Australian Government

Office of the National Data Commissioner

ONDC

Guide to developing a data inventory

A national approach to discovering and describing data assets

August 2023

GUIDE TO DEVELOPING A DATA INVENTORY

ISBN 978-1-925537-92-5
Guide to developing a data inventory (online)

Copyright Notice

Attribution

Use of the Coat of Arms

Contact us

Enquiries about the licence and any use of this publication can be sent to:
www.datacommissioner.gov.au/contact

Contents

Introduction

This Guide outlines a series of practical steps to support Australian Government agencies to develop an inventory of data they hold. A data inventory is a critical foundation for any organisation. It will help you use and re-use the data your agency collects – to maximise its value and better deliver on your organisation's mission. A data inventory is also key to protecting the data you hold. Your organisation will be better placed to meet its accountabilities, for example, under the *Public Governance, Performance and Accountability Act 2013*, the *Data Availability and Transparency Act 2022*, the *Privacy Act 1988* and the *Archives Act 1983*.

While this Guide is primarily for Australian Government agencies, it can be used by other organisations - big and small, and at all levels of data maturity – to develop their data inventory. The approach has been tested and refined after working with many Australian Government agencies to develop their data inventories. You can start at the step that best suits your organisation:

1. If your agency does not have a data inventory and it is your first time building one, you should start from step 1 of this Guide as it may be a large cultural change for your agency.
2. If you want to see where your agency is at now and where your agency wants to be at in the future, you can use the baseline assessment in step 1.3 and accordingly choose the relevant steps from the Guide based on the self-assessment.
3. If you have an incomplete data inventory, you can start at step 3 and continue with the rest of the steps.
4. If your agency already has an existing data inventory, you can proceed to step 6 to check that your metadata align with the ONDC Metadata Attributes Guide and if your data inventory is ready to populate the Australian Government Data Catalogue.

Feedback from agencies that have developed their own data inventories highlighted the importance of investing time in getting the foundations right to ensure the long-term success of a data inventory. This includes:

- having an understanding of your agency's data assets
- assigning a designated project team/individual responsible for establishing the data inventory
- establishing support and buy-in from senior leaders early
- partnering with the IT area and engaging with data stewards early
- designing your approach for the future from the beginning - think about how you will maintain the data inventory and 'next steps' for data inventory improvement and expansion
- the extent to which you engage, empower and upskill staff to ensure the data inventory is used, maintained, and remains relevant into the future.

How to Build a Data Inventory

Overview

Figure 1 below outlines a process to establish or enhance a data inventory and is designed to continually support any agency, regardless of data maturity level. The iterative nature of the process is demonstrated by how an agency should circle back from step 6 to step 4 after establishing or enhancing an existing data inventory to allow for continuous improvements and updates.

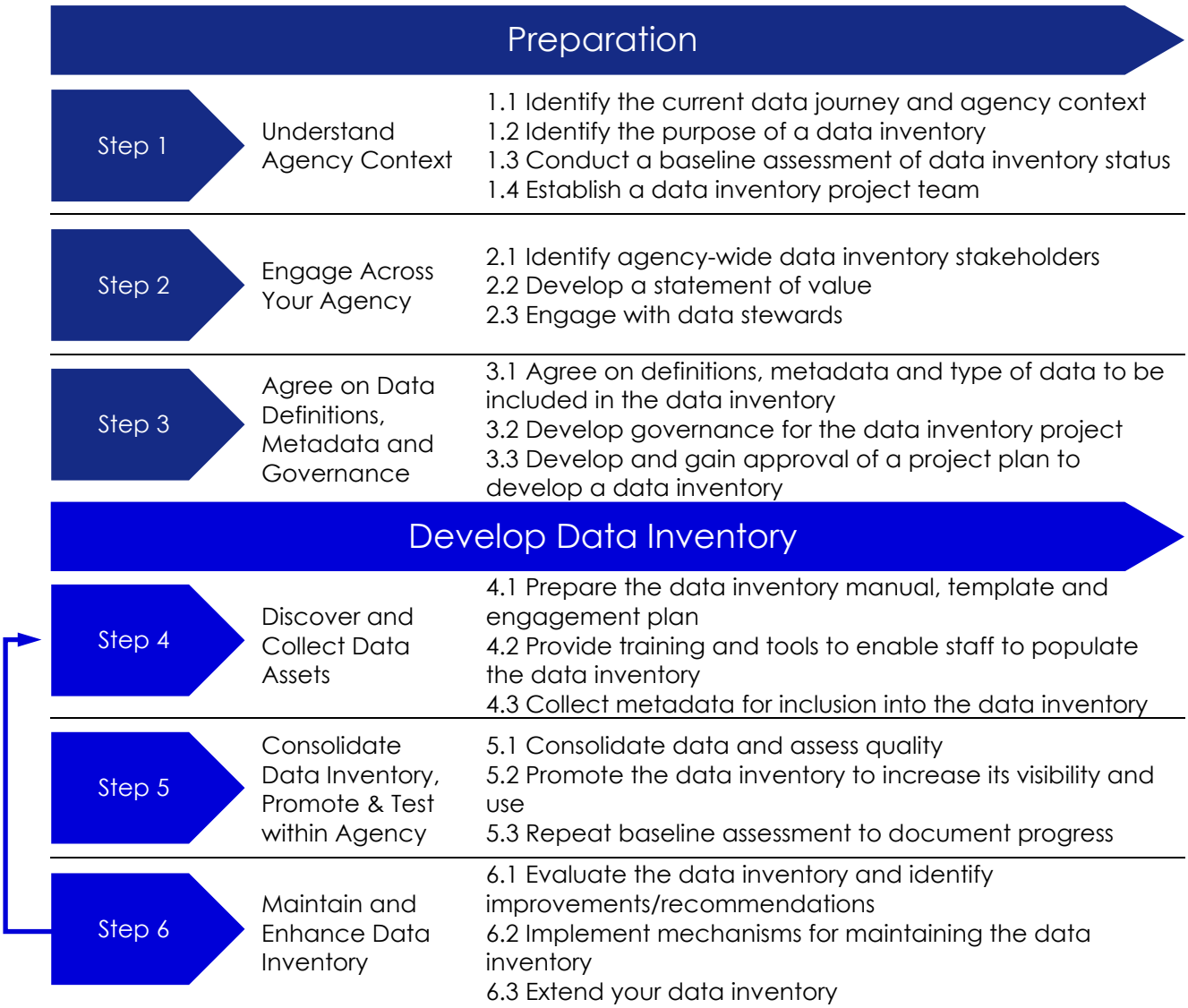
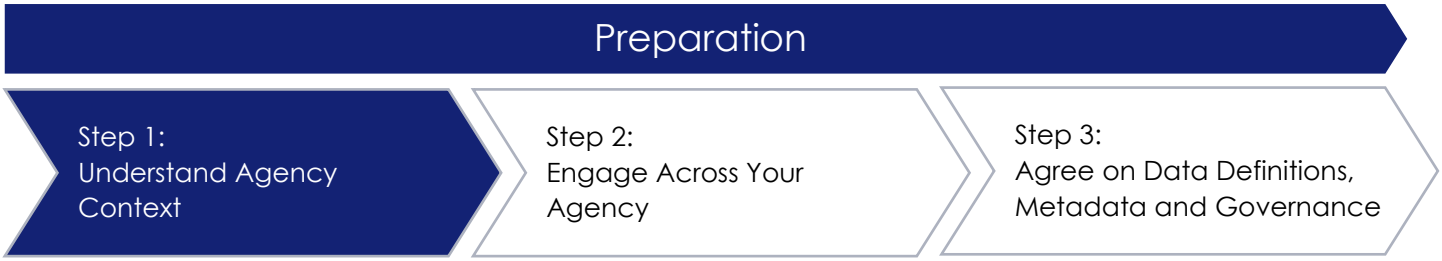


Figure 1: The six key steps to developing a data inventory.

Step 1: Understand Agency Context



Description

One of the most important outcomes of a data inventory project is improving the visibility of your agency's data assets. What products or systems you use to deliver this will vary depending on the size of your agency and its needs.

Understanding your agency's context and assessing your agency's baseline are key steps to demonstrate the need for a data inventory and to identify goals for the project. These will help secure the support of decision makers in your organisation for the project. This support will be important because you will need the resources and authority to request input from across the entirety of the agency to develop a data inventory and ensure the work is a business priority.

