

Step 1 - Identifying data for release

1.1 What data does my agency collect and manage?

The public sector creates, collects and manages a range of *data*ⁱ in the course of its everyday operations, including through administration, delivering and monitoring programs and services, and research. Agencies should review their data assets, considering:

- 1.3.1 Existing data (current and historical data).** Some of this data may already be made available to the public, for example on the agency's website, in annual reports and publications, and in databases.
- 1.3.2 Future data collection and creation (new data).** This may include upcoming or future programs, projects, services, reports and other initiatives. Agencies should consider and apply the Policy to the development of systems that collect or create data, and any modernisation projects that update existing data systems.
- 1.3.3 Data collected by the agency's contractors, consultants and grant recipients.** Agencies should consider including provisions surrounding the collection of data, consistent with the principles of the Policy, in their contracts and agreements with contractors, consultants and grant recipients.

1.2 What data should I consider for release?

Agencies are encouraged to adopt a position of "open data by default". That is, considering all data for release to the public. However, there are instances where data is not appropriate or suitable for release and, thus, access should be restricted or precluded (outlined in Step 1.3).

From an implementation perspective, the focus is high-value data. That is, where there is some value, outcome or benefit for the State from opening access to that data. This includes:

- 1.2.1 potential value and outcomes for government** – including reduced silos and improved collaboration, better decision-making, new and improved solutions and services, and efficiencies and savings (including reduced duplication and tools and service delivery more quickly and at lower costs); and
- 1.2.2 potential value and outcomes for industry and the community** – including supporting the development of new industries and business (and jobs), increased entrepreneurship and productivity, improved research outcomes, and better business and community decision making.

This assessment will help agencies evaluate against the cost and potential implications of making that data open, decide how best to manage the data (including whether to release it and under what conditions), and determine priorities for release.

In the first instance, agencies should improve the *discoverability* and *usability* of existing datasets, prioritising those that:

- **are already publically available**, for example on the agency website or in public reports;
- **are considered high-value** (as outlined above); and/or

- **are in high demand by the public** – front-line helpdesk requests, Freedom of Information (FOI) logs, and feedback and engagement with the public are useful ways of gauging this.

WA Government Open Data Fact Sheet: Where's the benefit?

Other resources:

- [Australian Government – Open Data Toolkit – Assessing what datasets agencies should publish to improve productivity](#);
- [Global Open Data Index – Dataset overview](#)
- [New Zealand Government – Guidance and Resources – Open Data Case Studies](#)

1.3 Is the data appropriate and suitable for release?

A large amount of data collected by the public sector will be suitable for public release. However, agencies should consider the potential implications and risks of releasing data, weighing them up against the potential value or benefits of openness. There are instances where agencies need to modify the data prior to release, or restrict or preclude access, including:

1.3.1 Privacy – where personal or sensitive information is involved that can be identified for an individual, or may be involved as an unintended result of *data linking*. Data that is released must not be (or be able to be) associated with any individual.

- a. Before releasing data, agencies should carefully consider and address any privacy concerns. Agencies may be able to mitigate privacy risks by de-identifying the data, which means removing anything that can identify a person. However, caution needs to be exercised where disparate datasets, individually de-identified could be linked to re-identify an individual. The Government of Queensland's guidance on [Dataset publication and privacy](#) and [Dataset publication and de-identification techniques](#) is useful.

1.3.2 Security – because of the nature of the data or information. For example, data that is associated with national security or State interests vital to stability and integrity.

1.3.3 Confidentiality – arising because of the nature of the data or information itself or because a contractual arrangement has been made in relation to the data or information. For example, data that reveals Western Australian Cabinet information, is subject to medical practitioner-patient privilege, or is commercial-in-confidence.

1.3.4 Legal privilege – attached to certain legal advice.

1.3.5 Commercial – such as commercial-in-confidence, patent pending or intellectual property considerations.

1.3.6 Public interest – if there are public interest considerations against disclosure and, on balance, those considerations outweigh the public interest considerations in favour of disclosure.

Agencies should also consider their own specific legislative provisions relating to the release of data and information.

Step 2 - Preparing data for release

2.1 How should I approach licensing and charging?

Agencies will need to license their data appropriately, with clear terms surrounding *copyright* and use. Where possible and appropriate, a non-restrictive licence should be used to promote maximum dissemination and reuse of the data.

- 2.1.1 Existing licensing arrangements** – where this makes the data available according to the principles of the Policy.
- 2.1.2 Standardised licensing frameworks, such as *Creative Commons*** – where existing licensing arrangements are not suitable, or there is no existing licence in place.
- 2.1.3 Other approaches to licensing** – there may be instances where agencies are required to set out special copyright arrangements; legislation stipulates a particular copyright position; or a more restrictive licence to those provided under Creative Commons is required.
- 2.1.4 Intellectual property** – data may have inherent intellectual property and deliver more value and better outcomes for the State when access to the data is managed appropriately. Agencies should consider the [Western Australian Government Intellectual Property Policy 2015](#).

Wherever possible, agencies should make data available at no cost to users in order to maximise the potential for reuse. However, there may be legitimate instances where agencies consider applying a reasonable charge for the data. Agencies are encouraged to discuss with the lead agency for implementation (Landgate).

WA Government Open Data Factsheet – Licensing and charging considerations

2.2 How should I format the data to make it usable?

Data should be published in internationally recognised, open standard formats that make it easy to use and transform.

- 2.2.1 Non-proprietary and machine-readable** formats are ideal.
- 2.2.2 Data as collected at the source, with a high level of *granularity*** is preferred, however some data may be best published in *aggregate* or modified forms, for example where there are privacy considerations.

WA Government Open Data Factsheet – Creating reusable machine-readable data

Other resources:

- [the NZ Government's guidance on the W3C's 5 Star Open Data Model](#);
- [Open Data Handbook – File Formats](#); and
- [the UK Government Service Design Manual – Choosing appropriate formats](#).

2.3 How do I make the data easy to understand?

Data should be published with:

2.2.3 *metadata* to better enable users to understand the data and how best to use it; and

2.2.4 a statement regarding the data's purpose and *quality*, including sufficient information about any caveats or limitations with regard to the dataset. This will better enable users to decide whether a dataset is suitable for their purposes.

Resources:

- the [ABS Data Quality Statement Tool](#) which helps with assessing and reporting on the quality of data; and
- the [Australian Government Open Data Toolkit – How to use data.gov.au – Metadata](#)

2.4 How often (i.e. frequency) should I publish data?

Data should be as up-to-date as possible and made available to users in a timely manner. As data is updated, agencies should aim to make it available as soon possible, or on a consistent periodic basis.

Next steps...

Publishing data and evaluation

The next step is to publish the data and make it publically available. Data should be accessible through the Government of Western Australia online data portal (data.wa.gov.au) – that is, published on the portal where possible or, alternatively, on the agency's website with a link from the portal.

Agencies should also assess their progress towards *open data maturity* and monitor the benefits and outcomes of implementing the open data policy.

WA Government Open Data Benefits Management Plan

The lead agency for implementation (Landgate) will provide further information and assistance to agencies.

ⁱ Refer to **Appendix A of the Western Australian Whole of Government Open Data Policy** for the definitions of terms used throughout (indicated in italics).