, collect, process, preserve, maintain, disseminate, or fund datasets. It is expected that making datasets available will be an ongoing (often scheduled) process and form a core part of departmental business activity.

Datasets that should be routinely published on the Data Directory include:

- datasets that are considered to be high value (see 3.2.1 below);
- datasets published in documents or reports that could be made more reusable (e.g. figures currently included in PDF reports, annual reports or reports already being provided to the Commonwealth);
- datasets which are already in machine-readable format;
- existing data catalogues (e.g. Spatial Datamart); and
- datasets currently made available on agency websites.

3.2.1 High value datasets

A high value dataset is defined as a dataset that is likely to be of interest to the Victorian community, and/or a dataset that has potential for valuable reuse. A dataset should be considered 'high value' if it:

- is central to the department/agency functions e.g. DTF and budget data;
- has been requested via 'suggest a dataset' (see 3.2.2 below);

- has previously/regularly been provided under the *Freedom of Information Act 1982*;[8]
- supports a major reporting process of government e.g. annual report data;
- planning data;
- spatial data;
- transport data;
- administrative data; and
- financial data.

Other types of data may also be high value. Agencies should consider which of its datasets are high value. Datasets that are not high value must still be released under the Policy, but agencies should prioritise high value datasets.

3.2.2 Dataset suggestions

The Data Directory includes a function called 'suggest a dataset'[9] that allows the community to make suggestions for datasets to be made available under the Policy. Once a suggestion has been received on the data directory, it will be forwarded to the relevant agency. It is the responsibility of the agency's dataset custodian (see Section 10 Accountability for datasets) to respond to the request within four weeks and if appropriate provide the relevant dataset in a timely manner.

Suggestions received on the Data Directory will be recorded in a central register on the website that will be displayed to the public. The agency custodian will be responsible for updating the register to record outcomes of suggestions relevant to their agency.

It is recommended that agencies report suggestions for datasets received, and the outcome of these suggestions, in its annual report. Agencies are also encouraged to provide a link from its agency website to the 'suggest a dataset' function on the data directory.

Further detail on the 'suggest a dataset' process including responsibilities, expectations and timeframes of responses is available on the DTF website.[10]

3.2.3 Establishing a process to make datasets available following freedom of information requests

Datasets made available under a freedom of information request must be considered for release under the Policy. The release of data via the Data Directory is consistent with the intent and language of the *Freedom of Information Act 1982.* Datasets will still need to be assessed as supporting the Policy intent and not breaching any restrictions. The time requirements stipulated under the *Freedom of Information Act 1982* do not apply to the Policy.

3.3 Collections of line agency data

Departments are encouraged to support its portfolio agencies in releasing data under the policy by collating data in meaningful collections e.g. Department of Health and Human Services making data available that it collects and compiles from hospitals.

---

[8] www.legislation.vic.gov.au/Domino/Web_Notes/LDMS/LTObject_Store/LTObjSt9.nsf/DDE300B846EED9C7CA25 7616000A3571/B74206A73D094529CA257D72001BD34D/$FILE/82-9859aa087%20authorised.pdf

[9] www.data.vic.gov.au/suggest_a_dataset_form

[10] www.dtf.vic.gov.au/Publications/Victoria-Economy-publications/IP-and-DataVic/DataVic-Access-Policy-Suggest-a-Dataset-Manual

3.4   Research data

The default obligation under the Policy is for agencies to make research datasets available. The advantages of sharing research data include that it:

- encourages scientific enquiry and promotes innovation;
- leads to new collaborations between data users and data creators;
- maximises transparency and accountability; and
- reduces the cost of duplicating data collection.

Allowing access to the data which underpins research significantly increases the potential benefits of the research. For example, it allows other researchers to test the findings in a research paper, or to take the research in a different direction.

Accordingly, where an agency owns or licences the data that underpins a research paper, the data should be released under the Policy if it meets the criteria for open data release.

Further, where an agency funds research which is likely to be underpinned by data, it should consider including appropriate terms in the funding or grant agreement to:

- allow the agency to release the data (see Chapter 6 Licensing datasets for more detail); and/or
- require or recommend that the researcher provide access to the data in a form that is compliant with the policy.

3.5   Creating an information asset register

Victorian Government Information Management Standards require agencies to create and maintain an information asset register. Using its registers agencies will be able to identify and manage datasets which can be made available to the public.[11]

3.6   Open data plan

Agencies are encouraged to develop and publish an open data plan. An open data plan describes the process for the release of an agency's data. The strategy should include:

- a description of the agency and the types of data the agency manages on behalf of the State;
- the agencies legislative and administrative framework;
- a statement of compliance with the Policy;
- the agencies governance and accountability framework;
- a list of short and long term goals;
- a plan for what data will be released and when;
- a discussion on what data will not be released and why; and
- a list of all data already released.

The supporting data to enable developing an open data strategy should be available in the agencies Information Asset Register (see section 3.5 for more information).

Publishing the agencies open data strategy should be done on the Data Directory.

---

[11] Guidance on information management, including asset management, can be found at: digital.vic.gov.au/policies-standards-guidelines/information-management/

4. Identifying which datasets must not be made available

> This chapter contains the following key action:
>
> 2. Datasets must be made available unless access is restricted for reasons of privacy, public safety, security, law enforcement, public health and compliance with the law.

The Policy stipulates that datasets will be made available unless access is restricted for reasons of privacy, public safety, security, law enforcement, public health and compliance with the law. These and additional restrictions are detailed below.

Datasets must not be made available if one of the following restrictions applies and a solution for removing the restriction cannot be found.

Section 4.1.   the dataset contains personally identifiable information and therefore does not support the Policy intent;

Section 4.2.   making the dataset available may have an adverse impact upon public safety;

Section 4.3.   the dataset attracts security restrictions;

Section 4.4.   the dataset contains information, the release of which is prohibited by court orders, legal proceedings or legal privilege, including legal advice;

Section 4.5.   making the dataset available may have an adverse impact upon public health;

Section 4.6.   the dataset contains third-party copyright and the State does not have permission from the copyright owner to make the data available;[12]

Section 4.7.   the dataset is subject to a contract/agreement that does not allow it to be made available;

Section 4.8.   making the dataset available is in breach of statutory or legislative requirements including the *Information Privacy Act 2000,*[13] or *Health Records Act 2001,*[14] or

Section 4.9.   the dataset contains confidential information.

Each of these restrictions are discussed below.

4.1 Personally identifiable information

The default obligation under the Policy is for agencies to make de-identified datasets available where possible. If a dataset contains personally identifiable information, and cannot be de-identified, it is not suitable for release under the Policy. Further information on de-identification can be found in section 5.2 below.

---

[12] Further guidance on agency ownership of copyright is provided in the whole of Victorian Government Intellectual Property Policy and Guidelines found at
www.dtf.vic.gov.au/Publications/Victoria-Economy-publications/Intellectual-Property-Guidelines-for-the-Victorian-Public-Sector-Version-1

[13] Guidance on the *Information Privacy Act 2000* can be found at
www.privacy.vic.gov.au/domino/privacyvic/web2.nsf/files/guidelines-to-the-information-privacy-principles/$file/guideline_11_11.pdf

[14] Guidance in the *Health Records Act 2001* can be found at www.health.vic.gov.au/hsc/legislation.htm

An example of a dataset containing personally identifiable information is unit record information about births, deaths or marriages. While aggregate datasets on this information, for example a dataset to measure the number of births in the State over a 12 month period may be appropriate for release under the Policy, individual birth data records include personally identifiable information and are not appropriate for release.

4.2 Public safety

Datasets must not be made available under the Policy that contain information that adversely affects public safety unless it can be cleansed and the risks mitigated.

4.3 Security classification

If data has been classified with a security classification and cannot be cleansed to allow the removal of the classification then it is unsuitable for release under the Policy.

The Victorian Government Information Security Management Policy requires agencies to assess the business impact of information release and classify information accordingly. For further information on dataset (information) classification, agencies should contact its Chief Information Officer or Information Security Adviser, and refer to:

- Agency Information Security policies and standards that are specific to the agency's risk appetite and relevant legislation;
- The Commissioner for Privacy and Data Protection website;[15]
- State of Victoria, Information Security Management Policy,[16] Standards and Associated Guidelines, 2012;
- Australian Government, Protective Security Policy Framework, 2012;[17] and
- Australian Government, Information security management guidelines – Australian Government security classification system, Version 2.1 - Commonwealth of Australia, 2014.[18]

4.4 Legal documents

Datasets containing information subject to court orders, legal proceedings or legal privilege including legal advice, must not be made available under the Policy unless it can be de-identified or appropriate approvals obtained.

4.5 Public health

Datasets under the Policy that contain information that may adversely affect public health must not be made available unless it can be cleansed and the risks mitigated.

4.6 Third-party copyright

Agencies must not make datasets available which contain third-party copyright materials where the State does not have permission from the copyright owner. This includes circumstances where an agency co-produces a dataset with another non Victorian state government agency at either the local, state or federal level.

---

[15] www.dataprotection.vic.gov.au/

[16] digital.vic.gov.au/policies-standards-guidelines/information-security/

[17] www.protectivesecurity.gov.au/Pages/default.aspx

[18] www.protectivesecurity.gov.au/informationsecurity/Documents/INFOSECGuidelinesAustralianGovernmentSecurityClassificationSystem.pdf

Under the *Copyright Act 1968*, the owner of copyright has exclusive rights to reproduce, publish and communicate the work, and to license others to do so. The State generally holds copyright in works created by its employees, but not in any third party material that is incorporated into those works unless otherwise specified under contract.

