Open Data Process Guide

Prepared by: Office of the Chief Information Officer
Version: 1.0
Date: November 2014

TABLE OF CONTENTS

1   INTRODUCTION

In September 2013, the Premier announced a Declaration of Open Data to make government data available for use by business and the community.
Download the Declaration of Open Data (PDF)

This process guide has been designed to assist agencies meet their obligations to release open data so that it is discoverable on Data.SA, the South Australian Government Data Directory, available at www.data.sa.gov.au. The guide recommends a series of steps to identify, assess, classify and mark datasets, plan the approach for release, seek approvals and ultimately release the datasets.

Agencies may customise the process to suit their requirements, resources and data management processes.

2   WHAT IS OPEN DATA

The South Australian Government is the custodian of a diverse range of data. Agencies create, collect, process, preserve, maintain and disseminate all sorts of data in the course of doing business. Open data is when certain data is made freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control. Open data that is released must be:

- available online so that it can be easily shared and traced to the source
- free where ever possible
- published using agreed open standards
- openly licensed for commercial and other re-use
- reliable and well managed.

A dataset is a collection of related or similar data. A dataset may include information collected at the same time, through the same process or on the same variables.

3   OPEN DATA GOVERNANCE

Agency Chief Executives are accountable for all material made publicly accessible from within their agencies. To ensure the government meets its obligations in making data open by default, roles and responsibilities within the organisation need to be established.

An Open Data Advocate has delegated responsibility by the Chief Executive to oversee open data activities and have oversight of data released openly by the agency.

It is recommended that an Open Data Coordinator be nominated to assist the Open Data Advocate coordinate open data activities, particularly where separate operating areas report to a lead agency. Large agencies may also consider establishing an open data working group that includes representation from business units and key stakeholders.

A Data Authority must be identified for each dataset. The Data Authority has the administrative or legislative responsibility for the datasets. In most cases they are a Director or Executive of a specific government business function to which the data is integral.

It is a requirement that the Open Data Advocate for the agency and the Data Authority of the dataset both approve a dataset for open release and ensure the data is sustainably managed, if it is a dataset that needs to be refreshed in the future. The following table (Table 1 – Open Data Roles and Responsibilities) is a recommendation on the roles an agency needs to establish. However an agency may alter this to suit their requirements depending on their size and resources. In some cases roles may be carried out by the same person e.g. the Subject Matter Expert may also be responsible for publishing the dataset on Data.SA. Care should be taken with regard to role segregation where an individual is responsible for multiple roles to ensure adequate representation of other parts of the business and independent oversight.

Table 1 – Open Data Roles and Responsibilities

Role	Characteristics	Suggested Responsibilities
Open Data Advocate	• member of the Executive • has an active interest in or their role has a natural fit with applying openness and transparency principles.	• oversees participation in open data activities including oversight of open data process • oversight and coordination of Business Managers and stakeholders • signs off open data plans, submissions and Open Data Process Worksheet • manages any associated risks of public data release • Open Data Advocate may also perform the activities of the Open Data Coordinator.
Open Data Coordinator	• role may be required for large organisations • understands the open data process and can coordinate activities.	• manages the process and participation in activities • technical and process support to Stakeholders and Data Authorities • assists with the approval through coordination of a peer review • collates Open Data Process Worksheets and maintains an open data inventory.
Data Authority	• has administrative or legislative responsibility for a specific government business function to which the data is integral e.g. A Director or Executive • has responsibility for the existence, protection and use of the dataset	• lead and champion the data prioritisation and release program for their Business Unit or Branch • consults and reports to Agency Open Data Advocate, Open Data Coordinator and Data Manager on open data candidates or activities • must approve the release of the data and provide due diligence oversight of the open data process for the dataset they are responsible for • each dataset may have a different Data Authority • the Data Authority may delegate tasks within the open data process to the Subject Matter Expert, Data Manager and/or Data Publisher.
Subject Matter Expert	• has technical/detailed knowledge of the data.	• the Data Authority must authorise the provision of the data, but may delegate tasks within the open data process or operational functions to the Subject Matter Expert • provides information and/or data to the Data.SA Publisher. This may include completing the Data.SA Publishing Content Summary and forwarding to the Data.SA Publisher • undertakes regular review activities and ensures open data is updated • responds to enquiries relating to the data once published.
Data Manager	• has technical knowledge of databases and information systems. • an Agency may have more than one Data Manager. This is likely to depend on the number of information systems where data is managed.	• enable and/or implement data release. They have primary accountability for the day-to-day management of the data, as delegated by the Data Authority • provide advice to Subject Matter Experts, Data Authorities and Data Coordinators on development the approach for the release of data and ongoing management • data management can encompass a diverse range of support roles and processes including data capture, cleansing, security, maintenance, backup and distribution • a Data Manager is assigned to each dataset • may also be responsible for supplying datasets to a Data Distributor.
Data.SA Publisher	• has operational knowledge of online publication. • has access to publish content for the agency on Data.SA	• publish data for reuse on appropriate websites, along with associated metadata and reuse license statement(s) • is responsible for publishing and maintaining datasets on Data.SA for the Agency. A Data.SA Publishing Content Summary must be provided with all datasets to the Data.SA Publisher • the Data Authority must authorise the provision of the data • an Agency may have several Data.SA Publishers depending on their business structure and data needs.
Data Distributor	• provides a data distribution service that collates and distributes data • may provide single point of truth benefits or scalable information management solutions	• a data distributor may be delegated the responsibility for distributing the information through a distribution channel or service • the Data Authority must authorise the provision of the data but may delegate the task of distributing the data to a Data Distributer • a Data Distributor may also take on the role as Data.SA Publisher, upon agreement. Other processes may be required when using a data distributor.
Legal Advisor	• legislative interpretation and advice.	• provides legal advice in relation to specific legal restrictions of data candidates • provides legal advice in regard to intellectual property rights that subsist in the third party material.
Information Technology Security Advisor	• agency Information Technology Security Advisor (ITSA)	• provides advice and guidance on agency specific information classification procedures • provides advice on the South Australian Government Information Security Management Framework (ISMF).

4   OPEN DATA PROCESS WORKSHEET

The Office of the CIO has developed an *Open Data Process Worksheet* to assist agencies work through the stages of identifying, classifying, marking, publishing and seeking approvals. The *Open Data Process Worksheet* also helps agencies meet other requirements:

Record Decisions

Agencies are required to keep a record of all decisions made through the open data process. It is recommended that decisions are recorded at each stage of the process to ensure stakeholders are informed.

An *Open Data Process Worksheet* should be completed for each dataset identified and kept as an official record. Datasets not approved for open data should also be documented as an official record.

Data Inventory

Agencies are encouraged to keep a dataset inventory of all datasets they consider through the open data process for accountability and reporting purposes. The *Information Security Management Framework (ISMF) Guideline 7 Asset Management (PDF)* requires agencies to maintain an Information Asset inventory. If your agency has an Information Asset inventory in place, you may consider expanding this to include key decisions for open data.

Refer to Appendix D for the *Open Data Process Worksheet*.

5   WHERE WILL DATA BE DISCOVERABLE

Datasets will be accessible and easily discoverable on Data.SA the South Australian Government Data Directory.

Data.SA helps citizens and industry locate openly licensed data so they can transform that data, creating new ideas and applications. Data.SA provides a central location to access data that can improve investment decisions, research and well-planned service delivery.

The Office of the CIO administers Data.SA and can be contacted through datasa@sa.gov.au

6   OVERVIEW OF THE OPEN DATA PROCESS

Identify
Identify Data
Identify the Data Authority
Third Party Rights
Value of the Data

Classify
Information Security Classification,
Responsible Information Sharing
Privacy of Personally Identifiable Information,
Creative Commons Licence

Approach
Data Format, Frequency of Release,
Method of Delivery, Metadata,
Funding, Cost Recovery

Approve
Identify, Classify and Approach decisions
Data Authority and Data Advocate Approval

Prepare and Publish
Prepare data and interpretative Metadata,
Data.SA Publishing Content Summary
Publish

Maintain
Data Maturity,
Data Maintenance and Amendments,
Customer Feedback, Apps & Ideas

7 IDENTIFY – STEP 1 OPEN DATA PROCESS

With the Declaration of Open Data, agencies need to consider what datasets are available for release. Agencies create, collect, process, preserve, maintain and disseminate all sorts of data in the course of doing business.

There are several activities that need to occur through the identify step of the open data process. Including identify data, confirm who is the Data Authority; establishing if there are third party rights; and considering the value and quality of the data identified.

7.1 IDENTIFY DATA

Agency data may be held in databases, reports, photographs, journals, videos, visualisations or exist in 3D, geospatial and numerical form. There are a number of ways agencies can identify existing and potential datasets for release.

7.1.1 Public Facing Information

Most agencies already publicly release information e.g. on their websites. This type of data can be considered first as minimal effort may be required to classify, mark and release the data in an open format and open licence.

Agencies should review and consider existing systems and public facing mediums:

- website content
- publications such as annual reports
- data published in tables
- public service announcements.

7.1.2 Inventory Method

Agencies should review their Information Assets for open data candidates; agencies are required under the ISMF to hold an inventory of Information Assets. Agencies may consider expanding this inventory to record open data process decisions. Agencies should also review other sources for datasets, which may include:

- KPI's and performance statistics
- research or administrative data records
- reports or data extracts generated from information systems
- reports or data used to make administrative or policy decisions
- data from commissioned research or consultation
- data supplied to a third party, including Commonwealth reporting
- databases or data catalogues.

7.1.3 Proactive Disclosure

Freedom of Information (FOI) requests, regular commonwealth reported data and across jurisdictional reported data should be considered as open data candidates. Resource efficiencies and significant savings can be achieved through the proactive disclosure of reported or requested data. Improved ability to make decisions can also be achieved through early access of data.

Agencies are encouraged to complete an annual review for dataset FOI requests to identify frequent requests for data. Significant resource efficiencies and expenditure savings could be realised through the proactive release of regular requested FOI data. Advice from your Agency FOI Officer must be received before considering this data for possible release.

Information about FOI is also available through the [Archives SA website](#).

7.1.4 Business Planning Method

Data can be identified through a business planning method where the agency identifies a specific goal or outcome to achieve through their strategic planning, business planning, or information communication and technology (ICT) planning processes. Agencies should consider the following through their planning process:

- business projects that are enabled by ICT should consider open data principles and plans. Business requirements documentation should identify data collection, consider the approach for release of data and automated release opportunities
- identify opportunities where third parties could add value to services if data was available for re-use e.g. could someone create an application from your data
- consider planned legislative or policy changes for open data opportunities
- review strategic programs to align activities with open data projects
- consider a problem you have that your agency does not have resources, funding and expertise to solve.

7.1.5 Community Demand

Agencies are encouraged to engage and collaborate with citizens and industry as per the [Better Together Principles for Engagement](). Channels such as [YourSAy]() and [Data.SA]() enable the community to request data. Demand from the community should be given a high priority if the release of data could in some way provide public benefit either directly or when value added by a commercial entity. Data that cannot be released immediately should be considered in future plans.

7.2 IDENTIFY THE DATA AUTHORITY

Once a dataset is identified, the Data Authority must be assigned. The Data Authority has administrative or legislative responsibility for a specific government business function to which the data is integral. The Data Authority has responsibility for the existence, protection, and use of this dataset. In most cases, the Data Authority will be a Director or Executive.

If data has been sourced from another State or Commonwealth Government agency, direction and approval to re-classify, re-mark and openly release the data must be received in writing from their Data Authority or information owner. Agencies are not permitted to alter the classification or markings of information they do not own.

In many cases the Data Authority will already be known, if you are unsure consider the following sources to identify the Data Authority:

Legislative Delegations	Review legislative delegations approved for the administration of the legislation that empowers the SA Government to collect the data.
Public Sector Delegations	Delegated responsibility for an operational function as per the Public Sector Delegations.
Policy and Accountability	Public sector policies or accountability mechanisms may define who is responsible for the data.
Agreements or contracts	Review any agreements or contracts with third parties during the collection of data to determine what authority exits for the publishing of data e.g. memorandum of understanding, terms and conditions of collecting data.

7.3 IDENTIFY THIRD PARTY RIGHTS

Data should be checked for third party rights. Third party rights are derived from:

- copyright of intellectual property
- terms and conditions of use specified when collecting data.

7.3.1 Intellectual Property and Copyright

Agencies must not make datasets available that contain Intellectual Property (IP) rights of a third party where the SA Government does not have permission from the copyright owner or term of use licence is not provided. This includes circumstances where an agency co-produces a dataset with another government agency (state or federal level).

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 government 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.

Copyright generally exists for at least 50 years. Accordingly, data, records or images older than this should be considered for release.

View The Government of South Australia Intellectual Property Policy (PDF). Agency specific IP policies may also apply.

It is recommended that agencies consider the following actions when using data from third parties:

- identify and document the input sources used in the creation of the work
- identify and review publishing restrictions that apply to the additional inputs
- obtain an agreement or terms of use license for using the data
- the source of original data is Attributed in works created
- any use, re-use or manipulation of external data should be documented
- when in doubt seek permission from the third party or seek legal advice.

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.

Existing contracts and agreements

Agencies may have entered into contracts or agreements with third parties who have sourced the data 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 for restrictions before releasing datasets.