Key Tasks

1.1 Identify the current data journey and agency context

Developing a holistic understanding of your agency's data requirements and data journey is critical to designing an approach and determining the resources required to undertake a data inventory project. This may involve:

- Conducting a scan of:
    - your agency's priorities and strategies
    - your agency's operational and business plans
    - the policy and legislative environment your agency needs to comply with including:
        - the *Data Availability and Transparency Act 2022*
        - the *Archives Act 1983*
        - the *Public Governance, Performance and Accountability Act 2013*
        - the *Privacy Act 1988*
        - Policy 8 of the Protective Security Policy Framework
        - other privacy and security provisions
    - relevant reviews and audits, for example those relating to data, ICT and cyber security
    - risks for the agency.
- Understanding whole-of-government directions, strategies and election commitments.
    - Is your agency meeting its obligations and aligning its practices with those outlined in government strategies, such as the DATA Scheme and Data and Digital Government Strategy (2023)?
- Identifying if your agency needs to create a data inventory or has an existing data inventory.

    - Building on existing work and leveraging what already exists in your agency rather than starting over will set you up for the best chance of success.
    - Insights from existing data inventories or data catalogues, or previous attempts to deploy them, are valuable to identify gaps and inform the deployment approach.
- Comparing the metadata used in your existing data inventory (if one exists) with the ONDC Metadata Attributes Guide.
- Meeting with counterparts in other departments and jurisdictions to consider how others have created their data inventories.
    - This can be helpful in persuading others in your organisation of the merits of a data inventory and your approach to building one by providing concrete examples.
- Drafting a summary of your agency's current strategic and operational context, what is the authorising environment for the project, and (with steps 1.2 and 1.3) who will be affected/benefit from the data inventory and what problems the data inventory will solve.
    - It may also be worthwhile to consider any upcoming organisational changes (such as Machinery-of-Government changes) that could impact the development of a data inventory and to plan the project around this.

1.2 Identify the purpose of a data inventory

A data inventory provides many benefits to agencies by providing a single 'source of truth' about their data holdings and enabling efficient and effective use and reuse of data. The increased visibility of data assets within an agency can also help better protect your data, for example, by mitigating cyber security threats and reducing data breaches which may involve personal information. Furthermore, a well-managed data inventory helps agencies meet their accountabilities such as those identified in step 1.1.

To gather evidence to support change and investment in a data inventory you could:

- Engage with your agency's data leaders, IT specialists, data analysts and policy or service delivery staff about:
    - Their experiences in metadata management and, more broadly, data governance and management.
    - Issues and barriers that staff face when trying to find, use, or manage data within the agency.
- Develop a clear problem statement which clarifies the barriers to making your agency's data discoverable, understood, and shared.
- Prepare an agency data "state of play" report. This brings together the known challenges in data discovery, sharing and use, the agency's data culture and practice, staff and systems capabilities, and legislative requirements.
    - Note: this work may have already been conducted through the agency's data strategy work.

1.3 Conduct a baseline assessment of data inventory status

A baseline assessment will measure the data inventory status within your agency, evaluate asset discoverability and assist in setting agency goals for the data inventory project.

You should conduct this assessment again at step 5.3 to measure the agency's progress and change, and to inform continuous improvement.

You can use the Data Inventory Baseline Assessment, found in Appendix D, for this purpose.

1.4 Establish a data inventory project team

Having a clearly visible and contactable team or individual that is responsible for coordinating and leading the agency's data inventory creates a central contact point to provide information and advice on the data inventory project and establishes clear accountability for the project. If no one is responsible for coordinating and securing support for the data inventory, the project risks confusion, inconsistency, inadequate authority and a lack of interest and direction.

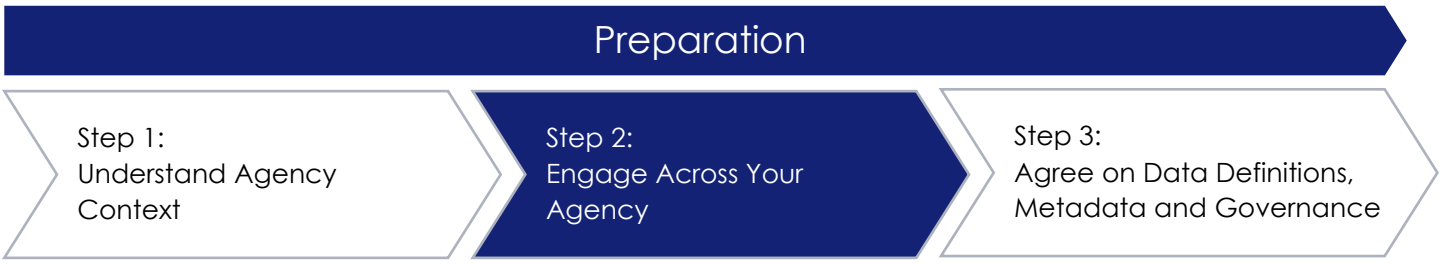
The role and responsibilities of the data inventory project team may vary. For example, it might sit with data governance, ICT or a knowledge management area. It is important to consider who is best placed in terms of relevant specialist capabilities and skills, authority and influence to successfully promote and deliver the data inventory project.

As part of this, you may need to seek agreement as to which team/individual will be responsible for the ongoing management and maintenance of the data inventory, and to agree on ongoing resourcing for the team.

Outputs and Outcomes from Step 1

- ☐ Understand your agency's context and needs
- ☐ Complete the Data Inventory Baseline Assessment and set goals
- ☐ Form a data inventory project team

Step 2: Engage Across Your Agency



Description

In planning the data inventory project, it is important to identify early those across your agency who will be affected by the project. Input from a range of stakeholder perspectives is a cornerstone of a well-designed project and will influence the design, implementation and future maintenance of the data inventory. Consider stakeholder requirements, issues, and concerns, their role in establishing the data inventory, and ongoing interaction after the data inventory is delivered. This will help you to decide how and when you engage with them, and what information and guidance you will need to provide to gain their buy-in and direct involvement.

If you are starting a data inventory for the first time, it can be useful to view this as a cultural change initiative. To be successful, a data inventory project will need to work across the entirety of your agency.

Key Tasks

2.1 Identify agency-wide data inventory stakeholders

Stakeholders for a data inventory can have varying data skills and interests and be diverse: from senior executives, data champions, data stewards, customers or users, IT, internal data groups and any area that holds data. To identify data inventory stakeholders, you could:

- Review your agency organisational chart.
- Hold working sessions within the data inventory project team to develop an initial list of agency teams.
- If required, hold workshops with key business and functional areas to gain a deeper understanding of their internal workings. Determine whether the team produces (gathers and collects) data, uses data, or both.
- Develop a stakeholder map.
- Utilise stakeholder analysis tools, such as a power-interest matrix, which can be a useful way to work out who are your stakeholders and the actions needed to align their goals with the project.

High power and high interest stakeholders, such as a senior data leader or Chief Data Officer, will be more important for getting buy-in for the project. In contrast, employees in areas which hold data might be considered low interest but can have high power. They may not be focused on the data inventory project but are critical to its success; you will need their input to ensure the data inventory represents your agency's collective data assets.

2.2 Develop a statement of value

There are many benefits of having a good data inventory. A statement of value will help articulate the advantages of a data inventory for your agency and why it should be prioritised.

For example, the Digital Transformation Agency (DTA) has identified the following benefits of having a data inventory:

- Helps identify key data assets across the agency, their locations, and owners. This will improve the quality of data and access to it for all stakeholders.
- Supports a core priority of the DTA Data Strategy around Governance, Asset Discovery, Data Landscape and Data Management.
- Helps improve management of data requests and data sharing agreements under the DATA Scheme.

The benefits articulated in a statement of value should be tied to the agency goals and priorities you have identified in step 1.1.

2.3 Engage with data stewards

If you are a senior data leader like a Chief Data Officer you may already have the authority and resources to initiate projects like a data inventory. Even so, it is still useful to gain the broad support of others across your agency so the work is seen as an organisational priority that requires ongoing maintenance and improvements.

Those who are responsible for managing data assets (such as data stewards) are key partners in the development and ongoing management and maintenance of an agency data inventory. To help gain their buy-in and direct involvement throughout the life of the data inventory, it is good practice to regularly engage with staff who work with data and/or are responsible for agency data assets (e.g. through workshops, meetings or via emails) including:

- Providing information of the upcoming data inventory project and the importance of a data inventory.
- Seeking their agreement to participate in the data inventory project, this may include their help in collecting and consolidating data in the first iteration of the data inventory project.
- Getting their feedback on the data inventory project plan and statement of value.
    - A schedule will need to be defined with the business areas participating in the data inventory project to minimise major clashes with their timelines and priorities.
    - A contingency plan should be developed in the event of major changes and significant priorities emerging in a participating business area. For example, are there other business areas and data stewards you could work with instead?
- Seeking data stewards who could become champions for the data inventory within the business areas.

