National Framework for the Sharing of Restricted Access Species Data in Australia

Date of Framework 20/02/2023

For bibliographic purposes, this publication may be cited as:

National Framework for the Sharing of Restricted Access Species Data in Australia 2023, Atlas of Living Australia, Publication Series No. 6, Canberra, Australia. https://doi.org/10.54102/ala.94894.

Supplements of this publications are regularly updated and may be cited as:
National Framework for the Sharing of Restricted Access Species Data in Australia 2023, Supplement Recognising Indigenous Data Sovereignty 20/2/2023, Atlas of Living Australia, Publication Series No. 6, Canberra Australia. https://doi.org/10.54102/ala.94894

DOI: 10.54102/ala.94894

Secretariat for the National Framework for the Sharing of Restricted Access Species Data in Australia

c/- Atlas of Living Australia (ALA)
GPO Box 1700
CSIRO Black Mountain
Canberra, ACT, 2601

More information is available at: www.rasd.org.au

The development of this Framework was supported by the Australian Research Data Commons (ARDC) and the Atlas of Living Australia (ALA). The ARDC and ALA are funded by the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS)

Contents

Acronyms

ABRS	Australian Biological Resources Study
AIMS	Australian Institute of Marine Science
ALA	Atlas of Living Australia
auNSL	National Species List
AusBIGG	Australian Biodiversity Information Governance Group
CARE	CARE principles for Indigenous data governance (Collective Benefit, Authority to Control, Responsibility and Ethics)
DwC	Darwin CORE
DOI	Digital Object Identifier
FAIR	FAIR data (findability, accessibility, interoperability, and reusability)
HISCOM	Herbarium Information Systems Committee
NCRIS	National Collaborative Research Infrastructure Strategy
RASD	Restricted Access Species Data
RASL	Restricted Access Species List
TDWG	Biodiversity Information Standards Group (also known by the abbreviation TDWG, previously known as the Taxonomic Databases Working Group)
TERN	Terrestrial Ecosystem Research Network

Statement of Intent

This framework is intended to enable consistent, effective sharing and management of Restricted Access Species Data (RASD) nationally in Australia. This includes:

- promoting responsible data sharing practices
- giving data custodians the confidence to safely share restricted access species data
- assisting data custodians to maintain the consistency and quality of data

What is Restricted Access Species Data?

Restricted Access Species Data (RASD) - often known as "sensitive species data" – are data about biodiversity that have features that mean there are reasons for withholding the data from public view or modifying the data before it is made public.

This can include, for example:

- Flora and fauna where knowledge about their exact locations make them sensitive to disturbance. This can be because they are extremely rare, attractive to poachers, their nest sites are highly sensitive to disturbance or simply because they are newly discovered and there is insufficient information known about them.
- Species locations where knowledge about the species can be seriously misinterpreted or cause damage without context. Examples include records of weed or pest control that the landowner might not be comfortable making publicly available. An important example is records of an invasive species that has been eradicated from Australia but whose historical records might cause confusion over Australia's pest-free status.
- Data that have been supplied under conditions that constrain what can be given to third parties or where the data can threaten research or an economic outcome
- Data that include personal identifiable information, where the person has not given approval for that data to be shared
- Data that has been gathered by or otherwise belongs to First Nations people and permission needs to be granted before the data are shared.

This data is needed for conservation management, decision-making and research.

An excellent overview of many types of sensitive data, the current state of play internationally, as well as the issues inherent in using these data are provided by the Global Biodiversity Information Facility's Current Best Practices for Generalizing Sensitive Species Occurrence Data by Arthur Chapman.

Who is this Framework For?

This framework is intended for use by any organisation that wants to gather, display, share or receive RASD (data custodians and data users). It provides a set of shared principles that are intended to guide users in sharing RASD responsibly and appropriately. It also provides specific guidance on procedures, formats and methodologies and is necessarily technical at some points.

While it is not intended for the general user, members of the public can use the framework to better understand how desensitised publicly accessible RASD that they are interested in may have been modified.

Who Manages this Framework?

This framework is maintained by the Atlas of Living Australia in consultation with the national community of practice, the Australian Biodiversity Information Governance Group (AusBIGG), which includes representatives from the jurisdictions around Australia as well as the Terrestrial Ecosystem Research Network, the Atlas of Living Australia, the Council of Heads of Australasian Herbaria, and the Council of Heads of Australian Fauna Collections.

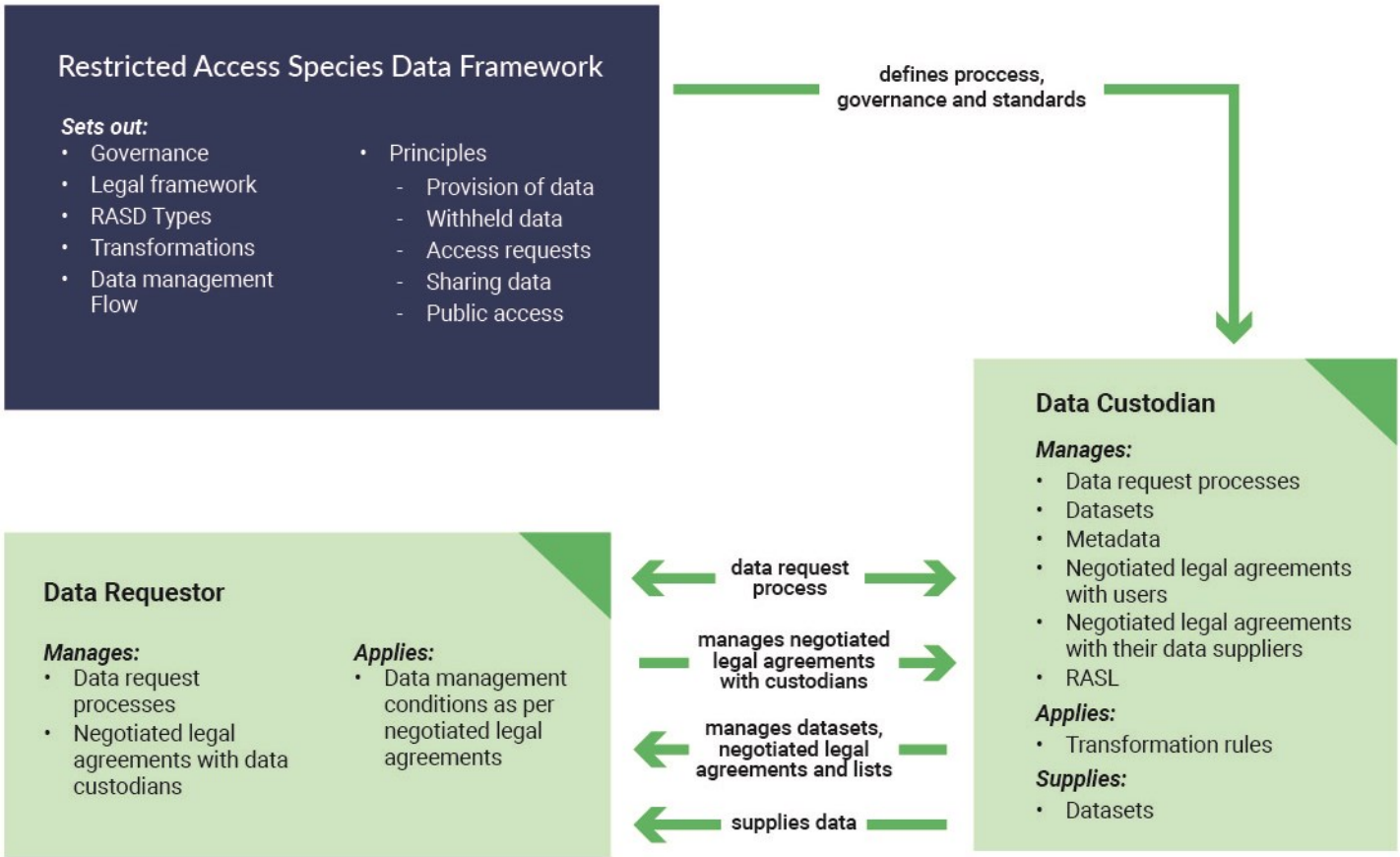
The framework will be reviewed after 12 months and then updated every three years following consultation processes.

A high-level view of how this framework interacts with data custodians and users is shown in Figure 1.

Organisations and individuals interested in commenting on the framework may address their comments to the Secretariat for the National Framework for the Sharing of Restricted Access Species Data in Australia:

c/- Atlas of Living Australia (ALA)
GPO Box 1700
CSIRO Black Mountain
Canberra, ACT, 2601

Figure 1- Restricted Access Species Data Governance



Who Produced the Framework?

The framework is the result of an agreed program of work by all the organisations listed in the section Organisations Involved in the Production of this Framework. Including the Australian Government, State and Territory conservation agencies, the Atlas of Living Australia, Australian Research Data Commons, Council of Heads of Australasian Herbaria, Council of Heads of Australian Faunal Collections, and the Western Australian Biodiversity Science Institute to improve the national handling of RASD.

The framework was also circulated nationally around a wide range of government and non-government stakeholders for comment during its development.

This framework relies on the principle that data custodians retain data sovereignty over their data.

Consistency

The framework draws upon the sensitive species policies of all Australian jurisdictions as well as the sensitive species position papers prepared by the Global Biodiversity Information Facility (2020) and the Atlas of Living Australia (2009).

The framework follows the Australian Government's Best Practice Guide to Applying Data Sharing Principles and is consistent with the principles of FAIR data (findability, accessibility, interoperability, and reusability), CARE (Collective benefit, Authority to control, Responsibility, Ethics) principles, the Research Data Alliance's Principles and Guidelines for Legal Interoperability of Research Data, and the Australian Research Data Common's Guide to Publishing Sensitive Data while also acknowledging the sensitive nature of the data covered.

Currently the framework does not fully cover restricted access Indigenous ecological data and knowledge. Many of the principles and best practices outlined in this framework could, however, be applied when handling access requests for these data. The framework is intended to operate within the CARE principles for Indigenous data governance to recognise Indigenous data sovereignty.

How Does My Organisation Show it is Consistent with the Framework?

The framework is voluntary and promotes consistency and transparency.

Organisations wishing to demonstrate that they are acting consistently with this framework should make their restricted access species data policy publicly available to demonstrate commitment to transparency and accessibility.

Purpose of the Framework - A Single Way of Handling Restricted Access Species Data for Australia

A single agreed way of sharing RASD aims to make data sharing between organisations easier and make the form of desensitised data that is made public consistent.

The advantages of this framework are:

- Standardised methods and processes for handling RASD
- Stream-lined access to and sharing of RASD
- Standardised methods of transforming RASD when displayed publicly

All the organisations involved in the development of this framework agreed that a more holistic approach to handling data sensitivities associated with species data was required. It was also agreed that a move away from the use of the term “sensitive” was helpful in understanding all the types of restrictions that influence data custodian decisions on sharing data as well as avoiding confusion with formal Australian government security classifications.

The purpose of this framework is to:

a) provide principles to facilitate a consistent national approach to RASD
b) provide overarching principles and best practice guidance to data custodians on identifying, managing, and sharing RASD
c) provide technical guidance on how to operationally deal with RASD
d) emphasise the principle that data custodians retain sovereignty over their data

The framework provides a national, structured, best-practice flow for:

a) the sharing of all RASD
b) the release of data to approved data requestors
c) the release of desensitised and modified datasets for use by the general community
d) agreed best-practice processes enabling a)-c)

National Principles for Restricted Access Species Data

RASD is invaluable for environmental and conservation management, decision-making, and research.

The national principles of Restricted Access Species Data are that:

- Handling of Restricted Access Species Data should be consistent with government requirements and FAIR and CARE principles
- Restricted Access Species Data should be consistently classified
- Restricted Access Species Data should be discoverable
- Restricted Access Species Data requests should follow a structured, transparent process
- Restricted Access Species Datasets provided in response to approved requests for Restricted Access Species Data should be as complete as possible
- Restricted Access Species datasets should be transformed consistently if made public or be as complete as possible if provided to approved data requestors
- Sharing of Restricted Access Species Data should be through negotiated legal agreement

More detail on each of the principles is detailed later in the framework.

Explanation of some important terms used in dealing with Restricted Access Species Data

There are several terms that need to be explained in this framework. A full list is provided at the back of the Framework under Definitions. Some particularly important terms are:

Dataset and Attribute Field

The terms dataset and attribute field are used widely in this framework. A dataset is a set of data held by a data custodian that includes information about species observations including the identity of the species, latitude, and longitude. Other fields that might be included are date, a location description, collector details etc.

An attribute field refers to one of the fields in the dataset used to describe the observations such as location, date, observer etc.

Transformation

Transformation is a term used widely in this document to describe the range of actions that a data custodian applies to a dataset to modify it from its raw state. Actions include:

a) Withholding attribute fields (removing attribute fields from a dataset to, for example, remove personal identifiable information about observers)
b) Obfuscating locations (modifying the latitude and longitude to make it difficult to identify the original location of the record)

Sensitive Species or Restricted Access Species Lists (RASLs)

RASLs, also known as sensitive species lists, are lists of species that are regarded as requiring protection from human interference for a variety of reasons. They can include both listed threatened species as well as unlisted species that have some aspect that makes them sensitive, for example, poaching from nest sites. Typically, the decision to classify a species as sensitive is balanced against the potential cost and disadvantages of hiding locality information and potentially impairing conservation efforts. Some RASLs only affect species records at certain times of the year or certain life stages, for example Birdlife Australia treats some RASL species records as sensitive only during breeding season.

Maintaining up-to-date, public, discoverable RASLs supports better and more efficient decision-making, and more informed use and management of data.

RASLs are maintained by all state and territory environmental agencies in Australia. Most jurisdictional RASLs are publicly available lists that delineate which species should have their geographic locations blurred (see obfuscation) or withheld entirely to prevent disturbance or exploitation of the species or the site.

Some custodians of large species-specific datasets such as BirdLife Australia, FrogID and Butterflies Australia also maintain RASLs derived from expert opinion.

