

Data Inventory Guide and Glossary

This guide and glossary document is a companion to the Data Inventory Template and part of the Data Inventory Pack.

How to use the Basic Data Inventory Template

Download the Data Inventory Pack

If you're reading this, then you've already completed this important step. Take note of the files included within this pack:

- » Basic Data Inventory Template (Excel Spreadsheet)
- » Data Inventory Guide and Glossary (PDF)
- » Staff Email Template (Word Doc)

Fill out the spreadsheet

Start by filling out the template for your own team to get a feel for the tool.

Refer to the glossary below for definitions

If you, or your colleagues, are unsure about the definitions or how to answer for any of the columns you can use the included glossary for more information.

Reach out to teams across your organisation

Next, you'll want to capture the data from across the organisation. To help you capture this information, we have included a **template email** within the **Data Inventory Pack** that you can use with suggested text to share with the data champion from each team.

Analyse your inventory

Once you have collated your completed spreadsheet, you will want to analyse your results for the following.

Address any discrepancies and gaps

Try to identify any gaps in information. Are there any data being overlooked? Engage the relevant team members as necessary.

Identify and limit access to sensitive data

Ensure the right team members have the right levels of clearance to access data. Some more sensitive data may be locked to certain team members per your own internal policies. Put appropriate barriers in place to remove access to sensitive data as required. Start [here](#) (*A Guide to Data Sharing Agreements and Data Privacy Resources* – Digital Transformation Hub) to help you understand your responsibilities under the Privacy Act and take practical actions.



Identify and support your data custodians

Ensure that all data custodians have the understanding and support to be responsible for the ongoing care and management of the dataset, and that they have the decision-making authority should questions about said dataset arise.

Understand data locations and use

Identify and clearly note where data is held and if any can be used for research or AI projects.

Understand your metadata needs

Do you collect data that could be reused or combined to generate new insights, or meet reporting requirements?

Consider a data policy

Now that you have oversight of all of your data, do you have a data policy in place that's sufficient to support and protect your organisation's data?

Keep the list up to date

Ensure that the list is kept updated and relevant leaders are aware of where the inventory is held in the event of data decisions, AI or research projects, cyber security or IT requirements. Set up an annual calendar entry for reviewing the inventory.

Important note about this inventory template

We have adapted the Inventory template developed initially by the [Office for the National Data Commissioner](#) so that it can be useful for NFPs. We have also added information about Indigenous data to help organisations improve their understanding and reporting and move towards CARE principles because how you classify data as sensitive can be different for Indigenous communities. The [CARE principles](#) are a vital structure for ensuring that your data supports Indigenous data sovereignty.



Glossary

If you're not sure about what any of the terms/fields mean, or how to use them, you can review those filed in our glossary below.

Identifier

Definition and Guidance

The identifier is used to distinguish the data asset as unique and different to another data asset. It is key to finding the data asset and ensuring the specific data asset can be referenced without confusion.

Ideally it is globally unique, such as a Digital Object Identifier (DOI). However, the identifier can start with an acronym relevant to your agency followed by letters, numbers or symbols.

Format and Recommended Values

Free text (max. 200 char)
e.g. FIN000077

CARE alignment

Metadata and cultural context

Team

Definition and Guidance

Fairly self-explanatory, but this field will help you identify which internal team holds the dataset and filter the inventory in the future.

Format and Recommended Values

Free text (max. 200 char)

Title

Definition and Guidance

The title is a unique, clear and culturally appropriate name for the data asset. People searching for data should gain a basic understanding of the business use/intent of the data asset from the title. It should help people understand the community, place or theme the data relates to.

If a new data asset record is required due to an error in the data, the title should note that it is a revised version.

Format and Recommended Values

Free text (max. 200 char)
e.g. Data Governance Workshop Outcomes - Central Desert 2024



CARE alignment

Metadata and cultural context

Description

Definition and Guidance

Easy to read information about the data asset to enable users to find and evaluate the data asset for their needs.

The Description attribute is typically several sentences long and is used to search for the data asset so keywords should be carefully considered.

Agencies are encouraged to include field names (e.g. gender/sex, age, address) collected within the data asset. Describe the dataset including the community, cultural context, and participant information (e.g. gender, age group, role), while respecting privacy and ICIP. This will help answer any specific research, policy or program questions a user may have, and help manage requests for additional information about the data asset received by your agency.

This attribute is supplemented by the Title, Keyword and Purpose attributes.

Format and Recommended Values

Free text (max. 500 char)

e.g. This dataset captures key themes and priorities identified during governance workshops co-led with Central Desert community organisations in 2024. It includes information such as participant role, gender, age group and organisation type to help analyse representation and perspectives across different governance contexts. It notes key discussion themes on governance priorities.

CARE alignment

Metadata and cultural context; Data source and cultural provenance

Data owner

Definition and Guidance

Data owners are senior leaders with ultimate authority and accountability for datasets. Their role is policy and strategic direction with decision-making authority for those datasets, ensure compliance with regulations, categorise data for sensitivity and approve access requests. Ownership reflects both legal authority and relational accountability under Indigenous Data Sovereignty principles.