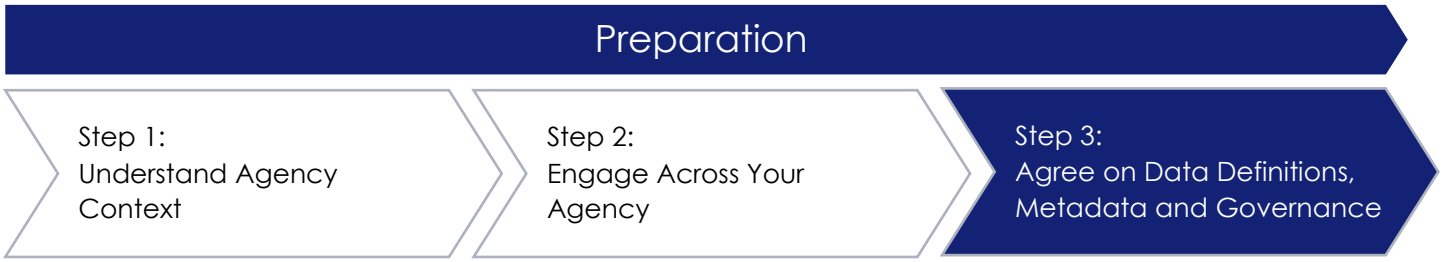
Outputs and Outcomes from Step 2

☐ List of stakeholders

☐ Statement of value

☐ Start early engagement with data stewards

Step 3: Agree on Data Definitions, Metadata and Governance



Description

This step involves agreeing on agency data definitions, metadata and governance arrangements for your data inventory so data stewards and other stakeholders have clarity of their roles and processes to support the data inventory, the scope of data assets for inclusion in the data inventory and the information that needs to be provided in the data inventory.

In addition, you will need to consider how you will go about developing a data inventory. This could take the form of a project plan.

A good project plan serves as a reference guide for your project. Your plan will help the data inventory project team prioritise their activities, remain on track and achieve clearly articulated objectives. It will establish project expectations for key stakeholders and set clear roles and responsibilities for the project team, senior executives and business areas.

Key Tasks

3.1 Agree on definitions, metadata and type of data to be included in the data inventory

To inform the direction and content of the data inventory, you will need agreement on definitions (including the definition of a data asset), the metadata requirements used to populate the data inventory, initial thinking on the form of the data inventory (e.g. it can be a simple list of data assets in SharePoint or Excel) and what data assets are within scope or out of scope for inclusion in your agency's data inventory.

For the definitions, it is important that they are clear, jargon-free and are tailored to your agency. These will provide clarity to data stewards and other in your organisation on topics such as what the characteristics of a data asset are, whether a data asset should be included in the data inventory, what information is required to comply with the agreed metadata requirements and so on. The ONDC Definitions Guide (Appendix A) may be useful as a starting point for this step.

Also, agencies should incorporate all 26 metadata attributes detailed in the ONDC Metadata Attributes Guide into their data inventory, as these are best practice for data management, data discovery and reuse of data. Your agency may choose to include additional metadata attributes as these may be required by legislative, program or reporting needs, or are specific to the type of data collected by your agency. Alongside this, you may develop a vocabulary and glossary for selected metadata attributes to drive consistency.

Most importantly, the metadata your data inventory should take into consideration the different needs of people across and beyond your organisation. Data users searching for data related to their work will want information on data assets that will help them quickly

identify whether the data meets their needs (e.g. based on the description, geographic coverage, etc.). Knowledge management or data governance areas will be interested in information that supports best practice data and records management (e.g. disposal class, date modified, etc.). Others may be interested in how the metadata support corporate reporting activities (e.g. knowing the business owners of data assets, etc.).

It is also critical to have very clear guidelines for determining which data assets held by your agency are within the scope of the data inventory:

- The definition of a data asset can be broad or as narrow as your agency requires.
- In the first iteration of the data inventory project, you may choose to include everything in the data inventory or you may prioritise data assets based on the criticality or frequency of requests of the data asset.
- To comply with the DATA Scheme, and respond to data sharing requests quickly, ensure you include data assets for which your agency is the Data Custodian. For example, if you value add or create new data assets from existing data assets, you become the Data Custodian of these products (even if you were not the creator of the existing assets) and so should include them in your data inventory. The definition of Data Custodian can be found in the ONDC Definitions Guide (Appendix A), as derived from the *Data Availability and Transparency Act 2022*.
- Third party or purchased commercial data assets should be considered within the scope of the data inventory to enable discovery and reuse of these data assets within your agency.
- Furthermore, consider which data assets may have value to other agencies, researchers and analysts or may be important to the broader public. Any data assets which your agency receives or is likely to receive data sharing requests for should be considered a high priority for inclusion in the data inventory.

When deciding on the definitions, metadata and scope of the data inventory, it may be helpful to run workshops or one-on-one meetings with a broad range of others across your agency (including key business areas and data stewards). This will help ensure the data inventory is inclusive of your agency's business needs, as each agency is different.

3.2 Develop governance for the data inventory project

A governance and oversight structure for the data inventory will ensure the data inventory project progresses and has accountability. You may need to:

- Establish who the data inventory project team will report to throughout the data inventory project and after it.
- Set up relevant steering committees, working groups and communities of practice to facilitate knowledge exchange.
- Ensure there is a secretariat function to support and arrange meeting and reporting requirements for developing the data inventory.
- Carefully consider arrangements for a permanent owner of the data inventory that would be responsible for the ongoing maintenance of the data inventory after it is established.
    - You may need to create additional materials (information, manuals, training materials) for handover of the agency's data inventory.
    - Develop maintenance policies such as time-based review or collection points to ensure the data inventory remains up to date.
    - Ensure you have an end-of-cycle policy in place that details the circumstances under which data is removed from the data inventory and approval processes.

If the initial project plan does not include resourcing or commitment to the data inventory's maintenance or improvement, your plan should include a timeframe to make another proposal or bid for these resources.

3.3 Develop and gain approval of a project plan to develop a data inventory

The process for securing support to develop a data inventory will vary from agency to agency, although it is likely you will need to develop a project plan and/or business case to present to the relevant decision maker. This could be a senior executive or a governance committee.

In developing the project plan, you should use the information gathered in step 1 and feedback from engaging with data stewards in step 2 to answer questions that senior executives and governance committees will have. The project plan should include the **objectives** of the project, the **key outcomes**, the **project team**, the **risks and how they are mitigated**, and the **ongoing maintenance** of the data inventory after it is established. A good data inventory is kept up-to-date and should have new data assets entered soon after they are created.

Outputs and Outcomes from Step 3

- ☐ Clear data definitions (for guidance see the ONDC Definitions Guide – Appendix A)
- ☐ Agreed metadata requirements for the data inventory
- ☐ Agency specific vocabulary glossary
- ☐ Project governance structure
- ☐ Maintenance and handover plans
- ☐ Project plan including a risk assessment

Step 4: Discover and Collect Data Assets

Develop Data Inventory

Step 4: Discover and Collect Data Assets

Step 5: Consolidate Data Inventory, Promote & Test within Agency

Step 6: Maintain and Enhance Data Inventory

Description

This step is where the data inventory project team will engage most closely with business areas to create a data inventory. It is also the step which involves the actual creation of the data inventory and collection of each data asset's metadata. The documentation used to support this step should be simple and accessible to help business areas understand the data inventory project and what it will achieve.

Key Tasks

4.1 Prepare the data inventory manual, template and engagement plan

Data Inventory Manual

The purpose of a data inventory manual is to provide staff with detailed information that they may need to engage with the data inventory project and ensure there is consistency in their understanding. The data inventory manual should answer questions and provide examples where relevant including:

- What is a data inventory?
- Why do we need a data inventory? What are the benefits?
- What data assets are in scope of the data inventory?
- How are data assets added to the data inventory? Who is responsible for this?
- What information is collected about each data asset (metadata)?
- How is metadata maintained in the inventory? Who is responsible for this?

Data Inventory Template

A data inventory template defines the structure of the data inventory and what metadata are included as decided upon in step 3.1. The ONDC has a Data Inventory Template (Appendix C) for agencies to use as a starting point. The agency data inventory template should include the 26 metadata attributes from the ONDC Metadata Attributes Guide but may also include additional agency specific attributes for collection. For some disciplines, such as health or geospatial data assets, recording additional metadata attributes would be useful to data users for finding the right data in the data inventory.