Lastly, large data aggregators also use these RASLs, for example the Atlas of Living Australia applies both jurisdictional and third-party RASLs to data.

Obfuscation

Obfuscation or "fuzzing" of data is the practice of transforming or obscuring the original geo-localities of an observation of flora or fauna to make it difficult to discern the original geo-locality by randomisation or generalisation. Randomisation has benefits for map display but is not a robust data management approach in a data ecosystem as it creates artificial points. In the following example (figure one taken from Chapman 2020), two alternative means of taking individual observations and generalising them to a grid are shown.

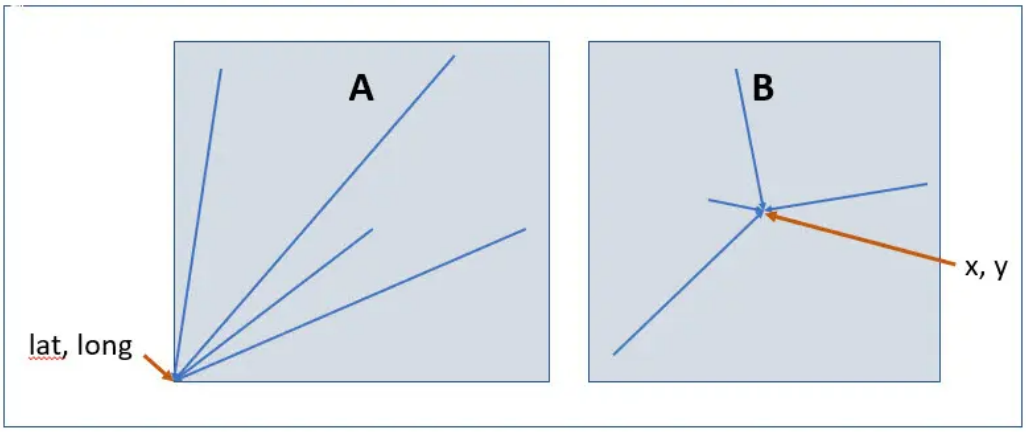


Figure 1. Two generalization methods: a) a geographic grid where all records are referenced to the bottom left-hand (SE) corner; b) a metric grid where all records are referenced to the centroid.

(Taken from Chapman 2020 – please note that this is *not* figure one in this framework)

Either of the above generalisation techniques, rounding (A) or the provision of records as grid square polygons (B) are preferred, if data are to be passed between systems that may then apply their own additional obfuscation. For the purposes of this framework, references to obfuscation imply generalisation.

Metadata

Metadata are data about data, helping a user to interpret data (or observations). For the purposes of this framework, metadata are either:

a) *Dataset metadata* – a concise description of a dataset which enables a user, not necessarily able to access a dataset, to gauge the relevance of the data for their purposes; or
b) *Record (or row)-level metadata* – these are documentation in an attribute field at the level of a record in a dataset. For the purposes of this framework, it refers to documentation of the sensitivity status of the record (or the species of which it is a part) along with access constraints pertaining to the record and details of any generalisation of the data.

Standard Form Data Licence Agreements and Negotiated Legal Agreement

This framework refers to both the above types of documents. Most data custodians make use of standard form data licence agreements. The difference between these types of agreements is:

a) *Standard Form Data Licence Agreement* – A license used between most data custodians and repositories, which uses standard (non-negotiable) terms and conditions to stipulate management of data including to control end user use of data.
b) *Negotiated Legal Agreement* – for the purposes of this framework, a legal agreement between data custodians and Approved Data Requestors, the terms of which are negotiated by the parties. This Framework encourages the use of these agreements instead of Standard Form Data Licence Agreements.

It is recognised that data custodians with existing processes that do not involve negotiated legal agreements may require time to transition but will ideally be working towards implementing a new process.

Principle: Handling of Restricted Access Species Data Should be Consistent with Government Requirements and the FAIR and CARE Principles

Sandalwood (Santalum spicatum), © David Muirhead CC BY NC

The Australian Government Best Practice Guide to Applying Data Sharing Principles encourages the use of a data-sharing purpose test. For the purposes of this framework, access to RASD shouldn't be considered automatic but governed by an access process (see **Principle: Restricted Access Species Data requests should follow a structured, transparent process**). Generally speaking, access to RASD should be dependent on a purpose test based on the following principles, where the purpose is to inform one or more of the following:

- government policy
- research and development with a public benefit
- conservation and management activities with a public benefit
- program design, implementation, and evaluation
- delivery of government services.

It is best practice to follow the CARE principles when dealing with Indigenous data or knowledge (see **Supplement 1**)

Availability of RASD is often dependent on regulatory and legislative controls. Data custodians must observe the requirements in force in their jurisdiction and where this framework conflicts with those requirements, those requirements hold sway. Freedom of Information and environmental approval legislation frequently legally require the publishing of raw data in particular circumstances. This may mean datasets that would otherwise be publicly available must be withheld or, conversely, datasets that a stakeholder may wish to be withheld or embargoed must be made available. Data custodians remain responsible for ensuring that any data that they are providing is consistent with the legislative requirements in the jurisdiction in which they operate.

This framework provides best practice guidance on sharing restricted access data relating to species. **Table 1** provides a risk impact matrix for the assessed risk of sharing data according to this framework. The high-level process for handling RASD data is shown in **Figure 2**. Further detail on the differences between requests from Full Access Users, Approved Data Requestors and Public Access is in the **Principle: Restricted Access Species datasets should be transformed consistently if made public or be as complete as possible if provided to approved data requestors.** Individual steps in the process of handling RASD are provided in **Supplement 6: Process for Release of Restricted Access Data, Supplement 7: Withheld Data,** and **Supplement 8: Process for Handling Embargoed Data.** A best practice example of a request form is available from **Supplement 9**.

Because best practice implies a paradigm shift in the handling of such data, the framework distinguishes between:

a) existing data – these are data containing RASD already held in repositories as at the date of a data custodian deciding to operate consistently with this framework
b) new data – these are data acquired after the data custodian decides to operate consistently with this framework

Existing Data

It is recognised that there are subsets of existing restricted access data held by most data custodians that are difficult to share with third parties because of formal or informal commitments with original suppliers.

Consistent with the Australian Government's Best Practice Guide to Applying Data Sharing Principles, this framework encourages data custodians to make best efforts to work towards minimising the practice of withholding datasets and information on this is available in **Supplement 7: Withheld Data and Restricted Access Species Data**. Recognising that some datasets need to be withheld, data custodians are encouraged to provide metadata on withheld datasets to provide transparency on what datasets exist nationally.

New Data

Data custodians should work towards the acquisition of and provision of data using the principles outlined in this framework. This is agreed best practice intended to promote FAIR and CARE principles, be consistent with the **Research Data Alliance Principles and Guidelines for Legal Interoperability of Research Data** and commitment to more efficient data-sharing

Principle: Restricted Access Species Data (RASD) Should be Consistently Classified

Common Slender Bluetongue (Cyclodomorphus branchialis), © Graham Armstrong CC BY NC

Restricted Access Species Data are biodiversity related datasets about species that are restricted for one or more of the following reasons. The examples that are provided are not necessarily the only type of data within a category, nor the only types of possible actions that may be taken:

1. **Personal Identifiable Information –**
   A dataset containing names or personal details about an individual, the release of which would be contrary to privacy legislation. For example, data held by a citizen science project contains personal information (names and addresses) of observers involved in the project. This constitutes personal information within the definition of most legislation and regulatory frameworks in Australia (see Personal Identifiable Information – legislative and regulatory instruments that affect RASD). While this applies to species data more generally, the use here refers to personal identifiable information associated with RASD only.

2. **Indigenous Data –**
   For the current coverage of the framework the application of the CARE principles for Indigenous data governance (Collective Benefit, Authority to Control, Responsibility and Ethics) relates to a species point dataset that has been gathered by, or contains knowledge of, Indigenous peoples. The CARE principles are intended to recognise Indigenous data sovereignty. For example: data gathered by the Australian Government Indigenous Ranger Program. This data requires permission to be sought from Indigenous peoples before use (see Supplement 1 for more information and guidance).

3. **Usage-Restricted Categories –**
   Represent species data where the dataset is constrained by third-party concerns. These are:

3.1 Legal Contract –

A dataset containing information that has legal conditions over its use. This might include a negotiated legal contract or a standard form data licence agreement that in effect is a legal contract limiting the transfer of the data to a third-party. Examples include data gathered under a signed contract with a third-party contractor or data from a research project provided under legal or licence agreement that relates to control of third-party data use. This limits who else the data can be shared with. It does not concern data where there are informal agreements in place.

3.2 Legal Financial –

A dataset that is commercial-in-confidence until an economic decision has been made or is part of current financial negotiations whereby transfer could compromise a financial process such as a tender. Such data are normally subject to an embargo period after which it may be released. For example, a company has access to data about endangered species in relation to a development they are involved with where third-party knowledge of that data (by competitors) might impact the economic viability of the project. This dataset might be embargoed until a certain date.

3.3 Non-legal –

A dataset that was acquired via an informal agreement that the information would not be transferred. Typically, this relates to third parties such as data from a landholder, data from a researcher who has not yet published or similar. As an example, a researcher working on a PhD project has a significant number of observations about a new species. The researcher is happy to share the data but does not want the data to be made public until after the publication of their results. An alternative example might be a researcher's data where the location can be revealed but other fields such as measurements must be withheld. This dataset might be embargoed until a certain date.

4. Species-Related Categories –

Datasets that represent locations of species and contain attributes where exposure of the location, or an attribute of a record has consequences. Species-related categories are:

4.1. Location data –

A dataset that contains species where access to information about their exact location causes sensitivity. This includes:

i. divulging exact location of a species where factors such as geographically restricted distribution, recent discovery, life history stage, reproductive habits, feeding habits or vulnerability to human interference may compromise conservation / management efforts where that information is not already in the public domain. Most jurisdictions and major data custodians maintain sensitive species lists containing species that meet this definition which require active record management. For example, the exact location of a critically endangered

plant or animal is withheld to prevent exploitation, theft for illegal trade, vandalism, the risk of disease posed by visitation and habitat destruction.

ii. management factors that make the record sensitive. This includes records on private land that interfere with privacy, or where the records are sensitive. To protect the plant, animal, habitat, or people, coordinates for the species are obfuscated.

For both i) and ii), an alternative to obfuscating or withholding the record, may be to accurately communicate the location but obfuscate the identification of the species, to avoid incidental damage such as to a nest tree in a forestry harvesting plan or roadside spraying and slashing.

4.2. Identification data –

Species whose identification has major economic, legal, or financial ramifications at a jurisdictional level, if the species name or location is exposed. This particularly refers to species which, if reported, are extremely sensitive, causing major issues at state, territory, or national scale. For example, there are highly invasive species that are spreading internationally that do not occur in Australia. These species present potential major biosecurity risks to Australian natural resource industries such as agriculture, forestry and fisheries and Australia's unique plants and animals. Incursions of these species are dealt with swiftly and decisively by eradication. However, publicly available historical records about an incursion that was wiped out can cause confusion over whether the species is active in Australia. Locations of incursions for these species may be withheld for the economic well-being of a jurisdiction.

4.3. Attribute data –

Data where metadata about the record in one or more attribute fields adds additional sensitivity. This generally means that the species record itself is not sensitive, but there is information attached to the record, such as breeding information, that make it easy to disturb the animal or important elements of its habitat. Other examples include culturally sensitive information or the location of pest control activities e.g. baiting / poisoning programs that the landowner might be worried about sharing because of tampering with baits. This may require similar approaches to location related data, where the location must be accurately communicated but species identity obfuscated.

The Risks of Using RASD

The Risk Matrix (Table 1) is a tool to assist data custodians to understand the risks posed by RASD and what actions might be undertaken to mitigate these risks, both for existing data and for data produced in the future.

Two Australian Government risk classes are recognised within these RASD types:

c) Official (Sensitive) – low to moderate impact, or
d) Official (Not Sensitive) – low impact

RASD risk classes generally fall within the equivalent of "Official - Government" risk classification. This means while action needs to be taken to manage RASD as per this framework, action only need be in line with normal institutional processes and thus does not require encryption or other high end security protocols.

The risk matrix for RASD shown in Table 1 shows:

a) the types of RASD listed in this framework
b) a summary of the reason why it is sensitive
c) how it should be handled if data are managed according to the framework
d) what the sensitivity of the data are if the data are handled according to the framework.

The framework suggests handling of both existing (legacy) datasets and new datasets.

Table 1- Risk Matrix for RASD

Note: This table only contains recommended handling and treatment for RASD based on recommended classifications under the Auditor General's guidelines on classifying data. Commonwealth, State and Territory governments may have their own classification systems. None of the RASD types listed here are security classified within the meanings of classifications used by Commonwealth, State and Territory governments. Most RASD require routine levels of protection (the equivalent of the Auditor General's class: OFFICIAL requiring dissemination limiting handling and marking (DLM)').

For the purposes of this table, entity refers to the data custodian and considers risks to their organisation.