Where a dataset has been developed by a third party under a contract or agreement, copyright ownership may be dealt with specifically and should be checked.

For further information on intellectual property and copyright, agencies may refer to the *Whole of Victorian Government Intellectual Property Policy Intent and Principles and supporting Guidelines*[19] or seek professional legal advice.

Where appropriate an agency should secure appropriate permission to release third-party copyright material particularly when negotiating a new contract.

4.7 Contracts and agreements

Agencies must consider the Policy when entering into contracts. Where possible agencies should secure appropriate permission to release data when new contracts are being developed by including a clause in contracts to facilitate making datasets available to the public. The Victorian Government Purchasing Board (VGPB) has developed standard contracts to support the release of open data.[20]

For more information on how to deal with copyright when establishing new contracts and agreements see Section 8 Developing and procuring datasets and the *Victorian Government Intellectual Property Policy Intent and Principles* and supporting Guidelines.[21]

Agencies may have already entered into contracts or agreements with third parties that specifically determine whether datasets can be made available and who owns copyright on material created under the arrangement. These contracts or agreements must be checked before releasing datasets and, where appropriate, agencies should negotiate with the contractor to have the data made available. This is best done when a contract is up for renegotiation.

4.8 Legislation

Agencies operate under various legislative provisions that specify conditions for restricting access or release of datasets. The Policy does not override existing legislation.

Legislation includes the following but is not limited to:

- *Public Records Act 1973*;[22]
- *Privacy and Data Protection Act 2014;* [23] and
- *Health Records Act 2001*.[24]

---

[19] www.dtf.vic.gov.au/Victorias-Economy/Victorian-Government-intellectual-property-and-data-policies

[20] www.procurement.vic.gov.au/Buyers/Market-Approach-Templates/Contracts

[21] www.dtf.vic.gov.au/Victorias-Economy/Victorian-Government-intellectual-property-and-data-policies

[22] www.legislation.vic.gov.au/Domino/Web_Notes/LDMS/LTObject_Store/LTObjSt6.nsf/DDE300B846EED9C7CA257616000A3571/D2B93E4380A56273CA257A0700083577/$FILE/73-8418aa036%20authorised.pdf

[23] www.legislation.vic.gov.au/Domino/Web_Notes/LDMS/LTObject_Store/LTObjSt8.nsf/DDE300B846EED9C7CA257616000A3571/B224F6F00FF339CDCA257D570023421D/$FILE/14-60aa001%20authorised.pdf

[24] www.legislation.vic.gov.au/Domino/Web_Notes/LDMS/LTObject_Store/LTObjSt7.nsf/DDE300B846EED9C7CA257616000A3571/015E6F0294580950CA257AC3007BCA5F/$FILE/01-2aa024%20authorised.pdf

Agencies may be subject to other legislation particular to its business that specifies conditions for restricted access and/or release of its datasets.

Additional legislation relevant to information management compliance can be found in Appendix 3: Related legislation.

4.9 Confidential information

Datasets containing confidential information must not be made available under the Policy. Confidential information may include, but is not limited to:

- Cabinet-in-confidence information;
- commercial-in-confidence information; and
- information protected by the *Protected Disclosure Act 2012*.[25]

---

[25] www.legislation.vic.gov.au/Domino/Web_Notes/LDMS/LTObject_Store/LTObjSt7.nsf/DDE300B846EED9C7CA2 57616000A3571/A30A0FA8D1561C1BCA257A7D0021E34A/$FILE/01-36aa024%20authorised.pdf

5. Preparing datasets before making them available

This chapter contains the following key actions:

3. Datasets must be released in a machine-readable, reusable and open format.

4. Personal, health and/or confidential information must be de-identified and aggregated.

5.1 Selecting a format

Datasets must be made available via the web in a machine-readable open format.

Machine-readable refers to a medium that stores data in a format that is readable by an electronic device and is therefore more usable, able to be manipulated and complies with accessibility standards. A machine-readable file has a structure which allows easy interrogation of the contents. Machine-readable open data can be used with spreadsheet software, statistics software, and custom-written code. By releasing raw datasets in open standard formats, more people will be able to use and reuse them without having to invest in special software.

Extracting data from unstructured documents (e.g. PDF's) is enormously time-consuming and can be the most laborious part of a data-driven research project, and it can also lead to unnecessary errors.

A dataset may also be made available in its original format (e.g. PDF) where that format makes the dataset easier to understand or reuse. It is often useful to have a well formatted document to cross-reference and sense check the machine-readable data. The original document should be referenced within the metadata description of the machine-readable file. The original document can also be referenced in the description of the dataset on the data directory.

5.1.1 Open standard formats

An open format is a specification for storing and manipulating content that any developer may use, that is usually maintained by a standards organisation, and is not locked into a propitiatory product.

The preferred open formats for datasets to be made available are:

* CSV (comma separated values) for simple spreadsheets and simple databases.
  (Note: CSV files can be previewed within the Data Directory without the need to download the file, allowing end users to decide it the file is suitable for its purposes)
* XML (extensible markup language) – a general-purpose markup language for complex datasets, standardised by the main international standards organisation for the World Wide Web.
* XBRL (extensible business reporting language) - XBRL is a freely available global standard, standards-based way to communicate and exchange business information between business systems.

Note: Excel and most databases allow users to export files to CSV or XML.

Open standards for spatial data are maintained by the Open Geospatial Consortium and include:

- KML (formerly Keyhole Markup Language) an XML language focused on geographic visualization, including annotation of maps and images;
- WMS (Web Map Service) a protocol that allows georeferenced map images to be served over the web;
- WFS (Web Feature Service) allows requests for geographical features to be drawn across the web;
- WCS (Web Coverage Service Interface Standard) provides access to coverage data in forms that are useful for client-side rendering, as input into scientific models, and for other clients; and
- ESRI Shapefile – geospatial vector data format for geographic information systems software.

5.1.2 Proprietary formats

A proprietary format is typically a file format that is restricted for use by a company or organisation. The restrictions can include the control of the specification of the encoding format or licences that only the company or licensees may use. Proprietary formats are usually discouraged for release and every effort should be made to release data in an open format. If however the main users require specific proprietary formats then they will be need to be considered for release.

The following formats have already been considered and deemed acceptable for release via the Policy due to the common usage of the format. An open format should always be considered as well for release.

- XLS and XLSX (Excel Workbook) Main spreadsheet format which holds data in worksheets, charts, and macros; and
- GTFS (General Transit Feed Specification) defines a common format for public transportation schedules and associated geographic information.

5.1.3 Application Programing Interface

An API is a set of procedures, protocols, and tools that are used in developing software applications. An API makes it easier to for the programmer to develop software by providing all the building blocks necessary to connect and interrogate the dataset. For developers who build upon web services it is easier to take advantage of external services and data as an API to enhance its offering.

API's are particularly suitable for real time or regularly updated information.

5.2 De-identifying and aggregating data

De-identified datasets are datasets where any information linking the data to an individual or business are deleted or modified to remove the capacity for identification.

The Public Records Office Victoria (PROV) provides guidance on personal and private records, as named in section 9 of the *Public Records Act 1973*. The relevant PROV fact sheet notes that information described as personal or private records 'covers such material as personnel records, medical records, police and prison records and case records concerning students, welfare recipients, children in government care or compensation claimants'.[26]

---

[26] Public Records Office Victoria Fact Sheet Closure of Public Records under Section 9 of the *Public Records Act 1973* prov.vic.gov.au/wp-content/uploads/2011/11/1110fs1.pdf

To ensure that datasets containing personal,[27] health[28] and/or confidential information are correctly and consistently de-identified and or aggregated in order to be made available under the Policy, a formal procedure must be documented and adhered to by agencies.

For simple datasets, the documented procedure could be in the form of a checklist to be completed when each dataset is made available.

This procedure must not be shared or released with the datasets to third parties as knowledge of the procedure could potentially allow re-identification.

An example of identified and de-identified datasets can be seen below. The difference between Table 3 and Table 4 is the removal of specific age and location detail.

Table 3: Identifiable data

Age (Years)	Gender	Location	Diagnosis
16	Male	3844	Broken Arm
20	Female	3170	Diabetes
44	Male	3166	Heart Disease
93	Female	3666	Arthritis

Table 4: De-identified data based on generalisation of the information

Age (Years)	Gender	Location	Diagnosis
Under 30	Male	Regional	Broken Arm
Under 30	Female	Metropolitan	Diabetes
30-59	Male	Metropolitan	Heart Disease
60+	Female	Regional	Arthritis

Note: The method used to de-identify the table above is not applicable to all circumstances. It is used as an illustration only.

5.2.1 How do I de-identify a dataset?

For detailed information on de-identification, agencies can refer to a number of expert sources including Australian Bureau of Statistics (ABS), National ICT Australia (NICTA) and the Commonwealth Scientific and Industrial Research Organisation (CSIRO).

ABS provides guidance on confidentialised unit record files (CURFs).[29] CURFs are confidentialised by removing name and address information, by controlling and limiting the amount of detail available, and by very slightly modifying or deleting data where it is likely to enable identification of individuals or businesses.