The data inventory template is critical because it determines your data inventory collection method. The template may be a questionnaire, Web-form or excel spreadsheet that is populated by business areas. Alternatively, you might use a customised data entry system such as SharePoint or a digital cataloguing software. Consider the data capability of different business areas when deciding what form your template will take. Clearly defining the format of each metadata attribute in the template will ensure consistency of metadata and makes the template more useful and user-friendly. The formats of specific metadata attributes may include controlled vocabularies, drop-down menus or standardised numeric formats, which should ideally be embedded in or enforced by the data inventory

template. Some metadata attributes may use a free text format, such as the description, and may benefit from a controlled vocabulary. A useful starting point for understanding or selecting a vocabulary to use is the Australian Research Data Commons' Research Vocabularies Australia.

Determining the formats and/or vocabularies may involve a co-design process with others in your organisation, or you may 'retro-fit' these to the metadata attributes after identifying patterns and commonalities across business areas' responses once the template has been filled out.

Engagement Plan

To facilitate continuous engagement with business areas, it is important to plan how you are going to engage them and seek their input throughout the project. An engagement plan will help formalise the involvement of different business areas during the project, including when you will approach them and the business area's contact information. This will help track the progress of the project. The engagement plan must also consider the data literacy across business areas to ensure communication is tailored to each business area.

Also consider compiling an engagement pack to give to business areas participating in the data inventory project. The pack should include the data inventory manual and template. It may also include FAQ sheets, information about the expected time and resource commitment needed for the project and a schedule of workshops or presentations that you may use for steps 4.2 and 4.3.

4.2 Provide training and tools to enable staff to populate the data inventory

Circulate the engagement pack prepared in step 4.1 and deliver workshops and presentations as required to ensure staff are adequately prepared to populate the data inventory in a consistent manner. The easiest and most reliable way to ensure that business areas are correctly identifying which data assets should be included in the data inventory and correctly populating metadata fields is to also meet with the business area to guide and assist them.

4.3 Collect metadata for inclusion into the data inventory

During the collection phase you will be dealing with business areas more directly.

This may mean emailing, calling or visiting business areas to guide them. Some approaches to collecting data inventory information include:

- circulating the engagement pack from step 4.1 and providing a live link or an offline copy of the data inventory template through managers and senior executives
- systematically working through the organisation contacting business areas directly
- holding meetings, presentations or workshops to fill out the data inventory and/or to train and support business areas to do it themselves
- deploying a member of the data inventory management team to business areas to support them.

Remember that the business area will have other competing work priorities. The data inventory project team will need to be proactive in engaging and supporting areas to complete the work. It may be useful to have an engagement tracker to track the progress of each team.

The project team should be visible and accessible to provide support with filling out the inventory template, as well as available for feedback so that updates can be made as needed. It is important to remember that the data literacy of business areas may vary so some business areas may require more support than others.

Outputs and Outcomes from Step 4

- ☐ Data inventory manual
- ☐ Data inventory template
- ☐ Engagement pack
- ☐ Populated data inventory

Step 5: Consolidate Data Inventory, Promote & Test within Agency

Description

The data inventory project team will be responsible for bringing together the information they receive from the different business areas into a coherent and consolidated data inventory. Once the data inventory has been consolidated, the data inventory should be promoted to the broader agency to encourage use of the data inventory and existing agency data assets.

Key Tasks

5.1 Consolidate data and assess quality

Consolidating the data in the data inventory may involve:

- Checking you have captured all in-scope agency data assets in the data inventory.
- Checking the consistency, accuracy and completeness of the populated metadata.
- Removing out-of-scope data asset entries.
- Ensuring all data asset entries comply with the metadata requirements agreed upon in step 3.1 (including the ONDC Metadata Attributes Guide).
- Reconciling duplicate data asset entries where a data asset is held by multiple areas. Also take the time to identify why the data was duplicated and address the underlying issue.
- Retro-fitting controlled vocabulary options to describe the data assets. You may also wish to align your vocabulary with an existing metadata schema, such as the DCAT schema.
- Finalising what the final data inventory product will be. As part of this, project team should consider:
    - Where the data inventory is stored (e.g. intranet, shared drive or folder)
    - Modification and access controls
    - Version control
    - Searching and filtering functionality within the data inventory
    - Inbuilt data asset hosting and transfer facilities
    - Whether it is compatible with an Application Programming Interface (API), for example, to enable automation or facilitate sharing more broadly to other initiatives (such as to the Australian Government Data Catalogue).

5.2 Promote the data inventory to increase its visibility and use

Once completed, the data inventory should be made available in a central location, such as your agency's intranet or a shared drive or folder. It will also need to be accompanied by appropriate explanatory documentation and contact details of the team who will be responsible for maintaining and improving it.

It is important that the data inventory is promoted internally within your agency to encourage the use of the data inventory. This could include:

- A 'launch' event and promotional activities by agency data leaders.
- Using internal tools to increase the visibility of the data inventory (e.g. ICT platforms, intranet, regular staff circulars, presenting at branch/divisional meetings).
- Documenting and promoting success stories and use cases to demonstrate the benefits of data sharing and integration to achieve organisational objectives. This ensures data inventory work is valued in the present and into the future, and positively contributes to a culture of data management and governance.
- Tailoring communications when engaging different business areas.
- Seeking feedback from stakeholders as part of this process, such as senior leaders and business areas. This can help improve promotion of the data inventory and inform steps 6.1 and 6.3.

5.3 Repeat baseline assessment to document progress

Repeat the same baseline assessment as in step 1.3 and compare your current assessment against your initial score and your agency's initial goals.

Repeating the baseline assessment will help your agency to acknowledge and celebrate achievements and improvements, as well as identify areas for further work (see step 6.1).

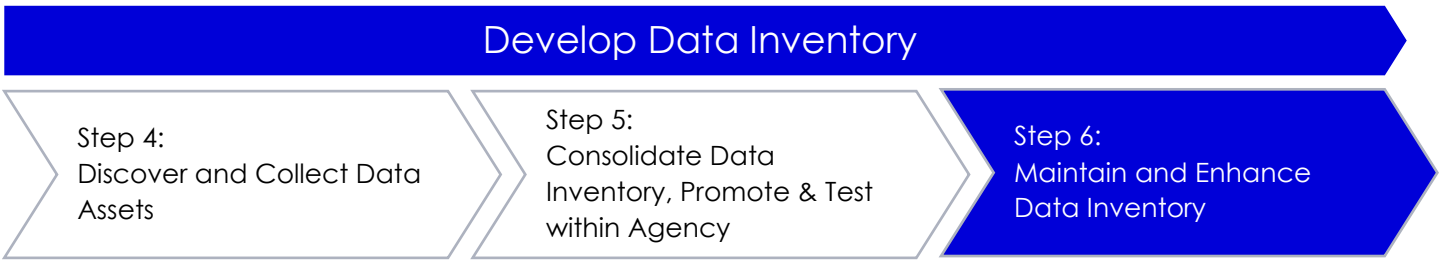
Outputs and Outcomes from Step 5

☐ A data inventory with consistent, high-quality metadata

☐ Completed a second Baseline Assessment

☐ Be able to participate and contribute to the Australian Government Data Catalogue

Step 6: Maintain and Enhance Data Inventory



Description

Evaluation, maintenance and enhancement of your agency's data inventory is an iterative and ongoing process. To keep pace with an ever-changing operating environment, it is important that the data inventory evolves to suit business needs.

Creating a data inventory for your agency is a significant achievement. Steps to ensure it is an enduring piece of business architecture that responds to agency needs include:

- Establishing mechanisms for maintaining the data inventory
- Analysing the data inventory and making recommendations to action any findings
- Reflecting on the effectiveness of the data inventory process
- Identifying next steps for continuing to expand and develop it.

Key Tasks

6.1 Evaluate the data inventory and identify improvements

Analysing the content of the data inventory will help to answer some basic questions about what data your agency holds, its characteristics and how your agency manages it. Sharing these insights will help reinforce the value of the data inventory and may inform improvements to agency data practices.

Questions to consider include:

- How many data assets are captured in the data inventory?
- How complete is the metadata of the individual data asset entries?
- Is the data inventory compliant with the 26 ONDC Metadata Attributes?
- Can the data asset entries from the data inventory be used to populate the Australian Government Data Catalogue?
- Does the data inventory capture all relevant data assets from business areas?
- Which themes or categories are most data assets associated with?
- What themes or categories appear to be missing or under-represented?
- Which data assets contain personal and/or sensitive information?
- What information appears to be collected in multiple data assets?
- What data assets do not have a clear business owner?
- Where are most data assets stored (e.g. in a system, on a shared drive, etc.)?
- Which data assets could be made publicly available or shared under restrictions?
- How is access managed for most data assets?
- What systems are required to support the data assets?