Key:
Orange = Official (Sensitive) – low to moderate impact
Green = Official (Not Sensitive) – low impact

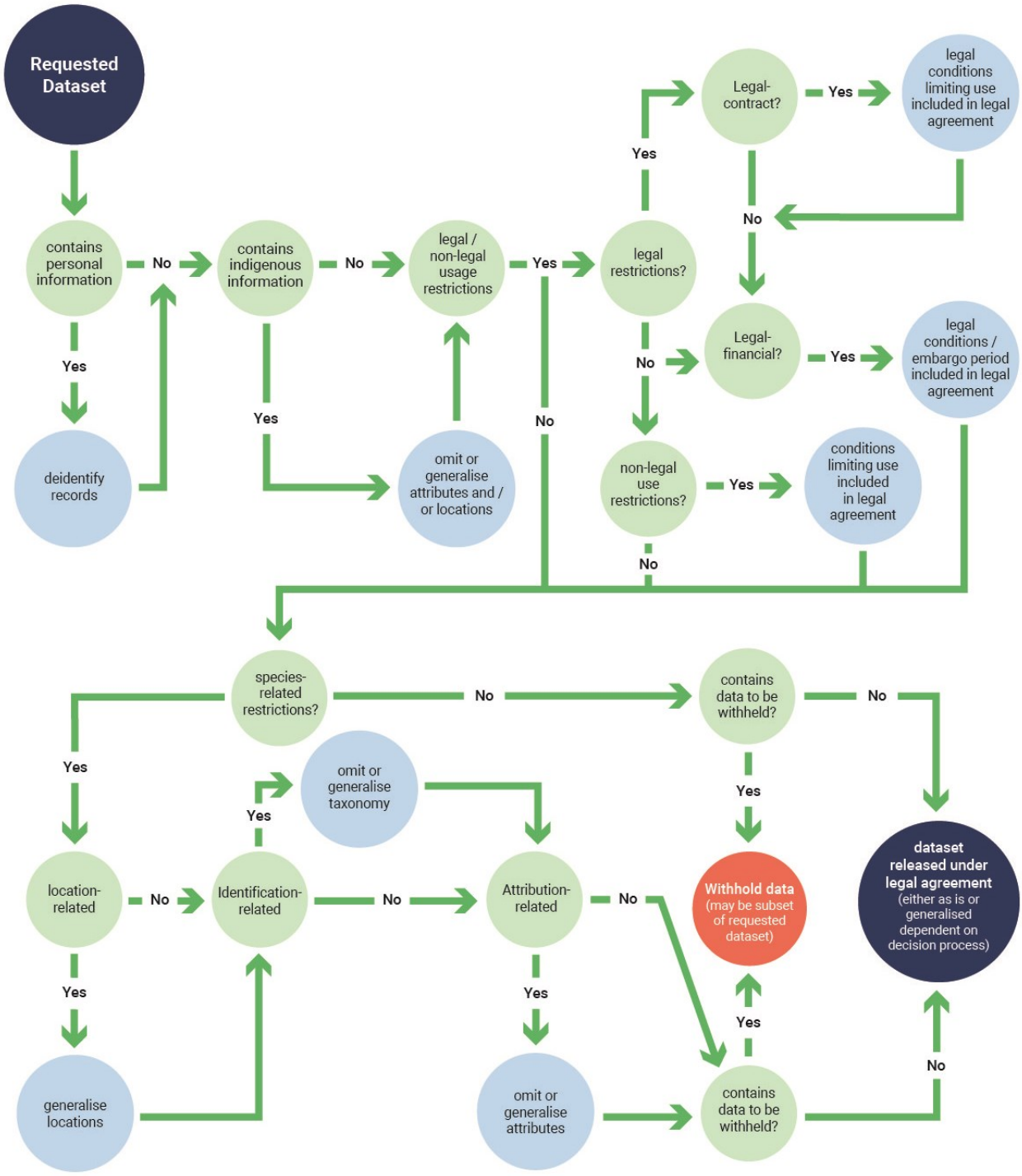
RASD Types	Categories	Reason	Handling of existing or new data	Classification (after handling)
Personal	**1 Personal Identifiable Information**	Privacy legislation requires the removal of information related to individuals from datasets and the control of de-contextualised datasets where an individual may be reasonably identified from the data.	**Handling of both data types:** Data should have fields including names and addresses excluded from the dataset before sharing. Dataset should be annotated to record the dataset has been modified. If both conditions met data can be shared publicly.	Official: (not sensitive – low impact) Information from routine business operations and services. Potential impact on individuals from compromise of the information. Personal information as defined in the Privacy Act (includes names and address). The information needs management but is not sensitive.
Indigenous	**2. Indigenous Data**	CARE principles require consultation with Indigenous parties about use of data before release.	**Existing Data** Data meeting the definition of Indigenous should be flagged as such in an attribute field. Data custodians should consult with Indigenous parties before releasing data. If consultation allows for specified use by third parties it can be shared as a restricted dataset but should be annotated "not for external use" and shared under negotiated legal agreement.	Official: (not sensitive – low impact) Information from routine business operations and services. Potential impact on communities from compromise of the information. The information needs management but is not sensitive.

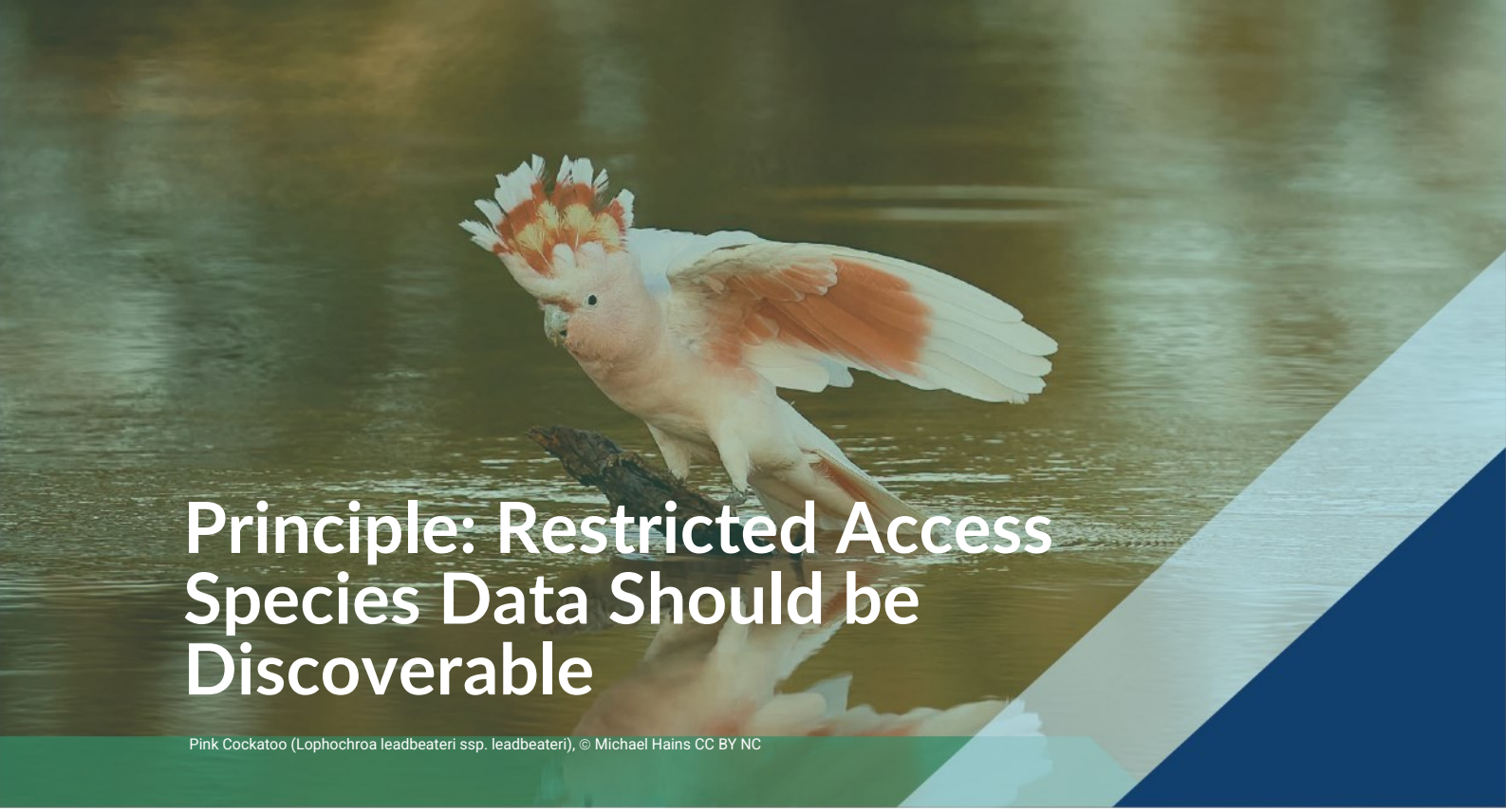
RASD Types	Categories	Reason	Handling of existing or new data	Classification (after handling)
			If consultation does not allow use by third parties then the data should be withheld, or agreement sought from the data custodian for its release. **New data** New data should be gathered in consultation with Indigenous communities and community wishes should be captured in an attribute field that defines how data are to be handled. Negotiated legal agreements to contain clauses consistent with the Framework enabling sharing within government and, under negotiated legal agreement, with Approved Data Requestors, but community consultation may well remain a requirement.	
Usage	**3.1 Legal Contract**	Conditions in a standard form data licence agreement or negotiated legal agreement contractually limit data-sharing.	**Existing Data (where data licence agreements exist)** If standard form data licence agreement allows for specified use by third parties it can be shared as a restricted dataset but should be annotated "not for external use" and shared under negotiated legal agreement. If standard form data licence agreement does not allow use by third parties then the data should be withheld, or agreement sought from the data custodian for its release. **New data** Future negotiated legal agreements to contain clauses consistent with the Framework enabling sharing within government and, under negotiated legal agreement, with Approved Data Requestors. Dataset should be annotated to record the dataset is restricted.	Official: Sensitive – low to medium impact Limited damage meets the definition of contract or agreement non-compliance
	3.2 Legal Financial	Written conditions on the dataset are part of current financial negotiations and transfer could compromise a financial process such as a tender.	**Existing Data (with data licence agreements in place)** If standard form data licence agreement allows for specified use by third parties it can be shared as a restricted dataset but should be annotated "not for external use" and shared under negotiated legal agreement. If the release of the data are subject to an embargo period and the data license permits use after this period only, data can be shared as a restricted datasets after the embargo period has	Official: Sensitive – low to medium impact Limited damage meets the definition of contract or agreement non-compliance

RASD Types	Categories	Reason	Handling of existing or new data	Classification (after handling)
			ended but should be annotated "not for external use" and shared under negotiated legal agreement. If standard form data licence agreement does not allow use by third parties then the data should be withheld, or agreement sought from the data custodian for its release. **New data** Future Negotiated legal agreements to contain clauses consistent with the Framework enabling sharing within government and, under negotiated legal agreement, with Approved Data Requestors. Dataset should be annotated to record the dataset is restricted.	
	3.3 Non-Legal	The dataset was acquired via informal agreement that the information would not be transferred (informal meaning no written Legal Contract)	**Existing Data** Data may be subject to restricted sharing under negotiated legal agreement or standard form data licence (where these already exist) with the data custodian or otherwise annotated as restricted. **New data** Future negotiated legal agreements to contain clauses consistent with the Framework enabling sharing within government and, under negotiated legal agreement, with Approved Data Requestors. Dataset should be annotated to record the dataset is restricted.	Official: (not sensitive – low impact) Information compromise would not result in legal and compliance issues. The information needs management but is not sensitive.
Species Related	**4.1 (a) Location**	The location of a species requires withholding to minimise human disturbance. Species locations are available publicly e.g., via published journal articles	**Handling of both data types:** Data should be made publicly available subject to obscuring locations consistent with the Framework. The dataset should be annotated to record the dataset has been modified. The restricted dataset with full locational data are available under legal agreement to all approved users under the Framework.	Official: (not sensitive – low impact) Information from routine business operations and services. Information is available from published literature such as original species descriptions. The information needs management but is not sensitive.

RASD Types	Categories	Reason	Handling of existing or new data	Classification (after handling)
	4.1 (b) Location	The location of a species requires withholding to minimise risk to species or people. Species locations are not available from any source	**Handling of both data types:** Data should be made publicly available subject to obscuring locations consistent with the Framework. The dataset should be annotated to record the dataset has been modified. The restricted dataset with full locational data are available only within jurisdictions on an as needs basis under negotiated legal agreement between parties.	Official: Sensitive – low to medium impact Limited damage to entity operations is a degradation in organisational capability to an extent and duration that, while the entity can perform its primary functions, the effectiveness of the functions is noticeably reduced and/or minor loss of confidence in government.
	4.2 Identification	Species whose occurrence has economic or legal ramification if disclosed	**Handling of both data types:** The restricted dataset with full locational data are available only within jurisdictions on an as needs basis under negotiated legal agreement between parties.	Official: Sensitive – low to medium impact Limited damage to entity operations is a degradation in organisational capability to an extent and duration that, while the entity can perform its primary functions, the effectiveness of the functions is noticeably reduced and/or minor loss of confidence in government. Limited damage to entity assets or annual operating budget is equivalent to $10 million to $100 million.
	4.3 Attribute	The metadata or attribution associated with a species record adds additional sensitivity	**Handling of both data types:** At the discretion of data custodians / agencies, the restricted dataset with full attribution is available under negotiated legal agreement to all approved users under the Framework.	Official: Sensitive – low to medium impact Information compromise may cause minor loss of confidence in government.

Figure 2 - RASD Data Handling Process (see also Table: Restricted Access Treatments and Metadata)





Principle: Restricted Access Species Data Should be Discoverable

Pink Cockatoo (Lophochroa leadbeateri ssp. leadbeateri), © Michael Hains CC BY NC

Data Catalogue

The Australian Government Best Practice Guide to Applying Data Sharing Principles specifies that a catalogue of held datasets is best practice, encourages transparency and enables decision-making on datasets to take place in an informed environment.

Recognising the limited resources available to most data custodians, working towards a publicly available catalogue of all Restricted Access Species Data (RASD) datasets, with metadata, to enable users to understand what datasets are accessible and not accessible, is strongly encouraged.

Metadata statements need to conform to the standards set out in the best practice metadata sample available in Supplement 4: Restricted Access Species Data Metadata Statement Template and row level metadata to conform to Darwin Core.

Provision of Data Catalogue

Data custodians should provide a copy of their RASD catalogues for inclusion in a centralised metadata repository such as Research Data Australia.

Principle: Restricted Access Species Data requests should follow a structured, transparent process

Queensland Whistling Tarantula (Selenocosmia crassipes), © Graham Armstrong

The suggested high-level data access request process is shown in Figure 3.