Legal rights transferred

Where the legal rights have been transferred through the terms of a contract, the assignment of copyright must be in writing. This is essential as it removes ownership of copyright and enables agencies to own the data and openly licence and release it where appropriate. If you are unable to determine if rights have been transferred, particularly in the case of historical data, seek to obtain agreement of terms of use from the third party.

Terms of use

In some cases the third party is the copyright owner but provides a 'terms of use licence'. These licenses may have conditions on how the information can be used.

A term of use license may exist in the contract of agreement for use or issued under a commonly accepted open licence that is published with datasets or information. For example:

- Attribution CC-BY licence
- Open Data Commons Open Data License (ODbl).

Generally a Terms-of-Use licence will require attribution for the work to the third party (IP holder). If you reuse this work ensure this attribution is published with the work or derivative of the work.

How to attribute data from a third person

If you need to attribute works from a third party, you may need to check the attribution requirements they have requested. Standard attribution required:

- organisation name
- title of works
- URL link to published location
- date sourced.

Attribution for third party material that is reused must be published with the work.

Unknown third party rights

If the rights to reuse data provided by a third party are not clear you must obtain agreement of terms of use.

The preferred method to seek agreement is a standard letter with the following elements:

- details of the data in question
- request an agreement for terms of use under an CC-BY licence
- detail why you will release the data (e.g. Public Service)
- provide a reasonable period for the person to respond and a due date
- disclose that a non-response will be taken as acceptance of CC-BY licence
- provide appropriate means to respond and/or discuss.

Creation of new datasets

New intellectual property may be created if:

- third party data are manipulated
- third party data are merged with other data or additional data is collected
- different data structures are applied to the data.

Agencies should deal with third party intellectual property in a manner that avoids infringing the IP rights of others and complies with the law. Agencies may only consider this new data for open data if you have clear rights to create derivative works from the data and authority to use.

New contracts and agreements to acquire data rights

Agencies are encouraged to consider open data principles when establishing contracts and agreements for research or consultation reports.

Data is often required and publicly funded in order to develop these works. This Data is often considering a background IP. Where IP inputs (Background IP and/or Third Party IP) will be brought to a procurement agreement, the agreement must address South Australian Governments use of those inputs. The default position is that:

- the Background IP of each party remains the property of that party, along with any improvements to such Background IP
- The contractor must grant the agency a licence over the Background IP and any Third Party IP to the extent needed to allow the agency to enjoy the full benefit of the agreement.

While the licence over Project IP must be restricted in scope to the purposes of the procurement, agencies should carefully consider what these purposes are and what they may be in the future. Depending on the circumstances, the licence may need to allow:

- other agencies (not involved in the specific procurement) to use the IP
- the IP to be used and adapted for subsequent projects
- the IP to be adapted later by another service provider engaged by government
- the IP or background IP to be released publicly.

Grants for research and data collection should encourage the open release of the background IP data.

The commercial impacts of openly releasing purchased data from commercial entities should also be considered. Data that is marked as Sensitive: Commercial, is information, whose compromise could affect the competitive process and provide the opportunity for unfair advantage including:

- information concerning the trade secrets of any person
- information (other than trade secrets) that has a commercial value to any person
- likely to harm the business's commercial advantage in the marketplace.

Agencies are encouraged to see legal advice in these circumstances.

For more information on how to deal with copyright when establishing new contracts and agreements [View the Government of South Australia Intellectual Property Policy](#) (under review).

7.3.2 Terms and conditions of use specified when collecting data.

Data collected through applications and information systems may have terms and conditions which let the user know how the data they provide will be used. The restrictions detailed in the 'Use of Information Disclosure' may prevent you from releasing the data.

Agencies should review existing terms and conditions and seek Crown Solicitor Advice when amending or creating new terms and conditions.

This is particularly important with the increase to digital services offered to the public and businesses. Upfront disclosure will ensure users of services are aware of how data will be used and a statement, which indicates the highest standards of privacy, security, confidentiality and integrity with respect with the data we collect.

Establishing new collection Terms and Conditions

If establishing a new data collection service you should ensure that the Information Use Disclosure in your terms and conditions, detail the following:

- your agency and contact information is provided
- how he or she is able to gain access to the information
- the purpose for which information is collected
- any laws that require particular information to be collected
- any information collected that will be protected (e.g. personally identifiable information or commercially sensitive); and
- how any information collected may be disclosed to other parties (e.g. under a licence for reuse with exclusion to personally identifiable information).

A privacy statement and disclaimer are also required when creating any terms and conditions. It is recommended that you seek Crown Solicitor Advice before finalising any terms and conditions.

7.4   DATA VALUE

Agencies must consider the value of data if released, as this may help determine the priority or approach for releasing the data or help determine if it is able to be released. Valuable data when re-used contributes to economic, social, cultural and environmental growth, illustrates government's performance and contributes to greater government efficiencies through improved information sharing.

There are 3 types of value that must be considered before releasing data:

- value to the Government – what is it worth to us? How could the release of data create value for the government? What protection and due care is required to prevent damage to government administration or reputation if sensitive information was released inappropriately or accidentally?
- value to criminals and perpetrators of terrorism – could the data be used by criminals for illegal profit or intelligence for harm?
- value to the consumer - Data can be useful in unexpected and unpredictable ways, which means the potential usefulness of any data set should not be pre-judged.

It is also important to consider the aggregate value of information; datasets (small or large) may be independently harmless, however when aggregated with additional datasets from the same agency, or datasets from other agencies (state or commonwealth) as a whole the information may have an entirely different value.

The following characteristics should be considered when considering the value of a dataset for release:

7.4.1   Publicly Available Information or Provides a Public Service

Publicly released information in tables, spreadsheets or reports on websites should be released as open data. Agencies should consider this data as a priority for release as the public need is already identified and minimal effort may be required to classify, mark and transform the data and release it in an open format and open licence.

Information that provides a public service is valuable and should be given a high priority to release in an automated and open format (e.g. Rich Site Summary (RSS) feed) to enable a more informed community. This will disseminate your information to a wider audience with the potential for new information services to be delivered by the digital community.

7.4.2 Supports Strategic Goals of Government

Open data that could support or aligns with a strategic goal should be considered as a priority as this has value to the Government e.g.

- [South Australia's Strategic Plan](#)
- [A Stronger South Australia](#)
- [Modern Public Service](#)
- [SA Connected](#)
- Aligns with Agency strategic plans.

7.4.3 Stimulate Development, Industry or Services

Data that has the potential to spur economic growth, innovate or create new products or services is highly valuable. South Australian industry has the skills to transform data into new digital services. The entrepreneurial community is agile and can take advantage of emerging technologies and trends and as such are responsive to the needs of South Australians. Often these services enabled by data are not the core business of government; however the data is valuable asset that can support the new services.

7.4.4 Empower Business or Citizens to Make Informed Decisions

If data can empower business or citizens to make more informed decisions it should be perceived as high value. Data on housing, services, safety, crime, transport, travel, sports, recreation and events could assist citizens make more informed decisions about their lives.

Data used to make industry, infrastructure or planning related policy is also likely to help business make decisions. Improved business decisions can lead to economic growth in our state.

7.4.5 Research or Industry Efficiencies

Statistical information that could assist researchers should be released as a high priority. Examples of research study fields include environment, agriculture, humanities, social sciences and health.

By releasing research data openly we allow others to build on data to further advance analysis and research without the need to fund duplicate data collection. Open release of research data can facilitate debate and communication of ideas to a broader audience.

Open data can reduce industry red tape and provide efficiencies through direct access to data collected from the source and reducing the need for agencies to collect or estimate value information required to make business decisions.

7.4.6 Comparable Across-Jurisdictional Open Data

Data that can be compared with other jurisdictions is beneficial for trend analysis and can improve business decision making. The business viability of a data enabled application (app) is also increased if the market for the app is national or mutli-jurisdictional. Agencies can check what comparable data has been released by visiting the other Australian Government Open Data Portals:

- [Australian Government data portal data.gov.au](#)
- [Queensland Government data portal data.qld.gov.au](#)
- [Queensland Government spatial data globe](#)
- [New South Wales Government data portal](#)
- [New South Wales Government spatial data portal](#)

- [Australian Capital Territory Government data portal](#)
- [Victorian Government data portal](#)

7.4.7 Government Sharing Efficiencies and Policy Development

Data that is often shared between government agencies or reported to other jurisdictions can create efficiencies if the data is released openly. Often data is reported to several different parties, at different times and in different formats. Openly releasing this data allows the users to access the required information when they require it and transforms the data into a format that suits their needs. Value to the Government can be achieved through reduced resources to process Freedom of information request. Open data can lead to improved public sector productivity or lower service delivery cost.

Open data can lead to improved data to decision policy making across the Government through improved access and currency of data required to make informed decisions.

Providing data openly to support the rationale for policy particularly through community engagement processes can improve the development and implementation of policies.

IDENTIFY SUMMARY

Identify Data		
☐	Public Facing Mediums	Review existing public facing data via agency websites, notifications & publications.
☐	Inventories	Review information assets and data inventories for possible datasets.
☐	Proactive Disclosure	Consider releasing data from frequent Freedom of Information Requests or regular across jurisdictional reporting.
☐	Business Planning	Identify opportunities for obtaining data in the future and obtaining intellectual property rights.
☐	Community Demand	Engage with the community and respond to requests for data.
Do Third Party Rights Exist?		
☐	Intellectual Property	Who has intellectual property rights of the data?.
☐	Copyright	Does a third party hold copyrights of the data?
☐	*If Yes, seek to Obtain Rights or else dataset is not eligible for release*	
Are there Terms and Conditions?		
☐	Use of Information Disclosure	Are there terms and conditions which prevent data collected from being released.
Value		
☐	Value	Consider the value of the data to the Government, Criminals and Public.
☐	Public Value	Consider whether a potential dataset will meet strategic goals, improve services, stimulate and empower business and industry, create efficiencies and improve public policy and enable more research or across jurisdictional comparisons.
Commence the Open Data Process Worksheet		
☐	Start a new worksheet for each dataset.	Record all decisions made about the dataset on the worksheet.
☐	Data Authority	Identify the Data Authority (usually a Director or Executive) who will ultimately approve the dataset release. Datasets may have a different data authority.

8 CLASSIFY - STEP 2 OPEN DATA PROCESS

Data can only be considered for release under open data after being assessed in terms of information confidentiality, integrity and availability classifications and applicable markings have been applied. Where data meets the requirements of open data, a licence classification using the AusGOAL licence framework is also required to ensure users know how they can use the data. The below discusses these requirements:

8.1 INFORMATION SECURITY CLASSIFICATION

Agencies must assess data to determine if it could be released publicly, based on their information classification procedures as required by the South Australian Government Information Security Management Framework (ISMF). Visit the Department of Premier and Cabinet website to the view the ISMF policies, standards and guidelines. The Information Privacy Principals Instruction (IPPI) also guides information privacy in the Government of South Australia. View the IPPI (PDF 241Kb).

Data can only be marked as Public if the release of this information does not cause any damage to the state, the government, an agency, commercial entities or members of the public.

Only data that has been assessed and marked as Public can be released as open data. Data that has a Protective Marking or Dissemination Limiting Marker (refer ISMF), for example **Secret**, **For Official Use Only** or **Sensitive**, cannot be released as-is, however may be re-classified if appropriate de-classification and/or risk mitigation activities are undertaken.

8.1.1 Governance and Accountability

Chief Executives
Agency Chief Executives are accountable for all security matters within their agency. The Chief Executive must authorise the disclosure of all official information (including data and datasets) to the public.

Data Authority
The Data Authority is responsible for ensuring that classification or reclassification of the data is undertaken and that appropriate input is obtained from the data subject experts and information users.

Executive Peer Review
Executive Peer Review provides risk mitigation to ensure the public release of a dataset does not inadvertently put another part of the business security at risk or disclose information that could lead to identification of a person (mosaic effect).

Agencies are encouraged to conduct an Executive Peer Review of all dataset security markings before a dataset that has been assessed and marked as Public, is released. Executive peer review is recommended if data has been manipulated to mitigate risks e.g. when personal information has been de-identified or when information is declassified.

The recommended approach for an Executive Peer Review is to engage your Information Technology Security Adviser or Advocate to circulate open data candidates to all Executives within an agency for review and comment. A sample of the data should be provided and summary of how any data risks that have been mitigated e.g. de-identification or redaction techniques applied.

Agency Information Technology Security Adviser (ITSA)
Agencies should consult with their Information Technology Security Adviser (ITSA) for further advice and guidance on agency specific information classification and marking procedures and guidelines. Agency ITSAs can assist with data assessment decisions and can provide advice on information security threats and risks.

8.1.2   Security classification and marking decisions

Information security is a complicated area and the Agency ITSA is best suited to both advise on how to properly assess information (including data) in terms of confidentiality, integrity and availability; and how to handle that information once it's been classified and marked in accordance with the ISMF and agency specific policies, standards and guidelines.

For Data to be released through open data, it must firstly be assessed and able to be marked as **Public** then authorised for release by an Agency Chief Executive.

Data marked as **Public** is authorised for unlimited public access and circulation such as agency publications, data download sites and websites.