In addition, you should reflect on the process of developing the data inventory. Since the inventory must evolve over time, identify effective or ineffective processes and how they can be improved. Consider:

- How effective was the data inventory project team?
- Was there sufficient internal support for the project? Were appropriate resources and skills available to the data inventory project team?
- Were the right governance processes used to secure project endorsement?
- Were key stakeholders effectively engaged early?
- Were the benefits and statement of value appropriately articulated and sufficiently targeted?
- Was the documentation developed for business areas appropriate?
- Was the data inventory template user-friendly and easily populated?
- Were the metadata requirements appropriate?
- Is the technical system supporting the data inventory fit-for-purpose?
- Did the final data inventory meet expectations?
- Were the benefits delivered?
- How effective was the process of identifying and classifying data assets?

It is also important to recognise that data management and governance practices will continue to develop and change over time. The data inventory project team should review information collected in the above steps and develop findings and recommendations, in consultation with business areas, to enhance the agency's data inventory processes and outcomes. Any findings and recommendations related to data inventory or business areas processes should be documented in an evaluation report to ensure there is continuous improvement.

6.2 Implement mechanisms for maintaining the data inventory

The ongoing usefulness of the data inventory relies on it remaining current and accurate. Once you have published the data inventory, you should review and adjust any maintenance arrangements or handover actions established in your project plan from steps 3.2 and 3.3 and begin implementing them.

- Identify the area responsible for maintaining the data inventory as part of their work program.
    - Establish processes and guidance for how updates to the data inventory are communicated across the agency.
    - Business areas will need to take responsibility for updating the data inventory and responding to questions about access to relevant data assets. Training or engagement may be required to help business areas understand expectations of their role.
- Engage with business areas to regularly review their data asset holdings to ensure the data inventory is accurate.
    - Establish timeframes for how often reviews are undertaken.
    - Reviews may be more extensive after significant events such as machinery of government changes where the data custodian for many data assets may change.

Recording and managing data assets in the data inventory should align with your agency's information and record management policies and strategies, as a data inventory is another form of record and information management.

It is also a good idea to gain your organisation's support to embed data inventory maintenance in delegation instruments, frameworks and performance agreements to ensure the data inventory remains up-to-date and that staff understand their responsibilities.

6.3 Extend your data inventory

A data inventory is the foundation for improving data discoverability, use, management and sharing for any agency. With a data inventory, agencies can begin to do more with their data.

Some improvements include:

- Improving data discoverability by installing an agency data catalogue to make it easy for staff to search and find data.
- Making data management more effective and cost-effective by enabling dashboards for better data oversight.
- Prioritising resources to data assets according to value by setting up a data valuation framework and using it to apply data quality and resourcing standards.
- Connecting to the Australian Government Data Catalogue.

Improving data discoverability

Many agencies will benefit from turning their data inventory into a data catalogue. A data catalogue's biggest benefit is that it makes it easier for others in and beyond your organisation to find and use the data.

There are many off-the-shelf and open-source data cataloguing options available. The best option for your agency will depend on factors such as existing data maturity, IT infrastructure, and budget; there is no one-size-fits-all option. There are benefits to purchasing a vendor supported product but there are also benefits to using flexible, cheaper open-source software that can be tailored by an agency in-house.

When selecting data catalogue software, some important considerations are:

- your agency's ICT and cybersecurity needs
- whether it can integrate with other business systems and/or data infrastructure, such as the Australian Government Data Catalogue
- whether it can appropriately extract and/or maintain metadata for individual data assets
- whether it is intuitive to use (this should be judged primarily by users of the data catalogue).

It may be helpful to speak to other agencies that have developed or purchased a data catalogue to understand the advantages and disadvantages of different systems and approaches.

A digital data inventory or catalogue is key to streamlining or automating data management and governance activities. For example, agencies could:

- Set review periods for data inventory asset entries and automate review notifications.
- Create a permissions system to allow business areas to add and maintain entries themselves.
- Enforce data asset metadata standards and requirements.
- Record details for data provenance.
- Track data use.
- Automatically identify duplicate data in data storage systems.
- Automatically identify new data assets to include in the data inventory.

Outputs and Outcomes from Step 6

☐ Evaluation report with recommendations for improvements for the data inventory

☐ Review and adjust maintenance and handover plans

☐ Ensure data inventory is continuously maintained and improved

Additional information is available on the ONDC website.

Appendices

Appendix A: ONDC Definitions Guide

This document provides definitions of common terms used by the ONDC when working with Australian Government agencies to develop their own data inventories. Please note that the examples given are not necessarily definitions from the *Data Availability and Transparency Act 2022* but are meant to be demonstrative of how a definition can be applied in a data context. Based on these definitions, agencies can make more informed decisions about the scope of their data inventory and may choose to include a broad or a narrow set of data items.

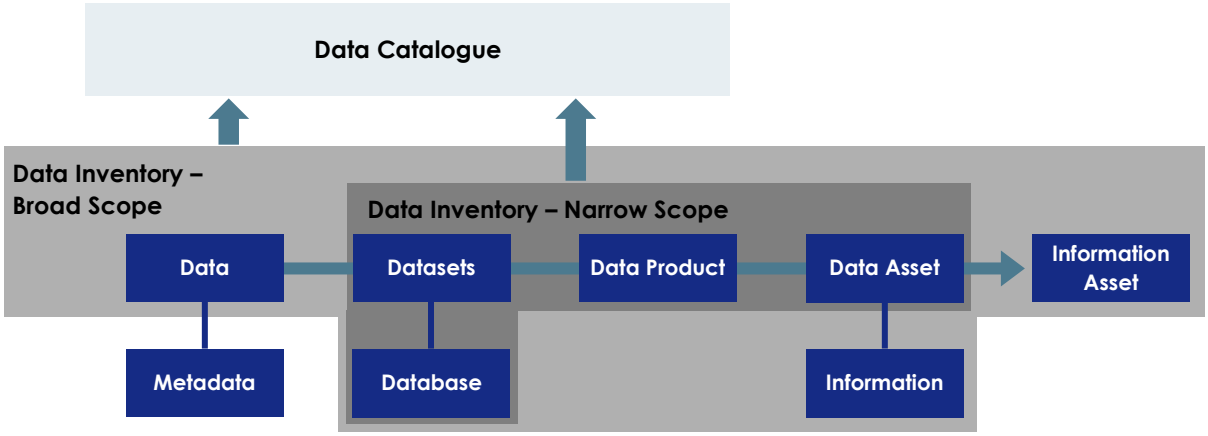


Figure 2: Illustrative example of how the scope of a data inventory may be determined.

Term	Description	Examples
Data	Data, as defined in the *Data Availability and Transparency Act 2022*, means "any information in a form capable of being communicated, analysed or processed (whether by an individual or by computer or other automated means)". More broadly, it refers to any collection of facts, figures, or information that can be processed, analysed and interpreted to derive meaning and insights.	Soil nutrient data collected from a specific location on a farm on a particular day. A consumer complaint against a trader.
Data Asset	A data asset is a collection of structured data developed for a purpose and that has inherent value to the agency or organisation. It may comprise of one or multiple more dataset(s) listed in the organisation's data inventory, deemed to be important and have the potential to create value for the organisation.	The soil nutrient data, rainfall data, temp data, etc. All complaints against all traders in the past year.
Data Catalogue	A data catalogue is a tool that enables a user to search and locate data for their specific needs. Its effectiveness relies on high quality metadata.	Research Data Australia is an example of data catalogue to find research data assets in Australia.