Classes of RASD Requests

This framework proposes three classes of typical data access:

1. raw data, as held in the data custodian's system (access after negotiated legal agreement)
2. structured full resolution data with Personal Identifiable Information data transformations and ability to download (access after negotiated legal agreement)
3. structured data with all transformations and obfuscations applied (open access)

The decision on which class of access to provide a requestor remains with a data custodian, ideally, consistent with the principles set out in this framework.

Requests Between Data Custodians Operating Consistently with this Framework

All reasonable access requests for internal use between data custodians who are operating consistently with the framework should be agreed and delivered.

It is best practice that Personal Identifiable Information (see Supplement 3: Personal Identifiable Information in Restricted Access Species Data) should be removed.

Providing Data to Approved Data Requestors

Full resolution data (being a full dataset less those records / datasets specified in the supplements relating to Indigenous Data (Supplement 1), Withheld Data (Supplement 7) and Personal Identifiable Information (Supplement 3) to Approved Data Requestors who have a) successfully applied to a data custodian; b) been approved consistent with this framework should be supplied by data custodians who operate consistently with this framework.

Principles of Providing Data

The central principles of this framework are to:

1. encourage data sharing between trusted parties so that the maximal amount of restricted data are available for important end-uses such as decision-making and research
2. where data are made publicly available, they are provided in an appropriate and consistent form to support conservation in Australia

Data custodians are responsible for the provision of their data and should make best efforts to apply all data transformations consistent with this framework, or, if supplying raw or structured full resolution data, under negotiated legal agreement.

Receiving Access Requests

The Australian Best Practice Guide to Applying Data Sharing Principles suggests the logical fields for a data request. Requests for RASD should meet the following criteria:

- define a geographic / taxonomic area of interest – the request form should ask data requestors to define the geographic / taxonomic extent of their request.
- demonstrate an appropriate aim – the request should clearly articulate the aim of the intended use. For research, the aim should include a statement on the topic of the research; for industry use, the request should specify the type of activity data are being sought for.
- demonstrate a public benefit – the request should clearly articulate what public benefit is derived by accessing the data and the expected outputs and outcomes.
- show that legal, ethical, and moral considerations (including CARE principles) have been addressed – how should the request ensure that the data are not misused? Particularly, the data requestor has to make a statement on how they should protect the data from being accessed inappropriately.
- state what data should be used and why it's required - the request should clearly articulate for what purpose RASD is required. For research, this should include a description of the project and expected outcomes; for industry use, the request should specify the type of activity and location.
- state the timeframes for which the user needs the data
- state who (both the entity requesting the data and nominated responsible individual within the entity) will be working on the project
- demonstrate feasibility – the data requestor needs to demonstrate that the requested data are suitable for answering the stated aims of the project.

Normally, requests meeting criteria and requesting access to a localised geographic area, or data for a single or small group of species should be approved. Requests for data at a jurisdictional or national level should receive greater scrutiny.

An objective of data requests under the framework should be to prevent restricted data being passed to third parties or entering the public domain.

A best practice guide to fields for an electronic form meeting these requirements is available in Supplement 9.

If a data requestor has identified as an Indigenous body or organisation on the data access request form, and they have not previously been subject to a breach, every effort should be made by the data custodians to facilitate the sharing of data with the requestor.

Assessment of Access Requests

Responsibility for approving access requests and releasing data remains with data custodians, except where the negotiated legal agreement between data custodian and a third-party has delegated access request assessment and approval.

Requests should be assessed according to the access request assessment process in Figure 3.

Data Custodians should take best endeavours to ensure the following:

a) assessments are undertaken consistent with the criteria in this framework.
b) reviews are undertaken in a timely manner.
c) responses to data requestors are in written form and, where a data request is rejected, provide a reason consistent with this framework.

It is best practice to recognise that datasets containing Indigenous data or knowledge may have longer assessment timeframes due to the consultation necessary for ensuring Indigenous data sovereignty.

Denial of Access Requests and Right of Appeal

Data custodians retain full rights to refuse a data request following an assessment using the processes outlined in this framework. For transparency, custodians should make best endeavours to provide a reason for a refusal to a data requestor in a written response.

It is the responsibility of the individual data custodians to make best endeavours to develop a transparent right of appeal process.

Negotiating Legal Agreements

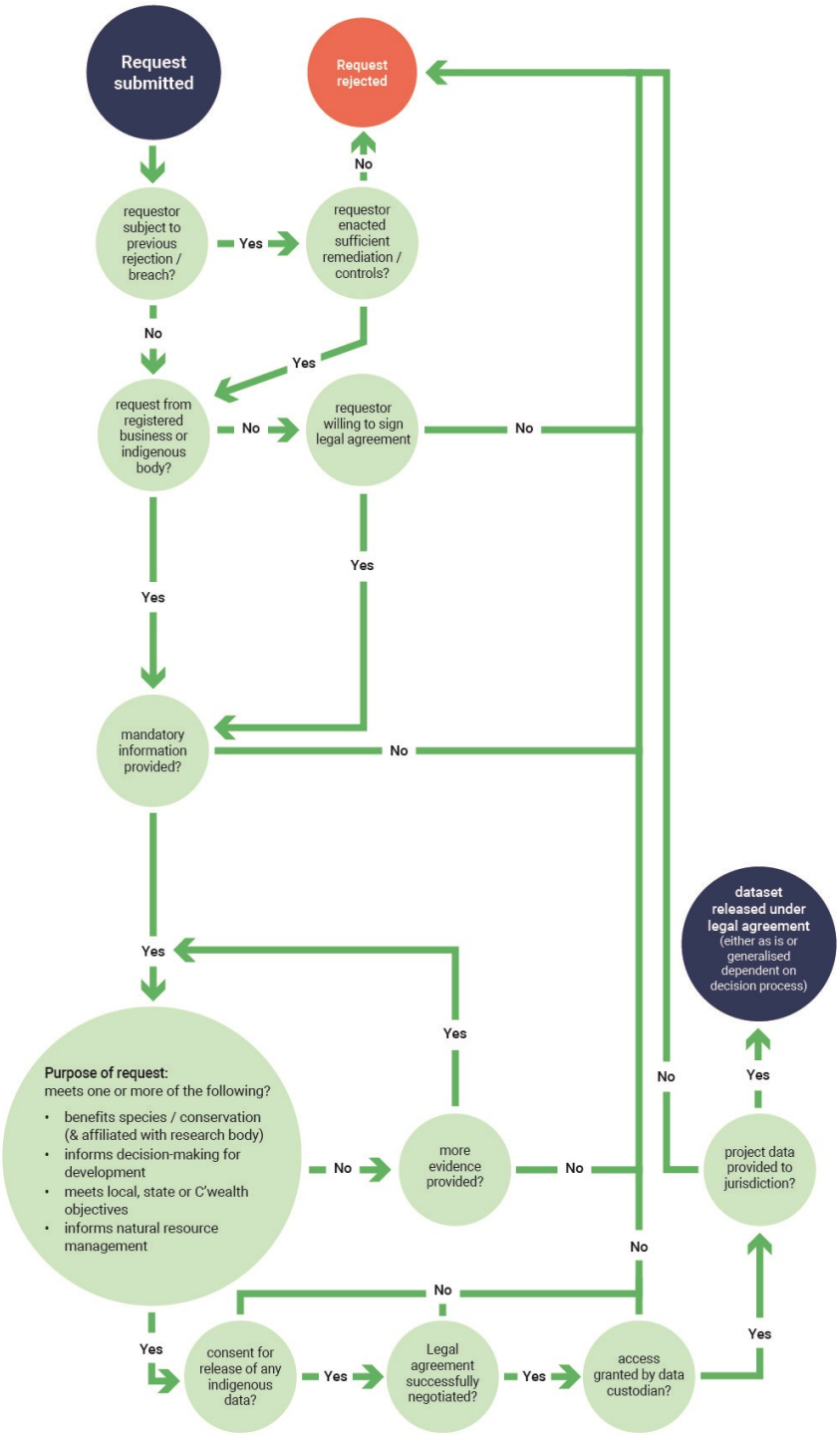
Following the assessment of the data request, the data custodian is responsible for negotiating their own legal agreements. A guide to what should be covered in a negotiated legal agreement about data sharing is provided in Supplement 2: What Legal Clauses Should be Included in a Restricted Access Species Data Negotiated Licence Agreement?

Releasing Data

Following the successful negotiation of a legal agreement, data are released to the data requestor consistent with this legal agreement.

It is best practice to enable tracking of data citations by minting a Digital Object Identifier (DOI) or equivalent for a dataset. DOIs linked to each data request mean that how the data are used can be reported back to the data custodian. Data custodians are strongly encouraged to apply similar techniques. Advice on citation and identifiers is available from the Australian Research Data Commons.

Figure 3- RASD Data Access Request Assessment Process



Principle: Restricted Access Species datasets should be transformed consistently if made public or be as complete as possible if provided to approved data requestors

Regent Honeyeater (Anthochaera phrygia), © Time Nickholds CC BY NC

It is good practice to recognise that there are different classes of data user making requests and the following provides guidance.

Full Access Users

Full Access Users are those requesting access to the full dataset with a minimum of withheld information. These are nominally likely to be data custodians operating consistently with and referencing this framework exchanging data with one another for the purposes of conservation, administration, research, or environmental decision-making. Data custodians sharing data between themselves should have signed a negotiated legal agreement for access to full resolution RASD.

Approved Data Requestors

Approved Data Requestors are users whose request has been assessed and granted. There are several levels of access for these users:

1. access to the entire dataset for a specified period – this normally covers high-level use such as contractors working for government on a specific project or research institutions conducting extended studies who need access either to the entire dataset or all data from a particular jurisdiction. Usage terminates at the end of the specified period
2. access to a subset of the dataset (e.g. a specified local area or all data relating to a species or group of species for a specified period) – this covers most use cases from research and commercial end users
3. access to a protected environment where data can be queried but not copied – this covers the alternative to level two, where commercial interests or research users wanting relatively fast access to large datasets can enter a protected environment, interrogate data, but not download a copy. Derived products such as models may be downloaded (this option is currently not widely available)

Generally speaking, individuals should not be recognised as Approved Data Requestors because of the legal complexities surrounding data access. This means a university school, rather than a PhD student, would be signatory, with the expectation that the university will take responsibility for ensuring that the student uses the data legally and responsibly.
All Approved Data Requestors should preferably only be able to acquire data if they sign a negotiated legal agreement (not a standard form data licence agreement) controlling data use (Supplement 2 and Principle: Sharing or Restricted Access Species Data Should be Through Negotiated Legal Agreement.

Instances where Approved Data Requestors breach legal agreement conditions may result in the Requestor having their agreement revoked. The breach should be considered in new data requests. Both the data custodian and the requesting organisation must comply with any jurisdictional breach reporting requirements.

Public Access

It is recognised that many data custodians provide publicly available data in various forms to the general public.

It is best practice to work towards transforming RASD that is exposed publicly following the process in Supplement 6. This ensures that data are transformed consistently, and the users understand how the data they are using has been transformed.

Principle: Sharing of Restricted Access Species Data Should be Through Negotiated Legal Agreement

(Ctenotus lancelini), © Robert Browne-Cooper CC BY NC

RASD is not publicly available and often has many restrictions around its use. Because of this, it is best practice to let two overarching principles guide the mechanism by which RASD is shared:

- RASD should be shared under some form of legal agreement (either standard form data licence agreements or negotiated legal agreements)
- Negotiated legal agreements are recommended rather than standard form data licence agreements. Negotiation allows the terms and conditions to be tailored to the specific data and use requirements, and should promote compliance by enhancing awareness of the terms

Most data custodians currently make use of standard form data licence agreements when sharing data. The difference between the two types of agreements is:

a) *Standard Form Data Licence Agreement* – This is a license used between most data custodians and repositories, which uses standard (non-negotiable) terms and conditions to stipulate management of data including to control end user use of data.
b) *Negotiated Legal Agreement* – for the purposes of this framework, this is a legal agreement between data custodians and Approved Data Requestors, the terms of which are negotiated by the parties.

It is recognised that data custodians with existing processes that do not involve negotiated legal agreements may require time to transition but will ideally be working towards implementing a new process.

What Should be Included in a Negotiated Legal Agreement?

A guide to what should be covered in a negotiated legal agreement about data sharing is provided in Supplement 2: What Legal Clauses Should be Included in a Restricted Access Species Data Negotiated Licence Agreement?

Once approval has been provided by data custodians, the intention of the negotiated legal agreement is to define between the data custodian and the entity requesting third-party data access what the expectations are about data use and to ensure that data are cited appropriately and removed after use.

Approval can only be granted by the custodian, either directly or as specified in a negotiated legal agreement.

The agreement should specify:

- the duration of the agreement
- citation of the dataset
- third-party Intellectual Property (IP)
- the requirement that data are held in a secure environment with role-based access controls where necessary
- whether data can be passed onto third parties with or without consent of the data custodian
- that data to be used for the purposes outlined in the agreements only and not for any other purpose
- the consequences if a breach of agreement conditions occurs including suspension of current and future access requests

What Happens if a Breach Occurs?

Breaches vary in seriousness. Some breaches may occur inadvertently and have at most, minor consequences (or potential consequences). Others may be more serious – for example they may result from the deliberate actions or negligence of the data recipient, and/or have significant actual or potential consequences. It is the responsibility of data custodians and users to familiarise themselves with any relevant breach reporting requirements in their jurisdiction, and for custodians to determine what constitutes a serious breach.

Data requestors that have been assessed consistently with the principles of this framework as Approved Data Requestors may, on occasion, breach the conditions of their negotiated legal agreement for data use. Breaches should be dealt with by the data custodian responsible for the agreement consistent with the conditions of the agreement.