The ABS has a number of factsheets to support de-identification, in particular from the National Statistical Service. These include:

- Statistical Data Integration involving Commonwealth Data – How to confidentialise data: the basic principles web page.[30]
- National Statistical Service – 11. Confidentiality and Privacy.[31]
- Frequently Asked Questions – About CURFs – How is CURF data confidentialised?[32]

---

[27] Personal information as defined by the *Victorian Information Privacy Act 2000*.

[28] Health information as defined by the *Victorian Health Records Act 2001*.

[29] abs.gov.au/websitedbs/D3310114.nsf/home/About+CURF+Microdata

[30] www.nss.gov.au/nss/home.nsf/pages/Confidentiality+-+How+to+confidentialise+data:+the+basic+principles

[31] www.nss.gov.au/nss/home.nsf/NSS/BA2EF335C739A1CECA25763F000B81A4?opendocument#11.2

[32] www.abs.gov.au/websitedbs/D3310114.nsf/4a256353001af3ed4b2562bb00121564/56936a48abd30f33ca2572c70021ab26?OpenDocument

5.2.2   How do I avoid the re-identification of data?[33]

As well as de-identifying information, effort should be made to ensure the risk of potential re-identification is low. This will include:

- dealing with small cell sizes and points at which data is aggregated; and
- ensuring that the information made available can't be linked to other publically available data which could then increase the risk of potential re-identification.

5.3   Preparing a data quality statement

When choosing to use a dataset, users need to assess if it is fit for purpose. A key factor in this decision is often the quality of the data. A data quality statement is an effective mechanism for communicating information about how data can be used, and can act as a qualifier or disclaimer for a dataset.

Data quality statements are an important risk management tool against datasets being misrepresented or misused, therefore the Policy recommends that one is created

A data quality statement can also be used to define key terms and identify any data standards which have been used.

5.3.1   How do I create a data quality statement?

It is recommended that the ABS, through its National Statistical Service (NSS) online data quality tool be used by agencies to create a data quality statement.

The data quality statement tool can be located at http://www.nss.gov.au/dataquality.

The tool outlines the seven dimensions that demonstrate fitness for purpose and provides users with a step-by-step guide to assessing the quality of their data. Additional information on these dimensions can be found at 1520.0 – ABS Data Quality Framework, May 2009.[34]

A template data quality statement produced by this tool is included at Appendix 2: Data quality statement.

It is important to note that other data quality statement processes are available for public sector use. Holders of spatial datasets may wish to use the ISO Standards approach to data quality statements, outlined in the Victorian Spatial Council's Spatial Information Data Quality Guidelines.[35] This document provides general principles for data quality, rather than a tool for agencies to apply to datasets. The Department of Environment, Land, Water and Planning has expertise in data quality for spatial datasets and can be contacted for information on this.

---

[33] DTF is developing a manual on the de-identification and re-identification of data. The manual, once developed, will be available via the DTF website at:
www.dtf.vic.gov.au/Victorias-Economy/Victorian-Government-intellectual-property-and-data-policies/DataVic-Access-Policy

[34] www.abs.gov.au/AUSSTATS/abs@.nsf/Latestproducts/1520.0Main%20Features2May%202009?opendocument&tabname=Summary&prodno=1520.0&issue=May%202009&num=&view=

[35] victorianspatialcouncil.org/cms/library/attachments/SIMF%20Data%20Quality%20Guidelines%20Edition%202_September%202009.pdf

6. Licensing datasets

This chapter contains the following key action:

5. Creative Commons Attribution (CC BY) is the default licence for datasets released under the DataVic Access Policy.

6.1 Licensing datasets under the Policy

When datasets are made available, they need to be done so under licence, to the extent they are protected by copyright (see 6.2 below). Public sector datasets should be released under terms allowing flexible public re-use without further permission. Applying Creative Commons (CC) licensing is the recommended way to achieve this. The default licence is CC-BY 4.0 as it is the least restrictive licence for datasets released under the Policy.

Detailed information on licencing can be found in the *Victorian Government Intellectual Property Policy Intent and Principles and supporting Guidelines.*[36]

6.2 Existence of copyright in a dataset

Copyright subsists in a dataset or database where the work of an author, in reducing that compilation to material form (including digital form) involves some intellectual activity that is directed not at collecting or inputting the data, but in expressing the work. Accordingly, a given dataset may or may not be subject to copyright.

As it is difficult to determine whether copyright subsists in a dataset, it is recommended that agencies apply a copyright licence. The licence should state that the data or dataset is subject to the terms of the licence to the extent that the data or dataset is protected by copyright.

Agencies may in some circumstances wish to seek specific legal advice about the State's capacity to licence particular intellectual property, including datasets.

6.3 Ownership of any copyright in datasets

Datasets and databases could include multiple copyright works, collated and or created by human labour or through the use of software programs. Datasets and databases may have a number of components that are protected by copyright including the model or schema for the database, data entry and output sheets and the data itself, to the extent to which it is, or includes, created or expressive works.

This chapter applies to State-owned datasets. Where datasets are likely to include third party data made available to the agency under agreements or arrangements with third parties or terms of consent, the agency will need to consider whether these permit it to licence the dataset using a non-restrictive licence.

---

[36] www.dtf.vic.gov.au/Victorias-Economy/Victorian-Government-Intellectual-Property-and-data-policies/Intellectual-Property-Policy

6.4 Form of licence

The most appropriate form of licence for State-owned data containing copyright material will vary depending on the type of material involved and the circumstances surrounding making the dataset available. Generally, the recommended form of licence for copyright material is a Creative Commons (CC) licence.

6.4.1 What is Creative Commons?

Creative Commons (CC) refers to international copyright licences which have been developed for use by the owners of copyright material, including governments. CC licences are designed to provide copyright owners with an efficient way to manage the rights contained in their copyright work, and to provide copyright users with simple and flexible terms for use.

A CC licence always requires that users attribute the work in the form specified by the licensor.[37] Attribution also requires the user to indicate if changes have been made to the work and to provide a link to the CC licence. CC also allows the copyright owner to apply one or more of three additional conditions, discussed below.

CC allows users of works to reproduce, communicate, distribute and perform those works without prior permission or payment. Some CC licences also allow users to make derivative works and to use works for a commercial purpose.

CC licences are applied by the licensor when the work is published. This process is automated via online tools. These tools include a version of the licence in 'human readable' form, the legal code which sets out the terms and conditions of the licence, and digital code which enables metadata to be attached to the on line version of the work.[38]

CC licences are designed to be used 'as is'. Each contains a non-exclusive licence which is stated to be worldwide, royalty free, non-exclusive and perpetual (subject to the termination provision). They do not provide for the imposition of licence fees, or licensing for a limited duration or purpose.

6.4.2 Creative Commons licences

The licence recommended for use by agencies under the Policy is CC BY 4.0.

There are six CC licences available. Each is identified by a combination of the symbols, an acronym and a descriptive label.

Table 5: CC licences available.

Image	Label	Acronym	Licence conditions
	Attribution	CC BY	• Attribution as required by licensor.[39]
	Attribution – Share Alike	CC BY-SA	• Attribution as required by licensor; and • any derivative works must be distributed under the same licence.
	Attribution – No Derivative Works	CC BY-ND	• Attribution as required by licensor; and • derivative works must not be made.

---

[37] For example, by attributing the name of the author or copyright owner.

[38] This metadata, which can be added to by the licensor, allows works to be readily searched online.

[39] A user of the work must attribute it as required by the licensor. For example, agencies should request that users of their copyright material attribute it to the State of Victoria. This condition also requires the user to indicate if changes have been made to the work and provide a link to the CC licence.

Image	Label	Acronym	Licence conditions
	Attribution – Noncommercial	CC BY-NC	• Attribution as required by licensor; and • work must not be used for commercial purposes.
	Attribution – Noncommercial – Share Alike	CC BY-NC-SA	• Attribution as required by licensor; and • work must not be used for commercial purposes; • any derivative works must be distributed under the same licence.
	Attribution – Noncommercial – No Derivatives	CC BY-NC-ND	• Attribution as required by licensor; • work must not be used for commercial purposes; and • derivative works must not be made.

CC BY should be used with most copyright material that is made publicly accessible. However, in some limited circumstances, an alternative CC licence may be more appropriate. For example, if an agency needed to prevent any commercialisation of copyright material (which would not normally be the case under the Policy), CC BY-NC may be the most appropriate licence. Agencies should carefully consider whether an alternative CC licence is appropriate, in consultation with its IP Coordinator, DTF at IPpolicy@dtf.vic.gov.au and, if required, legal advice.

The current version of CC licences is 4.0, which was released in November 2013. The original DataVic Guidelines recommended the use of the previous version 3.0, and some agencies have already released material under version 3.0. It is recommended that agencies now adopt version 4.0.

The following information in these Guidelines on CC relates to CC BY 4.0. The use of other CC licences may involve issues not addressed in the Guidelines. Legal advice should be sought before doing so.

6.4.3   How to apply a CC BY 4.0 licence to new material

Prior to making appropriate State owned copyright material available to the public, a CC BY licence should be applied as follows.

A CC BY licence may be applied to copyright material by attaching a notice to this effect to the work. Accordingly, departments and agencies are encouraged to apply CC BY licences to template documents.