Government information, including data, is required to be assessed and marked. Where government data is not able to be marked Public, some of this information can have the sensitive aspects redacted, amended or otherwise modified such that it could be re-assessed as meeting the requirements for Public and if still valuable to the public, released under open data.

The Data Authority (typically the business owner of the information) is responsible for re-assessment of information, where appropriate, and should engage their ITSA to assist with these decisions.

Refer to the [Data Security Marking Decision Diagram](#) (Appendix A) and [Open Data Guide to Security Classification](#) ([Appendix B)](#) for assistance with security decisions and actions required.

For a full list of the classification markings refer to [ISMF Guideline 8b](#) (PDF).

8.1.3   Privacy of personally identifiable information

In making government information publicly available, agencies must ensure that personally identifiable information regarding citizens is not released in accordance with the [Information Privacy Principles Instruction (IPPI)](#). This includes checking metadata for personally identifiable information.

Agencies may need to de-identify personally identifiable information, which means removing anything that can identify a person such as a person's name, address, gender, date of birth, ethnicity etc. Care must be taken to ensure information is properly de-identified and not able to be re-identified by linking with other information sources. Once the data is de-identified it needs to be reclassified.

The security classification assessment and marking decisions detailed above incorporate markings for personally identifiable information.

A number of techniques can be applied to properly de-identify the data and mitigate any risks of identification. Refer to the Privacy Committee of South Australia [Privacy and Open Data Guideline](#) (PDF) for more information.

This guideline also provides a [Privacy Risk Assessment Process (Refer appendix C).](#)

Executive Peer Review is recommended before release of privacy risk mitigated data. A sample data should be produced and peer reviewed as a means of testing and a summary of how any data risks have been mitigated e.g. de-identification techniques applied.

8.2   RESPONSIBLE INFORMATION SHARING

This process recognises that public information may require high degrees of integrity (accuracy) and availability (and by association that availability requirements may change based on calendar or event driven periods, a notion of 'peak-demand' for certain types of information). The benefits of responsible information sharing include:

- User experience both within and external to government
  - o   Information is easy to find and can be relied upon
  - o   Services are interactive and timely
- Elimination of delays
  - o   Information on demand (i.e. readily shared and accessible when required)
  - o   Accessible by leveraging the internet, including mobile devices and emerging technologies
- Reduction in costs and increase in organisational efficiencies
  - o   Lowers cost of service delivery
  - o   Fosters greater use of 'self-service' capabilities
  - o   Interactive government is an agile and consultative government
- Accurate and timely information
  - o   Timely information is achieved because availability requirements have been considered
  - o   Accurate information is more likely because integrity requirements have been considered
- Maintaining trust and confidence in government as a supplier and custodian of information
  - o   Reliable and secure information services engender trust and confidence. This is reflected by organisational capability, capacity and communication.

Sharing information with the public may require an adequate level of assurance to business users on the accuracy and availability of the data as per ISMF Guideline 8a.

Some examples of the varying degrees of responsible information sharing are illustrated below:

- Emergency management and crisis response information is generally distributed on a broad public scale, but requiring exceptionally high degrees of accuracy (integrity) and availability in order to inform the community and emergency services personnel in a timely and accurate manner
- Information Sharing Guidelines for Promoting the Safety and Wellbeing of Children, Young People and their Families - these guidelines deal with the legal and practical framework that supports appropriate information sharing for the provision of integrated support to children, young people and their families
- Data delivery through a web service for use in a developers business or application.

8.2.1 Availability Classification

Use the table below to determine how available data needs to be when made publicly accessible. Consult with your agency ITSA for advice to ensure that the data is not over classified and to advise of any additional controls required. A high level of classification (e.g. A4) will increase the costs to manage the data both internally and for public release.

Availability classifications should be discussed with you Data Manager to assist decisions with the approach for release of the data.

What is the tolerable outage for the information

Availability Classification

• Information is bound to become unavailable at certain points in time. Whether the information is unavailable due to system outages or planned maintenance windows or as a result of unintended and unplanned events. The determination of a 'tolerable outage' must be accepted by the business.

• Tolerable outages can drive an Availability Classification for the information. Factors to consider include the dependence the business, its customers or the community has on the information and the business impacts a disruption or loss of access to the information may cause. Availability requirements may change at certain times whether event or calendar driven.

Classification	Description
A4	ABSOLUTE requirement, meaning that the business would be crippled by the loss and recovery must be virtually instantaneous (no longer than a few minutes).
A3	HIGH requirement, meaning that loss would cause major disruption to the business and recovery must be achieved within a period measured in hours (typically same business day).
A2	MODERATE requirement, implying the loss would have a significant impact and recovery must be achieved within a period measured in days (typically three business days or less).
A1	LOW requirement, meaning that loss of the data would have only a minor impact on the business for an extended period (i.e. "best-effort" recovery).

8.2.2 Integrity Classification (Quality)

The integrity classification will assist you to determine what level of quality the data needs to be and is generally applied when misinformation could cause risk to a person or risk to the reputation of the government.

The quality of each datasets that will be released to the public may be considered when assessing integrity classification of the dataset. The below discuss elements of quality as per guidance provided by the National Statistical Service (NSS):

High Quality

- Timely, low errors, consistent collection, interpretable, authoritative source

Reasonable Quality

- Low unknown errors, consistent collection, interpretable, source clearly identified.

Poor Quality

- High errors, inconsistent collection, interpretability low and metadata required to interpret the data is not available.

Data that is of poor quality should not be publicly released until quality issues are resolved. If the data is considered to have a high public value, the Data Authority may need to investigate ways to improve the quality of the data in future plans.

8.2.2.1   Characteristics of quality data:

- Timeliness

  If there are lengthy delays between the reference period and data availability this can have implications for the currency or reliability of the data. However, historical data is still very valuable to publish, so timeliness in itself would not disqualify data from release.

- Accuracy

  Accuracy is an important component of quality as it relates to how well the data portrays reality, which has clear implications for how useful and meaningful the data will be for interpretation or further analysis. The major sources of errors that could cause inaccuracies should be assessed. Consider also the *Integrity Classification* the information has been marked with.

- Coherence

  Coherence is an important component of quality as it provides an indication of whether the dataset can be usefully compared with other sources to enable data compilation and comparison.

- Interpretability

  Interpretability is an important component of quality as it enables the information to be understood and utilised appropriately. Check the data to ensure it would have meaning to an external party. Are terms used in the dataset ambiguous, open to interpretation or likely to confuse a user? If so, detailed interpretation metadata with definitions or explanatory notes will need to be provided to ensure the data is understood as intended.

- Authoritative source

  Data is more valuable if it is collected at the source or has systems in place to ensure it is the single point of truth or authoritative source of information.

8.2.2.2   Quality Assessment Tool

  The ABS through its National Statistical Service (NSS) has developed a free online data quality tool that will help get agencies thinking about what makes data good quality. The Data Quality Statement online tool can be found at https://www.nss.gov.au/dataquality/

The tool takes you step by step through a series of 7 questions about your dataset.

The output is a data quality statement that can be saved as XML or RFT file and published with the dataset on Data.SA. The template can be saved in progress and revisited when the answers to the questions are available. This is an optional step for agencies, but is a good guide for data quality at a national standard level.

Ensure that personal contact information is not provided with any published version of the statement. If you will release the statement with your data ensure that you include it as an attachment for approval with the dataset.

8.2.2.3 Determine the Integrity Classification

Use the table below to determine the integrity classification of the data. Consult with your agency ITSA for advice to ensure that the data is not over classified and to advise of any additional controls required. A high level of classification (e.g. I4) will increase the costs to manage the data both internally and for public release.

How accurate must the information be

Integrity Classification

- Work with stakeholders to understand what inaccuracies to the information can be tolerated. The ultimate determination of how accurate the data needs to be is the responsibility of the Data Authority.

- Determine the extent of consequences if the information contains errors or contains omissions using the following Integrity classification and descriptions.

Classification	Description
I4	ABSOLUTE requirement, implying that no inaccuracies or omissions can be tolerated
I3	HIGH requirement, meaning that a loss of integrity would cause significant embarrassment and disruption and might be difficult to detect.
I2	MODERATE requirement, meaning that a user would be somewhat affected by a loss of integrity, but the situation could be easily detected and recovered.
I1	LOW requirement, such that there would be minimal impact if the data was inaccurate or incomplete

.

8.3   LICENCE CLASSIFICATION

The Government of South Australia supports and encourages the dissemination and exchange of public sector information, and endorses the use of the Australian Governments Open Access and Licencing Framework (AusGOAL) by its agencies.

Restrictions on the use of the data will be avoided where possible. Utilising the Australian Governments Open Access and Licensing Framework (AusGOAL) will support openness and improve usage.

A licence classification is only required for data that has been classified as Public and is intended for release as open data..

8.3.1   What is AusGOAL

AusGOAL is the Australian Governments Open Access and Licensing Framework which provides support and guidance to government and related sectors to facilitate open access to publicly funded information. AusGOAL makes it possible for organisations to manage their risks when publishing information and data in a way that drives innovation and entrepreneurial activities. The AusGOAL licensing framework exists in parallel with our copyrights, which protects our intellectual property, Government of South Australia brand and logo's.

8.3.2   Creative Commons

AusGOAL incorporates a suite of licences including the Creative Commons Version 4.0 International. The Government of South Australia also recognises work licensed under the Creative Commons Version 3.0 Australian licenses, which is an older version of the licence used in Australia only.

Key attributes of the Creative Commons V4.0 licences:

- internationally recognised
- defines how data can be used, shared or adapted
- provides clear instructions on the restrictions of using the data
- requires attribution to your Agency, title of work, date sourced and the URL used to access the data
- does not allow data users to make agencies endorse them or their data use
- requires users of the data to indicate if they modified the material
- notifies the user that no warranties are given
- material is offered as-is and as-available and makes no representations or warranties of any kind concerning the licensed material, whether express, implied, statutory or other
- limits the Government of South Australia's liability
- anonymity – can specify no attribution for derivatives
- 30 day period to cure breach before licence terminates.

8.3.3 The Creative Commons Licenses

	Attribution **CC BY** This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials. View License Deed | View Legal Code
	Attribution-ShareAlike **CC BY-SA** This license lets others remix, tweak, and build upon your work even for commercial purposes, as long as they credit you and license their new creations under the identical terms. This license is often compared to "copy left" free and open source software licenses. All new works based on yours will carry the same license, so any derivatives will also allow commercial use. This is the license used by Wikipedia, and is recommended for materials that would benefit from incorporating content from Wikipedia and similarly licensed projects. View License Deed | View Legal Code
	Attribution-NoDerivs **CC BY-ND** This license allows for redistribution, commercial and non-commercial, as long as it is passed along unchanged and in whole, with credit to you. View License Deed | View Legal Code
	Attribution-NonCommercial **CC BY-NC** This license lets others remix, tweak, and build upon your work non-commercially, and although their new works must acknowledge you and be non-commercial, they do not have to license their derivative works on the same terms. View License Deed | View Legal Code
	Attribution-NonCommercial-ShareAlike **CC BY-NC-SA** This license lets others remix, tweak, and build upon your work non-commercially, as long as they credit you and license their new creations under the identical terms. View License Deed | View Legal Code
	Attribution-NonCommercial-NoDerivs **CC BY-NC-ND** This license is the most restrictive of our six main licenses, only allowing others to download your works and share them with others as long as they credit you, but they cannot change them in any way or use them commercially. View License Deed | View Legal Code

8.3.4 The Preferred Licence

The preferred licence for data is **Creative Commons Attribution 4.0 Licence**. This licence is the most open with least restrictions. It supports research, analysis, data transformation and enables applications to be created from it. This licence is also recognised internationally.

Attribution CC BY
This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials.

View License Deed | View Legal Code

This licence must be used unless there is a reason a more restrictive licence should be provided such as:

- data was provided with a term of use
- there is a reason why you will not allow someone to re-use the material for commercial gain
- there is a reason you won't allow a user to adapt the work you require re-licence under identical terms.

8.3.5 Alterations and additions to a Creative Commons licence

- **Can I waive license terms or conditions?**

  Yes. You may choose to waive some license terms or conditions.
  Generally a request to waive any conditions will need to be reviewed on a case by case basis.

- **Can I change the license terms or conditions of a CC Licence?**

  No if you change the terms and conditions of any Creative Commons license, you must no longer call, label, or describe the license as a "Creative Commons" or "CC" license, nor can you use the Creative Commons logos, buttons, or other trademarks in connection with the modified license or your materials.

  Altering terms and conditions is distinct from waiving existing conditions or granting additional permissions than those in the licenses.

  If different licence terms are required a separate licence term agreement is required. You will require Crown Solicitor advice to develop new terms of licence. These agreements are discouraged as it creates confusion in expectations and creates additional burden of negotiating permissions on a case by case basis.

  If the preferred licence for data (**Creative Commons Attribution 4.0 Licence**) does not meet your needs you may require a more restrictive licence.

8.3.6 When to use a more restrictive licence

- **Data was provided with a term of use**

  When identifying third party property rights you may have identified data that was provided on a set term of use. Often these terms of use will define how you can re-license work. Check the term of use before selecting a licence.