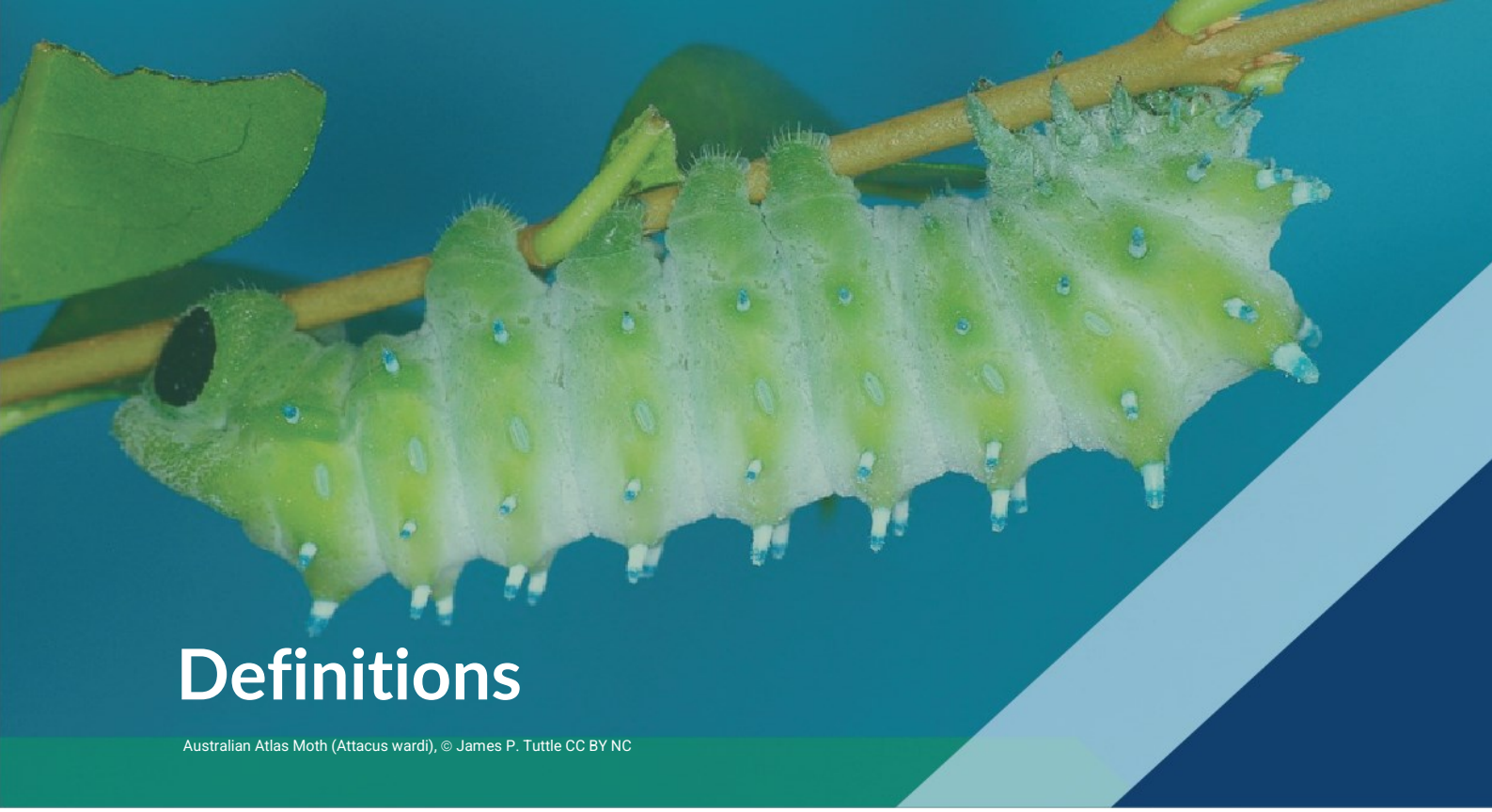
Where serious or multiple breaches occur, it is in the interests of data custodians to warn other data custodians of known offenders and take this into consideration in approving new data requests. In the case of a current negotiated legal agreement, the agreement should be cancelled immediately.

An appeals process should be set up by the data custodian for approved data requestors to appeal any breaches or rejections of data access requests.

New data Acquisition and New Licence Agreements

Data custodians operating consistent with and referencing this framework are committing to improve RASD access in future. A critical component of this should be to aim for future creation or acquisition of data to be consistent with the principles of this framework.

Data custodians should work towards new data acquisition or third-party data access negotiated to consistent with the **Principle: Sharing of Restricted Access Species Data Should be Through a Negotiated Legal Agreement** and the suggestions in **Supplement 2**: What Legal Clauses Should be Included in a Restricted Access Species Data Negotiated Licence Agreement? are recommended as best practice. This is to ensure transparency about data access and flow.



Definitions

Australian Atlas Moth (Attacus wardi), © James P. Tuttle CC BY NC

Approved Data Requestors – Registered businesses (government, non-government or research entities or businesses (not individuals)) who have applied under the processes stipulated in this framework and have been found to meet the requirements to be allowed a level of access to Restricted Access Species data.

Atlas of Living Australia – National species data aggregator funded by the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS).

Biodiversity Data Repository – database administered by the Commonwealth Department of Climate Change, Environment, Energy and Water providing highly defensible, highly curated national biodiversity data to governments, industry, and the community for decision-making, planning and reporting.

Darwin Core - Darwin Core (often abbreviated to DwC) is an extension of Dublin Core for biodiversity informatics. It provides a stable standard reference for sharing information on biodiversity. The terms used in the standard are a part of a larger set of vocabularies and technical specifications under development and maintained by the Biodiversity Information Standards Group (also known by the abbreviation TDWG). TDWG is an international reference group. Its previous title was the Taxonomic Databases Working Group (hence the acronym).

Data Custodians – The managers of data repositories. For the purposes of this Framework, data custodians comprise any data custodians that have decided to operate consistent with this framework.

Data Requestor – A user requesting data that has not yet been assessed under this framework. Data requestors that have been assessed and provided access are termed "Approved Data Requestors" (see separate definition).

Digital Object Identifier – A unique identifier used to permanently identify a digital artefact including a link to that artefact on the internet.

Embargoed Data – data that has a timestamp preventing release before a certain date. It is dealt with in the **Supplement 7: Withheld Data** and **Supplement 8: Process for Handling Embargoed Restricted Data**.

Existing (Legacy) Data – Existing or legacy data are data already held by a data custodian prior to the decision to act consistently with this framework.

Framework or National Framework – This document.

Full Access Users – Users who have access to all data within the RASD (excepting Personal Identifiable Information)

Full Resolution Data – The dataset as stored by the data custodian with Personal Identifiable Information removed.

Jurisdictions Australian, State and Territory Government environmental agencies.

Metadata – For the purposes of this framework, metadata are either:

a) **Dataset metadata** – a concise description of a dataset which enables a user, unable to access a dataset, to gauge the relevance of the data for their purposes; or
b) **Record (or row)-level metadata** – these are documentation at the level of a record in a dataset. For the purposes of this framework, it largely refers to documentation of the sensitivity status of the record (or the species of which it is a part) along with access constraints pertaining to the record and details of any generalisation of the data

National Framework for the Sharing of Restricted Access Species Data in Australia – This framework. A set of guidelines for data custodians providing best practice guidance

Negotiated Legal Agreement – for the purposes of this framework, a legal agreement between data custodians and Approved Data Requestors, the terms of which are negotiated by the parties. This framework encourages the use of these agreements instead of Standard Form Data Licence Agreements

New Data – Data acquired after a data custodian has decided to act consistently with this framework.

Obfuscation – The practise of rounding or randomising, according to an agreed standard, the geo-localities in a dataset to make it difficult to discern the original geo-locality. Randomisation has benefits for map display but is not a robust data management approach in a data ecosystem. Rounding or the provision of records as grid square polygons is preferred if data are to be passed between systems that may then apply their own additional obfuscation. For the purposes of this framework and the best practice guide, references to obfuscation imply rounding or the provision of records as grid square polygons, not randomisation.

Organisations operating and referencing this framework – Organisations that state publicly or reference that their processes are consistent with this framework.

Personal Identifiable Information – Data that meets the definition of private information under any of the Commonwealth, State or Territory legislative or regulatory definitions. For the purposes of Restricted Access Species Data, generally names and/or contact details held in third-party datasets

where the record contributor has not given permission for the sharing of their data. Discussed in full in Supplement 3.

Raw Data – Restricted Access Species (RASD) as stored by the data custodian including Personal Identifiable Information.

Registered Business – A legal entity with a registered Australian Business Number. May be a government, non-government, or research organisation, or a business.

Restricted Access Species Data (RASD) – For the purposes of this framework, RASD are any dataset or record related to species that is determined to contain restricted access or sensitive information (data class or geolocational) as defined by this framework that should not be shared openly. RASD categories are defined in Table 1

Restricted Access Species List (RASL) – RASLs, also known as sensitive species lists, are maintained by each jurisdiction in Australia. A jurisdictional RASL is a publicly available list that delineates which species should have their geographic locations blurred or obfuscated to prevent disturbance or exploitation of the species or the site. In addition, Third-party data custodians (including individuals) may have their own RASLs.

Species – in the context of this framework, species refer to all species, subspecies or infraspecies, including (where identified under RASLs) populations, varieties, and phrase name taxa.

Standard Form Data Licence Agreement – A license used between most data custodians and repositories, which uses standard (non-negotiable) terms and conditions to stipulate management of data including to control end user use of data.

Transformation – The practice of modifying the fields or data in a dataset. For the purpose of this framework:

Transformation involves:

(i) the bulk manipulation of data to desensitise for exchange or sharing, including obfuscation

Removals include one or more of the following:

(i) removing Personal Identifiable Information
(ii) removing fields containing sensitive data (other than Personal Identifiable Information)
(iii) removing rows containing sensitive data (other than Personal Identifiable Information)
(iv) temporarily removing data held under an embargo
(v) removing data with legal or identification constraints

Flagging includes one or more of the following:

(i) adjustment of "local" taxonomy to an agreed national standard (while retaining local taxonomy as well)
(ii) flagging of suspect records

Withheld Data – Data which does not leave a Data Custodian's database at all UNLESS relevant transformations metadata and removals outlined in this framework are applied. May be a full dataset or fields within a dataset. Withheld data are dealt with in Supplement 7.

Organisations Involved in the Development of this Framework

Organisations overseeing the compilation of this framework:

- Atlas of Living Australia, CSIRO
- Council of Heads of Australian Faunal Collections
- Council of Heads of Australasian Herbaria
- Department of Climate Change, Environment, Energy and Water (the Commonwealth)
- Department of Biodiversity, Conservation and Attractions (WA)
- Department of Energy, Environment and Climate Action (VIC)
- Department of Environment and Natural Resources (NT)
- Department of Environment and Science (QLD)
- Department of Environment and Water (SA)
- Department of Planning and Environment (NSW)
- Department of Natural Resources and Environment (Tas)
- EcoCommons Australia
- The Environment, Planning and Sustainable Development Directorate (ACT)
- Western Australian Biodiversity Science Institute

Supplement 1: Recognising Indigenous Data Sovereignty in Restricted Access Species Data
(Version 1/12/2022)

The CARE principles for Indigenous data governance (Collective Benefit, Authority to Control, Responsibility and Ethics) are intended to recognise Indigenous data sovereignty.

Suggestions on how biodiversity data may be treated or gathered using these principles are contained in Robinson et al (2021)[1].

The CARE principles apply to Indigenous data and knowledge generally, but for the purposes of the current framework, this attachment specifically references species point data, including knowledge about those species or data points gathered by or about Indigenous peoples.

Data Custodians are encouraged to adopt the following principles:

a) **Existing Data**
All species point data, including knowledge about those species or data points gathered by or about Indigenous peoples should carry an attribute field that defines a record as this type of data. The attribute field should indicate what individual or organisation should be contacted to discuss data release.

The consultation should cover whether data can be released, whether derived products may be produced from the released data and how long the data may be shared. If data are not to be released, agreement should be sought to a metadata statement that describes, but does not contain, the withheld data.

Negotiated legal agreements on data to be released should stipulate any conditions requested in Indigenous consultation.

b) **New Data**
New species point data should be gathered with full consultation with Indigenous communities. This consultation should cover whether data can be released, whether derived products may be produced from the released data and how long the data may be shared. If data are not to be released, agreement should be sought to a metadata statement that describes, but does not contain, the withheld data. Ideally this information should be included in the attribute field relating to Indigenous data.

The result of consultation may well be that communities still wish to be consulted before data are released, in which case the attribute field should indicate what individual or organisation should be contacted to discuss data release.

Negotiated legal agreements on data to be released should stipulate any conditions requested in Indigenous consultation.

1 Caring for Indigenous Data to Evaluate the Benefits of Indigenous Environmental Programs, Cathy J. Robinson, Taryn Kong, Rebecca Coates, Ian Watson, Chris Stokes, Petina Pert, Andrew McConnell, Caron Chen (2021) Environmental Management 68:160–169 https://doi.org/10.1007/s00267-021-01485-8

Supplement 2: What Legal Clauses Should be Included in a Restricted Access Species Data Negotiated Licence Agreement?

(Version 1/12/2022)

This supplement provides sample clauses to be included in legal agreements either between data custodians or between data custodians and data receivers under the RASD Framework. Data receivers may be Approved Data Requestors or, in the case of data access requests between data custodians, the requesting data custodian.

These clauses address:

- the duration of the agreement
- citation of the dataset
- third-party Intellectual Property (IP)
- the requirement that data are held in a secure environment with role-based access controls where necessary
- whether data can be passed onto third parties with or without consent of the data custodian
- that data to be used for the purposes outlined in the agreements only and not for any other purpose
- the consequences if a breach of agreement conditions occurs including suspension of current and future access requests

In these sample clauses capitalised words are terms that should be defined in the negotiated agreement. It is important to clearly identify the “Data” that are the subject of the agreement.

The sample clauses use the terms “Supplier” and “Receiver”. For agreements between a data custodian and an Approved Data Requestor, the Supplier is the data custodian and the Receiver is an Approved Data Requestor. Text in yellow needs to be replaced when the agreement is being prepared.