In many cases, the following notice may be appropriate:

*© State of Victoria ([name of agency, for example, '(Department of Treasury and Finance)'])
[year]*

This work, [insert title of work], is licensed under a Creative Commons Attribution 4.0 licence *[link to http://creativecommons.org/licenses/by/4.0/]. You are free to re-use the work under that licence, on the condition that you credit the State of Victoria ([name of agency]) as author, indicate if changes were made and comply with the other licence terms.*

[Optional addition to exclude aspects of the work from the CC licence:] The licence does not apply to [specify which aspects of the work the licence does not apply to, for example. 'any branding' or 'any images or photographs']

The 'optional addition' at the final sentence of the notice enables agencies to exclude aspects of a work from the CC licence. For example, it can be used to ensure that the State does not provide third parties with a licence to use its branding, in order to protect the State's reputation. Images and photographs may also be appropriate to exclude because they are often owned by third parties. However, this exclusion should only be used in appropriate circumstances, and should not unduly limit the scope of the licence. Agencies are encouraged to consult with its IP Coordinator, DTF at IPpolicy@dtf.vic.gov.au and, if required, seek legal advice.

These Guidelines are licensed under CC BY 4.0. See the rear side of the cover page above for the copyright notice. DTF has also applied CC licences to its template documents and publications. More information is available at http://www.dtf.vic.gov.au/Victorias-Economy/Victorian-Government-intellectual-property-and-data-policies/Intellectual-Property-Policy.

It is recommended that the notice be accompanied by a symbol indicating the CC licence applied. The symbols are available on the CC website at

http://creativecommons.org.au

The CC BY symbol is as follows:

For online materials, the agency should include the Digital Code provided by CC.

The CC Australia website contains useful material relating to the application of CC licences by government.[40]

---

[40] creativecommons.org.au/sectors/government

7.   Publishing datasets

This chapter contains the following key actions:

6. Metadata must be created for datasets released under the Policy.

7. All datasets made available under the Policy are to be linked to the Data Directory, www.data.vic.gov.au.

8. Agencies must make a determination about how often published datasets must be updated.

7.1   Where will data be published?

Most Victorian data is published and hosted on agency managed websites and web portals. All publically released datasets must also be linked to the Data Directory at www.data.vic.gov.au.

7.2   Publishing on the Data Directory data.vic.gov.au

The Data Directory has been created to simplify discovery and access of Victorian data. All datasets made available under the Policy must be linked to the Data Directory. Data is not usually directly published on the Data Directory.

To publish a dataset, a data record, with appropriate metadata completed by the dataset custodian, is created on the Data Directory. The dataset is considered to be 'published' once the data record and metadata are on the Data Directory.[41]

The data record on the Data Directory will point back to either:

* the dataset itself, such that the user can directly download the dataset from the agency website via the Data Directory without having to navigate to the agency website; or
* a page on the agency website that provides direct access to datasets via a data tool.

Datasets are not published on the Data Directory unless there is an agreement between the Data Directory technical support and the agency to host the dataset on the Data Directory.

All activities relating to hosting and publishing are the responsibility of the dataset custodian (see section 10 Accountability for datasets for further information on the role of dataset custodian).

A dataset publishing manual detailing the technical elements of publishing on the Data Directory is available on the DTF website.[42]

7.2.1   Agency responsibility for datasets published on the Data Directory

Once the dataset is published on the Data Directory, agencies must take responsibility for managing its presence on the site. This includes dealing with user feedback and moderating content (for further information on agency roles see section 10 Accountability for datasets).

---

[41] When a dataset is 'hosted', it is placed on a website server in preparation for being made available on the World Wide Web. A dataset is 'published' when the dataset is linked to the data.vic web portal.

[42]

www.dtf.vic.gov.au/Publications/Victoria-Economy-publications/IP-and-DataVic/DataVic-Access-Policy-Dataset-Publishing-Manual

7.2.2 Terms and conditions of use[43]

Agencies should be aware of the legal compliance risks associated with hosting and publishing datasets online.

Information on legal compliance can be found in the Website Management Framework Guideline: WEB GUIDE 07 Managing Legal Risks Online.[44] This guideline provides general advice on identifying and managing legal risks within an online environment. It includes standard terms and conditions which all agencies must use in relation to its datasets and online content.

Agencies should also be aware that when they publish datasets on the data directory, the website is governed by terms and conditions of use.

The Data Directory terms of use include guidance on:

- lawful use;
- links to and from the website;
- data and systems integrity;
- the website code of conduct;
- copyright and attribution;
- copyright in user comments;
- the website moderation policy and process;
- disclaimers on the site, links to external websites and security; and
- limitations on the use of datasets.

For specific terms and conditions, refer to www.data.vic.gov.au/terms-of-use.

7.3   Metadata for the creation of a record on the Data Directory

Metadata provides contextual information to users of government data. It helps them find information through searches and helps them understand how the information can be used.

The following description of metadata must be provided when creating a data record. Detail on how to provide the metadata statement is found in the Policy - Dataset Publishing manual on the DTF website.[45] This standard is derived from the *Victorian Government Guideline: Information Management Roles and Responsibilities*.[46]

Table 6 Metadata template

Field name	Description
Agency (if applic.)	Organisation with authority and accountability for the information asset.
Department	Organisation with authority and accountability for the information asset.
Data Administrator	Officer managing agency presence on data.vic (will not be made public)
Data Administrator Contact Details email	Email address (will not be made public)

---

[43] For more information refer to the whole of Victorian Government Intellectual Property Policy Intent and Principles and supporting Guidelines, available at:
www.dtf.vic.gov.au/Victorias-Economy/Victorian-Government-Intellectual-Property-and-data-policies/Intellectual-Property-Policy

[44] www.digital.vic.gov.au/wp-content/uploads/2014/07/WEB-GUIDE-07-WMF-Managing-Legal-Risks-v2.1.pdf

[45]
www.dtf.vic.gov.au/Publications/Victoria-Economy-publications/IP-and-DataVic/DataVic-Access-Policy-Dataset-Publishing-Manual

[46] www.digital.vic.gov.au/policies-standards-guidelines/information-management/

Field name	Description
Data Administrator Contact Details phone	Phone number (will not be made public)
Approver Contact	Officer authorised to release data to the public)
Approver Contact Details email	Email address (will not be made public)
Approver Contact Details phone	Phone number (will not be made public)
Dept./Agency URL	URL
Release Date	Date the data source was first released for public use
Format	The file format of the information asset. (i.e. XML, excel)
Data Source URL	Direct link to data source location i.e. http://www.premier.vic.gov.au/images/stories/datasets/brigades.csv
Program URL	Can be used to link to specific agency programme website
Data Frequency	Frequency that the data source is updated i.e. weekly, monthly
Temporal Range	Date range the data covers i.e. Branch Expenditure 2012 - start date to end date - 01/01/2012 to 31/12/2012
GeoData (if applic.)	Granularity (extent of drill down level - i.e. to LGA level)
GeoData (if applic.)	Coverage (extent of data coverage - i.e. all of Victoria, all LGAs)
Data Type/Category	DataVic data category
Licensing	A licence giving official permission to use the information asset
Site Tags	Two or three tags to support search functions

7.4 Frequency of update

Agencies must make a determination about how often published datasets must be updated. Criteria to consider include:

- the resources required for data extract and preparation;
- how much of the date has been de-identified or aggregated;
- whether the dataset is an important input to other processes; and
- regular reporting requirements that could be aligned to making particular datasets available, for example for quarterly, end of calendar year or end of financial year reports.

7.5 Legal compliance risks associated with hosting and publishing datasets

Agencies should be aware of the legal compliance risks associated with hosting and publishing datasets online. Under the *Whole of Victorian Government Standard Website Management Framework*,[47] the minimum information requirements dictate the use of copyright and disclaimer statements.[48] Agencies are to ensure compliance with the *Copyright Act 1968 (Cth)*, taking into account the purpose and functionality of the website. For specific terms and conditions, refer to www.data.vic.gov.au/terms-of-use.[49]

[47] State of Victoria, Whole of Victorian Government Standard Website Management Framework, 'Minimum Information Provision' (2011) available at www.digital.vic.gov.au/wp-content/uploads/2014/07/WEB-STD-09-WMF-Minimum-Information-Provision-v3.1-1.pdf

[48] Other notices such as security and privacy are also required, but these are not considered in this section

[49] State of Victoria, *Whole of Victorian Government Standard Website Management Framework, Legal Compliance* (2010) available at: www.digital.vic.gov.au/wp-content/uploads/2014/07/WEB-STD-03-WMF-Legal-Compliance-v2.1.pdf

7.5.1 Copyright statements

Copyright statements for agency websites can take a number of forms. It should be noted that datasets downloaded through the Data Directory are subject to the terms and conditions of the Data Directory at the point of download, as well as the applicable creative commons licence. Agencies must ensure that where datasets are available for download from its own website, the terms and conditions on the agency website are identical to those of the Data Directory. More information can be found on the Data Directory at:

- www.data.vic.gov.au/terms-of-use

See also information on copyright and attribution at:

- www.data.vic.gov.au/copyright-and-attribution

7.5.2 Website disclaimers

Website disclaimers are used to draw to the attention of users that the State or agency is not responsible for loss that arises from use or reliance on information or material on the website.

The State makes no statements, representations or warranties as to the accuracy or completeness of the content of a website and does not accept any liability for the information on a website.

The State also assumes no responsibility or liability for the condition or content of any third party website or for the operation or function of any service or facility offered on any third party website.