Term	Description	Examples
Data Custodian	Data Custodian, as derived from the *Data Availability and Transparency Act 2022*, refers to Commonwealth Government bodies, that are not excluded entities, who control public sector data. The Data Custodian has the authority to deal with the public sector data.	An Australian Government agency with a mandate to collect data.
Data Governance	Data governance is defined in The Foundational Four as the 'exercise of authority and control (planning, monitoring and enforcement) over the management of data assets'.	Agency guidance material, checklists and templates that inform how to manage privacy and respond to incidents.
Data Inventory	A list of data assets that records information (metadata) about the data assets. In some organisations the term **Data Asset Register** is used.	A data inventory can be a simple table or a specialised software to capture agency data holdings with complete metadata descriptions.
Data Management	The practice of applying data governance procedures in collecting, organising, protecting, and storing an organisation's data so it can be analysed for business decisions. More information is provided in The Foundational Four.	A data management plan covering all aspects of data usage including data acquisition, processing, storage, security, release, archiving and disposal.
Data Product	As defined by the Productivity Commission Inquiry Report on Data Availability and Use, 'transformed data products' are the result of extensive data processing and curation to increase the value of the data and prepare it for specific users or use cases.	Soil nutrient map based on spatial interpolation of data collected from several sites over a large area. A dashboard showing complaint trends against traders in the past year by geographical area.
Data Quality	The Australian Bureau of Statistics Data Quality Framework refers to seven features – institutional environment/context, relevance, timeliness, accuracy, coherence, interpretability and accessibility – to assess data quality so users can decide if a given data is fit for purpose to solve a specific problem.	Availability of key material to support correct interpretation of soil nutrient data such as data collection method, accuracy/sensitivity of instruments used, type of analysis, etc.
Data Repository	A central location in which data is stored, managed and made available to internal or external users. Unlike a Data Inventory or Data Catalogue, a Data Repository contains the actual data, not just metadata. In some organisations the terms **Data Lake** or **Data Stores** are used to refer to systems where data is held.	Internal or external web sites such as data.gov.au where users can search and download data or derived products, generated by an agency.

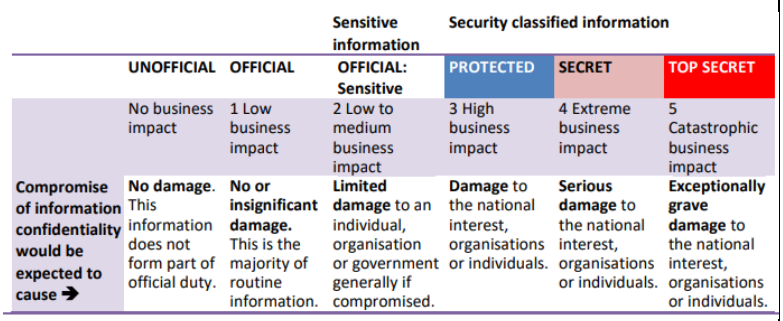
Term	Description	Examples
Data Steward	A defined role within agencies that is accountable for overseeing day-to-day management of data collected and held by an agency within a defined data domain or business area. In some organisations the term **Data Owner** is used.	Designated person who is responsible to manage particular data assets in an agency.
Data Vocabulary	Data Vocabulary refers to standardised terms with consistent semantic definitions. In some organisations the term **Data Glossary** or **Data Dictionary** is used.	The DCAT Vocabulary defines a set of vocabulary options useful for describing data assets in a consistent and standardised manner.
Database	A database is a specialised software capable of storing, querying, analysing, visualising and reporting data in a structured format. Typically, it is stored electronically in a computer system.	A relational database where soil nutrient data, temperature data, etc. are captured and managed. A Customer Relations Management (CRM) system that stores all consumer complaints including follow up actions taken.
Dataset	The National Archives of Australia defines a dataset as a structured collection of data generally associated with a unique body of work, a particular subject, or created for a specific purpose.	The soil nutrient data collected periodically over many years. All complaints against a trader in a year.
Information	Information is data that has been processed into a form (physical, oral, or electronic) that is meaningful to the recipient.	A report summarising the findings from a soil nutrient data analysis. A document that summarises the nature of the complaints against a certain trader in a year.
Information Asset	The National Archives of Australia defines information assets as records, information and data that are created, collected, received and kept as part of government business. Information assets listed in a register are typically grouped content that has been identified as being of significant value, risk or priority.	Current and archived reports of cornfield soil nutrient levels of all farmer's in a particular region.

Term	Description	Examples
Metadata	Metadata, in the context of data management, is additional information linked to a data asset which describes its important features. Metadata can be organised into **Metadata Attributes**. The ONDC defines 26 metadata attributes which are critical for data management and data asset discovery.	Metadata for the above examples may include: title (e.g. Soil nutrient on farm A), who can access the data asset/security classification (e.g. OFFICIAL). format of the data asset (e.g. CSV). Describes the period in which the consumer complaints were collected, the language of the customer complaints, etc.
Metadata Schema	A set of metadata components for a dataset, including metadata attributes, attribute types, definitions, vocabularies and data structures. Schemas provide procedural rules which ensure a standardised approach to metadata creation and management.	The DCAT schema designed to facilitate interoperability of open data by standardising metadata language.
Personal Information	Personal information, as derived from the *Privacy Act 1988*, means "information or an opinion about an identified individual, or an individual who is reasonably identifiable."	Name and contact information of the farm owner where the soil nutrient data was collected from. Name and address of complainant against a trader.

Appendix B: ONDC Metadata Attributes Guide

Core Attributes

Name	Definition and Guidance	Content (Format and Recommended Values)
Identifier	**The identifier of the data asset is specific and unique to the agency.** The identifier distinguishes the data asset as unique and different from another agency data asset. Ideally it is globally unique, such as a Digital Object Identifier (DOI), it may be a unique label used within the agency. It is key to finding the data asset and to ensuring that the specific data asset can be referenced without confusion.	Free text (max. 200 char)
Title	**The most common useful name by which the data asset is known** by your agency and by your largest perceived audience. The title should be relatively unique and can be created by following a naming convention used by your agency.	Free text (max. 200 char)
Description	**A descriptive statement of the data asset.** Easy to read information about the data asset; the purpose is to enable users to find, categorise and evaluate the fitness of a data asset to their needs. The Description attribute is also searchable and is typically several sentences long. It is useful to consider what keywords your potential audience may use to search for the data asset. This field could be supplemented by the attributes **Keyword** and **Purpose**.	Free text (max. 500 char)
Point of Contact	**The relevant contact for the data asset.** Identified contact to provide additional information related to the data asset. Ideally, a group email address or contact web page is provided because it is generic and enduring (preferable to an individual's contact). This minimises the need to regularly update metadata records.	**Email** (or **URL** to web form) for the point of contact

Name	Definition and Guidance	Content (Format and Recommended Values)
Access Rights	**Specifies access to the data asset.** Access will be based on the agency's privacy, security, or other policy approaches that apply to this data asset. Access can be: • **Open** - Data is publicly accessible online (registration may be required) • **Conditional** - Data is publicly accessible subject to certain conditions. For example: a fee applies; or the data is only accessible at a specific physical location • **Restricted** - Data access is limited. For example: during an embargo period; to a particular group of users; or where formal permission is granted. This attribute relates to **Security Classification** and **Sensitive Data**.	Choose term from: Open Conditional Restricted
Security Classification	**The security classification applied to the data asset as specified by the Australian Government Protective Security Policy Framework.** The originator of the data asset is responsible for applying the relevant Security Classification.  Protective Security Policy Framework (PSPF) policy 8: Sensitive and classified information (Attorney-General's Department) This attribute relates to **Sensitive Data** and **Access Rights**.	Choose term from: UNOFFICIAL OFFICIAL OFFICIAL: Sensitive PROTECTED SECRET TOP SECRET
Data Custodian	**The custodian(s) of the data asset.** The custodian is the agency who has the control of the data asset and has the authority for sharing and disclosure. The custodian may not be the publisher (see **Publisher** attribute). According to the *Data Availability and Transparency Act 2022: "An entity is a **data custodian** if the entity:* (a) *is a Commonwealth body; and* (b) *is not an excluded entity; and*	Free text e.g. The Office of the National Data Commissioner (ONDC) **Name of Agency/Department or Non Government Organisation**

Name	Definition and Guidance	Content (Format and Recommended Values)
	(c) *either:* (i) *controls public sector data (whether alone or jointly with another entity), including by having the right to deal with that data; or* (ii) *has become the data custodian of output of a project in accordance with section 20F."* The default may be your agency. If the asset's custodian is another agency, use a term from the Government Directory, NGO List or Research Organisation Register.	Choose Government Agency from: https://www.directory.gov.au/departments-and-agencies Choose NGO from: List of Australian accredited non-government organisations (NGOs) Choose Research Organisation identifier from: https://ror.org/
Keyword	**Word(s) or terms that describe the data asset subject matter.** These word(s) or terms describe the topic(s) covered by the data asset. It answers the question "what is this data asset about?" and supports data discovery. When selecting keywords, consider what search terms your users may choose when searching for the data asset. It is recommended to include at least one term from the Australian Governments' Interactive Functions Thesaurus (AGIFT) that covers words and terms related to Australian Government agencies' core business functions and activities. Also include words such as Indigenous, Disability or Gender if appropriate to better support the Government's priority data activities. The top-level terms in AGIFT are: Business Support and Regulation Civic Infrastructure Communications Community Services Cultural Affairs Defence Education and Training Employment Environment Finance Management Governance Health Care Immigration Indigenous Affairs International Relations Justice Administration	Free text e.g. Education and Training, Early Childhood Education, ... Australian Governments' Interactive Functions Thesaurus (AGIFT)