Condition	Example Clause(s) to Add	Guidance
Provision of Data		
How and in what timeframe will the Data be delivered to the Receiver?	The Supplier will transfer the Data to the Receiver on [insert the date when the Data will be provided] by [describe the manner in which the Data will be transferred to the Receiver].	A clause should be included specifying when and how the Data are to be provided.
Who is responsible for ensuring that the Data are de-identified and that there are no -party restrictions on the transfer of Data?	1. The Supplier must ensure that the Data are de-identified, and that it has the right to transfer the Data to the recipient. 2. Notwithstanding clause 1, if any personal information is incorporated in the Data, each party must, in dealing with that personal information: (a) promptly notify the Supplier that the Data includes personal information; (b) comply with all laws applicable to that party (including, without limitation, privacy laws and data protection laws) which regulate the collection, storage, use and disclosure of Personal Information; (c) promptly notify the other party of any complaint or investigation under, or relating to, any breach of the foregoing laws in relation to that information; and (d) reasonably cooperate with the other party in resolving any such complaint or investigation.	This reflects the onus outlined in the Framework.
Is the Data to be held longer term and updates provided periodically?	(a) At the Receiver’s written request, which may only be made [stipulate agreed frequency of these requests], the Supplier will provide updates of the Data to the Receiver. (b) After receiving the written request under paragraph (a) and subject to paragraph (c) below, the Supplier will use reasonable efforts to provide the updates of the Data promptly. (c) The Receiver acknowledges that the updates of the Data may contain data (such as Indigenous data or -party data) that requires changes to the conditions placed on the use	This clause can be used to specify the frequency of updates e.g. annually or two times in a calendar year.

Condition	Example Clause(s) to Add	Guidance
	of Data under this Agreement. The parties agree to negotiate in good faith to vary this Agreement before the updates to the Data are released.	
Use of Data		
What can the Data be used for?	1. The Receiver will only use the Data for the Purpose. 2. No right, licence or ownership is granted to the Receiver in relation to the Data and any Intellectual Property in the Data, except as set out in this Agreement.	The purpose of this clause is to specify how Data can be used and to make clear that there are no additional rights to the Data or any IP in the Data. "Purpose" should be defined in the agreement. Thought should be given to whether Derived Products may be produced from the Data.
Are there restrictions on the use of the Data?	The Receiver will not use the Data for the following purposes: (a) [e.g. commercial purposes; (b) aggregation with other datasets for the purpose of publication; (c) teaching; (d) producing Derived Products; and (e) publication].	The purpose of this clause is to emphasise how Data cannot be used. Note that this clause operates in addition to the stipulation of what Purposes the data can be used for. As such, there is no need to be exhaustive. It can be used to stipulate uses that are of particular concern. If there is Indigenous data contained within the Data, any conditions requested by Indigenous peoples should be stipulated here. Note regarding "commercial purposes": if commercial purposes are to be excluded, consider

Condition	Example Clause(s) to Add	Guidance
		whether this needs to be defined. In particular, if the Receiver is a for-profit it may be necessary to carve out specific commercial uses that are not permitted.
Is there Indigenous data within the Data with discreet restrictions on use?	The Receiver acknowledges that the Data contains Indigenous data. The Receiver agrees that it will not use the Data for the following purposes without first speaking to the relevant Indigenous communities: (a) [stipulate any matters for which Indigenous peoples requested to be consulted e.g. commercial purposes or aggregation with other datasets for the purpose of publication]	Data Custodians will need to consult with Indigenous peoples prior to release of Data containing Indigenous data. In some cases, Indigenous peoples might agree to the use of Indigenous data for certain purposes. If so, those conditions can be listed in the clause in the row above ("Are there restrictions on the use of the Data?") However, in most cases, it will be appropriate for the Receiver to consult with Indigenous peoples before certain uses are made. Such conditions should be addressed in this clause.
Does the Data contain third-party datasets?	The Receiver acknowledges that the Data includes third-party data that can only be used for the Purposes defined in this Agreement.	A clause similar to this should be included in cases where the Data includes third-party data.

<table>
<tr><th>Condition</th><th>Example Clause(s) to Add</th><th>Guidance</th></tr>
<tr>
<td>Is access to the Data required for more than one individual?</td>
<td>
1. The Receiver must create the following roles to manage access to the Data:
<table>
<tr><th>Name of role</th><th>Description of role</th><th>Access Rights</th></tr>
<tr><td>Administrator</td><td>Assigns roles to other Authorised Users and manages Authorised User access</td><td>Full access to Data including read / create / delete / update.</td></tr>
<tr><td>General user</td><td>Uses the Data for the Purpose (or aspects of the Purpose)</td><td>Able to query and analyse but not edit the Data.</td></tr>
<tr><td>Developer</td><td>Uses the Data for the Purpose (or aspects of the Purpose)</td><td>Able to query and analyse but not edit the Data.

Able to create and publish Derived Products.

Able to integrate and query Data with other datasets.</td></tr>
</table>
2. The Receiver must maintain a record of Authorised Users, roles assigned to each Authorised User (as provided in clause 1) and what Data are accessible to each Authorised User.

3. The Receiver must ensure that each Authorised User is only given the Access Rights and the Data that the Authorised User needs for their role.

4. Before the Receiver provides Authorised Users with access to Data:

(a) the Receiver must ensure that Authorised Users understand and are able to comply with the Data use obligations in this Agreement; and
(b) the Receiver must maintain a documented record (co-signed by the Administrator and the general user or developer, as applicable) of the Authorised User's understanding, including a description of the Data use obligations.
</td>
<td>In some cases, more than one individual within a Receiver's organisation will need access to the Data. In such cases a clause should be included to specify how Data should be managed by the Receiver.

A possible definition of 'Authorised User' is "An employee, director, officer or student of the Receiver that needs, and is granted, access to Data by the Receiver and who is assigned a role under clause X of this Agreement."</td>
</tr>
</table>

Condition	Example Clause(s) to Add	Guidance
Are there restrictions to Data being passed to third parties by the receiving Approved Data Requestor?	Where Data must not be provided to third parties The Receiver acknowledges and agrees that the Data are being provided for use by Authorised Users only. The Receiver agrees that it has no right to provide the Data or any sub-set of the Data to any third-party and that it has no right to sub-licence any of its rights under this Agreement. Where Data can be provided to third parties [Option A] (a) If the Receiver wishes to pass Data received under this Agreement to a third-party, the Receiver must first obtain the express written permission of the Supplier. The Supplier may withhold permission at the Supplier's discretion. [Option B] (a) The Receiver may pass Data received under this Agreement to a third-party provided that the Receiver first notifies the Supplier of the third-party. (b) Before any Data are provided to a third-party, the Receiver must: (i) enter into a written agreement on terms that are similar to, consistent with and at least as onerous as the terms of this Agreement; and (ii) ensure that: A. no legal restrictions on the use of the Data are breached by providing the Data; B. the Data are cited and attributed correctly; and C. the Data does not contain any Personal Information.	A clause should be included to address whether Data – that is, the **untransformed** Data received from the Data Custodian – can be provided to third parties. If Data can be provided to third parties, the clause should specify the circumstances under which Data may be transferred. Two potential approaches are suggested. Another approach would be to permit Data to be passed on to certain categories of third parties or for certain usages only.
Are there restrictions to Derived Products being passed to third parties by the Approved Data Requestor?	Where Derived Products must not be provided to third parties The Receiver agrees that it has no right to provide Derived Products to any third-party. Where Derived Products can be provided to third parties (a) [Option A] Subject to clause (b) and (c) [amend as appropriate], the Receiver may provide Derived Products to a third-party. (b) [Option B] If the Receiver wishes to provide Derived Products to a third-party, the Receiver must first obtain the express written permission of the Supplier. The Supplier may withhold permission at the Supplier's discretion.	A clause should be included to address whether Derived Products can be provided to third parties. Here Derived Products means any dataset or product that is produced using or derived from the Data

Condition	Example Clause(s) to Add	Guidance
	(c) Before any Derived Product is provided the Receiver must ensure that: (i) the location of restricted access species or sensitive attributes relating to them cannot be derived from the Derived Product; (ii) no legal restrictions on the use of the Data are breached by providing the Derived Product; (iii) the Data are cited and attributed correctly in the Derived Product; and (iv) the Derived Product does not contain any Personal Information. [Where an embargo period applies prior to release of a Derived Product] (a) The Receiver will not release any Derived Products to the public or to third parties prior to DD/MM/YYYY.	including transformed data and models. If Derived Products can be provided to third parties, this clause should stipulate the conditions of release including any embargo periods, what transformations are to be applied and what output checks are required before the Derived Product is released.
What must the Receiver do to protect the Data?	The Receiver must maintain reasonable administrative, technical, and physical safeguards to: (a) protect the Data from loss or misuse; (b) ensure that the Data are only accessible by Authorised Users; and (c) protect the Data from unauthorised use, modification, or disclosure [, including (without limitation) through the measures stipulated in clause [refer to "Is access to Data required for more than one individual at the Receiver?"] of this Agreement].	The approach in the Example Clause does not stipulate that the Data are held in a certain location; rather it stipulates that however the Data are held, the Data must be afforded certain protections. Where there are multiple users of the Data at the Receiver, the bracketed text should be included, which provides a cross reference to the provisions under the Condition "Is access to Data required for more than one individual at the Receiver?".
Term and termination		
Is the Data for a specified project with an agreed defined timeframe in which the Data are to be used?	This Agreement will commence on the Delivery Date [the agreed date for delivery of the Data by the Data Custodian to the Receiver] and end on DD/MM/YYYY, unless terminated earlier under clause x of this Agreement.	

Condition	Example Clause(s) to Add	Guidance
Is the Data for an agreed range of uses that are ongoing, such that an option to extend the term of the agreement should be included?	This Agreement will commence on the Delivery Date [the agreed date for delivery of the Data by the Data Custodian to the Receiver] and end on DD/MM/YYYY, unless terminated earlier under clause x of this Agreement. This Agreement may be extended for a further [stipulate time period] by mutual written agreement.	Clause specifying option to extend the term of the Agreement.
What are the consequences of termination of the Agreement (either because the term has expired, or the Data Custodian has terminated the Agreement)?	1. If this Agreement is terminated or the term has expired, [except to the extent the Data has been incorporated into a Derived Product consistent with the terms of this Agreement]: (a) the Receiver must immediately stop using the Data; (b) any right or licence granted to the Receiver in relation to the Data and any Intellectual Property in the Data, expires; and (c) [subject to clause 2,] the Receiver must destroy all copies of the Data on the Receiver's systems and devices within 7 days of the date of termination. [2. The Receiver may retain one copy of the Data with the project file for a period of [specify period] after the end date of this agreement for the sole purpose of [specify the reason that the Data can be retained]. The Receiver must maintain the safeguards outlined in clause ["What must the Receiver do to protect the Data?"] for the duration of this period.]	This clause should address what the Supplier requires the Receiver to do with the Data upon termination of the agreement. If relevant, this clause will need to take into account Data that has been legitimately incorporated within Derived Products. Consider whether the Receiver will need to retain a single copy of the Data for a specified purpose and a specified period of time.
What are the consequences of breach of agreement conditions?	5. If the Data Custodian reasonably determines that the Receiver has: (a) used the Data for any purpose other than the Purpose [; or (b) passed the Data to a third-party; or (c) passed a Derived Product to a third-party,] the Data Custodian may immediately terminate the Agreement by written notice. 2. The Receiver acknowledges and agrees that in the event that the Data Custodian reasonably determines that the Receiver has committed a serious or repeated breach of the conditions of use outlined in this agreement: (a) the Data Custodian may notify other Data Custodians that are signatories to the Framework that the Receiver has committed a serious or repeated breach of the conditions of Data use contained in this Agreement; and	A clause should be included to specify what circumstances warrant immediate termination of the agreement. The text in brackets will only be appropriate if the agreement restricts third-party access to Data and/or Derived Products.

Condition	Example Clause(s) to Add	Guidance
	(b) the Receiver's behaviour may be taken into account by other signatories when assessing future requests from the Receiver to access restricted access species data from other signatories.	
General provisions		
What is the mechanism for handling any disputes that arise under this Agreement?	(a) Any dispute, controversy, difference, or claim arising out of or in connection with this Agreement and/or the subject matter of this Agreement, including its existence, breach, validity, or termination (**Dispute**), must be dealt with in accordance with this clause. (b) Nothing in this clause prevents any Party from seeking urgent injunctive or similar interim relief from a competent court. (c) Any of us claiming that there is a Dispute must notify each other in writing and give details of that Dispute to each other's contact person. (d) Within 30 days of the date that the written notice of the Dispute is received, the Director, CEO or equivalent of each Party, or their delegates who have appropriate authority to resolve the Dispute, will conduct a meeting by telephone or video conference in an effort to resolve the Dispute. Each Party will bear its own costs of that meeting. (e) If the Dispute is not resolved within 90 days from the date that the written notice of the Dispute is received, then the Dispute must be submitted to mediation in accordance with, and subject to, the Resolution institute Mediation Rules. The mediation must take place in Sydney, Australia and be administered by the Resolution Institute.	A dispute resolution process should be included in the agreement.
What assurances does the supplier provide in relation to the Data?	1. The Receiver acknowledges its use of the Data is at its own risk and accepts that the Supplier, to the fullest extent permitted by law, is not liable (whether in tort (including negligence), contract, statute or otherwise) for any direct or indirect loss or liability arising from Receiver's use of the Data. 2. The Supplier makes no warranty or any representation in respect of the Data, including any warranty or representation that the Data are accurate, is of a certain quality, or is fit for any particular purpose. 3. To the extent permitted by law, the parties exclude all warranties or other terms which otherwise might be implied in this Agreement.	In most cases, it will not be appropriate for the Supplier to provide any warranty as to the quality or fitness for purpose of the Data. The Data are provided 'as is'.

Supplement 3: Personal Identifiable Information in Restricted Access Species Data

(Version 1/12/2022)

The Australian Privacy Principles (APP) usefully define Personal Identifiable Information (PII) as a broad range of information, or an opinion, that could identify an individual, including an individual's name, signature, address, phone number or date of birth. PII may be included in a biodiversity dataset for a variety of reasons – for instance, property information collected for operational purposes, or a record of the individual who sighted or identified an organism.