- **There is a reason you won't allow someone to re-use the material for commercial gain**

  Restrictions on data for commercial re-use should be avoided as open data could stimulate economic growth in South Australia through using the data to support a commercially viable business. If a commercial licence restriction is in place on the data, then the user cannot even recover the costs of using data.

- **There is a reason you won't allow a user to adapt the data (Derivative)**

  Restrictions on the data to adapt the work should be avoided, as this will limit the user's ability to display the data in a different form, innovate, analyse, and create services. Note that all Creative Commons Licenses require the user to indicate if they modified the material and retain an indication of previous modifications.

  The Data Authority should consider if there are any reasons that derivative work should not be allowed due to a perceived risk or duty of care to citizens if data was used in a way that could cause harm. For example, health advice should not be licensed for derivative reuse, however non-personal health statistics may not have a perceived risk if displayed in a different form or used to advance research. Personal information should never be licensed for derivative use, even if there is a legal requirement to publish information.

- **If you require Re-licence under identical terms (Sharealike)**

  If you require users of the data to share the data under an identical licence you can choose a ShareAlike licence. This licence restricts the use of data and requires users of the work to licence any derivative works under the same terms.

If a more restrictive licence is required, use the following ***Open Data Licence Decision Diagram***

8.3.7 Open Data Licence Decisions Diagram

1. Is this data sourced from a 3rd party? — Yes → **Does the data have existing licensing terms?**

No →

2. Will you allow someone to re-use this material for commercial gain and/or to adapt the work and you do not require the work to be licensed under identical terms?
Material allowing commercial gain will give users the opportunity to recover the costs encountered while using the data. If you require the works to be licensed under the same terms, users won't be able to adapt your work so that they can display of data in a different form, innovate, analyse and create service delivery. (A perceived risks to users may prevent derivative work e.g. health advice. If this is the case then it is recommended that you licence under share-a-like.)

Does the data have existing licensing terms?

Yes →

No →

Reasonable contact must be made, in writing, to seek permission to openly licence the material (it is recommended to set a response period of 30 days). Once a reply is received or the 30 day period is over, which ever comes first, restart the process.

Is the licensing term copyright?

Yes (left branch of Step 2):

CC BY

No (right branch of Step 2):

Is there any reason you wont allow someone to re-use this, or part of this material for commercial gain? Innovation, building on research and commercial activities using our data are encouraged. Applying commercial restrictions will limit re-use and should be avoided unleass adequate reason is demonstrated.

Yes →

Is there any reason you won't allow derivatives of the material? Generally this may be the case is there is a perceived risks or duty of care to the public if data was displayed in a different form. E.g. Health advice.

Yes → CC BY NC ND

No → **Do you require share-a-like licence?** (This will require any new creations to be licenced under the identical terms.)
Yes → CC BY NC SA
No → CC BY NC

No →

Is there any reason you won't allow derivatives of the material? Generally this may be the case is there is a perceived risks or duty of care to the public if data was displayed in a different form. E.g. Health advice.

No → **Do you require share-a-like licence?** (This will require any new creations to be licenced under the identical terms.)
Yes → CC BY SA
No → CC BY

Yes → CC BY ND

Is the licensing term copyright?

No → **Is a share-a-like licence required?**
Yes → Select the licence that is the same as the existing licensing terms.
No → Check the terms of the licence and complete the licence selection from Step 2.

Yes → **STOP** this data is copyright and can not be published.

CLASSIFY SUMMARY

Determine Data Security Marking		
☐	**Public**	A dataset with a marking of **Public** can be released as open data.
☐	**Reclassification**	Some data can be reclassified as **Public**. The data will require some redactions, amendments or manipulation to ensure sensitive elements are removed so that there is no damage (or potential damage) to the government, business or members of the public. The Data Authority is responsible for reclassification of information and should engage their ITSA to assist with these decisions. An Executive Peer Review is also encouraged.
Determine Responsible Information Sharing		
Availability classification		
☐	**A4**	ABSOLUTE requirement
☐	**A3**	HIGH requirement
☐	**A2**	MODERATE requirement
☐	**A1**	LOW requirement
Integrity Classification		
☐	**Quality**	Consider the quality of the data to assist the integrity classification
☐	**I4**	ABSOLUTE requirement
☐	**I3**	HIGH requirement
☐	**I2**	MODERATE requirement
☐	**I1**	LOW requirement
Determine the Creative Commons Licence		
☐	**CC BY**	The preferred and most open licence
☐	**CC BY-SA**	ShareAlike
☐	**CC BY-ND**	No Derivs
☐	**CC BY-NC**	Non-Commercial
☐	**CC BY-NC-SA**	Non-Commercial-ShareAlike
☐	**CC BY-NC-ND**	Non-Commercial-NoDerivs
Complete Open Data Process Worksheet		
☐	Record decisions	Update the Open Data Process Worksheet.

9   APPROACH FOR RELEASE–STEP 3 OPEN DATA PROCESS

Data needs to be in a format that makes it easy to use, transform and reuse. It is important for the community to have confidence that the data they are accessing is current, reliable and well managed.

We will favour approaches that securely automate the release of regular or live open data services direct from our systems in order to generate sustainable value.

It is the Data Authorities responsibility to ensure a sustainable approach to the release of data is developed and approved.

It is recommended that the Data Manager of the dataset is engaged to advise and recommend the most suitable approach for the release of data. The Data Manager will have primary accountability for the day-to-day management of the information systems where the data is managed.

The approved approach will commit the agency to ongoing support and management of the released data in the format, frequency and delivery method determined. For this reason it is recommend that decisions in this section are included for Approval by the Open Data Authority and Advocate.

The approach for releasing your data includes:

- the format(s) the data will be released in
- the frequency of release
- the method of delivery
- consider data distribution services
- determine metadata and supporting information
- funding open data investment (automated delivery methods)
- cost recovery (cost recovered data is not in scope until further policy is developed).

9.1   FORMAT

For data to be open, it must be released in an open format and machine readable.

Open format: An open format is one, which is platform independent or non-proprietary, which means it does not matter what operating system or licensed software you have access to.

Machine readable: means that a computer without human aid can read data both in its format and in structure. Machine readable data is structured and easy to query using software code. Developers can consume the data from their programs/applications and re-use it.

Data may need to be transformed into another format to make it open. Releasing data in an open format may be as simple as saving or exporting system generated data in csv format instead of Microsoft Excel or PDF which is proprietary. Alternatively tools may be required to convert the data.

To understand what open formats are available refer to the Table 2 Examples of Common Open Data Formats.

Table 2: Examples of Common Open Data Formats

Format Name	Definition	Type of data to use this for
Comma Separated Values (CSV)	Comma Separated Values (CSV) is a great way of storing large amounts of data with just commas separating the data values. Often the csv file will contain a header with names describing what data is populating the file.	Tabular data e.g. Use instead of Excel
Tab-Separated Values (TSV)	TSV is a very common form of text file format for sharing tabular data and is highly machine readable.	Tabular data Use instead of Excel
JavaScript Object Notation (JSON)	JSON uses human-readable text to transmit data objects consisting of attribute–value pairs. It is used primarily to transmit data between a server and web application, as an alternative to XML. The file size will be more compact or smaller than XML	Complex structured data Multidimensional data Tabular
Extensible Markup Language (XML)	XML is a widely known markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable. Users create and define their own tags.	Complex Structured data Multidimensional data Tabular data e.g. database extract metadata
Rich Site Summary (RSS)	RSS (originally RDF Site Summary), often dubbed Really Simple Syndication, uses a family of standard web feed formats to publish frequently updated information: blog entries, news headlines, audio, video. An RSS document (called "feed", "web feed" or "channel") includes full or summarised text, and metadata, like publishing date and author's name.	Use for announcements or events e.g. on websites
ATOM	The Atom Syndication Format is an XML language used for web feeds. The Atom format was developed as an alternative to RSS. Note RSS is the preferred standard	Use for announcements or events e.g. on websites
Open Document Format for Office Applications (ODF)	The Open Document Format for Office Applications (ODF), also known as OpenDocument, is an XML-based file format for spreadsheets, charts, presentations and word processing documents. It was developed with the aim of providing an open XML-based file format specification for office applications.	Non-system generated metadata or additional information you release with your dataset. (replaces Excel, Word, PDF)
HTML	Used for formatting information on the web	Non-system generated metadata or additional information you release (replaces PDF, Word)
Keyhole Markup Language (KML)	KML is an XML language focused on geographic visualization, including annotation of maps and images.	Spatial/location data
Geography Markup Language (GML)	GML is the XML grammar defined by the Open Geospatial Consortium (OGC) to express geographical features. GML serves as a modelling language for geographic systems as well as an open interchange format for geographic transactions on the internet.	Spatial/location data
GeoJson	GeoJSON is an open standard format for encoding collections of simple geographical features along with their non-spatial attributes using JavaScript Object Notation.	Spatial/location data

9.1.1 Format Decision Considerations

When deciding on what format to release the data, agencies may also consider:

- which open formats are compatible with current system infrastructure
- which formats are widely accepted by developers
- which formats maintain the highest degree of integrity when converted
- if a dataset is proprietary, are there tools available to convert it to an open format e.g. Microsoft Excel file can be saved as a csv
- can the data be released in multiple formats to meet the preferences of the community
- open data maturity.

9.1.2 Native Format Data

A dataset may also be made available in its Native (original) format. Although native format data or raw data may not be in an open format, it can provide interpretative benefits and provide users with greater knowledge for development. Agencies can provide native data as an additional file which will be saved alongside the dataset as a resource.

9.1.3 Multiple Data Formats

Agencies should consider providing multiple data formats. Additional data formats provide greater scope for developers and the public to re-use the data in format that suites them.

A dataset may also be made available:

- in a widely used proprietary format (e.g. .xls) that encourages re-use; or
- via an existing data tool that allows users to explore, manipulate, and reuse the data.

9.1.4 Open Data Maturity

As agencies adopt 'open data' practices, the maturity level of their data will develop over time. The World Wide Web Consortium (W3C) has developed a 5 Star Open Data Model that agencies can aspire to.

3 Star Level of Open Data Maturity (Machine-readable and Non-proprietary) is the minimum standard for release of government's public data for re-use.

It is recognised that agencies need to build capability and maturity in open data. 2 Star Level of Open Data Maturity (machine-readable but proprietary format) may be released in the first release of the data.

The following table provides a summary of the 5 Star Open Data Model:

The 5 Levels	Description of the 5 Levels	Examples
⭐	ON THE WEB, LICENSED FOR REUSE Data is visible, licensed for re-use, but requires considerable effort to re-use	Data is licensed for re-use Tables on a website or data in a PDF document.
⭐⭐	MACHINE READABLE Data is visible, licensed, machine readable and easy to reuse but not necessarily by all	On the web and in proprietary formats.
⭐⭐⭐	NON-PROPRIETORY FORMAT Data is visible, easy to reuse by all and non-proprietary format, which means it is not restricted to specific software.	On the web, machine readable and non-proprietary e.g. CSV file, XML, ATOM; JSON, KML.
⭐⭐⭐⭐	RDF STANDARDS Data is visible, easy to use and described in a standard way. It uses open standards from W3C to identify things so that people can point at your data.	Resource Description Framework (RDF) is a framework for describing resources on the web. It breaks down data into a series of facts.
⭐⭐⭐⭐⭐	LINKED RDF Data is visible, easy to use, described in a standard fashion and its meaning is clarified by being linked to a common definition e.g. linked to other people's data to provide context.	It provides a link to the meaning via a Uniform Resource Identifier (URI). Any other reference to same source means the same meaning can be assumed with confidence.

9.2   FREQUENCY

The Data Authority must make a determination about how often the published datasets will be refreshed.

Consider who is responsible for the datasets ongoing release and maintenance.

This is especially important if data delivery/publishing methods are manual. You may need to establish a process to ensure data is refreshed.

Criteria to consider include:

- how often is the data collected
- the value of data is highest at the point of collection
- will timeliness affect the quality of the datasets
- the resources required for data extract and preparation
- availability classification (refer to availability classification decision)
- will digital services be reliant on the data
- the delivery method for the data.

The frequency for release of the data will be published with the dataset on Data.SA to inform users of the currency of the data.

Frequency options include:

- daily
- weekly
- monthly
- yearly
- quarterly
- As required.

9.3   METHOD OF DELIVERY

The Data Authority and Data Manager will need to determine what the best method to deliver data to the user from an information management system. Where possible the Data Authority should consider data delivery methods that securely automate the release of regular or live open data services direct from our systems.

Open data will be easily discoverable through Data.SA the Government of South Australia Data Directory. Each dataset will have a unique entry on Data.SA that will allow users to search and understand the content of the dataset.

Data.SA currently provides two ways for users to access data:

- hosted on Data.SA (Manual Publishing)
- linked data to an agency portal, data service or automated data delivery.

9.3.1   Hosted on Data.SA

Datasets can be stored and hosted on Data.SA. In this instance, files are manually uploaded to the site. Refresh of this data will require a new data file to be manually uploaded to the dataset each time the data is available
.

When to host data on Data.SA:

- data collected periodically where automation costs would exceed benefits (e.g. annual, in frequent or once of data)
- historical data
- system infrastructure does not support automated delivery

- where data protection amendments required cannot be system generated (privacy, secrecy, legislate protected data)
- to release data initially while automation processes are established
- if data is linked and updated automatically then an annual version of the data should be published to preserve the data for historical, research and analysis purposes.