For more information on website terms of use and disclaimers see the *Victorian Government's Intellectual Property Policy Intent and Principles and supporting Intellectual Property Guidelines*.[50]

---

[50] www.dtf.vic.gov.au/Victorias-Economy/Victorian-Government-Intellectual-Property-and-data-policies

8.   Developing and procuring datasets

This chapter contains the following key action:

9.   Agencies must consider the Policy when developing and procuring datasets and databases.

8.1   Developing databases and datasets

When developing a database or dataset consideration should be given in the design phase to enabling public access to the data that is suitable for release under the Policy. There are many ways this can be achieved. It may be as simple as designing a report suitable for releasing to the public. It could also be as complex as developing a mirror image of the database on a separate server and developing an API that gives the public direct read access to a separate secure and stable instance of the dataset.

8.2   Procuring databases and datasets

The State enters into many contracts that result in the use, creation or assignment of data in some form. The data generated can be the purpose of the procurement (e.g. research), or arise incidentally. In both of these cases data IP rights and issues must be addressed.

Procurement processes and contracts must reflect the requirements of the Policy and the IP Policy for all State procurement.

When procuring databases and datasets agencies should secure appropriate permission to release the data when contracts are being developed or renegotiated. This should be done by including a clause in the contract to facilitate making datasets available to the public. The VGPB has developed standard contracts to support the release of open data.[51] The following is an extract of a clause within the VGPB standard contract:

Data

The ownership of all Data, including any Intellectual Property Rights, shall vest in the Organisation upon the time of its creation.

This is an intended departure from how other Intellectual Property is dealt with and consistent with both the DataVic and IP policies. For all other IP the Intellectual Property clause limits the states reuse of the IP and gives ownership of the IP to the supplier as noted below.

Contract Intellectual Property

(a) The ownership of any Contract Intellectual Property shall vest in the Supplier upon the time of its creation.

(b) The Supplier hereby irrevocably and unconditionally grants to the Organisation, free of additional charge, a non-exclusive, worldwide perpetual, transferable licence (including the right to sublicense) to use, reproduce, adapt, modify, publish, distribute and communicate any Contract Intellectual Property only to the extent necessary to achieve the purposes of the procurement.

---

[51] www.procurement.vic.gov.au/Buyers/Market-Approach-Templates/Contracts

If an agency does not wish to maintain ownership of a dataset a contract to licence the data to use, reuse and share should be negotiated.

Note: within the VGPB contracts the following terms are defined as:

Data: any information, data, datasets or databases created by or on behalf of the supplier in the course of providing the services.

Organisation: Agency of the State of Victoria.

9. Commercialising datasets

This chapter contains the following key action:

10. Datasets will not be commercialised unless an agency has a statutory function to do so, or Ministerial approval is granted.

9.1 Commercialisation background

Releasing government datasets at no or minimal cost will create more opportunities for the community to engage with data and for the private sector to reuse and add value to datasets. Making datasets freely available is the State's default position and where possible agencies must make data available with minimum restrictions, including the removal of cost barriers.

Research and economic modelling has found that releasing public sector data at no or minimal cost maximises economic benefits.

Broadly, intellectual property including datasets must not be commercialised unless there is a clear net benefit to the community. The government is not in the business of commercialising datasets. **Approval to commercialise will only be granted in exceptional circumstances.**

Detailed information on commercialisation of data and IP can be found in the *Whole of Victorian Government Intellectual Property Policy and supporting Guidelines*.[52]

9.2 When may a dataset be commercialised?

There are times when the decision will be made to commercialise government data. In these cases the decision will be based on a number of considerations.

An agency may commercialise, or apply the *Cost Recovery Guidelines* to, data if:

a) it has an explicit statutory function to do so; or
b) it has been explicitly authorised by the Minister for Finance[53] to do so because of a clear net benefit to the Victorian community.

Each basis for commercialisation is discussed further below.

Accordingly, the State discourages the commercialisation of data except where provided for by legislation or in other exceptional circumstances.

9.2.1 Commercialisation under explicit statutory function

An agency may commercialise or apply the *Cost Recovery Guidelines* to data if it has an explicit statutory function to do so. In such cases there is no need for the agency to seek authorisation for commercialisation or cost recovery.

---

[52] www.dtf.vic.gov.au/Victorias-Economy/Victorian-Government-Intellectual-Property-and-data-policies/Intellectual-Property-Policy

[53] The IP Policy refers to the Treasurer in this Principle, but in practice the Minister for Finance is responsible for authorisation requests under Principle 8(b) of the Victorian Government Intellectual Property Policy and supporting Guidelines.

Some statutes may provide an agency with a power, rather than a function, to commercialise data. Whether this allows commercialisation under the Policy will depend on the surrounding circumstances.

Agencies that intend to commercialise data should carefully consider whether there actually is an explicit statutory function to do so. If there is any doubt about whether this applies, an agency should consult with DTF at IPpolicy@dtf.vic.gov.au.

9.2.2 Commercialisation with Ministerial authorisation

If an agency intends to commercialise or apply the *Cost Recovery Guidelines* to data without an explicit statutory function to do so, it requires explicit authorisation by the Minister for Finance[54] under Principle 8(b) of the IP Policy.

This Principle applies when:

- an agency proposes to enter a new commercialisation arrangement in relation to existing data (even where that data has previously been commercialised);
- an agency proposes to develop data to be commercialised (including developing new data for the purpose of an existing commercialisation arrangement); and
- an existing commercialisation arrangement comes up for renewal or amendment.

Commercialisation or application of the *Cost Recovery Guidelines* may be authorised where the agency has demonstrated that it will result in a clear net benefit to the Victorian community. A 'net benefit' is an overall positive impact on the community. It takes into account the costs and benefits related to:

- public health and safety;
- social and community impact;
- environmental impact;
- competition; and
- economic impact.

Approval to commercialise or apply the *Cost Recovery Guidelines* will only be granted where a clear net benefit to the Victorian community has been demonstrated, having regard to all of the circumstances. For example, the possibility of a financial return from commercialisation will not, of itself, necessarily constitute a clear net benefit to the Victorian community.

Agencies are strongly encouraged to consult with DTF at IPpolicy@dtf.vic.gov.au prior to making a request for authorisation of commercialisation. Doing so is likely to reduce the time required for the agency to develop the request and for the request to be considered.

Detailed information on gaining Ministerial authorisation for commercialisation of data and IP can be found in the *Whole of Victorian Government Intellectual Property Policy and supporting Guidelines*.[55]

---

[54] The IP Policy refers to the Treasurer in this Principle, but in practice the Minister for Finance is responsible for authorisation requests under Principle 8(b).

[55] www.dtf.vic.gov.au/Victorias-Economy/Victorian-Government-Intellectual-Property-and-data-policies/Intellectual-Property-Policy

9.3   Freemium commercialisation model

There are occasions when a 'freemium' commercialisation model may be required to support the publics need when releasing data to the public. This is usually only required if the public require higher levels of support beyond the access to the dataset.

In the freemium model the dataset continues to be made available at no or minimal cost. The supporting elements that improve access to the dataset, including improved levels of reliability and service, contracted service level agreements and access to technical and specialist support may be considered suitable of commercialisation.

In such cases there is no need for the agency to seek authorisation for commercialisation or cost recovery as long as the data is still available at no or minimal cost.

9.4   What other policies must be considered when commercialising datasets?

9.4.1   Cost recovery

Where seeking approval to commercialise, (or continue to commercialise) datasets, a cost recovery model must be considered. This allows an agency to comply with the principle of making government data available at no or minimal cost.

Agencies must use the *Victorian Government Cost Recovery Guidelines*[56] to propose an appropriate charging model. The guidelines provide a framework for use by government entities when considering, developing and reviewing user charges and regulatory fees. They ensure that cost recovery arrangements in Victoria are transparent, efficient, effective and consistent with legislative requirements and government policy.

9.4.2   Competitive neutrality policy

Decisions to commercialise datasets are subject to the *Victorian Government Competitive Neutrality Policy*[57] which is concerned with transparent cost identification and pricing in a way that removes advantages arising from public ownership. Competitive neutrality ensures that the significant business activities of publicly owned entities compete fairly in the market.

---

[56] www.dtf.vic.gov.au/Publications/Victoria-Economy-publications/Cost-recovery-guidelines

[57] www.dtf.vic.gov.au/Publications/Victoria-Economy-publications/Competitive-neutrality-policy

10. Accountability for datasets

This chapter contains the following key action:

11. The agency head has overall accountability for implementing the Policy within their agency.

12. Each dataset made available must have an assigned custodian to ensure the dataset is managed through its lifecycle.

13. The progress of agencies compliance with the Policy will be reported to the responsible Minister and to Cabinet.

10.1 Who has accountability?

The agency head[58] has overall accountability for implementing the Policy within their agency. This would typically be the Secretary of the Department or the CEO of the Agency. It is recommended, however, that the agency assign responsibility for the implementation of the Policy to an agency-wide committee reporting to the agency head. The committee would monitor, report and steer the agency through making datasets available.

An agency's Information Management Governance Committee (IMGC) would be the recommended body to perform this function.[59]

The IMGC or equivalent would be responsible for:

- providing leadership and fostering an organisational culture of releasing datasets;
- building organisational capability in developing, releasing and managing datasets;
- monitoring and reporting compliance with the Policy;
- identifying information assets with the potential for high public value;
- collaborating with other agency IMGCs to share lessons learnt and promote intra-departmental initiatives;
- collaborating with the whole of Victorian Government IMG's to identify opportunities to achieve consistencies and efficiencies; and
- considering dataset suggestions as provided by the public.