Use and on-sharing of PII may be restricted by legislation, particularly when collected by a government agency. These restrictions are detailed in the privacy legislation and instruments that apply in different Australian jurisdictions, which are listed in Personal Identifiable Information – legislative and regulatory instruments that affect RASD (current at date of publication). The legality of sharing PII may also vary depending on the date a record was collected; for example, a government agency may only have started to obtain informed consent from individuals to on-share their PII after privacy legislation was introduced in its jurisdiction. Additionally, use and on-sharing of PII in a dataset may be restricted by contractual arrangements – for example, an agreement between the owner of the dataset, and individuals engaged to undertake surveys or identifications on their behalf.

Any organisation that shares data is responsible for understanding and fully complying with relevant legislative or contractual obligations relating to PII, noting that these obligations may vary considerably from case to case (for instance between government and non-government entities, or between organisations that collect data and those that operate data aggregation services).

Where a data custodian cannot be certain that they are legally or contractually permitted to share PII, it is best practice to remove PII fields from a dataset before sharing that dataset with third parties. If the PII provides additional utility (for instance, where there is value in knowing that the individual named in one record is the same individual named in another), this may be preserved by substituting a unique identifier (e.g. 'Contractor001' rather than 'Jane Smith'). Data custodians required to comply with the APPs should note that restrictions may also apply to the sharing of unique identifiers and should consult the APPs to understand these limitations.

Data custodians should include metadata to indicate that PII fields have been removed and/or that an identifier code field has been substituted. A sample metadata statement is available in Supplement 9.

Data custodians should also be aware that Personal Identifiable Information may appear in attribute fields not intended to contain such information, such as comment fields or similar. These fields are often essential to the integrity of the dataset. Recognising that these constitute a small number of records, it is the responsibility of data custodians to remove these references where they are encountered, on a case-by-case basis, rather than withholding these fields altogether.

Personal Identifiable Information – legislative and regulatory instruments that affect RASD

The following list relates to the legislation and instruments within which this framework must operate (as at February 2023):
Data Availability and Transparency Act 2022 (Commonwealth)
Privacy Act 1988 (Commonwealth)
Information Privacy Act 2009 (Qld)
Privacy and Personal Information Protection Act 1998 (NSW)
Victorian Charter of Human Rights and Responsibilities Act 2006 (Vic)
Information Act 2002 (NT)
Information Privacy Act 2014 (ACT)
Personal Information Protection Act 2004 (Tas)
Cabinet Administrative Instruction (Information Privacy Principles Instruction) Reissued 2020 (SA)
Freedom of Information Act 1992 (WA) (includes privacy principles relating to sharing of data with third parties)

Supplement 4: Restricted Access Species Data Metadata Statement Template (Version 1/12/2022)

It is best practice for all data, even withheld data, to have a publicly available metadata statement. This is an important principle of the RASD Framework. Data custodians should share metadata for all datasets including those datasets that cannot be shared (i.e. withheld) with any other user, including other jurisdictions.

The metadata statement fields are shown in the Table below. It is recommended that asterisked fields be treated as mandatory and required for all datasets including withheld datasets.

Table: Metadata Fields for RASD

Metadata Field	Description	Example content
*Title	The name of the dataset	NT Herbaria Flora Specimen Records
Abstract	A brief description of what the data contains including the purpose for which it was collected if relevant	Database of all flora records that are managed by the Department of Natural Resources and the Environment. Records are derived from the NT Vegetation Site Database and HOLTZE (N.T. Herbarium specimen) database and NT Weeds database.
*Key Word (s)	Key words to enhance discoverability and searchability of the dataset	Flora, Fauna, species, sensitive, NSW
*Date Range	Start and end dates of data collection – may be able to be generated from the data	01/12/2010-present 31/07/1788-21/10/1983
*Geographic Extent	Bounding box generated from the data	Bounding Coordinates: • North Bounding Coordinate: -10 • South Bounding Coordinate: -26 • East Bounding Coordinate: 138 • West Bounding Coordinate: 129
*Taxa covered	List of unique species covered by the dataset or if contains all taxa under a higher level classification then the taxonomic group(s) covered (may be able to be generated from the data)	All Flora Mammalia Orchidaceae *Wollemia nobilis*
*Collection methods	Method of data collection (drop down list) – includes unknown /	Incidental observations Systematic survey

Metadata Field	Description	Example content
	other where the method of collection is not known)	
*Data Source	URI / DOI for the dataset	doi:10.4227/05/508637F997933
*Embargoed?	Yes / No field	Yes
Embargo Release Date	If embargoed, the date that the data may be released	30/06/2022
*Custodian	Organisation / body who is the data custodian (may be different from contact organisation)	NT, Department of Environment, Parks and Water Security
*Contact Organisation	Organisation who is the point of contact for the dataset	Department of Agriculture Water and the Environment
*Contact Position	Position for the point of contact for the dataset	Data Manager
*Contact Email	Email address for the point of contact for the dataset	data@awe.gov.au
*Stored Format	Primary format of the dataset	ArcGIS shapefile
Available Formats	Other formats which the data can be made available	csv, Blob
*Access Rights	Any access restrictions based on privacy, security or other policies that apply to how this data are accessed	Controlled vocabulary e.g. 11 - Open 12 - Open with Exemptions 13 - Closed 14 - May be Released Under FOI 15 - Not for Release 16 - May be Published 17 - Limited Release 98 – Unspecified 99 – Not Applicable
*Use Restrictions	Any specific use constraints on the dataset including legal, restricted access species, embargo periods etc in addition to access rights	Internal use only, not to be passed onto third parties without consent
*Security Classification	The security classification applied to the dataset as specified under the Australian Government Protective Security Policy Framework or jurisdictional equivalent	Official, Official: sensitive
*Generalisations †	Any rulesets which have been applied at a dataset level to obfuscate or generalise either species attributes or location	Locations obfuscated to 2 decimal places and attribute Habitat removed from dataset.

* Mandatory Fields

† In addition to the dataset level metadata, row level metadata regarding generalisations applied to individual rows in a dataset is required for transparency and custodians should look at the suggested treatments and metadata in Supplement 5.

Supplement 5: Restricted Access Treatments and Metadata for Restricted Access Species Data
(Version 8/12/2022)

The intent of the Restricted Access Species Data Framework is to encourage the maximum possible sharing of restricted data under the auspices of a negotiated legal agreement. However, in certain circumstances such as before a dataset is made public, or (in some cases) shared with an approved user, the data may need to be transformed to protect or remove RASD. The following explains why an individual record may be restricted, what the treatment method is for that category and what metadata are required.

It is important to note that RASD classifications and treatments may be relevant to only a certain geographic area e.g. a state or territory and, as such, a treatment may only be required to be applied to data within that area or to a dataset e.g. where a third-party Restricted Access Species List (RASL) applies to a particular dataset only. In addition, some records / datasets may require more than one of the treatments outlined in the Table below. This is particularly true in the case of data requiring geographic obfuscation, where generalising the collection date and collector (for example) as well as obfuscating the geo-locality will prevent reverse engineering of the location from associated points. In these cases, the dataset metadata / row-level metadata (attribute metadata) should reflect this.

The reasons outlined below for restricting access to species data are intended to be descriptive rather than exhaustive, and so are not aligned with reasons for listing species as threatened under international conventions nor Commonwealth or State and Territory legislation. Not all restricted access species are listed as threatened nor are all threatened species considered restricted access species.

NB: Currently state and territory sensitive species categories do not align directly with the restricted access reasons in this Table. A national effort is needed to align restricted access terminology and classifications for species records

Text in yellow is intended to be replaced by the data custodian when clause is in use.

Table: Restricted Access Treatments and Metadata

Sensitivity Reason	Treatment	Metadata*	RASD Restriction Type(s)
Species under extreme risk of exploitation / harm	Withhold records of species	*Dataset Metadata:* Records may have been withheld for XX species at XX level (resolution at which the policy is being applied e.g. jurisdiction)	Location data
Dataset contains information about species regarding their habitat or physical attributes which could result in exploitation or increased risk to populations e.g. nesting sites, reproductive status etc.	Remove the dataset attribute containing the sensitive information for this species	*Dataset Metadata:* The attribute XX has been removed from this dataset as it contains potentially sensitive information regarding a species	Attribute data
Species for which the release of precise locations would subject it to a high risk of exploitation and disturbance	Location coordinates for records of this obfuscated to 1 decimal place **AND** Ensure that locality information is removed	*Row-Level Metadata:* coordinates generalised to 1 decimal place and location information removed for sensitivity reasons ***AND (in the case of where a whole dataset contains these species)*** *Dataset Metadata:* All coordinates have been generalised to 1 decimal place and location information removed for sensitivity reasons	Location data
Species for which the release of precise locations could subject it to a moderate risk of exploitation or disturbance	Location coordinates obfuscated to 2 decimal places **AND** Ensure that locality information is removed.	*Row-Level Metadata:* coordinates generalised to 2 decimal places and location information removed for sensitivity reasons ***AND (in the case of where a whole dataset contains these species)*** *Dataset Metadata:*	Location data

Sensitivity Reason	Treatment	Metadata*	RASD Restriction Type(s)
		All coordinates have been generalised to 2 decimal places and location information generalised for sensitivity reasons	
Record contains information regarding the management of species that the land manager considers mildly sensitive e.g. pest control on private property	Location coordinates obfuscated to 2 decimal places **AND** Ensure that locality information is removed **AND** Attribute information removed where it pertains to a sensitive activity	*Row-Level Metadata:* coordinates generalised to 2 decimal places and location information removed for sensitivity reasons ***AND (in the case of where a whole dataset contains these species)*** *Dataset Metadata:* All coordinates have been generalised to 2 decimal places and location information removed for sensitivity reasons	Location data and / or Attribute data
Record contains species information that the landholder or land manager considers sensitive from a privacy perspective	Location coordinates for records of this obfuscated to 1 decimal place **AND** Ensure that locality information is removed	*Row-Level Metadata:* coordinates generalised to 1 decimal place and location information removed for sensitivity reasons ***AND (in the case of where a whole dataset contains these species)*** *Dataset Metadata:* All coordinates have been generalised to 1 decimal place and location information removed for sensitivity reasons	Location data
Dataset contains personal identifiable information e.g. names of individuals or any personal identifiers	Remove the dataset attribute containing the personal identifiable information	*Dataset Metadata:* The attribute XX has been removed from this dataset as it contains personal identifiable information	Personal identifiable information
Species at a high to moderate risk of exploitation that have been recorded as	*For datasets which contain multiple surveys:* remove survey identifier from the record	*Row-Level Metadata:* coordinates generalised to 1 (high risk) or 2 (moderate risk) decimal places and location	Location data / attribute data

Sensitivity Reason	Treatment	Metadata*	RASD Restriction Type(s)
part of a survey where a survey identifier can be resolved to a precise locality	**AND** Obfuscate location coordinates to 1 (high risk) – 2 (moderate risk) decimal places dependent if moderate or high risk **AND** Ensure that locality information is removed *For datasets which include only an individual survey:* Withhold records for the species of high to moderate risk	information removed for sensitivity reasons ***AND (in the case of where a whole dataset is for an individual survey)*** *Dataset Metadata:* Some records have been withheld from this dataset for sensitivity reasons	
Data contains records of declared biosecurity species outside of exclusion zones e.g. fruit fly location recorded inaccurately	Withhold records of species	*Dataset Metadata:* Records withheld for biosecurity reasons	Identification data
Data contains declared biosecurity species which may be of concern to biosecurity or trade e.g. a species of concern recorded as a quarantine intercept	Withhold records of species	*Dataset Metadata:* Records withheld for biosecurity reasons	Identification data
Data contains potentially culturally sensitive information relating to species	Pending national consultation with first nations people	Pending national consultation with first nations people	Indigenous data

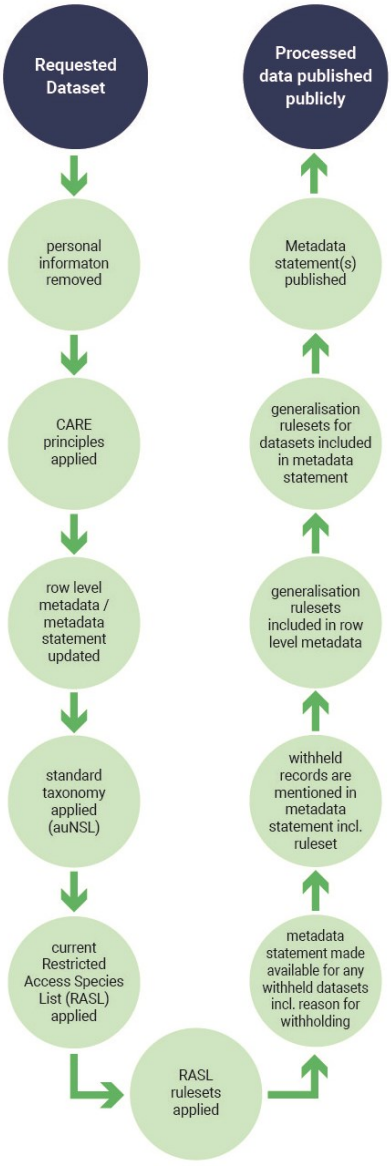
* Where possible, it is desirable to represent obfuscated records as polygons rather than points e.g. a 0.01 x 0.01 degree square or a 0.1 x 0.1 degree square with the centroid defined

Supplement 6: Process for Release of Restricted Access Species Data Publicly (Version 1/12/2022)

This document provides guidance to data custodians on the best practice process flow for making Restricted Access Species Data suitable for public release. It outlines the data transformations that need to be performed on datasets before making them public.

Data custodians should generally retain responsibility for applying the following process, except where that has been delegated under a negotiated legal agreement. The process in the Figure below is recommended.