Format and Recommended Values

Free text

CARE alignment

Custodian and governance



Data steward

Definition and Guidance

Data stewards have operational responsibilities and day to day management of data quality, integrity and compliance without the overarching decision-making. They enforce the standards set by the owners and resolve quality issues in alignment with the data owner decisions.

Format and Recommended Values

Free text

CARE alignment

Custodian and governance

Access rights

Definition and Guidance

Specifies access (or restrictions) to the data asset.

Access is based on the agency's privacy, security, or other policy approaches that apply to the data asset. Access is determined not only by agency policies but also by the rights, interests and decisions of the relevant Aboriginal and Torres Strait Islander communities or custodians.

Access can be:

Open - Data that is publicly accessible online, while still respecting Indigenous governance decisions and cultural sensitivities. Account registration may be required.

Conditional - Data is accessible only when certain conditions are met, which may include community consent or adherence to cultural protocols, as well as technical or administrative conditions such as fee-for-service models, verified credentials, or the data is only accessible at a specific physical location.

Restricted - Data access is limited for reasons such as legal, privacy and sensitivity. For example, data that is culturally restricted or subject to community protocols, such as information that is gender-specific (e.g., women's or men's knowledge), secret-sacred knowledge, or site-specific information; or data that can only be accessed with approval from the relevant Aboriginal and Torres Strait Islander custodians or through an agreed governance framework.

This attribute is supplemented by the Security Classification and Sensitive Data attributes.

Format and Recommended Values

Choose term from:

Open

Conditional

Restricted IDS: Public with community consent or Shared; Community-restricted or Culturally sensitive or Closed

CARE alignment

Access and permissions



Keyword

Definition and Guidance

Word(s) or terms that describe the data asset subject matter.

It answers the question “what is in this data asset?” and supports the discovery of the data asset. This is a critical component in helping users find your data asset. Careful selection of keywords is critical to ensure the data asset is findable, appropriately classified, and respectful of Aboriginal and Torres Strait Islander rights, cultures and priorities. Use as many keywords as you can. [The AIATSIS Thesaurus](#) is a useful reference tool terms, place names and other terminology used in Indigenous research.

As an absolute minimum, agencies must include:

Indigenous related data are tagged with the keywords ‘First Nations people’ and ‘Aboriginal and Torres Strait Islander’.

Include information about the cultural or community origin of the data. For example, record whether the data relates to a specific Aboriginal or Torres Strait Islander community, nation, or language group, and note any Indigenous Cultural and Intellectual Property (ICIP) considerations.

Tag data with terms that reflect Indigenous data stewards, governance bodies or community decision-making structures responsible for the dataset.

Use keywords that signal how the data aligns with community-defined priorities or benefits. Examples include community health, self-determination or land management. Include terms that reflect Indigenous knowledge systems, languages, or worldviews represented in the data. For example, include Yolŋu language, cultural sites.

Disability related data are tagged with the keyword ‘Disability’

Data assets containing sex (information collected on sex characteristics observed at birth or infancy) or gender (information collected as a result of gender identity, expression and/or experience) are tagged with the relevant term. Further information can be found on ABS Standard for Sex, Gender, Variations of Sex Characteristics and Sexual Orientation Variables.

Where multiple keywords apply, separate terms with commas.

Format and Recommended Values

Data assets with information on people should include (if applicable): Language group, Country, Aboriginal and Torres Strait Islander, First Nations people, Gender, Sex, Disability.

e.g. A data asset containing attendance, outcomes, and feedback from programs designed to strengthen Indigenous leadership skills could have the keywords: Governance, Leadership development, Capacity building, Community organisations, First Nations people, Aboriginal and Torres Strait Islander, Workshop participation, Feedback, Community-led program, Cultural protocols

CARE alignment

Metadata and cultural context



Resource type

Definition and Guidance

Specifies the type of data asset.

The most common types of data assets 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.

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.

physical object

Tangible items of cultural, scientific, or historical significance, such as artefacts or community-held materials.

place

a geographic area, such as community lands, sacred sites, or areas of cultural significance

party

a person, organization, cultural group, or institution.

event

a non-persistent, time-based occurrence. Metadata for an event provides descriptive information that is the basis for discovery of the purpose, location, duration, responsible custodians or organisations, cultural protocols, and links to related events and resources. The resource of type event may not be retrievable if the described instantiation has expired or is yet to occur.

Examples include: conference, workshop, open-day, forum, program, ceremony.

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, chat services, virtual reality.

This attribute could be supplemented by attribute Format.



Format and Recommended Values

Choose a term from:

collection
dataset
image
model
physical object
place
party
event

CARE alignment

Metadata and cultural context

Date modified

Definition and Guidance

The most recent date the data asset record was either created, changed, updated or modified. Record the most recent community-reviewed or officially modified date of the dataset. This date refers to the date in which the metadata of the data asset changes or is first recorded in the data inventory, not a date pertaining to the underlying data asset itself. This attribute is critical for agencies in managing their data assets and supplemented by the Publish Date attribute.

Format and Recommended Values

Date/Time in format: AS/NZS ISO 8601.1:2021
e.g.
2023-09
2023-09-17
2023-09-17T23:20:30+04:00

CARE alignment

Metadata and cultural context

Purpose

Definition and Guidance

A descriptive summary of the intentions which the data asset was developed and proposed to be used for.

This might include how the data aligns with community-defined priorities, supports self-determination and delivers community benefits.

Format and Recommended Values

Free text (max. 500 char)

CARE alignment

Purpose and community benefit



Location

Definition and Guidance

The geographic area, country, community that the data asset covers.

Location represents the geographic area of the entire data asset (e.g. “Australia”). It is not intended to represent specific location values contained within the data asset itself.

For example, a dataset about community governance workshops could use 'Yuin Country, South Coast NSW' or 'Kowanyama community, Cape York, Queensland', or a mesh block from the Australian Statistical Geography Standard. Austlang is a useful resource to help identify language groups to include in metadata: <https://aiatsis.gov.au/austlang/search>

Location values contained in the data asset such as specific suburbs or regions can be captured within the Keyword, Description or Purpose attributes.

Format and Recommended Values

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 area from:

Australian Statistical Geography Standard

(ASGS) Edition 3, July 2021 - June 2026 .

Use Indigenous place names where appropriate

CARE alignment

Data source and cultural provenance

Sensitive data

Definition and Guidance

The type of sensitivity of the data asset, where applicable.

If the Security Classification is “OFFICIAL: Sensitive” or above, the type of sensitivity should be provided.

For further guidance, refer to section 9 of the PSPF Annual Release 2025 | Protective Security Policy Framework



In addition to formal classifications (e.g. OFFICIAL: Sensitive or above under the PSPF), you should consider Indigenous cultural and community-defined sensitivities, such as information that is sacred, gender-restricted, or community-controlled under Indigenous Cultural and Intellectual Property (ICIP) rights. Where relevant, describe the basis of sensitivity (e.g. cultural protocols, consent agreements, ICIP restrictions, privacy considerations) and reference any community governance processes or permissions required for access or use.

This attribute is supplemented by the Security Classification and Access Rights attributes. See Protocol Labels <https://localcontexts.org/labels/traditional-knowledge-labels/>

Format and Recommended Values

Choose term from:
N/A [e.g. open data]
Legislative secrecy
Personal privacy
Legal privilege

CARE alignment

Risk & Compliance/ Protection and ethics

Language

Definition and Guidance

Refers to the language used within the data asset - e.g. "English".

The default value may be set to "English".

Format and Recommended Values

Drop down or free text.

Record the language(s) of the data, including Indigenous languages using ASCL codes.

We have included some of the more commonly used languages as a drop down, however your organisation may have assets containing languages other than English, in which case the [Australian Standard Classification of Languages \(ASCL\)](https://www.abs.gov.au/statistics/classifications/australian-standard-classification-languages-ascl/latest-release) <https://www.abs.gov.au/statistics/classifications/australian-standard-classification-languages-ascl/latest-release> can be used.

The ASCL has a comprehensive list of Australian Indigenous languages.

CARE alignment

Metadata and cultural context

Disposal

Definition and Guidance

Information on the correct retention or disposal action of the data asset.



For datasets containing Aboriginal and Torres Strait Islander information, this may also include community-approved timelines, ICIP protections, and culturally sensitive disposal practices.

This is important because agencies are legally required to appropriately manage and dispose of their data. NFPs should check the ACNC data guidelines and OAIC for additional guidance Chapter 11: APP 11 Security of personal information | OAIC

Where multiple disposal actions exist within the data asset, provide the longest retention period.

For further guidance, refer to “18.3 Disposal Action” within Australian Government Recordkeeping Metadata Standard | naa.gov.au

This information may be sourced through the agency’s legal department.

If information is not yet available, fill in “To be determined” and update the disposal date as soon as it becomes known.

Format and Recommended Values

Free text:

e.g.

“Destroy 3 years after contract is terminated”

“Destroy 7 years after last entry”

“Destroy 7 years after last recorded community consultation”

“Destroy 75 years after date of birth of employee”

“Retain in archives under Indigenous custodianship”

“Repatriate or securely destroy following community-approved protocol”

CARE alignment

Retention and disposal

Consent

Definition and Guidance

Is consent required and has been obtained for this data (e.g. Client data). Provide date consent obtained, and any restrictions on use (e.g. Cannot be used for any other purpose than program reporting requirements)

Format and Recommended Values

Yes/No.

Dates, funder requirements.

CARE alignment

Basic individual Indigenous-person consent (see progressive version for more detail on Indigenous consent requirements)