Information Management Standards and Guidelines include recommended membership and sample terms of reference for the IMGC. For more information see:

- Whole of Victorian Government Standard – Agency Information Management Governance[60]
- Whole of Victorian Government Guideline – Information Management Roles and Responsibilities[61]

---

[58] The head of a department, public body, public statutory body, body, office or trust body as described by the *Financial Management Act 1994*.

[59] Under whole of Victorian Government Information Management Standards, it is now mandatory to establish and maintain an Information Management Governance Committee.

[60] www.digital.vic.gov.au/policies-standards-guidelines/information-management/

[61] www.digital.vic.gov.au/policies-standards-guidelines/information-management/

10.2 Custodianship of released datasets

Each agency must have an assigned custodian to ensure that datasets released on the Data Directory are managed through their lifecycle. It is recommended that the custodian's role include responsibility for such coordination activities as:
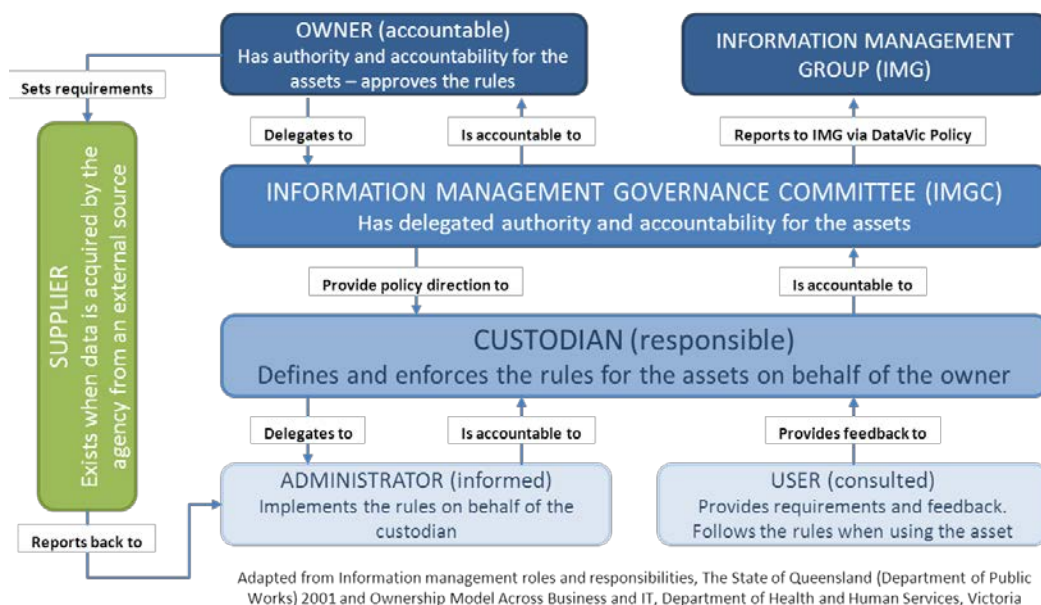
- maintaining a data asset register;
- identifying and coordinating datasets to be made available;
- defining the appropriate data quality statements;
- uploading datasets to the data directory;
- managing dataset suggestions and feedback via the data directory; and
- reporting the progress of making datasets available to the IMGC and DataVic Access Policy team.

10.2.1 Custodianship framework

Existing Victorian Government Information Management Standards require agencies to develop a custodianship framework.

A recommended model for data custodianship follows. There are six key roles: Owner, Information Management Group, Information Management Governance Committee, Custodian, Administrator and User. This model also includes a seventh user (Supplier) for data acquired from an external source. This model (developed based on the *Whole of Victorian Government Chief Information Officer's Roles and Responsibilities Guide*) is represented graphically below:

Figure 1: Custodianship framework

Adapted from Information management roles and responsibilities, The State of Queensland (Department of Public Works) 2001 and Ownership Model Across Business and IT, Department of Health and Human Services, Victoria

For more information see:

- Whole of Victorian Government Standard – Information Asset Custodianship[62]
- Whole of Victorian Government Guideline – Information Management Roles and Responsibilities[63]

---

[62] www.digital.vic.gov.au/wp-content/uploads/2014/07/VG-Information-Asset-Custodianship-IM-STD-01.pdf

[63] www.digital.vic.gov.au/wp-content/uploads/2014/07/Information-Management-Roles-and-Responsibilities-IM-GUIDE-1.1.pdf

10.3 DataVic Access Policy accountability checklist

The following checklist for accountability compliance requirements can be used by agencies to assist implementation. (Note: Some activities may not apply to your agency).

Table 7: DataVic Access Policy accountability checklist

Activity	Yes/No
Assign implementation responsibility to the IMGC or equivalent	
Applicable agency policies (e.g. information management, records management, website policy, privacy policy etc.) have been updated to include the Policy	
Agency has developed and communicated a custodianship framework	
Agency staff have been informed of Policy changes	
Agency has started an information asset audit	
Priority datasets have been identified for a Policy transition program of works	
A program of works is developed including delegating tasks to appropriate resources	
Agency has established a process for managing dataset suggestions and feedback	

Once requirements have all been marked 'Yes', Policy accountability elements are in place.

10.4 Dataset feedback

The Data Directory includes a function that allows users to provide feedback on datasets published on the website.

Agencies will be notified when a feedback request is made. Agencies are responsible for moderating and managing the treatment of feedback. The custodian nominated by the agency should undertake this role. The timeframes for responding to feedback are the same as the 'suggest a dataset' process.

Further detail on the 'suggest a dataset' process including responsibilities, expectations and time frames of responses is available on the DTF website. [64]

10.5 Reporting

10.5.1 Compliance reporting to Government

There are two ways in which implementation of the Policy will be reported. These are:

- quarterly reports from DTF to the Minister for Finance on the progress of the Policy, which include an update on the total number of datasets published; and
- annual reporting to Cabinet by the Minister for Finance which will detail how many datasets agencies have published on the Data Directory. This will involve a chart comparing the number of datasets made available by each agency as well as requests for datasets and progress on release.

DTF may (at its discretion) request status reports from agencies on their implementation of the Policy for the purpose of accurately reporting how many datasets have been made available.

---

[64]

www.dtf.vic.gov.au/Publications/Victoria-Economy-publications/IP-and-DataVic/DataVic-Access-Policy-Suggest-a-Dataset-Manual

10.5.2 Data Directory reporting

The Data Directory has an analytics page supplying usage data on all data released via the portal.[65] This page has been set up to support agencies and the public to gain an understanding of the usage of the site. Statistics may include:

- total published datasets;
- number of datasets updated in period;
- number of new datasets updated in period;
- number of downloads; and
- time since datasets were last updated.

On request the technical support team can also supply agencies with tailored web statistics reports from the data portal.

10.5.3 Annual reports

Agencies are encouraged to release the data supporting the development of its annual report and to incorporate policy achievements into its annual report. The 'Model Report'[66] template developed by DTF to assist Victorian government departments and public sector entities with the planning and preparation of disclosures in its annual reports includes a section on DataVic compliance to support this process.

Suggested items include:

- outcomes from the previous year's program of works targets;
- a list of categories of datasets that have been made available;
- suggestions for datasets received and the outcome of the suggestions;
- feedback for datasets received and the outcome of the feedback;
- any known benefits of making datasets available achieved to date; and
- strategies for the coming year, including a list of proposed categories of datasets to be made available.

On request the technical support team can supply statistics to support this process.

Example from the Model Report:

Compliance with the DataVic Access Policy

The Victorian Government's DataVic Access Policy enables the sharing of Government data at no, or minimal, cost to users. The policy intent is to support research and education, promote innovation, support improvements in productivity and stimulate growth in the Victorian economy as well as enhance sharing of, and access to, information-rich resources to support evidence-based decision making in the public sector.

Government data is progressively published in a machine-readable format on www.data.vic.gov.au, to minimise access costs and maximise use and reuse.

[Department/agency]'s release of government data in accordance with the policy is proceeding at a steady rate, with (number of datasets) of [Department/agency]'s datasets and data tools now available on the Victorian Government Data Directory. As part of the Department's (Information, Communications or Technology) Strategy, (number of datasets) [Department/agency] datasets are to be published for the calendar year ending 31 December 2015; as at the end of June 2015, (number of datasets) datasets have been published.

---

[65] www.data.vic.gov.au/data/stats

[66] www.dtf.vic.gov.au/Government-Financial-Management/Financial-reporting-policy/Model-report

The release of these datasets has made a valuable contribution to the policy. In addition, [Department/agency] has made the (name high value datasets) available online in machine readable format.

Consistent with the DataVic Access Policy issued by the Victorian Government in 2012, the financial statements, performance statements and tables included in this Annual Report will be available at www.data.vic.gov.au in machine readable format.

10.6 Risk management

Risks associated with the Policy must be considered and managed as part of an agency's risk management framework.

Failure to properly manage intellectual property risks was identified in the *Victorian Auditor-General's Report Managing Intellectual Property in Government Agencies* (2003)[67] as posing a number of risks to the public sector. For a list of the potential risks identified by the Victorian Auditor-General, agencies may refer to the *Whole of Victorian Government Intellectual Property Intent and Principles*.

It is recommended that agencies consider processes for managing the risk of infringing third party intellectual property rights, such as record keeping management of any third party copyright arrangements, or licensing terms (see Chapter 6: Licensing datasets).

---

[67] www.audit.vic.gov.au/publications/2005/20050720-Managing-Intellectual-Propertyl.pdf

Glossary

Application Programming Interface (API): A specification intended to be used as an interface by software components in order to communicate with each other.