9.3.2   Linked Data

Linked data is when data is discoverable on Data.SA and the dataset entry is linked to the source of the data, either an agency portal, data service or through automated data delivery.

When to link data on Data.SA:

- automated data delivery is established
- if an existing data portal or website exists where the agency will maintain the data (either manually or automatically)
- to a data tool that allows users to explore, manipulate, and reuse the data
- if data is of a considerable size (Terabyte+) and a hosting service is used.

9.3.3   Automated Data Delivery

It is recommended that agencies consider automating data delivery. Examples of automated data delivery include:

- **Application Programming Interface (API)**
  An API allows your product or service to talk to other products or services. In this way, an API allows you to open up data and functionality to other developers and to other businesses. It is increasingly the way in which agencies and companies exchange data and services, both internally and externally. API's allow developers to build applications that use data. Data that changes rapidly is often delivered through an API. An example of an API is the Australian Tourism Data Warehouse.

  As API's are external to Data.SA agencies need to provide information on how to use the API. Usually instructions and a sample key are provided and published on Data.SA with the link to the API.

  For more information on API's refer to Online Tools and Resources.

- **File Transfer Protocol (FTP)**
  FTP is a standard network protocol used to transfer files. An FTP can securely transfer files from a system to an external location. Agencies may use FTP to deliver data to a FTP address that is linked on Data.SA. The user will still need to download or copy the data when it is delivered. FTP is often used when regular overnight, weekly or monthly data is collected and can be extracted. A web browser can connect to FTP addresses exactly as you would to connect to HTTP addresses. Using a web browser for FTP transfers makes it easy for you to browse large directories, read and retrieve files. A FTP has the potential to scale the automated delivery service to other datasets from the same information system.

- **Web Services**
  A web service is a method of communications between two electronic devices over the World Wide Web and use via standard request-response protocol which allow the exchange of messages in which a requestor sends a request message to a replier system which receives and processes the request, ultimately returning a message in response.

- **Real-Time Data**
  Real-time denotes information that is delivered immediately after collection. There is no delay in the timeliness of the information provided. Real-time data is often used for navigation or tracking. Agencies are encouraged to research real-time standards and protocols that exist that are appropriate for the data.

  View the Government of South Australia's Real time water data and Adelaide metro real time passenger information.

- **Really Simple Syndication. RSS**
  Also called Rich Site Summary or web feeds, RSS is an automated content delivery vehicle in a standard XML file format. RSS feeds benefit users who want to receive timely updates from many sites. As the data is fed out only when data is refreshed it is a popular and simple solution to automation that includes notification.

- **XML Data Extraction**
  Agencies can export a copy of their database into an XML file on a regular basis (nightly or weekly). As the XML file can be quite large it can be compressed into a Zip file. The file is then made available on a web site server. A working example of this method can be viewed at the United States Grants website:

9.3.4   Selecting the Automated Method

Before making delivery decisions the following needs to be considered:

- frequency of the data release
- availability classification
- how the data will provide public value
- long term costs savings of investment in automation compared to manual data extract and refresh.

For automated delivery of data, you may also need to consider the following
- additional consideration for webservices, API's, real time data
- developer engagement and support
- relevant standards and protocol
- API basics (refer to Online Tools and Resources.)
- cyber security (consult with your ITSA and/or StateNet Services)
- framework and infrastructure required to maintain the data delivery
- the potential to scale the automated delivery service to other datasets from the same information management system to create efficiencies
- stability of the data structure or information management systems (consider the future of legacy systems or legislative changes that may alter how the content of data delivered)
- emerging technologies
- Automated scripts to remove protected data.

9.3.5   Additional Consideration for Webservices, API's, Real Time Data

Reliable automated and sustainable data services (Request-response web services, API's and Real time data) can support economically viable and sustainable digital services that have a great benefit to our community however agencies should also consider:

- testing the concept and benefits
- stakeholder consultation (you must include StateNet Services, your corporate ICT unit and your ITSA)

- determine the approach and design such as protocol, standards (internal and external),hosting design, developer engagement, formats
- consider a single protocol that can be used on several platforms
- Infrastructure and service levels required to support a request-response web services
- establishing gateway limits/ pressuring testing (how many hits on the server to access data)
- developer support such as the framework, documented outputs, technical documentation, notification of updates, forums and terms of service
- risks and mitigation controls (pressure testing, security testing required to protect the reputation of the Government if data services fail.
- system improvement/review of frameworks, protocols, security and pressure test.

9.3.6 Developer Engagement and Support Eco System

Where the release of data is through a web service or real-time application early engagement with the community and developers and the establishment of support and communication channels is highly recommended. This may assist you to identify formats, frequency, delivery, and support required to create sustainable value form data release.

The benefits of setting up a developer eco system include:

- you will know who is using your data to develop
- test the concept, improve design and maximise use
- channel to communicate changes, fix problems and find solutions through the community
- developers can act as a first point of call for issues with services and then relay them to you
- emerging standards or technology changes can be discussed.

Simple blogs and online forums can provide low maintenance communication channels that allow two way conversations. These mediums can also be used to supply documentation, frameworks or change notification.

View the Adelaide Metro developer website for a working example of a developer support ecosystem. .

9.3.7 Licensing your Linked or Automated Data

If you are publishing data on an agency managed website/tool, ensure that you display a licence logo and statement with the dataset on your website. Note that all dataset entries on Data.SA will be licensed. The licence appears on the bottom left hand side of the dataset page.

Example of CC-BY

This Government of South Australia website is licensed under a Creative Commons Attribution 4.0 License. © Copyright 2014

If you publish several datasets on the same webpage and there are different licence terms required, each dataset will require a licence logo to be displayed in a way that clearly indicates to the user of the dataset what the terms of use are.

Host a Copyright Statement that Supports an Open License

A current open licence Copyright Statement is available on
http://www.sa.gov.au/copyright

This statement has been approved for use on South Australian Government websites. You are required to reference your own agency and contact detail in this statement. If you alter any other elements of the statement you will need to seek Crown Solicitor advice.

9.3.8 Publishing Supporting Resources

Where ever possible supporting information should be published with your datasets. Agencies often do this through a Zipped file to download.

If you are publishing material to be printed or downloaded as a resource with your dataset you should also ensure this work is licensed appropriately and the material includes a licence logo and statement.

Example provided for CC-BY licensed material.

With the exception of the Government of South Australia brand, logos and any images, this work is licensed under a Creative Commons Attribution 4.0 Licence. To attribute this material, cite the >>Agency>, Government of South Australia, >>title of work<<,>> date the content was sourced<<, >>dataset URL<<

9.4   METADATA

Metadata is the information that defines and describes data. It provides data users with information about the purpose, processes and methods involved in the data collection. There are two types of metadata that the Data Authority should consider in their plans.

9.4.1 Discovery Metadata

Discovery metadata is the minimum metadata required for all datasets to make it discoverable on Data.SA. It helps a user decide whether the dataset suits their needs before downloading it.

Discovery metadata is captured when an agency completes the *Data.SA Publishing Content Summary Sheet* (Appendix E).as part of the publishing process (covered in section 9). It includes the name of the dataset, a description, data format, licence, keywords etc.

The fields on the Data.SA Publishing Content Summary are mandatory. Fields include temporal coverage, geospatial coverage, jurisdiction and frequency of release. The metadata fields are defined on the back of the *Data.SA Publishing Content Summary Sheet* (Appendix E).

9.4.2 Interpretative Metadata

Agencies should also provide interpretative metadata to accompany their datasets. This type of metadata provides users with more information about the dataset such as its purpose and the methods involved in its collection to help the user understand and interpret the data correctly.

Often industry based metadata standards will be used.

Metadata could include:

- metadata standards requirements
- definition of terms e.g. average weekly earnings
- accuracy e.g. the number of errors
- response rate

- explanatory notes - detailed contextual information, purpose, processes, and methods involved in the data collection

Interpretive metadata needs to be published in an open and machine readable format such as TXT or HTML as separate file to the dataset. Often interpretation metadata will be compressed with the data for download (e.g. zipped). Ensure metadata does not contain information that may breach privacy or security restrictions.

Engage your Data Manager to discuss if interpretation metadata is available and or automated from your information management system.

Consider the *National Statistical Service* Principles for managing metadata View the NSS Principles of metadata and metadata standards that apply to your field of data. The Australian Bureau of Statistics (ABS) offers assistance and advice on metadata. Contact the ABS representative if it is known or contact datasa@sa.gov.au.

9.4.3 *Additional Resources*

Additional resources should also be made available with the data to encourage use or understanding of the data. Any additional files provided will be published alongside the dataset as a resource:

- research reports
- website links which provide context and additional information
- photographs or images of the subject matter e.g. flora.

These resources should be licensed for reuse or copyrights clearly stated to reduce any confusion on how these additional resources can be used.

9.5 DATA DISTRIBUTION SERVICE

Datasets are discoverable on Data.SA, however, your dataset may also be a candidate for a Data Distributor service. A Data Distributor is another data distribution service that collates and distributes data as a single point of truth for a specific type of SA Government data.

If another data distribution service exists, it is recommended that they are engaged to discuss options for data delivery as a **Data Distributor**.

Contact that data distribution service to discuss and confirm the delegation of responsibility as the Data Distributor. They may also take on the role as Data.SA Publisher for the dataset to streamline the process, if they will not then clear process between the parties will need to be defined to ensure that Data.SA is maintained.

These services may also have other processes that will need to be considered, however it is necessary to ensure the open data process has been completed and approved before the Data Distributor releases data on your behalf.

Another Agency may also be your Data Distributor, this often happens as a result of machinery of government changes and infrastructure is maintained by another agency.

9.6 FUNDING OPEN DATA INVESTMENT

The implementation of automated data delivery may require upfront investments depending on the maturity of existing information life cycle management processes at individual agencies. Agencies are encouraged to evaluate current processes and identify implementation opportunities that may result in more efficient use of taxpayer dollars. Where possible open data delivery options should be considered in new ICT enabled projects.

Effective implementation should result in downstream cost savings for the enterprise through:
- information sharing efficiencies

- reduced impact of resources due to freedom of information
- reduction of costs associated with manual publishing open data
- improved policy development based on evidence based data
- scalable automatic delivery service
- maintenance of the data.

Therefore, these potential upfront investments should be considered in the context of their future benefits and be funded appropriately through the agency's capital planning and budget processes.

A detailed business case or cabinet submission may be required for automated data delivery to fund open data investment that requires additional tools or resources.

9.7 COSTS RECOVERY

Data should be available on a non-discriminatory basis to anyone where practicable to encourage its widespread use and to achieve maximum value. Government data will be released free of charge unless:

- the benefits of accessing and using the data are predominantly private in nature rather than creating broader public benefits
- there is a statutory requirement for charges to apply
- Cabinet has approved that charges be applied.

Agencies should refer to the Governments Cost Recovery Guideline (Once approved by Cabinet) for policy direction and guidance on how to develop and review cost recovery arrangements so that they are consistent, transparent and effective.

The Department of Premier and Cabinet in consultation with stakeholder is developing the Governments Cost Recovery Guideline. Until this guide is finalised data that is predominately and exclusively private in nature or has current pricing arrangement in place is out of scope for open data.

APPROACH SUMMARY

Prepare		
☐	Format	Ensure dataset is released in an open format (machine readable and non-proprietary). Multiple formats of data can be released, including native format Identify how data will be transformed into the format for release.
☐	Frequency of Release	Determine frequency of release and what process will be implemented to refresh the data.
☐	Data Delivery	Determine how dataset will be delivered: • hosted on Data.SA (agencies will need to send files to datasa@sa.gov.au for loading) • linked via existing agency portal (enter links to the dataset on the *Data.SA Publishing Content Summary Sheet*) • automated e.g. for an API prepare sample key and any instructions.
☐	Interpretative Metadata & additional resources	Prepare any other material which will help users understand what your dataset is about. This includes interpretive metadata files (csv or txt format), website addresses, reports or images. This will be emailed with the *Data.SA Publishing Content Summary Sheet* to datasa@sa.gov.au.
☐	Data Distributor	Consider if dataset is a candidate for a Data Distributor Service and engage the data distributor in planning the approach for release. Data Distributor may have additional processes. Determine who's responsibility it will be to publish data on Data.SA
☐	Costs	Consider funding or cost recovery implications.
Open Data Process Worksheet		
☐	Record decisions	Record all decisions made about the dataset on the *Open Data Process Worksheet.*

10 APPROVE - STEP 4 OPEN DATA PROCESS

The approval process for open data is necessary to protect both the government and citizens of South Australia. The approval process covers all aspects of the open data process from identifying data, third party rights, classification, marking, and ongoing commitment to the approach for release of the data.

The Open Data Process Worksheet is designed to capture decisions and ensure that data when released has been cleared of risks and will be managed in the future. This approval process does not cover:

- open data investment, this will need to be approved through the agency's capital planning and budget processes
- cost recovery.

10.1 PROCESS TO APPROVE

Subject Matter Expert	Completes an **Open Data Process Worksheet** (or similar records) that detail all decisions and plans.
	A printed sample of the data and any other supporting documentation must be attached to the Open Data Process Worksheet.
	If data is reclassified as meeting the requirements of Public, a sample of data with amendments, redactions, de-identification or actions that are applied to the data to make it ready for public distribution must also be attached.
Data Manager	The Data Manager is consulted in decisions relating to the approach for release of data. This will ensure the data release is sustainable. A Data Distributor may also be consulted.
Data Authority	The Data Authority is required to approve the dataset for release and is responsible for the existence, protection, and use of this dataset.
	The Data Authority is responsible for the re-assessment and re-marking of data to Public and should engage their ITSA to assist with these decisions. An Executive Peer Review is also encouraged.
Open Data Advocate	The agency Open Data Advocate will provide final approval as the delegated authority to oversee open data for the agency.
Chief Executive	Chief Executive approval is required if data already marked as 'Public' is ready for public distribution e.g approval to publish on websites or Data.SA.
Open Data Coordinator	Completed *Open Data Process Worksheets* to be filed by the Data Coordinator, includes worksheets for datasets not approved for release. Agency record management practices should be followed as per agency policies and procedures.

11 PUBLISH - STEP 5 OPEN DATA PROCESS

Once data is approved a Data Authority will need to ensure the data is prepared and published in accordance with approved open data decisions.

The approved Open Data Process Worksheet will guide the preparation and publishing steps required.

The Open Data Process Worksheet can also be used to capture actions completed, this is highly recommended as different parties may be responsible for different steps in the Publish step of the process.

11.1 PREPARE

11.1.1 Prepare data

The most current datasets is required to be sourced. Refer to the Open data processing worksheet to see what formats are required to be published, there may be more than one.

Perform format transformation and protection techniques if they have not yet been undertaken as detailed in the approved open data process worksheet.

Format transformation may be as simple as selecting the appropriate open format to extract and save the data. E.g. Extract a report from the system as export it in CSV format instead of xls format. Refer Online Tools and Resources for assistance to transform tabular data (XLS) into useable machine-readable format.

Compare the data with the sample provided with the open data approval to ensure all amendments have been actioned.

If data is *linked or automated* these services need to be established and publicly accessible. The Data Manger may assist the Subject Mater Expert to perform the open data preparation.

11.1.2 Metadata Preparation

Prepare a current version of the interpretation metadata required to be published with dataset. Interpretation Metadata can be published in an open format so that it is machine readable and non-proprietary format. This metadata can be published as a separate resource, compressed (e.g. zipped) with the dataset if it is linked or a web address provided to a linked location. Before publishing interpretation metadata ensure that the metadata does not contain information that may breach privacy or security restrictions.

11.1.3 Data.SA Publishing Content Summary

All datasets published to Data.SA must have a completed *Data.SA Publishing Content Summary* sheet. This summary sheets details exactly what should be entered on the portal for the dataset including location of data and supporting resources.

The publishing content contains mandatory discovery metadata which assists users to determine whether the dataset is fit for purpose and provides a snapshot of the dataset before they download the data. Plain English should be used and no acronyms where possible so that all users can understand what the dataset is about.

It is recommended that the Subject Matter Expert complete the *Data.SA Publishing Content Summary* sheet. The definition of all fields are available on the back page of the *Data.SA Publishing Content Summary Sheet, refer to Appendix E*.

11.2 PUBLISH

Once datasets are prepared for release they can be published by your agency on Data.SA the South Australian Open Data Portal.

The Data.SA publisher will require:

- Data.SA Publishing Content Summary
- data file or URL of Linked dataset
- interpretation metadata – separate open format file
- other resources such weblinks, reports, images.

For datasets to be published on Data.SA a *Data.SA Publishing Content Summary* needs to be completed and submitted to datasa@sa.gov.au.

11.2.1 Publishing Organisations

Each Agency has their own publishing organisation on Data.SA where you can view all your datasets published.

Agencies not already established as publishing organisation on Data.SA can be added at any time. South Australian Local Councils are welcome to become a publishing organisation on Data.SA.

Publishing organisation will need to provide the following information to datasa@sa.gov.au:

- agency name
- description of the agency/organisation
- website address
- linked logo - The logo must be available online and have a web address (url). The URL can often be obtained from the homepage of a website. From the homepage of a website, right-click on the website logo, click on properties and copy the displayed web address e.g. http://www.sa.gov.au/__data/assets/image/0010/370/logo_sagovau.png. Logos are usually jpg or png files.

11.2.2 Checking the Published Dataset

Once a dataset has been published on Data.SA the open data team will email the Subject Matter Expert. The Subject Matter Expert will need to check the accuracy of the dataset entry and that all links work. Each dataset created will have a unique Data.SA web address which may be added to your data inventory.

11.2.3 Publishing linked data

If you are publishing data on an Agency managed website/tool ensure that you display the approved licence statement and logo.

12 MAINTAIN - STEP 6 OPEN DATA PROCESS

The approval of open data commits the agency to maintain the data accordingly to decisions. Ensuring users of data have confidence in the supply and maintenance of the data is a critical element that affects the value of the data.

It is the Data Advocates responsibility for ensuring data is maintained. It is recommended the Data Advocate have oversight to ensure that the maintain process is implemented across the agency. An open data register may assist this process.

The Data Authority is responsible for the management and review of the data they publish including:

12.1 Maturity

Once datasets are released agencies should consider open data maturity of the dataset. This can be done by releasing new open formats for users. A review of the delivery method should also be considered if automation could improve the useability and reduce the manual resources required to maintain the data. The data quality should be considered including improvements to the data collection and timeliness of release of the data.

Section 11 provides a list of online resources relating to data fitness, quality and metadata.

12.2 Data Maintenance

Datasets must be refreshed according to their update frequency. This ensures user confidence and reliability in our data. If the Dataset is manually maintained processes to update the data should be in place, the open data process worksheet should document who has responsibility for performing refresh.

All datasets have a minimum required annual review of the data to check links, metadata and add new collected data files to the dataset.

If a new collection of data is to be published for an existing dataset, it is recommend that either:

- additional data files be added to the existing dataset, such as figures for 2014 are added to a time series data file
- a new file covering the complete refreshed dataset from inception to current can be published. E.g. Extract containing 2008-2012 data can be replaced with a new file containing 2008-2014 data.

This is not an amendment but additional refreshed data.

An open data inventory capturing key decision and responsibilities may assist the data maintenance process.

Note, once a dataset has been marked as Public and approved for public distribution, subsequent series of the dataset collected do not need to be re-approved on the provision that all elements of the data remain consistent (e.g. variables of data collected and structure of data). Care should be taken where the original data was reclassified for release through open data, so that only data purposely marked as Public is released.

12.3 Dataset Amendments

Once a dataset has been published, agencies need to be mindful that members of the public may be using the dataset so the impact of any changes needs to be considered. Once a dataset is published it cannot be removed. There is significant value to accessing historical data even if it is no longer maintained. Amendments to datasets will be treated as below:

- Minor changes to data content if errors are fixed or metadata changes can replace the dataset. Discovery data can be updated any time if it improves the users ability to understand the data
- Where a dataset has major structural changes, the old dataset needs to remain but can be marked as inactive. The new dataset can be added as a new dataset record or else added to the existing dataset as a new data resource (classification and Approval process will apply)

For major updates complete a new *Data.SA Content Summary Sheet* and include text to indicate the dataset is a replacement

Agencies should develop a process to ensure any system changes that affect the location of data e.g. changes to the URL of an agencies portal are communicated to datasa@sa.gov.au so that the dataset can be updated.

Dataset updates can be emailed to datasa@sa.gov.au until an agency publisher is established.

12.4 Notification of Changes

Before making changes consider if any applications or ideas have been created using the data and how changes may affect this data. If an agency has created an online forum etc. to communicate with developers, dataset changes should be communicated via these channels.

12.5 Community Feedback

Members of the public may contact the Data.SA team with requests or issues concerning datasets. This could include:

- requests for data
- requests for different dataset formats
- requests for additional information about the dataset
- questions about the meaning of the data
- errors encountered accessing the data.

Mechanisms will therefore be required to ensure that the feedback relevant to your agency datasets is directed to the Data Advocate for response. A response where required is to be issued within an appropriate time period.

12.6 Applications and Ideas

Data.SA provides mechanisms to share information about applications and ideas inspired by data. If South Australian open data is used in an application or idea it can be published on Data.SA. Visit the Apps and Ideas section of the open data portal.

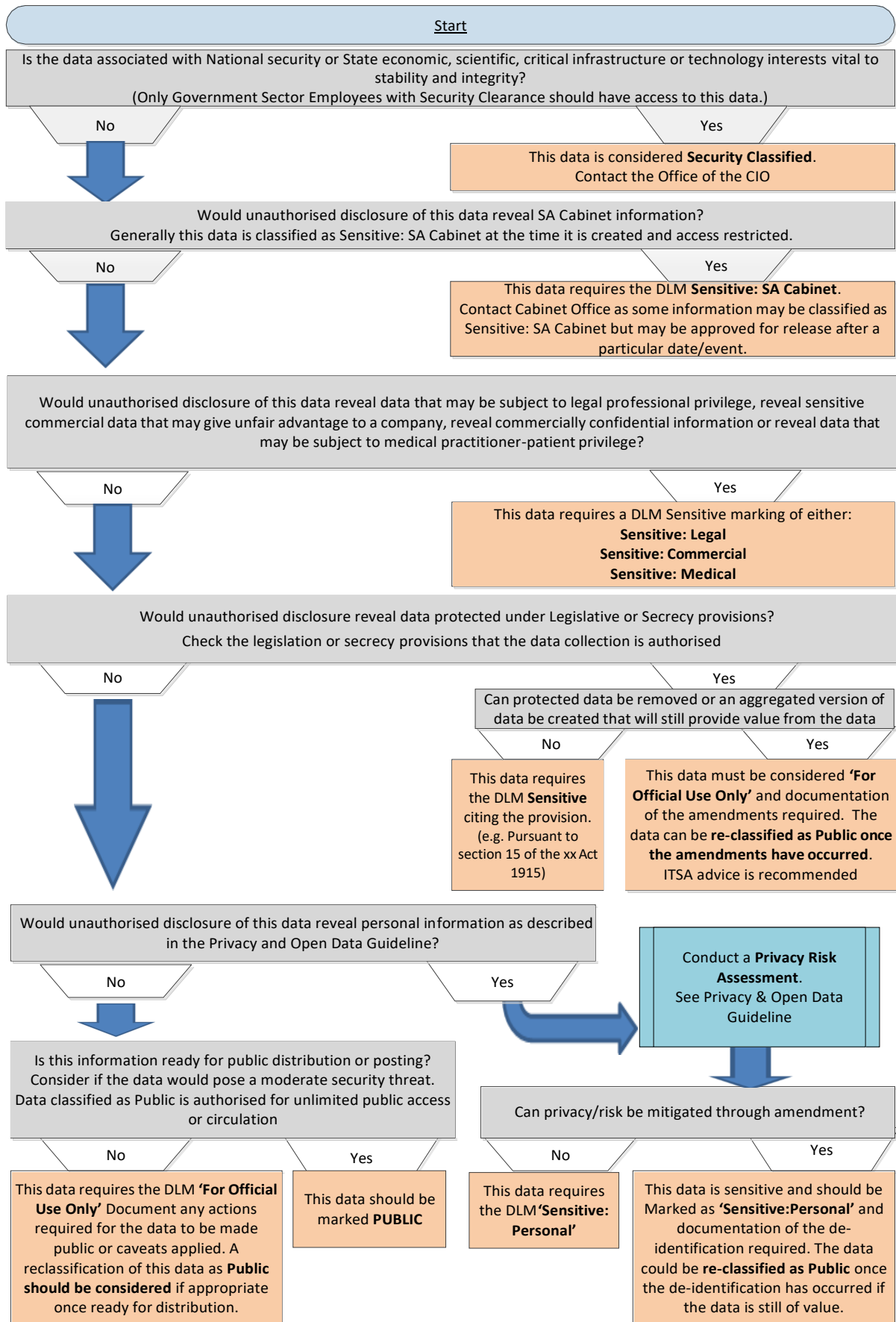News articles can also be published on Data.SA. Contact datasa@sa.gov.au

PREPARE, PUBLISH AND MAINTAIN SUMMARY

Prepare		
☐	Prepare Data	Extract current data or establish data delivery. Provide multiple formats of the datasets including the native form where approved.
☐	Transformation	Ensure data format is transformed and protection techniques performed. Compare outputs with approved sample.
☐	Interpretative Metadata	Provide in a separate file that is in an open format (e.g TXT HTML). Send with the *Data.SA Publishing Content Summary Sheet* to datasa@sa.gov.au.
☐	Additional Resources	Provide any other material which will help users understand what your dataset is about. Links to website, research papers, Images. Provide as links or separate files.
☐	Data.sa Publishing Content Summary Sheet	Complete a Data.SA Publishing Content Summary Sheet
Publish		
☐	Data.SA Publisher	Provide Data.SA publishing content summary, data file, interpretive metadata file and supporting resources to the Data.SA Publisher
☐	Published	Check published dataset on Data.SA. to ensure dataset l discovery metadata, data downloads, links and resources are correct. Email datasa@sa.gov.au with any amendments.
Maintain		
☐	Maturity	Improve dataset quality, metadata and increase the star rating over time.
☐	Data Maintenance	Refresh data according to update frequency.
☐	Data Amendments	Create a new dataset for major amendments
☐	Community Feedback	Respond to customer enquiries about the data, make improvements where applicable or provide additional metadata or dataset formats based on requests.
Open Data Process Worksheet		
☐	Record Decisions	Record all decisions made on the *Open Data Process Worksheet* which will become a record of your open data inventory.