Name	Definition and Guidance	Content (Format and Recommended Values)
	Maritime Services Natural Resources Primary Industries Science Security Sport and Recreation Statistical Services Tourism Trade Transport Other subject matter specific words or terms can be found in: • Vocabularies used by your agency • ANZSRC Field Of Research Code 2020 • Description of functions and sub-functions (Department of Finance). Where multiple keywords apply, separate the terms with a comma ‘,’.	
Resource Type	**The type of data asset being described.** This attribute specifies the type of data asset. The most common types of data asset applicable are listed below with their definitions. **collection** an aggregation of items. The term collection means that the resource is described as a group; its parts may be separately described and navigated. **dataset** structured information encoded in lists, tables, databases, etc., which will normally be in a format available for direct machine processing. For example - spreadsheets, databases, GIS data, midi data. Note that unstructured numbers and words would be considered as text. **image** the content is primarily symbolic visual representation other than text. For example - images and photographs of physical objects, paintings, prints, drawings, other images and graphics, animations and moving pictures, film, diagrams, maps, musical notation. Note that image may include both electronic and physical representations. **interactive resource** a resource which requires interaction from the user to be understood, executed, or experienced. For example - forms on web pages, applets, multimedia learning objects, virtual reality.	Choose a term from the following list. Use the most specific relevant type: collection dataset image interactive resource model physical object place service software sound Terms above are a subset of https://www.dublincore.org/specifications/dublin-core/resource-typelist/

Name	Definition and Guidance	Content (Format and Recommended Values)
	model an abstraction of the real thing, i.e. some generalisation and interpretation. Models could be considered a symbolic representation. Examples include performance models, cost models, mechanical models, etc. **service** a system that provides one or more functions of value to the end-user. Examples include: a photocopying service, a banking service, an authentication service, interlibrary loans, a Z39.50 or Web server. **software** a computer program in source or compiled form which may be available for installation non-transiently on another machine. For software which exists only to create an interactive environment, use interactive instead. **sound** a resource whose content is primarily audio or intended to be realised in audio. For example - music, speech, recorded sounds. This category includes musical notation, including score, which is unrealised in sound. This attribute could be supplemented by attribute **Format**.	
Date Modified	**The most recent date the data asset record was either created, changed, updated or modified.** This refers to the *registration* of the data asset in the inventory, not a date pertaining to data asset itself.	**Date/Time** in format: AS/NZS ISO 8601.1:2021 e.g. 2023-09 2023-09-17 2023-09-17T23:20:30+04:00 [dateX]["T"][timeX][shiftX]

Additional Attributes

Name	Definition and Guidance	Content (Format and Recommended Values)
Publish Date	**The date on which the data asset was formally issued or made available.**	**Date/Time** in format: AS/NZS ISO 8601.1:2021 e.g. 1973-09 1973-09-17 1973-09-17T23:20:30+04:00 [dateX]["T"][timeX][shiftX]
Temporal coverage from	**The start period for which this data asset is applicable.** Temporal coverage refers to the time period that a data asset covers, which may be broader than a single collection. E.g., a data service or data portal. This field is related to the attribute **Temporal coverage to**.	**Date/Time** in format: AS/NZS ISO 8601.1:2021 e.g. 1973-09 1973-09-17 1973-09-17T23:20:30+04:00 [dateX]["T"][timeX][shiftX]
Temporal coverage to	**The end period for which this data asset is applicable.** The data asset may not have an end date if it is being continually added to, in which case, a value is not required. E.g. a data service or data portal. This field is related to the attribute **Temporal coverage from**.	**Date/Time** in format: AS/NZS ISO 8601.1:2021 e.g. 2023-09 2023-09-17 2023-09-17T23:20:30+04:00 [dateX]["T"][timeX][shiftX]
Update Frequency	**The frequency at which new, revised, or updated versions of this data asset are made available.** For data assets regularly released, one data asset record will represent a series; separate records will not be required per update. Agencies will determine when a new record is required for a data asset, based on changes in methodology, collection and related policies.	Choose term from: Triennial Biennial Annual Semiannual Quarterly Bimonthly Monthly Semimonthly Biweekly Weekly Semiweekly Daily Continuous Irregular Terms above are a subset of https://www.dublincore.org/specifications/dublin-core/collection-description/frequency/

Name	Definition and Guidance	Content (Format and Recommended Values)
Purpose	**A descriptive summary of the intentions which the data asset was developed and proposed to be used for.** This field supplements the attribute **Description**.	Free text (max. 500 char)
Location	**The scope of the geographic area or location that the data asset covers.** Location represents the geographic scope of the *entire* data asset (e.g. "Australia") and is not intended to represent location values contained within the data asset, for example *street*, *suburb* or *region* which could be captured within **Keyword**, **Description** or **Purpose**.	Choose term from: Australia New South Wales Victoria Queensland South Australia Western Australia Tasmania Northern Territory Australian Capital Territory Other Territories* International *Other territories include Jervis Bay Territory, Territory of Christmas Island, Territory of the Cocos (Keeling) Islands and Norfolk Island **OR** Provide at least one of: Australian Statistical Geography Standard (ASGS) Edition 3, July 2021 - June 2026
Access URL	**The Uniform Resource Locator (URL) that links to the data asset.** If the **Access Rights** of the data asset is "open", this could be a publicly accessible permanent URL that provides (direct/mediated) access to the data asset. If the **Access Rights** of the data asset is "conditional" or "restricted", the URL could be a permanent file path to an internal system location.	URL
Licence	**A legal document under which the data asset is made available.** This information may be sourced through the agency's legal department.	Attach (or provide a URL to) the licence document

Name	Definition and Guidance	Content (Format and Recommended Values)
Sensitive Data	**The type of sensitivity of the data asset, where applicable.** If **Security Classification** has value "OFFICIAL: Sensitive", provide type of sensitivity. Where multiple sensitivity types exist within the data asset, provide the most restrictive dissemination limiting marker (DLM). Refer to Protective Security Policy Framework (PSPF) policy 9: Access to information (Attorney-General's Department) for guidance. This attribute relates to **Security Classification** and **Access Rights**.	Choose term from: None Commercial Cultural Environmental Government Health/Medical Legal Personal
Legal Authority	**All legal mandates under which the data asset was collected, created, received, used or disclosed.** Legal mandates could include Memorandum of Understanding; Legislation; Machinery of Government; Government policies or acts; etc. It could include the authority, e.g. (Australian Government) Federal Register of Legislation or Data Availability and Transparency Act 2022. Where multiple legal mandates exist, separate their URLs with a comma ','. This information may be sourced through the agency's legal department.	Free text (max. 200 char) No URL (or file path) to relevant mandate (e.g. https://www.legislation.gov.au/)
Disposal	**The disposal action to which the data asset is subject to.** Where multiple disposal actions exist within the data asset, provide the longest retention period. Refer to "**18.3 Disposal Action**" within AGRkMS V2.2 (June 2015) for guidance. This information may be sourced through the agency's legal department.	Free text: e.g. "Destroy 3 years after contract is terminated" "Destroy 7 years after last entry" "Destroy 75 years after date of birth of employee" "Retain as national archives" ...
Data Status	**A status that describes the state of progression or registration of the data asset.** This refers to the status of the data asset registration within the inventory, not the status of the data asset itself. • **Planned** - Registration is planned • **Under Development** - Registration in progress (not all attributes populated) • **Completed** - Registration complete	Choose term from: Planned Under Development Completed

Name	Definition and Guidance	Content (Format and Recommended Values)
File size	**The volume of the data asset.** For digital assets, provide a number and units. This field may not be relevant, for example, if your data asset is a data service or interactive resource. This information may be sourced through the agency's IT or data management departments.	Free text e.g. N/A 2KB 4MB 5GB 1TB
Format	**The distribution format of the data asset.** It is recommended that an agency determines a relevant set of terms to be used consistently when collecting. This information may be sourced through your agency's IT or data management departments. **Format** is related to the attribute **Resource Type**.	Free text e.g. CSV Database DataCube GeographicData JPEG MP4 Spreadsheet WebPage WebApplication
Language	**The language used within the data asset.** This refers to the language used within the data asset - e.g. "English". The default value may be set to "English". Some agencies may have assets containing languages other than English, in which case the Australian Standard Classification of Languages (ASCL) can be used.	Free text e.g. English **OR** Choose term from: Australian Standard Classification of Languages (ASCL), 2016 | Australian Bureau of Statistics (abs.gov.au)