Confidentialised Unit Record Files (CURF): A record in which name and address information, are controlled and limited, or data are deleted or modified to disable identification of individuals or businesses.

Custodian: Individual responsible for managing the lifecycle of a dataset.

Data: Datasets and databases stored in formats including hardcopy, electronic (digital), audio, video, image, graphical, cartographic, physical sample, textual, geospatial or numerical form.

Data quality statement: A presentation of information about the quality of a statistical collection or product.

Data record: A collection of data items related in some fashion and usually connecting.

Data tool: A tool included with a dataset which can load and/or manipulate data.

Dataset: Data that is machine-readable, reusable and open format. In this Policy, the term also applies to data made available in the form of an API, web services and existing data tools.

De-identify: The process by which data is stripped of information which would allow the identification of the source of the data.

Metadata: Listed information that describes an information resource, or helps provide access to an information resource.

Net benefit: having an overall positive impact on the relevant community. A net benefit takes into account the costs and benefits related to:

- Public health and safety;
- Social and community impact;
- Environmental impact;
- Competition; and
- Economic impact.

Open format: An open format is a specification for storing and manipulating content that is usually maintained by a standards organisation.

Unit record: A collection of data elements for a given object. Also a row in a database.

Uniform Resource Identifier (URI): A string of characters used to identify a name or a resource.

Victorian Government data: Datasets and databases owned and held by the Victorian Government and stored in formats including hardcopy, electronic (digital), audio, video, image, graphical, cartographic, physical sample, textual, geospatial or numerical form.

Web service: A method of communication between two electronic devices over the web.

Appendix 1:   Key actions

These Guidelines specify a set of key actions that will help agencies comply with the Policy and are identified by text boxes at the start of relevant chapters. The list of key actions are:

1. The Government's default position is that departments and public bodies datasets must be made available to the public.
2. Datasets must be made available unless access is restricted for reasons of privacy, public safety, security, law enforcement, public health and compliance with the law.
3. Datasets must be released in a machine-readable, reusable and open format.
4. Personal, health and/or confidential information must be de-identified and aggregated.
5. Creative Commons Attribution (CC BY) is the default licence for datasets released under the Policy.
6. Metadata must be created for datasets released under the Policy.
7. All datasets made available under the Policy are to be linked to the data directory, www.data.vic.gov.au.
8. Agencies must make a determination about how often published datasets must be updated.
9. Agencies must consider the Policy when developing and procuring datasets and databases.
10. Datasets will not be commercialised unless an agency has a statutory function to do so, or Ministerial approval is granted.
11. The agency head has overall accountability for implementing the Policy within their agency.
12. Each dataset made available must have an assigned custodian to ensure the dataset is managed through its lifecycle.
13. The progress of agencies compliance with the Policy will be reported to the responsible Minister and to Cabinet.

Appendix 2:   Data quality statement

This attachment sets out an unpopulated data quality statement as discussed in section 5.3.1 How do I create a data quality statement? and illustrates the categories contained within the ABS data quality statements, that the ABS data tool will populate.[68]

Data source	
Institutional environment:	Data collector(s):
	Collection authority:
	Data compiler(s):
	Additional information:
Relevance:	Data topic:
	Level of geography:
	Key data items:
	Numerator/denominator source:
	Additional information:
Timeliness:	Data collected:
	Data available:
	Reference period:
	Additional information:
Accuracy:	Method of collection:
	Data adjustments:
	Sample size:
	Collection size:
	Standard errors:
	Under counts:
	Over counts:
	Additional information:
Coherence:	Consistency over time:
	Consistency of jurisdictions:
	Numerator/denominator:
	There is a consistent time series for this data:
	There is not a consistent time series for this data:
	Additional information:
Interpretability:	Context:
	Other supporting information:
	Additional information:
Accessibility:	Unpublished data can be requested:
	No further data is available:
	Contact details:
	Additional information:

---

[68] www.abs.gov.au/AUSSTATS/abs@.nsf/Latestproducts/1520.0Main%20Features2May%202009?opendocument
&tabname=Summary&prodno=1520.0&issue=May%202009&num=&view=

Appendix 3: Related legislation

This list is not exhaustive. It is recommended that agencies take care to consider legislation relevant to their portfolio.

Audit Act 1994 (Vic)[69]

The Audit Act establishes the legislative framework governing the ongoing role and functions of the Auditor-General.

Copyright Act 1968 (Cwlth)[70]

The Copyright Act defines the legally enforceable rights of creators of creative and artistic works under Australian law.

Electronic Transactions (Victoria) Act 2000 (Vic)[71]

The Electronic Transactions (Victoria) Act gives electronic transactions and communications the same status, for legal purposes, as paper-based transactions and documents.

Evidence Act 2008 (Vic)[72]

The Evidence Act sets out the State rules of evidence. Generally speaking, the Act applies to proceedings in State courts and before other persons or bodies required to apply the laws of evidence.

Financial Management Act 1994 (Vic)[73]

The Financial Management Act lists agencies requirements in keeping proper financial accounts, risk management, audit requirements, financial reporting, annual reporting to Parliament and responding to Ministerial requests for information.

Freedom of Information Act 1982 (Vic)[74]

The Freedom of Information Act lists agencies requirements in giving the public access to requested information held by public sector entities.

Health Records Act 2001 (Vic)[75]

The Health Records Act is a framework designed to protect the privacy of individuals' health information. It also regulates the collection and handling of health information.

---

[69] www.legislation.vic.gov.au/Domino/Web_Notes/LDMS/LTObject_Store/LTObjSt6.nsf/DDE300B846EED9C7CA257616000A3571/802581D0C934FE4DCA25794A0073AB71/$FILE/94-2aa052%20authorised.pdf

[70] www.comlaw.gov.au/Details/C2010C00476

[71] www.legislation.vic.gov.au/domino/Web_Notes/LDMS/LTObject_Store/LTObjSt6.nsf/DDE300B846EED9C7CA257616000A3571/7A96404D4512CDA2CA25795800046CB8/$FILE/00-20aa003%20authorised.pdf

[72] www.legislation.vic.gov.au/domino/Web_Notes/LDMS/LTObject_Store/LTObjSt6.nsf/DDE300B846EED9C7CA257616000A3571/9E0D437610121663CA2578B7001601BA/$FILE/08-47aa011%20authorised.pdf

[73] www.legislation.vic.gov.au/Domino/Web_Notes/LDMS/LTObject_Store/LTObjSt6.nsf/DDE300B846EED9C7CA257616000A3571/995A145D169B4B9CCA2579F30005072B/$FILE/94-18aa062%20authorised.pdf

[74] www.legislation.vic.gov.au/Domino/Web_Notes/LDMS/LTObject_Store/LTObjSt7.nsf/DDE300B846EED9C7CA257616000A3571/5D20A33EFCCA5BD3CA257AC400145537/$FILE/82-9859aa080%20authorised.pdf

[75] www.legislation.vic.gov.au/Domino/Web_Notes/LDMS/LTObject_Store/LTObjSt7.nsf/DDE300B846EED9C7CA257616000A3571/015E6F0294580950CA257AC3007BCA5F/$FILE/01-2aa024%20authorised.pdf

Privacy and Data Protection Act 2014 (Vic)[76]

The Privacy and Data Protection Act provides for the responsible collection and handling of personal information in the Victorian public sector, and for the establishment of a protective data security regime.

Public Administration Act 2004 (Vic)[77]

The Public Administration Act reflects the State Government's belief that the fundamental role of the public service is to serve the public interest and defines a comprehensive approach to supporting good public administration.

Public Records Act 1973 (Vic)[78]

The Public Records Act details the procedures to manage the holding, disposal and transfer of public records.

---

[76] www.legislation.vic.gov.au/Domino/Web_Notes/LDMS/LTObject_Store/LTObjSt8.nsf/DDE300B846EED9C7CA2 57616000A3571/B224F6F00FF339CDCA257D570023421D/$FILE/14-60aa001%20authorised.pdf

[77] www.legislation.vic.gov.au/Domino/Web_Notes/LDMS/LTObject_Store/LTObjSt7.nsf/DDE300B846EED9C7CA2 57616000A3571/9995806A403028CCCA257ABD0076497C/$FILE/04-108aa032%20authorised.pdf

[78] www.legislation.vic.gov.au/Domino/Web_Notes/LDMS/LTObject_Store/LTObjSt6.nsf/DDE300B846EED9C7CA2 57616000A3571/D2B93E4380A56273CA257A0700083577/$FILE/73-8418aa036%20authorised.pdf

Appendix 4:   Related policies and standards

Datasets created or managed by public sector agencies are public records and should be managed accordingly. Related policies and standards to support this include:

Victorian Government Intellectual Property Intent and Principles[79]

The Whole of Victorian Government Intellectual Property Policy Intent and Principles is the State's framework for the ownership and management of its IP, as well as for the use by the State of IP belonging to other parties.

Recordkeeping standards[80]

The Public Record Office Victoria (PROV) sets recordkeeping standards for the efficient management of public records under section 12 of the *Public Records Act 1973*. The standards apply to all records created by the Victorian Government and detail requirements for the creation, maintenance and use of these records.

Information management standards[81]

The whole of Victorian Government Information management standards describe the way in which an organisation should plan, identify, create, receive, collect, organise, govern, secure, use, control, disseminate, exchange, maintain, preserve and dispose of its information.

Victorian Protective Data Security Standards (VPDSS)[82]

The Data and Privacy Protection Commissioner is responsible for the whole of Victorian Government protective data security framework which includes protective data security standards, protective data security plans prepared by public sector bodies to implement the standards, and specific law enforcement data security standards.

Victorian Government Website Management Framework[83]

The Website Management Framework provides standards and guidelines on website content, services and processes, including content ownership approval and review. The 15 inner government agencies must comply with the Content Ownership Approval and Review standard which requires that all website content has an owner. It is recommended that agencies seek to align the process of assigning dataset custodianship and website content ownership where possible.

Cost Recovery Guidelines[84]

These Guidelines clarify the Government's policy principles underpinning cost recovery arrangements. The guidelines provide a rigorous framework for use by government entities when considering, developing and reviewing user charges and regulatory fees.

---

[79] www.dtf.vic.gov.au/Victorias-Economy/Victorian-Government-intellectual-property-and-data-policies/Intellectual-Property-Policy

[80] www.prov.vic.gov.au/government/standards-and-policy/all-documents/recordkeeping-standards-framework

[81] www.digital.vic.gov.au/resources/information-management/

[82] www.dataprotection.vic.gov.au/

[83] www.digital.vic.gov.au/resources/online-and-mobile/

[84] www.dtf.vic.gov.au/Publications/Victoria-Economy-publications/Cost-recovery-guidelines

Competitive Neutrality Policy[85]

Competitive neutrality ensures that the significant business activities of publicly owned entities compete fairly in the market. It is about transparent cost identification and pricing in a way that removes advantages arising from public ownership. This policy outlines a process for lodgement and investigation of complaints against public entities.

Data integrity manual[86]

The data integrity manual supplements the Standing Directions of the Minister for Finance on Information Management and Information Technology Systems (Standing Direction 3.4.13).

---

[85] www.dtf.vic.gov.au/Publications/Victoria-Economy-publications/Competitive-neutrality-policy

[86] www.dtf.vic.gov.au/Publications/Government-Financial-Management-publications/Financial-Management-Compliance-Framework/Data-integrity-manual

Appendix 5:   Quick reference checklist

This checklist provides guidance on the process involved when making datasets available. Agencies may not have to undertake every step in this checklist in order to publish a dataset.

Colour code reference	Go to next step	Complete step before continuing	Data unsuitable for release

Step 1	Yes	No/Not applicable
Identify a dataset (the default is that a dataset must be made available) - Refer to Chapter 3 and 4 of the Guidelines.		
Is the dataset considered a high value data?	Prioritise dataset for release, subject to other restrictions.	Make dataset available, subject to other restrictions.
Does the dataset contain Personally identifiable information not yet aggregated?	De-identify or aggregate dataset prior to making dataset available.	Make dataset available, subject to other restrictions.
Does the dataset contain information that may adversely affect public safety	Check whether it can be cleansed and the risks mitigated and seek approval for this.	Make dataset available, subject to other restrictions.
Does the dataset have a security classification?	Check whether it can be cleansed and the risks mitigated or if reclassification is possible and seek approval for this.	Make dataset available, subject to other restrictions.
Does the dataset contain information subject to legal proceedings?	Do not make dataset available.	Make dataset available, subject to other restrictions.
Does the dataset contain information that may adversely affect public health	Check whether it can be cleansed and the risks mitigated and seek approval for this.	Make dataset available, subject to other restrictions.
Is the dataset subject to third-party copyright?	Where appropriate obtain permission.	Make dataset available, subject to other restrictions.
Is the dataset part of a contract or agreement?	Check details of contract. Dataset may be eligible for release. Negotiate to have the data made available.	Make dataset available, subject to other restrictions.
Is releasing the dataset in breach of legislative requirements?	Do not make dataset available.	Make dataset available, subject to other restrictions.

Step 2	Yes	No/Not applicable
Prepare the dataset for publication - Refer to Chapter 5 of Guidelines		
Is the dataset in an open and machine-readable format?	Go to next step.	Convert dataset into a more user-friendly format.
Have you prepared a data quality statement?	Go to next step.	Preparing a data quality statement is recommended to qualify rules of dataset use.

Step 3	Yes	No/Not applicable
Select a licence - Refer to Chapter 6 of Guidelines		
Does the State own the copyright in the dataset?	Go to next step.	Obtain copyright.
Can you apply a Creative Commons licence to your dataset?	Go to next step.	Dataset may be subject of more restrictive licence. See steps below.
Can you apply the Creative Commons Attribution (CC BY 4.0) licence? [CC BY 4.0 is the State's preferred licence]	Go to next step.	Apply alternative CC licence or restrictive licence.
Have you applied the relevant licence symbol and notice?	Dataset has met licensing process.	Apply notice and symbol. Refer to http://creativecommons.org to choose a licence.
Is the dataset subject to a restrictive licence?	Seek further legal advice.	Apply appropriate CC licence or go to next step.
Is the dataset subject to a bespoke licence?	Seek further legal advice.	Apply above steps. Datasets must be made available under a flexible licence.

Step 4	Yes	No/Not applicable
Publish the dataset - Refer to Chapter 7 of Guidelines		
Is the dataset hosted on your agency's website?	Go to next step.	Host dataset on agency website before proceeding to next step.

Step 5	Yes	No/Not applicable
List the dataset at data.vic.gov.au - Refer to Chapter 7 of Guidelines		
Have you provided a data record to data.vic.gov.au?	Go to next step.	Provide data record then proceed to next step.
Have you provided the metadata for this data record?	Go to next step.	Provide metadata in relevant fields on data.vic.gov.au. Metadata should comply with Information Management Roles and Responsibilities standard.

Step 6	Yes	No/Not applicable
Manage the dataset - Refer to Chapter 10 of Guidelines		
Has a governance structure been established to oversee accountability arrangements?	Go to next step.	Establish a governance structure to oversee the process for making datasets available
Have your agency's dataset custodians been identified?	Go to next step.	Select dataset custodian to manage making datasets available
Is this dataset included in the asset register?	Ensure dataset is also published on data.vic.gov.au, subject to other restrictions.	Include dataset in asset register
Have risk management principles been applied to assessing datasets being made available?	Make dataset available.	It is recommended that agencies review the release of datasets in light of risk management plans
Do you have a process in place for managing dataset suggestions and feedback?	Dataset meets publishing requirements.	Create this process to ascertain which datasets have high value for the public

Appendix 6:   DataVic Access Policy intent and principles

The Victorian government recognises the benefits associated with mandating a whole of government approach to the availability of Victorian government data for the public good.

The DataVic Access Policy provides direction on the release, licensing and management of Victorian Government data so that it can be used and reused by the community and businesses.

The Victorian government holds, creates and collects a vast amount of data, ranging from demographic and economic to geospatial data.

> **Victorian government data** refers to datasets and databases owned and held by the Victorian government and stored in formats including hardcopy, electronic (digital), audio, video, image, graphical, cartographic, physical sample, textual, geospatial or numerical form.
>
> Victorian government data does not include software.

Not all Victorian government data is suitable for release under the policy. Access to some data will need to be restricted for reasons of privacy, public safety, security and law enforcement, public health and compliance with the law. Only data owned by the State of Victoria or sufficiently licenced to the State of Victoria will be released under this policy.

Expected benefits of the policy

The policy is expected to achieve the following benefits:

- stimulate economic activity and drive innovation and new services to the community and business;
- increase productivity and improve personal and business decision making based on improved access to data;
- improve research outcomes by enabling access to primary data to researchers in a range of disciplines; and
- improve the efficiency and effectiveness of government by encouraging better management practices and use of the data.

Relationship to other policies

This policy operates in conjunction with the whole of Victorian government Intellectual Property Policy, and, like the Intellectual Property Policy, applies to all departments and public bodies (agencies).

DataVic Access Policy intent

To enable public access to government data to support research and education, promote innovation, support improvements in productivity and stimulate growth in the Victorian economy.

To enhance sharing of, and access to, information-rich resources to support evidence-based decision making in the public sector.

DataVic Access Policy principles

Principle 1

Government data will be made available unless access is restricted for reasons of privacy, public safety, security and law enforcement, public health, and compliance with the law.

Principle 2

Government data will be made available under flexible licences.

Principle 3

With limited exceptions, government data will be made available at no or minimal cost.[87]

Principle 4

Government data will be easy to find (discoverable) and accessible in formats that promote its reuse.

Principle 5

Government will follow standards and guidelines relating to release of data and agency accountability for that release.

Governing framework for this policy

The DataVic Access Policy is administered through the Department of Treasury and Finance. The Department of Business and Innovation maintains the primary gateway to government data – www.data.vic.gov.au.

The DataVic Access Policy Intent and Principles will be supported by mandatory standards and guidelines approved by the Assistant Treasurer. Standards and guidelines will cover a range of topics including:

- metadata standards
- copyright licence tools (in conjunction with the Intellectual Property Policy)
- the assessment of data quality and appropriate data quality statements for datasets
- other terms and conditions of use, including disclaimers and limitation on use of data
- data governance and roles and responsibilities

Departmental Secretaries and agency heads are responsible for authorising the release of datasets maintained by their agency, subject to this policy.

---

[87] An agency may commercialise, or apply the Cost Recovery Guidelines to, government data if:

a)   it has an explicit statutory function to do so; or

b)   it has been explicitly authorised to do so by the relevant Minister after consulting with the Treasurer, because of a clear net benefit to the Victorian community.

DataVic Access Policy Guidelines
for the Victorian public sector (version 2, August 2015)

VICTORIA
State
Government

**Department of
Treasury & Finance**