13 ONLINE TOOLS AND RESOURCES

Tools and Resources
Open Data
What is Open Data
Understanding statistical language
Licensing
Australian Governments Open Access and Licensing Framework (AusGOAL)
Creative Commons - The Licenses
Frequently Asked Questions about the Creative Commons licences
Standards
List of Metadata Standards
AGLS Metadata Standards
ANZLIC Metadata Standards for Australian and NZ geographic data (PDF)
Statistical Data and Metadata Exchange (SDMX)
Quality
Data Fitness – A guide to keeping your data in good shape (PDF)
Data Quality Online Tool
Open Format
Examples of open formats
A guide that describes how to best format a resource (csv, xls, etc.) so that it will be machine-readable.
New Zealand Government, Applying the 5 Star Open Data high value public data.
API's
API Basics
7 Ways to make your API more successful
Geoscience
Location SA
Gazetteer of Australia Place Names
National Map
Geographical Tools
Real Time
Adelaide Metro Developer Information
Tabular Data Tools
Unstructured Text Data Tools
Graph (relationships and networks) Data Tools
Open Data Annual Competition
Unleashed Competition

APPENDIX A - DATA CLASSIFICATION AND MARKING DECISION DIAGRAM

Start

Is the data associated with National security or State economic, scientific, critical infrastructure or technology interests vital to stability and integrity?
(Only Government Sector Employees with Security Clearance should have access to this data.)

No →

Yes → This data is considered **Security Classified**.
Contact the Office of the CIO

Would unauthorised disclosure of this data reveal SA Cabinet information?
Generally this data is classified as Sensitive: SA Cabinet at the time it is created and access restricted.

No →

Yes → This data requires the DLM **Sensitive: SA Cabinet**.
Contact Cabinet Office as some information may be classified as Sensitive: SA Cabinet but may be approved for release after a particular date/event.

Would unauthorised disclosure of this data reveal data that may be subject to legal professional privilege, reveal sensitive commercial data that may give unfair advantage to a company, reveal commercially confidential information or reveal data that may be subject to medical practitioner-patient privilege?

No →

Yes → This data requires a DLM Sensitive marking of either:
Sensitive: Legal
Sensitive: Commercial
Sensitive: Medical

Would unauthorised disclosure reveal data protected under Legislative or Secrecy provisions?
Check the legislation or secrecy provisions that the data collection is authorised

No →

Yes → Can protected data be removed or an aggregated version of data be created that will still provide value from the data

No → This data requires the DLM **Sensitive** citing the provision. (e.g. Pursuant to section 15 of the xx Act 1915)

Yes → This data must be considered **'For Official Use Only'** and documentation of the amendments required. The data can be **re-classified as Public once the amendments have occurred**. ITSA advice is recommended

Would unauthorised disclosure of this data reveal personal information as described in the Privacy and Open Data Guideline?

No →

Yes → Conduct a **Privacy Risk Assessment**. See Privacy & Open Data Guideline

Is this information ready for public distribution or posting? Consider if the data would pose a moderate security threat. Data classified as Public is authorised for unlimited public access or circulation

No → This data requires the DLM **'For Official Use Only'** Document any actions required for the data to be made public or caveats applied. A reclassification of this data as **Public should be considered** if appropriate once ready for distribution.

Yes → This data should be marked **PUBLIC**

Can privacy/risk be mitigated through amendment?

No → This data requires the DLM **'Sensitive: Personal'**

Yes → This data is sensitive and should be Marked as **'Sensitive:Personal'** and documentation of the de-identification required. The data could be **re-classified as Public** once the de-identification has occurred if the data is still of value.

For assistance and explanation of each decision point contact your Agency ITSA.

APPENDIX B – OPEN DATA GUIDE TO SECURITY CLASSIFICATION

Agencies should assess the data to determine if it could be released publicly, based on their Information Classification and marking procedures as required by the South Australian Government Information Security Management Framework (ISMF). Visit the Department of Premier and Cabinet website to the view the ISMF Policies and Standards. The Information Privacy Principals Instruction (IPPI) also guides information privacy in the Government of South Australia. View the IPPI (PDF 241Kb).

Data can only be marked as Public if the release of this information does not cause any damage to the state, the government, an agency, commercial entities or members of the public.

Only data that has been assessed and marked as Public can be released as open data. Data that has a Protective Marking or Dissemination Limiting Marker (refer ISMF), for example **Secret**, **For Official Use Only** or **Sensitive**, cannot be released as-is, however may be reclassified if appropriate declassification and/or risk mitigation activities are undertaken.

Responsibility and accountability

Chief Executives
Agency Chief Executives are accountable for all security matters within their agency. The Chief Executive must authorise the disclosure of all official information to the public.

Data Authority
The Data Authority is responsible for ensuring that classification/marking or re-classification/re-marking of the data is undertaken and that appropriate input is obtained from the data subject experts and information users.

Executive Peer Review
Executive Peer Review provides risk mitigation to ensure the public release of a dataset does not inadvertently put another part of the business security at risk or disclose information that could lead to identification of a person (mosaic effect).

Agencies are encouraged to conduct an Executive Peer Review of all dataset classifications and markings to confirm the dataset meets the requirements of Public, prior to release. Executive peer review is recommended if data has been manipulated to mitigate risks e.g. when personal information has been de-identified.

The recommended approach for an Executive Peer Review is to engage your Information Technology Security Adviser or Data Advocate to circulate open data candidates to all Executives within an agency for review and comment. A sample of the data should be provided and summary of how any data risks that have been mitigated e.g. de-identification techniques applied.

Agency Information Technology Security Adviser (ITSA)
Agencies should contact their Information Technology Security Adviser (ITSA) for further advice and guidance on agency specific information classification and marking procedures and guidelines. The ITSA should be consulted where appropriate in classification and marking decisions and can provide advice on threats and risks.

Security Classification and Markings Decisions

Data than can be made open must be assessed as meeting the requirements to be marked **Public.** Data marked as **Public** is authorised for unlimited public access and circulation such as agency publications, data download sites and websites.

Some data can be re-assessed and re-marked as **Public**. The data may require redactions, amendments or manipulation to ensure sensitive elements are removed so that there is no damage or potential damage to the government, business or members of the public. The Data Authority is responsible for re-classification and re-marking and should engage their ITSA to assist with these decisions and activities.

Security classification and marking decisions may identify data that contains some elements that needs to be protected or de-identified, please [view the Privacy and Open Data Guideline](#) for information on how to mitigate these risks.

Detailed instruction on how to protect sensitive elements of the data being considered for re-classification, is required to be documented in the open data process approval. An Executive Peer Review is recommended for data once these risk mitigation techniques are identified and applied in a sample of the modified dataset.

A [Data Security Marking Decision Diagram](#) based on the ISMF is provided (refer appendix A) to assist you to classify and apply appropriate markings. This following guidance is to be used in conjunction with this diagram to assist the user to understand each decision.

Security Classification and Markings Decision Guidance

1. Is the Data Security Classified Information?

National Security Classified information is any official information about, or is associated with Australia (including states and territories):

- security from espionage, sabotage, politically motivated violence, promotion of communal violence, attacks on Australia's defence system or acts of foreign interference
- defence plans and operations
- international relations, significant political and economic relations with international organisations and foreign governments
- national or state interests, that relate to economic, scientific or technological matters vital to Australia's stability and integrity.

This type of information will be classified Confidential, Secret, or Top Secret as per *S19.7 ISMF. Only Government employees with a Security Clearance and need to know can access this information.*

> Where a Data Authority considers a dataset to require a Security Protective Marking they should consult with their ITSA about the classification and appropriate controls to be applied.
> **Not eligible for open data**

2. Would unauthorised disclosure of this data reveal South Australian Cabinet information?

Sensitive: SA Cabinet information includes:

- any document including but not limited to business lists, minutes, submissions, memoranda and matters without submission that is or has been:
  - submitted or proposed to be submitted to Cabinet
  - official records of Cabinet.
- any other information that would reveal:
  - the deliberations or decisions of Cabinet
  - matters submitted, or proposed to be submitted to Cabinet.

> Where a Data Authority considers data to be Cabinet information they must apply a marking of
> ***Sensitive: SA Cabinet***
> **Contact Cabinet Office or State Records** for more information on the release of this data.
> Certain Cabinet decisions or information may be releasable after a caveat date or media announcement (e.g. budget papers, major infrastructure initiatives etc.)

3. Would unauthorised disclosure of this reveal Sensitive Information?

Sensitive information includes:

- sensitive legal
- sensitive commercial
- sensitive medical - practitioner-patient privilege

a. Sensitive: Legal

Datasets containing information subject to court orders, legal proceedings or legal professional privilege, must not be made available under the Policy unless appropriate approvals are obtained.

Legal Professional Privilege protects all communications between a professional legal adviser (a solicitor, barrister, or attorney) and his or her clients from being disclosed without the permission of the client.

> Where a Data Authority considers data to be subject to legal proceedings they must apply a marking of ***Sensitive: Legal***
> **Not eligible for open data**

b. Sensitive: Commercial

Commercially sensitive data is any information, whose compromise could affect the competitive process and provide the opportunity for unfair advantage including:

- information concerning the trade secrets of any person
- information (other than trade secrets) that has a commercial value to any person
- any other information concerning the business, professional, commercial or financial affairs of any person
- likely to harm the business's commercial advantage in the marketplace.

Sensitive data should not preclude the release of other valuable data from being released from a dataset. It may be possible to remove data that is commercially sensitive from a dataset and this new dataset could be assessed for open data purposes. Controls should be put in place to ensure sensitive information is removed. Refer to Section 27 Freedom of Information Act (SA)—Documents affecting business affairs for more information.

If classification and marking decisions identify data that contains some elements that needs to be protected or de-identified please view the Privacy and Open Data Guideline for information on how to mitigate privacy risks.

Note: you are assessing the content of the data and not any contractual arrangement for the provision of data when determining the sensitive nature of the data.

> Where a Data Authority considers data to be commercially sensitive they must apply a marking of ***Sensitive: Commercial***
> **Not eligible for open data**
>
> Document whether data elements can be protected or de-identified as per the Privacy and Open Data Guideline.
> Once protection techniques are applied, the Data Authority can re-asses and re-mark the data as Public.
> ITSA advice and an Executive Peer Review of this data may be required.

c.   Sensitive: Medical

Health professionals and services are under a strict ethical and legal duty to keep patient information confidential. A health professional may only provide information to a person other than the patient for reasons of significant public interest or when required by *legislation.*

Healthcare enactments or other medical industry legislation may also apply.

> Where a Data Authority considers data to be subject to *medical practitioner patient privilege* they must apply a marking on the data as ***Sensitive: Medical***
> **Not eligible for open data**

4.   Would unauthorised disclosure of this data reveal sensitive data protected under legislative or secrecy provisions?
Agencies operate under various legislative provisions or secrecy provisions that specify conditions for restricting access or release of data e.g. *Public Sector Act* 2009 or the *Taxation Administration Act* 1996. The Declaration of Open Data does not supersede existing legislation. Agencies may be subject to other legislation particular to their business that specifies conditions for restricted access and/or release of their datasets. Data Authorities are encouraged to seek advice from their in-house legal advisors in relation to specific legislation they administer and any restrictions that may preclude the release of data.

Consider any third party rights concerning the use of information collected that was identified in the Open data process - identification stage.

It may be possible to remove the restricted data from the dataset and this new dataset could be assessed for open data purposes. Transformation and data amendments controls should be put in place to ensure restricted data is not released.

If security classifications or marking decisions identify data that contains some elements that needs to be protected or de-identified please view the Privacy and Open Data Guideline for information on how to mitigate privacy risks.

> Where a Data Authority considers data to reveal sensitive data protected under legislative or secrecy provisions they must apply a marking on the data as ***Sensitive citing the provision***
> (e.g. Pursuant to section 15 of the xx Act 1915).
> **Not eligible for open data**
>
> Document whether data elements can be protected or de-identified as per the Privacy and Open Data Guideline.
> Once protection techniques are applied, the Data Authority can re-asses and re-mark the data as Public.
> ITSA advice and an Executive Peer Review of this data may be required.

5. **Would unauthorised disclosure of this data reveal personal information as described in the Privacy and Open Data Guideline?**

Information privacy in the Government of South Australian is guided by the Information Privacy Principals Instruction (IPPI) issued as Premier and Cabinet Circular No 12 to regulate the way personal information can be collected, used, stored, and disclosed by State Government agencies.

The primary risk to privacy in the release of government data is the identification of individuals or data that can be made into personally identifiable information through easily linking with other information.

Privacy and Open Data Guideline

The Privacy Committee of South Australia has released the Privacy and Open Data Guideline to assist the government to maintain high standards of privacy when making its data open by default. Personal information of private citizens will not be released through open data. View the Privacy and Open Data Guideline.

Privacy Risk Assessment

The Data Authority is responsible for ensuring that a Privacy Risk Assessment has been conducted to identify privacy risks, consider de-identification and detail transformation requirements.

Executive Peer Review is recommended if data has had protection techniques applied to mitigate risks e.g. when personal information has been de-identified.

The following *Privacy Risk Assessment Process* will assist agencies to identify risks.

PRIVACY RISK ASSESSMENT PROCESS

Does the Dataset contain personal information? Not clear?

- Is it reasonably likely someone can be identified from the data?
- What other data is available that could be linked to your dataset?

Yes – individuals can be identified.

No – the data does not relate to individuals.

Data can be published.

It is likely the IPPI prevents you from disclosing the data.
You will need to consider whether the data can be de-identified.

Assess the identification risk:

1. Undertake an initial assessment of the privacy risk that considers:

   - Specific unique identifying variables, such as name
   - Cross-tabulation of other variables to determine unique combinations, such as age income and postcode
   - Availability of other publicly available datasets and information that could be used for list matching
   - The likelihood of an individual being identified from the data ;and
   - The consequences of such identification.

2. Consider what methods can be employed to reduce the risk, such as

   - Removing identifiers
   - Pseudonymisation
   - Reducing the precision of the data; and
   - Aggregation.

3. Test selected method of de-identification on the dataset.

4. Re-assess the privacy risk.

Can individuals be identified from this data?

Yes – consider making further adjustments to the data and reassess risk

If it is not possible to mitigate the risk to an acceptable level do not publish the data.

Apply an appropriate ISMF Classification to the data Sensitive Private

No – data can be published.

Regularly review published dataset for privacy risks

For more information about a **Privacy Risk Assessment** refer to the Privacy and Open Data Guideline

> Conduct a **Privacy Risk Assessment** as per the Privacy and Open Data Guideline.
> Where a Data Authority considers data to include personally identifiable information they must apply a marking on the data of *Sensitive: Personal*
>
> Document whether data elements can be protected or de-identified as per the Privacy and Open Data Guideline.
> Once protection techniques are applied, the Data Authority can re-asses and re-mark the data as Public.
> ITSA advice and an Executive Peer Review of this data may be required.

6. **Is the information ready for public distribution or posting (Public)**

The Open Data Declaration requires Public data to be open by default and therefore "Official Use" of data should only be applied to information that is not for use or of benefit to the public at large.

Data may require caveats for confidentiality before release (e.g. Budget papers). In this case, the point at which the information will be entered in the public domain should also be specified. When this information ceases to need confidential treatment, agencies must continue to consider reclassification to public. It is recommended to continue the open data process for this type of data to plan the approach and seek approval for future release.

Some data may pose a moderate security threat. A moderate security threat is when government data, whose compromise could affect:

- the government's capacity to make decisions or operate
- the public's confidence in government
- the stability of the market place
- law enforcement operations, whose compromise could hamper or inhibit crime prevention strategies or particular investigations or adversely affect personal safety.

The government holds considerable information that was created for an official purpose, however on further review could be valuable information to stimulate innovation, the economy, and open access of information to the community. Data marked *For Official Use Only* can be reclassified as Public. Agencies should consider what steps are required to make data fit for use.

> Data that could pose a moderate security threat must be classified as *For Official Use Only.*
> Data that is classified as *For Official Use Only* may define a caveat date for publication or steps that would be required to make data fit for public use. For this data it is recommended to continue with the open data process to determine if it should be released as open data at a later stage.
> Once protection techniques are applied, the Data Authority can re-asses and re-mark the data as Public.
> ITSA advice and an Executive Peer Review of this data may be required.

Marking data as Public

This data is authorised for unlimited public access and circulation, such as agency publications and web sites and as open data. Marking data as Public indicates that the data is a candidate to consider for open data and the open data process should be progressed.

> Where a Data Authority considers data can be released for unlimited public access they must apply marking of *Public*.
> **Eligible for Open Data**

APPENDIX D - OPEN DATA PROCESS WORKSHEET

Government of South Australia

1. Identify	
Dataset Name	
Business Unit	
Description	

Roles	
Identify Data Authority (Name & Title)	
Identify Subject Matter Expert (Name & Title)	

Value

Government Value	Public Value	☐ Criminal Value
☐ supports strategic goals	☐ public demand	
☐ comparable with jurisdictions	☐ stimulates development, industry	
☐ sharing efficiencies & policy	☐ empowers business/citizens	
☐ other	☐ research or industry efficiency	

Third Party Rights – Are there

☐ Intellectual Property Rights   ☐ Copyright   ☐ Terms and Conditions

If yes, what terms can data be released, if any ?

2. Classify

Data Security Markings – is data marked as

☐ Security Protected   ☐ Sensitive: SA Cabinet   ☐Sensitive: Legal   ☐Sensitive: Commercial

☐ Sensitive: Medical   ☐Sensitive: Personal   ☐Sensitive.. (Detail Act)

☐ **Public** – Eligible for open data

☐ **Reclassified as** Public– subject to the following

    ☐ Release of data caveat until   /   / 20   or……………………………………………………...

    ☐ Protection techniques mitigate security or privacy risks (sample provided)

Protection instructions or decisions to note :

    ☐ Executive Peer Review completed

    ☐ Agency ITSA consulted in reclassification

Responsible Information Sharing

Availability classification:	☐ A1	☐ A2	☐ A3	☐ A4

Integrity: ☐ High Quality   ☐ Reasonable Quality   ☐ Poor Quality

Quality improvement required in: ☐ timeliness   ☐ coherence   ☐ accuracy   ☐ interpretability

 Decisions to note:

Integrity Classification

☐ I1	☐ I2	☐ I3	☐ I4

AusGOAL Classification

☐ CC-BY - Preferred	☐ CC-BY ND	☐ CC-BY NC ND
☐ CC-BY SA	☐ CC-BY NC	☐ CC-BY NC SA

Reason for more restrictive license:

3.	Approach

☐ Identify Data Manager (Name & Title)
☐ Data Manager consulted to ensure ongoing support

Open Format

Formats data is to be released in:

☐ csv	☐ xlm	☐ kmz	☐ API	☐ Other…………………………
☐ txt	☐ shp	☐ Json	☐ Native/raw data	…………………………..
☐ html	☐ kml	☐ gml	☐ RSS	

Format Transformation instructions: Notes to Action

Frequency of Release

| ☐ Daily | ☐ Weekly | ☐ Monthly | ☐ Yearly | ☐ Quarterly | ☐ As required |

Person (include position title) or process responsible for datasets ongoing release and maintenance:

Data Delivery Method

☐ Hosted on Data.SA (Manual publishing)
☐ Linked to agency portal, data service or automation service
Refresh process ☐ Automated delivery   ☐ Manual delivery
Details/comments of data delivery and refresh method:

Data Distribution Services

Is dataset a candidate for a data distribution service? ☐ Yes ☐ No
Detail/comments data distribution agreement:

4.	Approve

Data Authority

Name:
Title:
Date:
Signature: ☐ Approved / ☐ Not Approved

Data Advocate

Name:
Title:
Date:
Signature: ☐ Approved / ☐ Not Approved

5.	Publish and Maintain

☐ Prepare current version of data. Apply format transformation and/or protection techniques

☐ Interpretative Metadata Provided

☐ *Data.SA Publishing Content Summary* prepared

☐ Native Data Provided

☐ Dataset checked once published

Decisions to Note:

Data Coordinator to file completed Open Data Process Worksheet

APPENDIX E – DATA.SA PUBLISHING CONTENT SUMMARY

data.sa PUBLISHING CONTENT SUMMARY

Government of South Australia

Agency	
Dataset Title	
Description	
Business Unit	
Contact Email	*Generic Mailbox preferred*
Expose User Contact Info	☐ Yes ☐ No
Licence	*CC-BY preferred*
Data Status	☐ Active ☐ Inactive (no longer maintained/once off)
Keywords (Tags)	

Groups – *Select which group your dataset should be published to.*

☐ Arts, culture & history
☐ Business industry and trade
☐ Community information and citizen services
☐ Crime, justice and law
☐ Education, skills and learning
☐ Emergency and safety
☐ Finance and employment

☐ Government reporting and policy
☐ Health and wellbeing
☐ Housing, land, planning and infrastructure
☐ Sports and Recreation
☐ Transport, travel and motoring
☐ Water, energy, resources, environment & science

File Format(s)	
Linked Data URL	
Update Frequency	☐ Daily ☐ Weekly ☐ Monthly ☐ Quarterly ☐ Yearly ☐ As required
Temporal Coverage From	
Temporal Coverage To	
Geospatial Coverage	
Data Granularity	
Interpretative Metadata	*Metadata files to be attached as separate txt or HTML format. Do not detail here*
Other Material	*Website address, reports, images etc attached as separate file or links*

Agency Contact

Subject Matter Expert/Publisher	*Date:*	*Signature:*
Contact Details	*Phone:*	*Email Address:*

Instructions

Agency	Add your Agency Name. Datasets are grouped by Organisations on Data.SA e.g. all datasets released by the Department of Planning, Transport and Infrastructure can be found under their Organisation name.
Dataset Title	Provide the name of the dataset. The name should use plain language, is easily understood by the public and identifies the content of the data. Do not include Agency names, technical names or acronyms.
Description	Provide a detailed description on what the dataset is about e.g. why the data was collected, what it was designed to measure, any sources of errors, field descriptions. Acronyms should not be used. The description should be easily understood by members of the public.
Business Unit	This is the Business Unit that administers the dataset for the data authority.
Contact Email	Provide an email address for the Business Unit that will be the point of contact for all enquiries. A generic email address is recommended.
Expose User Contact Information	Should the user's contact details be made public? Researchers and Scientists may want their contact details published.
Licence	The AusGOAL licence [Creative Commons Attribution 4.0 Licence](#) (CC-BY) is the preferred licence as it is the most open. Licence restrictions should be avoided where possible but should not prevent release of data. Refer to the *Open Data Process Guide* for more information on the alternative licences.
Data Status	Select Active or Inactive. Active data has a program in place to ensure it is maintained and refreshed in the future. Inactive data is no longer being refreshed and/or has been replaced.
Keywords	Keywords are Tags which help people find your dataset via the search engine. Tags represent the key topics of your dataset. E.g. a dataset about Fire Hazards could have keywords such as fire, hazards, emergency, bushfires, CFS, MFS, which help discover a dataset.
Groups	Datasets are grouped on Data.SA by common themes or groups. For example datasets which contain historical information can be found under 'Arts, Culture & History'. Datasets about bushfires or earthquakes can be found under the group 'Emergency and safety'.
File Format	The data must be released in at least one open format e.g. CSV, XML, HTML, JSON, KML, GML, TXT, API. You can also release the native version of the data.
Linked Data URL	If you are linking data then please provide the URL of the data location. Leave blank if the data is being Hosted on Data.SA.
Update Frequency	Stipulate how often the dataset is updated e.g. daily, weekly, monthly, annually, as required etc.
Temporal Coverage From	Start of a temporal series in dataset. If only a point in time, then no need to fill in the Temporal Coverage To.
Temporal Coverage To	End of temporal series used in dataset.
Geospatial Coverage	Input one of the following 1) a point/polygon/bounding box 2) an administrative boundary API or 3) a reference ID from www.ga.gov.au/place-names.
Data Granularity	The level of granularity of data in time and / or place e.g. incident reports on a daily basis by location (optional).
Interpretative Metadata	Metadata files or field definitions should be provided in an open format such as csv or txt.
Other Material	Any other material that can be published with the dataset to provide users with more context on the dataset e.g. research papers, reports, websites.

APPENDIX F – REFERENCES

The Department of Internal Affairs, New Zealand, Prioritisation and Release of High Value Public Data for Reuse, Process and Guidelines, 24 August 2012.

NSW Government, Open Data Policy, Consultation Draft

Queensland Open Data strategy

The White House developed Open data project sourced 6/3/2014

Intellectual Property Guidelines for the Victorian Public Sector (Working Draft)

Copyright issues, Copyright on terminology and similar data, Christian Galinski, Infoterm, Vienna

Intellectual Property Policy, Government of South Australia, Principles, What is Intellectual Property.

South Australian Government Information Security Management Framework (ISMF), the Office of the Chief Information Officer, Government of South Australia, Information Security Management Framework, version 3.1.1

ISMF Guideline 7 – Asset Management

ISMF Guideline 8a – An approach to classification using the ISMF (PDF, 220KB)

ISMF Guideline 8b – New classification scheme for confidentiality of information and associated assets (PDF)

Legal Services Commission, Medical treatment an related issues Confidentiality, Sourced March 2014

Australian Governments Open Access and Licensing Framework

Creative Commons Org, Licenses

Office of the information commissioner, Government of Ireland, Commercially sensitive information, sourced 10 March 2014

Privacy and Open Data Guideline, the Privacy Committee of South Australia, Government of South Australia, 2014.

DataVic Access Policy Standards and Guidelines, State of Victoria through the Department of Treasury and Finance 2013

Australian Copyright Council, An introduction to copyright in Australia (Fact Sheet) May 2007

Wikipedia – open data formats