Name	Definition and Guidance	Content (Format and Recommended Values)
Publisher	**The agency that made the data asset formally available.** The publisher is the agency that formally produced and released the data asset and controls any future version release. The Publisher may not be the custodian (see **Data Custodian** attribute), e.g. Office of the National Data Commissioner (ONDC) The default value may be your agency. If the asset is published by another agency, use a term from the Government Directory, NGO List or Research Organisation Register. This field is related to attribute **Data Custodian**.	Free text e.g. Office of the National Data Commissioner (ONDC) **Name of Agency/Department or Non Government Organisation** Choose Government Agency from: https://www.directory.gov.au/departments-and-agencies Choose NGO from: List of Australian accredited non-government organisations (NGOs) Choose Research Organisation identifier from: https://ror.org/

Appendix C: Data Inventory Template

The Data Inventory Template has been developed for agencies to customise to their business needs and requirements. The example shown below is only one of a possible plethora of template solutions, and the template should be tailored to the chosen Data Inventory solution and metadata requirements (as described in steps 3.1 and 4.1). The template is available for download here.

	CORE										ADDITIONAL																
Click heading for instructions	Identifier	Title	Description	Point of Contact	Access Rights	Security Classification	Data Custodian	Keyword	Resource Type	Date Modified	Publish date	Temporal coverage from	Temporal coverage to	Update Frequency	Purpose	Location	Access URL	Licence	Sensitive Data	Legal Authority	Disposal	Data Status	File size	Format	Language	Publisher
Data Asset 1	ISBN:98765-43-21-0	Example Data Asset	This dataset is an example.	data.custodian@agency.gov.au	Conditional	PROTECTED	https://www.directory.gov.au/departments-and-agencies	Education and Training	dataset	2022-12-13	2023-01-16	2000-09-27	2023-09-27	Annual	To illustrate a potential Data Inventory Template	New South Wales	www.url.au	no	Personal	no	Retain	Under Development	5MB	CSV	English	Agency
Data Asset 2																											
etc...																											

Appendix D: Data Inventory Baseline Assessment

The purpose of this baseline assessment is to determine agency's current state of data asset management. It should be completed at steps 1 and 5 of the Guide.

Focus Area	*Guiding Questions*	Level 1 – Initial Process is ad hoc & not established	Level 2 – Developing Process is under development	Level 3 - Well Defined Process is standardised, communicated	Level 4 - Well Managed Process is managed, measured	Level 5 – Optimised Focus in on continuous improvement
Data Inventory	*Does your agency have a central, enterprise-wide data inventory, data asset register?*	Enterprise-wide data asset register does not exist, agency business areas register data assets in silos, specific to their work/projects.	Agency data asset register/s are managed in silos and only have selected critical/high value data assets, with incomplete metadata.	Agency enterprise-wide data asset register exists containing some or most of the agency's high-value data assets.	Agency enterprise-wide data asset register contains a large proportion of the agency's data holdings and is manually updated and maintained including complete metadata	Agency enterprise-wide data asset register is updated and maintained regularly and consistently, and supported by appropriate tools and services (data catalogue)
ONDC Metadata Attributes	*Does the agency accurately describe its data assets within the enterprise-wide data asset register using the 26 ONDC Metadata Attributes and a standard vocabulary?*	Agency data asset is not accurately described using the 26 ONDC Metadata Attributes, and/or other metadata standard. Metadata relating to data assets is poorly understood and/or not widely known.	Agency is working to accurately describe data assets using the 26 ONDC Metadata Attributes and consistent vocabulary in data asset register/s.	Agency data asset is accurately described in the agency data register using all core metadata attributes (10 out of 26 ONDC Metadata Attributes) and may also apply agency specific metadata attributes.	Agency data asset is accurately described in the agency data register using all 10 core and several of the 16 additional metadata attributes using standard vocabulary and agency-specific metadata attributes.	Agency data asset is described in the agency data register using the 10 core and all of the 16 additional metadata attributes using standard vocabulary along with agency-specific metadata attributes, using automated tools where applicable.
Data Discovery (catalogue)	*Can agency staff discover agency data holdings in a streamlined, efficient and consistent way?*	Agency data holdings are not visible to agency staff.	Data asset register is built on Excel spreadsheet or SharePoint, captures datasets from selected business areas, visible to some agency staff outside selected business areas, and not widely publicised across the agency.	Agency enterprise-wide data asset register is findable by agency staff in streamlined, efficient and consistent way.	Agency enterprise-wide data asset register has additional functionality to make it simple and easy for agency staff to access the register and request data holdings from data stewards.	Agency has an enterprise-wide data catalogue which is automatically maintained and updated. Accessible to internal and external users.
Data Accessibility	*Is there a process where agency staff can request and gain access to the data asset identified in the data asset register with security and controls in place?*	Teams cannot easily identify data assets within their business areas, agency data stewards have ad hoc processes to enable agency data users to access their data assets/holdings.	Teams can identify data assets within their business areas and may know how to seek out data from other business areas. Agency data stewards have data request and data accessibility processes which can be manual and time consuming.	Agency data stewards are supported by standardised and streamlined data request and data accessibility processes including security and controls.	Agency data request and data access processes and tools are streamlined and include security and controls managed by data stewards (including through automated tools such as APIs).	Agency data request and data access processes including security and controls are being reviewed and measured for continuous improvement.

Focus Area	*Guiding Questions*	**Level 1 – Initial** Process is ad hoc & not established	**Level 2 – Developing** Process is under development	**Level 3 - Well Defined** Process is standardised, communicated	**Level 4 - Well Managed** Process is managed, measured	**Level 5 – Optimised** Focus in on continuous improvement
Data Inventory defined in Agency Data Strategy	*Does the Agency Data Strategy include data inventory and/or metadata management and improvement?*	Agency does not have a Data Strategy and/or does not include data inventory and/or metadata management.	Agency is developing its Data Strategy which will include data inventory and/or metadata management.	Agency's Data Strategy includes data inventory and/or metadata management and improvement is developed and endorsed for implementation.	Agency's Data Strategy includes data inventory and/or metadata management and improvement is implemented.	Agency's Data Strategy includes data inventory and/or metadata management and improvement is being reviewed and measured for continuous improvement.
Process, control and governance	*Does the agency have procedures, processes, controls and governance to manage their internal data inventory and metadata?*	Data inventory and metadata management processes, procedures, guidelines do not exist.	Data inventory and metadata management processes, procedures, guidelines in development and no formal governance roles and responsibilities defined/assigned.	Data inventory and metadata management processes, procedures, guidelines developed, promulgated, and applied in practice. Roles and responsibilities assigned for the management and maintenance of the data inventory.	Data inventory and metadata management processes, procedures, guidelines accepted as part of the culture and embedded in data management practices, data governance steering/working groups and subject to audit.	Data inventory and metadata management processes, procedures, guidelines are well maintained with ongoing reviews and continuous improvement.
Data Management Proficiency	*Does the agency apply the APS Data Capability Framework and provide training to build data steward capability? In particular, capability areas:* *1 (value data assets)* *4 (data governance)* *6 (enabling data access)* *8 (data collection)* *13 (data quality)* *16 (data classification)* *19 (Metadata)* *20 (enabling data use and re-use)* • *Metadata capability is highlighted in this example.*	Agency is not consistently applying the APS Data Capability Framework.	Agency is working to establish an organisation capability and uplift support as defined in the APS Data Capability Framework.	Agency has established foundational proficiency as defined in the APS Data Capability Framework. • *Understands there are different ways to summarise data and has a basic understanding of commonly used metadata options.* • *Understands the concept of metadata, including its purpose and benefits.*	Agency has established intermediate proficiency in metadata management as defined in the APS Data Capability Framework. • *Can use various summary options to effectively describe data and explain and justify those choices.* • *Follows organisational standards and procedures relating to metadata creation, storage, and use.* • *Can access metadata and use the descriptors to better understand existing data and effectively use it.*	Agency has established advanced proficiency in Metadata management as defined in the APS Data Capability Framework • *Has expertise in metadata. Can establish standards for metadata and provide oversight and advice to others.* • *Maintains knowledge of metadata best practice, including standards and applications.* • *Can use a range of tools for storing and working with metadata. Keeps metadata refreshed and updated and can repair items that are incorrect or out of date.*