Figure: Public Data Release Process



Data Removal

Data custodians are responsible for applying rulesets removing data prior to sharing data. If data are removed, it is important that data custodians provide a metadata statement covering the ruleset used to remove data and the fields that have been removed.

Standard Taxonomy

The Australian National Species List, maintained by the Australian Biological Resources Study, provides agreed national accepted taxonomic concepts for all key organism groups.

Data custodians should consider using this taxonomic framework when sharing their data to standardise taxonomy.

Restricted Access Species (Currently Known as Sensitive Species) Lists

Defines Restricted Access Species and provides best practice advise on sharing restricted access species lists and data transformations to be applied to them.

a) Newly Described Species

These are species which have been recently described, have very high levels of uncertainty about their vulnerability to disturbance, and have not yet been assessed for inclusion on a RASL. In the rare event where a species meets these criteria, they may be dealt with consistently with the processes outlined for RASL species below.

b) Definition of Restricted Access Species (Currently Known as Sensitive Species)

Restricted Access Species are species identified by a jurisdiction or a third-party data custodian in Australia as requiring restricted access to geolocational or identity information on the species. These are described in the Principle Restricted Access Species Data (RASD) Should be Consistently Classified and are generally species that have undergone expert assessment according to a defined process and are held as lists (RASLs). RASLs may relate to either conservation-related species or biosecurity-risk related species. This framework currently only covers conservation-related RASLs.

c) Management of Jurisdictional Restricted Access Species Lists (RASLs)

RASLs are maintained by each jurisdiction in Australia. A RASL is a publicly available list that delineates which species should have their geographic locations obfuscated or withheld to prevent disturbance of the species or for management reasons. Users and data custodians should always check with the originating data custodian that the list is up to date.

RASLs serve an essential purpose in ensuring that data from all sources relating to a restricted access species are treated in a similar fashion in that jurisdiction. This is essential, firstly to ensure the management intent behind the RASL is achieved and secondly to ensure that a particular observation that may enter an aggregated data source from several sources, when obfuscated, appears as the same point.

d) Management of Third-party RASLs

In some instances, non-government data custodians may maintain a RASL with similar intent to jurisdictional RASLs but at variance to jurisdictional lists. An example might be a bird dataset where nest tree locations must be withheld, location obfuscated, or the species identity obfuscated. An alternative might be the records of a local orchid society, where obfuscation of all records is a prerequisite of access.

Third-party RASLs are the prerogative of an organisation, however, organisations intending to create or maintain these lists should be cognisant of the risks:

a) The effectiveness of independent third-party RASLs may be compromised where such lists are at odds with jurisdictional lists. A mixture of obfuscated and un-obfuscated records may allow a user to identify a locality by triangulation or data linkage.
b) Because third-party datasets are frequently sought by many aggregators or projects, there is an implicitly higher risk of third-party RASLs resulting in a particular observation entering an aggregated data source from several sources, and when obfuscated, appearing at a different point, confusing analysis.

Nevertheless, third-party RASLs serve an important purpose in giving organisations sufficient reassurance to share data. The importance of third-party RASLs are recognised as an important mechanism for ensuring that the maximum amount of data are included in management and research. Third-party data custodians who work consistently with the principles in this framework are strongly encouraged to seek amendments to jurisdictional RASLs rather than maintaining separate RASLs.

Where third-party RASLs are inevitable, either

a) these lists are provided to other data custodians so that the best-practice rules identified under this framework are applied consistently but are flagged so that users can discern that obfuscation of this data may divert from other data
b) the third-party data custodian applies the best-practice rules identified under this framework but are similarly flagged so that users can discern that obfuscation of this data may divert from other data

e) Location generalisation (Obfuscation) on Jurisdictional RASLs

All states and territories in Australia manage data on species in RASLs by obfuscation of locality information or by preventing queries on data below a minimum radius of 1km.

Where RASD needs to be transformed spatially as per jurisdictional or third-party RASLs it should be transformed via obfuscation. Obfuscation ideally needs to be conducted so that the same observation point, regardless of source, is spatially moved to a consistent spot to avoid confusion.

Best practice for obfuscation, therefore, needs to:

a) be deterministic and repeatable so that a transformed point, regardless of source, will end up at the same point
b) minimise the flow-on risk of double obfuscation
c) allow modellers to use the obfuscated data with confidence, provided that their grid cell is larger than the obfuscation algorithm

There are two levels.

Level 1 – round latitude and longitude to nearest 1 decimal place
Level 2 – round latitude and longitude to nearest 2 decimal places

Where possible, it is desirable to represent obfuscated records as polygons rather than points i.e.. a 0.01 x 0.01 degree square or a 0.1 x 0.1 degree square with the centroid defined by the obfuscation treatment above. For example: where the RASL dictates that location coordinates should be obfuscated to 1 decimal place, and returns a value of latitude -30.5 longitude 148.7 this would be represented by a 0.1 x 0.1 degree square with centroid at -30.5, 148.7. Accordingly, the polygon would have minimum Latitude -30.55, maximum Latitude -30.45, minimum Longitude 148.65, maximum Longitude 148.75.

Jurisdictions should advise whether Level 1 or Level 2 is required for each species via RASLs.

The instances where obfuscations are applied are outlined in Supplement 5.

f) Attribute Generalisation on Jurisdictional RASLs

The rulesets which apply to RASLs may include generalisation / withholding of some attributes for particular species as outlined under Species-related categories in Principle Restricted Access Species Data (RASD) Should be Consistently Classified. Where information for a particular attribute has been removed, row level metadata should reflect that this has occurred and the ruleset for this change. Reasons should be standardised and align with the species-related categories in Principle Restricted Access Species Data (RASD) Should be Consistently Classified.

The instances where generalisations are applied are outlined in Supplement 5.

g) Generalisation on Third-Party RASLs

Generalisation rulesets which apply to Third-Party RASLs should follow the same methodologies as outlined in e) and f) for jurisdictional RASLs above.

Data custodians are expected to apply a third-party RASL ruleset to that third-party's dataset.

Supplement 7: Withheld Data and Restricted Access Species Data
(Version 1/12/2022)

The intent of this framework is specifically to encourage data custodians to maximise the sharing of higher risk data within a trusted environment by constraining access to that environment and ensuring that public access to all higher risk data are either obfuscated or withheld.

There is a category of data – or datasets – that carries an implicitly higher risk of consequences when shared due to management, access concerns, policy, legal or financial requirements. This excludes Personal Identifiable Information, and information is available about PII in Supplement 3.

This type of data are best handled by curtailing access and is therefore called withheld data.

The Australian Government Best Practice Guide to Applying Data Sharing Principles notes that it is best practice for data custodians to explore how they can share data legally rather than simply dismissing a request to access data due to perceived legislative restrictions.

There are four levels of withheld data recognised by this framework and Table 1 (referred to below) is available in the Principle Restricted Access Species Data (RASD) Should be Consistently Classified.

a) cannot be shared at all with third parties, even other jurisdictions – typically this should be data relating to a small subset of species under extreme risk of exploitation or harm whose distribution is so restricted or sensitive that knowledge shared unnecessarily places the species at risk (Category 3.1 (b) in Table 1)
b) can be shared with jurisdictions and approved data requestors under negotiated Legal Agreement but not with other users. Typically this should be Legal Contact or Financial issue datasets (Category 2.1 and 2.2 in Table 1), non-Legal Third-party issue datasets (Category 2.3 in Table 1), data relating to "sensitive" species (Category 3.1 (a) in Table 1), where the implications of sharing species data has ramifications at a national scale such as an incursion of an extreme biosecurity risk species (Category 3.2 in Table 1), or where the attribute of a species record makes a record sensitive (Category 3.3 in Table 1)
c) can be shared with approved data requestors under negotiated Legal Agreement provided the source data are not duplicated and not accessible by general users. This is the same as (b) but access to the data are constrained to allow analysis but not data downloading or transfer
d) can be viewed by the public. This data has had all possible data transformations applied including systematically obscured geolocations

There is a 5th category of withheld data, Embargoed Data, being data that has a timestamp preventing its release before a certain date. As the intention is to release this data in due course, it is dealt with by providing metadata reflecting this (see Supplement 8). Levels a), b) and embargoed data fall within the definition of usage-restricted data.

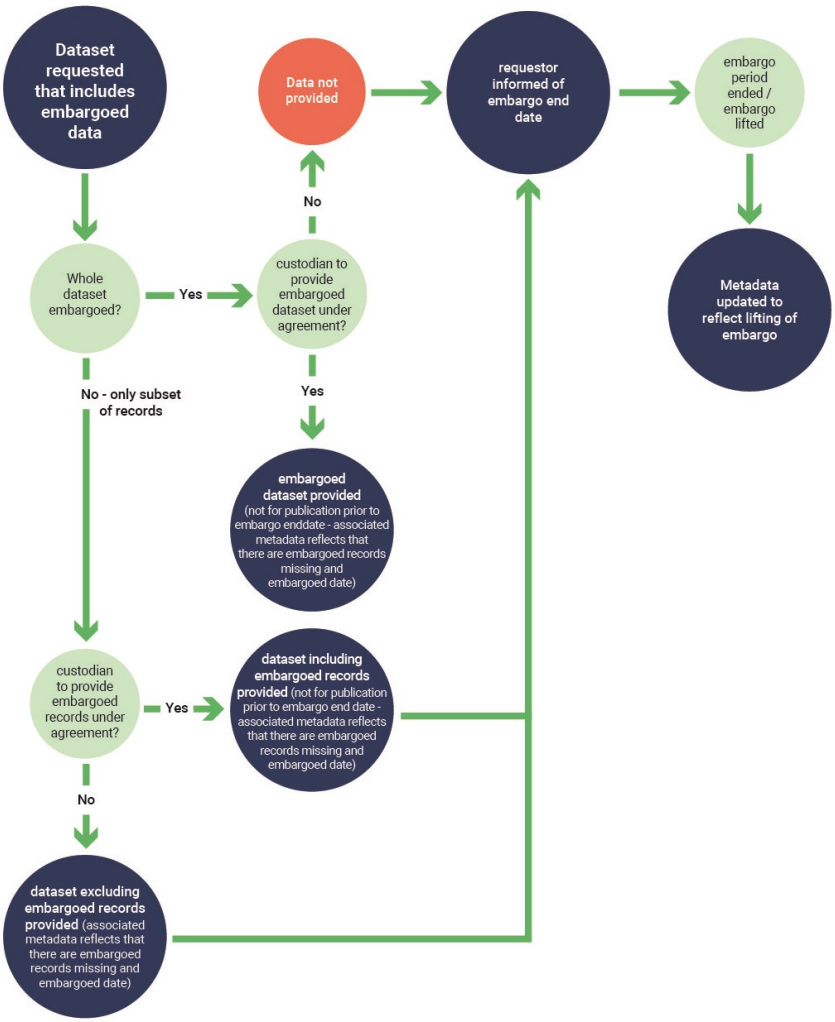
Data custodians are encouraged to work towards the provision of full resolution data (comprising all data except (a) above) when providing data under negotiated legal agreements and the application of data transformations consistent with this framework when providing public data.

Supplement 8: Process for Handling Embargoed Restricted Access Species Data
(Version 1/12/2022)

This supplement outlines the best practice process for handling requests for embargoed Restricted Access Species Data and provides some common scenarios for illustration purposes.

The figure below provides a high-level process for handing embargoed data.

Figure: Process for Handling Requests for Embargoed RASD



Example scenarios

A data custodian may wish to embargo either an entire dataset or a subset of records within a dataset for a variety of reasons. These may include, but are not limited to:

- Embargoing data for a specified period which identify a previously unknown population of a species or a previously undescribed species prior to the data custodian publishing a paper documenting these new findings
- Embargoing data that are commercial-in-confidence prior to the end of legal contract to develop an area of land
- Embargoing data which was part of a research project prior to researchers exclusively being able to analyse and publish their results/ findings
- Embargoing attributes in a dataset that may contain information on the dates and locations of questionably obtained specimens for collections e.g. items acquired from a deceased estate
- Embargoing data on species which may impact on trade e.g. for a period of 10 years
- Embargoing data on species which are undergoing assessment for threatened species listing

Supplement 9: Example Restricted Access Species Data Access Request Form Template
(Version 1/12/2022)

Organisation / Research Institution Name

Click or tap here to enter text.

Address Click or tap here to enter text.

Indigenous Body / Organisation? ☐

Requester (must be an accountable representative of the organisation)

First Name: Click or tap here to enter text.
Last Name: Click or tap here to enter text.

Email: Click or tap here to enter text.
OrcID: Click or tap here to enter text.

Project

Project Title: Click or tap here to enter text.

What is the purpose of the data request? Choose an item.

Topic of Research (if applicable): Choose an item. **OR**

Industry Type: Choose an item.

Is the data to be used for commercial purposes? Yes: ☐ **No:** ☐

Public Benefit Statement (is there a public interest in the outputs and outcomes of this project?):

Data

Data Requested (provide exact name of dataset if known, otherwise describe the data you need):

Relevance of data to project:

How often do you require this data?

Single once off? ☐ **Defined period?** ☐ **Ongoing?** ☐

Date required from: Click or tap to enter a date.
Date required to: Click or tap to enter a date.

What frequency and time period do you need the data for (if ongoing provide reasons to support this)?

Area:
Whole Dataset: ☐ **Specific Area*:** ☐
* (provide bounding box – i.e. max and minimum latitudes and longitudes in decimal degrees using datum GDA94)

What physical and IT controls will you use to ensure that the data are safe and secure during storage and use?

Who will be able to access the data?
Choose an item.

Name the group of people accessing the data (may be a business / organisational group / project team etc):

Does your proposed use of the data involve distribution of products (including publications) created from the data outside your organisation – describe how and to whom these products will be distributed and / or presented including any data transformations:

If your access to full resolution data is not approved, would you like to receive a transformed dataset?
Yes: ☐ No: ☐

The development of this Framework was supported by the Australian Research Data Commons (ARDC) and the Atlas of Living Australia (ALA). The ARDC and ALA are funded by the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS)