Atlas of Living Australia
Our secrets are not your secrets
Sensitive data report

John Tann
Paul Flemons

October 2009

ATLAS OF LIVING AUSTRALIA
sharing biodiversity knowledge

Australian museum
nature culture discover

Atlas of Living Australia

Our secrets are not your secrets

Proposed national policy and sensitive data report

John Tann
Paul Flemons

Version 1.1 for comment
October 2009

Atlas of Living Australia sensitive data report

Executive Summary

This report and proposed national policy addresses a barrier to sharing biodiversity information – sensitive data, data that might cause harm if made public.

- Managers of biodiversity data face a significant difficulty when preparing their datasets for shared access. They are concerned that some of their data may be considered sensitive; however, currently there is no straightforward way to find out.
    - Data may be sensitive for different reasons: conservation status, quarantine, biosecurity, trade, taxonomy, privacy
- Each State and Territory conservation agency treats the sensitivity of information about threatened species in their jurisdiction differently
    - Each agency uses a different method to determine which species are sensitive
    - Each agency restricts sensitive information in different ways
    - No State or Territory agency promotes its sensitive criteria in a form that can be readily discovered or used by others
- Flora and fauna specialists are aware of sensitivities associated with the organisms with which they work. Their fields of expertise extend beyond the regulations of conservation agencies, and they can recognise species distributions not constrained by political boundaries. They can also respond rapidly to new information.
- Information about specific plant pests and diseases of quarantine concern is highly sensitive. Australian trade depends on the belief that certain pests are not present in the country. The perception of a particular pest being present could trigger commercially difficult, international trade problems.
- Information about some microorganisms can be sensitive for security or trade reasons. Legal obligations apply to organisms that carry a high risk to animals and humans. As with plant pests, the perception that exotic animal pathogens or diseases are present in Australia may cause considerable trade problems.
- Sensitivities to data can be locally determined or confidential, and include such information as personal identities, unpublished work, uncorrected data, and data owned by others. Identifying and separating this information from other data before it is released publicly can be difficult.
- Currently no register is available of sensitive taxa that is applicable at a national level. Restricting public access to specific information only makes sense if those restrictions are also followed by others with similar information.

Key recommendations

The Atlas of Living Australia should develop and implement a suite of appropriate tools and services so that data managers could apply those tools to their institutional data to filter and flag sensitive records. The ALA should:

1. Develop and implement a rules-based National Register of Sensitive Species, which will include sensitive threatened species, quarantine-sensitive species and sensitive microorganisms. The rules for sensitivity would be determined by interested agencies and specialists concerned with the consequences of releasing sensitive data. The rules will vary depending on geographic area, time, species, authority, and so on.
2. Develop and implement a Sensitive Data Service to enable data managers and custodians to determine which of their records refer to sensitive species. This service will use as its reference standard, the National Register of Sensitive Species.

3. Develop and implement a support service for data managers. Include standards, schemas, tools for managing data sharing, and checklists to help detect locally sensitive information.

Proposed national policy

A national policy is proposed to create a framework for incorporating these recommendations. The policy will apply to data managers and custodians as well as those agencies and specialists who can determine sensitivity rules.

Atlas of Living Australia

Our secrets are not your secrets

Part 1. Proposed national policy
Assessment of sensitive biodiversity data

Proposal

Your comments are invited on this proposal.
Please respond to:

John Tann
Australian Museum
6 College St
Sydney NSW 2010

john.tann@austmus.gov.au
(02) 9320 6449
0448 000 909

Document Control

Author	John Tann and Paul Flemons Australian Museum
Date of original endorsement	<date>
Date last modified	<date>
Date for review	<date>
Publisher	Atlas of Living Australia http://www.ala.org.au/

Proposed national policy
Assessment of sensitive biodiversity data

Contents

P1 Overview

This policy establishes a method for assessing the sensitivity of data relating to plants, animals and microorganisms in Australia for public release.

The Atlas of Living Australia will aggregate information from a wide variety of sources including museums, herbaria, and other natural history collections as well as ecological, observational and molecular data. Institutions, researchers and others holding datasets will be able to share their biodiversity data through the ALA where it will be made available world-wide – freely, and openly.

Without appropriate controls in place, the large amount of data that the ALA will release will include some records which, if they were made public, may cause harm. For example, records might contain location information of collectible and commercially attractive rare or threatened species; or personal details; or incorrect or misleading data.

A major barrier faced by custodians wishing to share their data is the lack of a straightforward way of confidently discovering within their own datasets, information that could be considered sensitive.

Being able to flag and potentially restrict data that may be generally agreed to be sensitive, removes an obstacle to release of data and gives confidence to both the provider and the user about data they share through the ALA.

P2 Scope

This policy describes a framework that allows assessment of biodiversity data prior to public release. This framework for identifying sensitive biodiversity data:

- discriminates between public and restricted data based on rules about the data itself; rules that are extensible, and can be readily modified in a timely manner
- is not dependent on the source of the data
- applies exclusively to data about organisms found in Australia, its States and Territories including marine areas and Australian Antarctic Territory
- does not apply to data on organisms in other countries, or places outside Australia; this policy however, could be adopted or adapted for use in other jurisdictions
- can be used by institutions outside Australia that may hold sensitive data about organisms found in Australia
- can incorporate some international conventions, such as those of IUCN[1] and WTO[2]

This policy applies to managers and custodians of datasets wishing to share their data relating to biodiversity in Australia. Their datasets may hold records of a sensitive nature, such that releasing those records to the public may cause harm. This policy is especially applicable to datasets that may contain records from more than one State or Territory, include a number of species, or have been recorded over a broad time range.

This policy is designed to be used as data is prepared for sharing, before it leaves an organisation. The policy applies to people within an organisation who need to make decisions about data restriction.

[1] IUCN – International Union for Conservation of Nature
[2] WTO – World Trade Organization

This policy also applies to agencies and specialists who are aware of the impact of the release of sensitive information relating to plants, animals and microorganisms, and the harm it could cause. These include:

- Agencies responsible for conservation of rare and threatened species
- Agencies responsible for plant pests and biosecurity
- Agencies responsible for microorganisms, animal health and biosecurity
- Wildlife specialists working with rare and threatened species
- Curators and taxonomic specialists working with undescribed organisms
- Other specialists who appreciate the potential harmful impact of releasing information about specific organisms

This policy acts to guide the Atlas of Living Australia in managing sensitive data belonging to, and used by, other parties. It is designed to be used when creating procedures and services for data providers, and when preparing data for public access.

Although this policy applies specifically to data sharing through the Atlas of Living Australia, it can also apply to those who would like to share their data either independently, or elsewhere through a third party aggregator.

P3 Policy Principles

This policy is based on the principles of

- making scientific data readily accessible
- minimising harm by explicitly restricting access to selected sensitive information
- assisting State and federal authorities with their obligations under Freedom of Information
- assisting State and federal authorities with Australia's international trade obligations
- assisting with the sharing of data held in trust by a custodian, where there is an agreement or expectation that this data will not be misused
- respect for the differences of approach to sensitive data in jurisdictions across Australia
- respect for privacy and restrictions to personal information

P4 Authorities

This policy draws on the expertise and authority of State, Territory and Commonwealth conservation agencies, departments of primary industries, plant and animal health and biosecurity, museums, herbaria, CSIRO, and other organisations. This policy also draws on the expertise of specialists from across Australia, who have an understanding of the sensitivity of specific biodiversity information.

P5 References

Guide to Best Practices for Generalising Sensitive Species Occurrence Data, Arthur Chapman and Oliver Grafton, 2008, published by GBIF.
http://circa.gbif.net/Public/irc/gbif/pr/library?l=/webfiles/bpsensitivedata/_EN_1.0_&a=d

Our secrets are not your secrets, Atlas of Living Australia sensitive data report, John Tann and Paul Flemons, 2009.

P6 Framework for assessment of sensitive biodiversity data

In order for data custodians to be able to assess their institutional data to filter and flag sensitive records, the Atlas of Living Australia aims to create a suite of tools and services that includes three linked entities:

1. **National Register of Sensitive Species** – an authoritative and up-to-date reference of sensitivity criteria for biodiversity data
2. **Sensitive Data Service** – an accessible service where datasets can be assessed for sensitivity against the National Register of Sensitive Species
3. **Sensitive Data Toolbox** – standards, schemas and checklists to help assessment of datasets subject to local in-house sensitivities prior to preparing those datasets for shared access

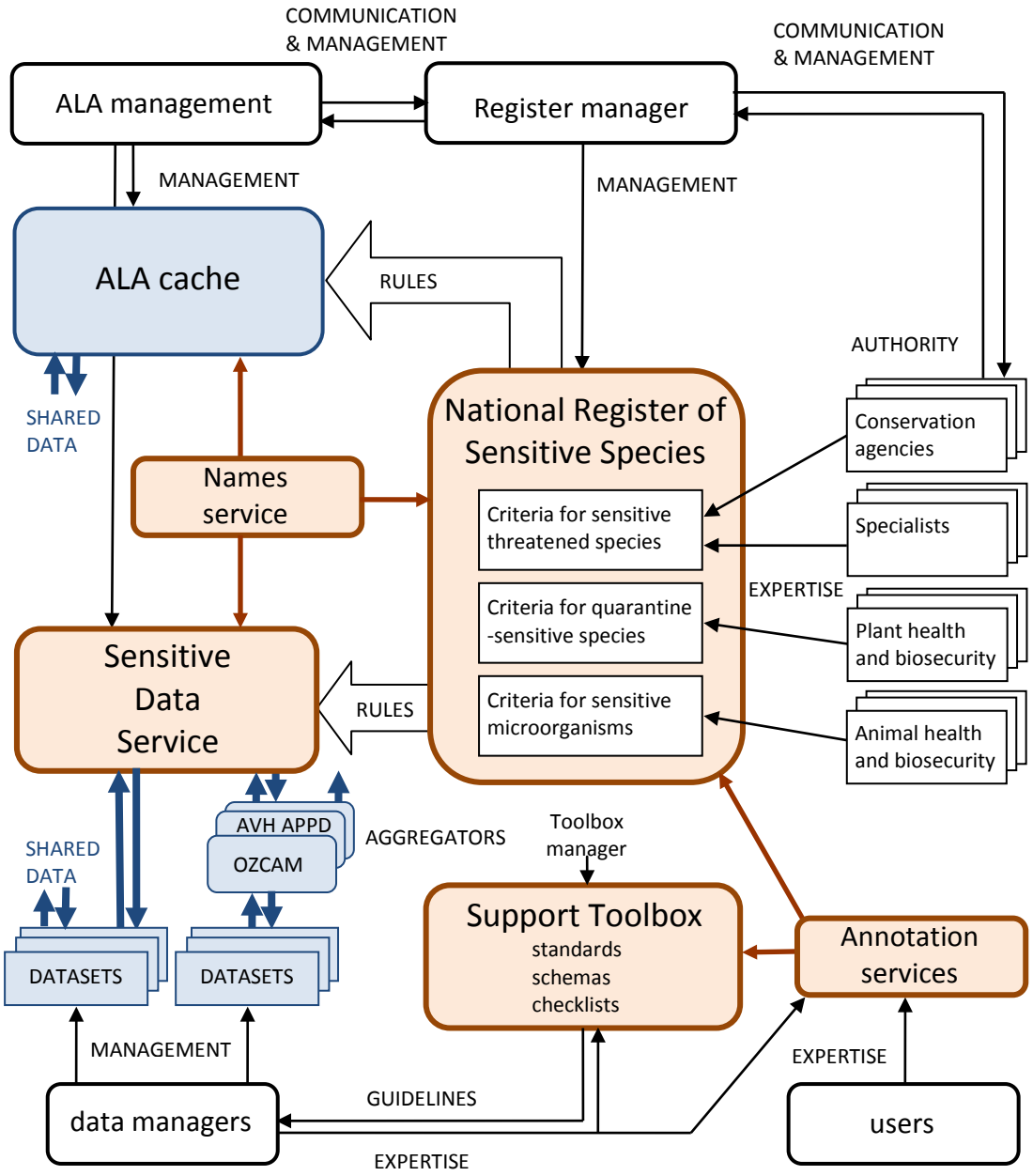

Figure 1. Diagram illustrating a suggested operating framework for assessment of sensitive biodiversity data. The National Register of Sensitive Species, the Sensitive Data Service and the Sensitive Data Toolbox (shown here as the Support Toolbox), are the three linked entities proposed for this model.

P6.1 National Register of Sensitive Species

There is no single agency in Australia that either defines, or takes responsibility for, what is, or what is not, sensitive biodiversity data. Desirably, a National Register of Sensitive Species would be established with the cooperation and support of agencies and specialists who understand the impact of inappropriate release of sensitive information.

This National Register of Sensitive Species should be comprised of three separate registers:

1. **National register of sensitive threatened species** – a set of rules that define sensitivity according to criteria of State and Territory conservation agencies, plus criteria recommended by specialists and experts
2. **Register of quarantine-sensitive species** – a set of rules that define sensitivity according to criteria of those agencies responsible for plant pests and biosecurity
3. **Register of sensitive microorganisms** – a set of rules that define sensitivity according to criteria of those agencies responsible for animal health and biosecurity

Contents of the National Register of Sensitive Species

The National Register of Sensitive Species is designed to include, for example: the recommended precision to be given for the location of collectible and commercially attractive threatened species; species names of trade sensitive pests and invasives; and rules about restrictions for the locations of institutions holding pathogens of biosecurity interest. The register may hold references to graded sensitivity.

The register is designed to be expandable, able to include other conditions as required.

The reach of the above criteria could also be constrained. For example, criteria might be limited to a restricted geographic area (eg the Wimmera), or cover a restricted time (eg nest sites observed in the last 10 years) or to data from a particular collector/collection (eg to questionably obtained specimens), or to a particular dataset (eg licence restrictions).

Documenting criteria

Each rule in the National Register would need appropriate documentation that adequately explains the reason for the rule. The explanations could then be attached to any data that is restricted as a consequence of the rule.

Register currency

One of the important reasons for compilation of a National Register is that the perceived sensitivity of species is continually changing. Contributors to the National Register would be expected to update their criteria as, and when, they change.

Users of the National Register would know that they are using the most recent sensitivity criteria. This same set of rules would be applied to data sourced through the ALA itself, and potentially other data aggregators, at the time of public data release.

National Register creation process

Defining the sensitivity rules for the National Register is expected to be collaborative and iterative, with some natural lines of demarcation. The current Australian federated legislative system allows for delineation along State and Territory borders. Quarantine operates at national, State and regional borders. Specialists are aware of sensitive species, distributions of which are not constrained by political boundaries, or are not acknowledged by other agencies. Curators may make suggestions about personal or intellectual sensitivities not considered by other authorities.

National Register of Sensitive Species management
The ALA Management Committee will nominate an agency or institute to manage the National Register of Sensitive Species. The ALA will support register management with appropriate tools and standards for managing the National Register and for creating, submitting and maintaining criteria for assessment of sensitivity. The register management appointment will be reviewed every 2 years.

The role of the register manager will be to invite and liaise with contributors to the registers, accept and vet contributions, compile the sensitivity rules list and make it accessible to the Sensitive Data Service, apply changes in a timely manner, and regularly review the currency of the criteria. The register manager would report regularly to the Atlas of Living Australia management.

Contributions
Relevant agencies, organisations and specialists will be invited to contribute to each of the separate registers that comprise the National Register of Sensitive Species. Contributions may also be nominated directly to the register manager.

Agencies and specialists will be selected to contribute based on those organisms and areas that fall within their jurisdictions or expertise.

Conflicts
The National Register of Sensitive Species would work most effectively in a spirit of collaboration and cooperation. If they occur, conflicts and disputes about the sensitivity rules for the National Register would be decided by a tribunal or Expert Advisory Committee appointed by an appropriate and respected body, at the request of either the register manager or ALA management.

P6.2 Sensitive Data Service
The Sensitive Data Service will use the National Register of Sensitive Species rules to filter all data for public access. The service will be open and accessible to those wishing to share their biodiversity data.

Conceptual function of Sensitive Data Service
Data managers and custodians will be able to send their datasets over the internet to the Sensitive Data Service, where each record will be assessed for sensitivity. Using the National Register of Sensitive Species as a reference standard, the Sensitive Data Service would compare each data record against the rules of the National Register, and mark and grade appropriately those records that are sensitive. The Sensitive Data Service would then return to the data manager, the marked and graded sensitive records in a suitable form that will allow review. In order to address concerns about data security this would be a stateless system, and neither the data nor the results would be stored.

By using and accepting the results of the Sensitive Data Service, a data manager or custodian would know that their dataset complies with the same rules for sharing sensitive data as applied by the ALA, other data aggregators, and other data managers and custodians when sharing data. By being given the opportunity to review their sensitive records, data managers would also be in a good position to note discrepancies and provide feedback to the Sensitive Data Service.

Data managers would, of course, have the option of withholding or further restricting their data if they choose.

The Sensitive Data Service and the National Register of Sensitive Species would be openly accessible to other data aggregators such as OZCAM, AVH, APPD, etc. These aggregators would be able to implement the same rules automatically against public views of all records, whether or not their data providers have already tested their data using the Sensitive Data Service.

Documenting changes
Any data that is altered, generalised or removed due to its sensitivity will need to be documented accordingly. Restricted text fields would need to be replaced with appropriate explanations; comments would need to be given for data that is denatured.

The following minimum information will accompany all data filtered due to its sensitivity:

- Reason for sensitivity
- Comments
- Review date
- Precision for denatured locations

This information will be included as comments in accompanying metadata. Ideally this would be included as record-level metadata.[3]

Sensitive Data Service Management
The ALA will manage and maintain the Sensitive Data Service. The ALA will work with and help data managers and custodians with their task of sharing data, while caring for their sensitive data.

P6.3 Sensitive Data Toolbox

The datasets of many institutions may include a broad array of types of data that may be confidential or locally determined as sensitive. Examples are working names, third party datasets, the identity of specific people, data in preparation, and so on. The National Register of Sensitive Species is expected to be invaluable for defining those criteria that are determined by others; however, it would be limited in its ability to help with in-house peculiarities.

In order to assist data managers with datasets that may hold confidential or locally determined sensitive information, the ALA should set up and support a Sensitive Data Toolbox that would provide resources and tools such as standards, schemas, and checklists.

The checklists may be compiled in a wiki-type environment, where data managers can help each other with the intricacies of preparing data for sharing, while paying attention to local and in-house sensitivities.

The ALA would create additional tools where required, to help automate this process. These tools may extend, for example, to a locally implemented version of the Sensitive Data Service with an option for data providers to add rules of their own.

Sensitive Data Toolbox Management
Components of the Sensitive Data Toolbox may need to be moderated. The ALA Management Committee will nominate an agency or specialist group to manage the Sensitive Data Toolbox, which would act in the interests of data managers. The ALA will support Toolbox management with appropriate tools for managing the Toolbox and moderating discussions and checklists.

[3] However neither Darwin Core nor ABCD standards have appropriate fields for this type of metadata.

P7 Responsibilities

ALA Management Committee	○ Responsible for nominating National Register of Sensitive Species management ○ Responsible for developing and managing a Sensitive Data Service ○ Responsible for ensuring data shared through the ALA complies with National Register of Sensitive Species criteria ○ Responsible for developing appropriate tools and services to assist the operation of the National Register of Sensitive Species ○ Responsible for developing appropriate tools and services to assist the operation of the Sensitive Data Toolbox ○ Responsible for coordinating the review of this policy
National Register of Sensitive Species manager	○ Responsible for managing the National Register of Sensitive Species ○ Responsible for inviting and liaising with contributors ○ Reports to ALA management
Sensitive Data Toolbox manager	○ Responsible for managing the Sensitive Data Toolbox ○ Responsible for appointing moderators for some components of the Sensitive Data Toolbox ○ Reports to ALA management
Contributors	○ Responsible for suggesting and preparing sensitivity criteria for their jurisdiction or expertise ○ Responsible for maintaining the currency of their sensitivity criteria ○ Communicate with register manager

P8 Review

This policy and its practices will be reviewed every 5 years; the review to be coordinated by the ALA Management Committee.

Atlas of Living Australia

Our secrets are not your secrets

Part 2. Sensitive data report

Foreword

Over the past 15 years, international initiatives such as the Global Biodiversity Information Facility and numerous national networks have applied themselves to a herculean goal – that of making the data in the world's biodiversity collections readily and universally available. Progress towards this goal has been impressive, as evidenced by the millions of records now accessible by means of specialised, web-based, search tools. However, significant challenges remain.

Technical challenges abound and still hinder some would-be data providers and users. The vicissitudes of the web and local systems, the idiosyncrasies of some biodiversity databases and limited internet access, are just some of the persistent frustrations.

The scope of data available continues to disappoint many users. In the past 15 years, huge numbers of records have been digitised. However, the volume of data still to be entered into local systems is immense and there is no matching enthusiasm among funding bodies to support the effort required. As a scientific community, we are well on the way to digitising all available names and integrating the systems that will form a comprehensive, nomenclatural resource – but that work also is still incomplete.

The old taxonomic impediments remain – taxonomists are in short supply and many are needed to identify specimens, update old determinations and deal with complex, systematic problems. Some biological groups remain as taxonomically intractable as ever.

Finally, many data custodians remain reluctant or unable to share records in the 'global commons'. To label these institutions or individuals as modern-day, digital curmudgeons, is to overlook some of the most stimulating issues confronting biodiversity scientists today. The present report on 'sensitive' data addresses these issues.

I join my colleagues in welcoming the following examination of what constitutes 'sensitivity' and how to deal with it. I urge those responsible for data sets, whether the sets include endangered species, trade-critical pests or data of commercial or social importance, to consider the report's analysis and recommendations. Is the analysis consistent with your own assessments and are the recommendations helpful? Do the recommendations seem reasonable and practicable to those with unwelcome gatekeeper roles? I commend this report to readers. The ideas in the following pages are critical to achieving the goal described in the first paragraph of this foreword.

IAN NAUMANN
Chair, Australian Plant Pest Database Steering Committee
Canberra Australia
September 2009

Contents

1 Background

Australia is a signatory to the OECD declaration on access to research data from public funding. This declaration recognises the importance, value and benefits of open access to, and unrestricted use of, scientific data.[1] The Atlas of Living Australia has been established by the Australian Government National Collaborative Research Infrastructure Strategy and will develop a biodiversity management system.

The Atlas of Living Australia will bring together and organise information from a wide variety of sources including museums, herbaria and other significant collections in Australia, as well as ecological, observational and molecular data.

Data shared through the ALA will be accessible to anyone in the world – openly, publicly and freely. However, amongst the large volume of data will be some records of a sensitive nature which, if made public, could cause harm.

In 2008, the ALA conducted a user needs analysis which found that sensitive data is an area of significance for users.[2] Data managers and custodians responsible for the release of data, when preparing their datasets for sharing, have difficulty determining which data is sensitive and how to deal with it.

Specialists and agencies in particular fields are commonly aware of sensitivities relevant to the information and organisms with which they work. For example, quarantine services are acutely aware of sensitivities relating to information on invasive and pest species. However, there is currently no straightforward way that these sensitivities can be either identified or honoured by others.

Determinations of sensitivity are nationally inconsistent. For example, some conservation agencies suggest that the locations of all species with rare or threatened conservation status should be significantly restricted; other agencies have followed a formal process to determine and list those species which are highly sensitive to threats such as disturbance and collection.

Figure 1-1 The critically endangered Spotted Handfish, *Brachionichthys hirsutus*, lives in the Derwent River in Tasmania. It has an extremely restricted distribution due in part to its unusual life cycle. Data on locations of handfish in Tasmania are considered to be sensitive. Photo Phil Malin.[3]

[1] *OECD Declaration on Access to Research Data from Public Funding*, January 2004
[2] *Atlas of Living Australia user needs analysis*, John Tann, Lynda Kelly, Paul Flemons, 2008
[3] See http://australianmuseum.net.au/Spotted-Handfish-Brachionichthys-hirsutus-Lacepede-1804/

Restrictions flowing from sensitivity determinations are also nationally inconsistent. For example, some conservation agencies have implemented categories of sensitivity for occurrence data; some agencies hide the location information; whereas others may not release the record at all.

These complications act as a barrier to releasing of data; and over-restriction of access is a barrier to effective data discovery and use.

The ALA is interested in reducing barriers to sharing of biodiversity data. It aims to assist data managers and custodians by creating ways and means to resolve many of their concerns about sensitive data. In order to help it determine the best way to assist, the ALA commissioned and provided resources for the Australian Museum to undertake this study.

2 Who is this report for?

This report is for the Atlas of Living Australia. It will help the ALA to determine how it can assist those who wish to share their data.

This report accompanies a proposed national policy for assessment of sensitive biodiversity data.

This report makes recommendations that will involve the cooperation of agencies and specialists within Australia who are able to determine whether the release of specific biodiversity data may cause harm. It discusses the attitudes toward sensitive data of State and Territory conservation agencies, and those responsible for plant pests, animal health and biosecurity.

This report makes recommendations that will involve the cooperation of managers of biodiversity data and institutions who wish to share their datasets, and are concerned that some of their data may be sensitive.

3 Scope

This study investigates the types of biodiversity data that are considered sensitive in Australia. It covers information about plants, animals and microorganisms found within Australia, its off-shore territories and marine areas. The information in this report is relevant to holders of data about Australian biodiversity, and can be applied to datasets that may exist in other countries. This study does not, however, apply to data about organisms that are found overseas, even when that data may be held in Australia.

This report addresses public access to data. Sensitive information is commonly dealt with by imposition of restricted access. Restriction methods, means of access, passwords, details of data security and security techniques, although important for restricted and sensitive data, are not considered in this report.

4 Objective of this study

The object of this study is to:

Determine the difficulties associated with working with sensitive data and how they act as a barrier to data sharing.

Create a sensitive data policy that will allow biodiversity data to be shared publicly and will be applicable nationally. It needs to have broad acceptance by data custodians, conservation and biosecurity agencies, and other specialists throughout Australia who understand the need to treat information associated with particular organisms as sensitive. It will offer a workable course of action and address the needs and difficulties of those wishing to share their data publicly.

5 Consultation

This project was conducted from April to June 2009. Submissions were accepted from, and interviews undertaken with, 65 groups and individual specialists from organisations in all States and Territories of Australia.

Discussions were held with the Chairs of the collection councils, CHAH, CHAFC, CHAEC and CHACM; representatives of State and Commonwealth agencies responsible for conservation, biosecurity, quarantine, primary industries, and animal and plant health; and specialists, curators and data managers at museums, herbaria, universities, CSIRO and other research organisations in Australia. See Appendix A for a list of agencies and organisations consulted and an explanation of their acronyms.

Preliminary findings were presented and discussed at a Faunal Collections Informatics Group (FCIG) meeting in Launceston in April 2009.

Interviews and discussions were conducted in person and over the phone. The objectives of these interviews and discussions were to determine:

- barriers associated with sharing data
- the means by which sensitive data is currently managed by those who share data
- the types of information that are considered to be sensitive
- the impact of a third party (perhaps inadvertently) releasing data that could be considered to be sensitive
- whether those aware of the harm caused by inappropriate release of sensitive biodiversity data would be prepared to contribute to nationally agreed criteria for sensitive data

6 What type of data is sensitive?

This table provides an overview of the types of data that are considered sensitive by people working with biodiversity data. Also included are the specific types of sensitive data that could be assessed by using either externally determined or locally determined criteria for sensitivity. Externally-determined criteria are decided by specialists and agencies responsible for conservation, plant and animal health, and biosecurity. Locally determined criteria are decided within an institution.

Sensitivity	**Sensitive element with examples**	**Criteria for assessing data sensitivity**
Location latitude/longitude	**threatened species** – frogs, risk of disturbance and disease	external
	collectible species – bird eggs, snakes, large spiders, scorpions, beetles	external
	commercially attractive species – orchids, ferns, cycads	external
	private property – weeds, rare and uncommon species	external
	commercial in-confidence – commercial fishing	external & local
Species names	**manuscript names** – create long-term taxonomic problems	local
	misleading names – cause confusion	external & local
	threatened species – *in this very well defined area (roadside) there is an endangered habitat*	external
Personal information	**donors** – often confidential by agreement; questionable legitimacy	external & local
	collectors – confidentiality, privacy *(mostly vertebrate collectors)*	external & local
	landholders – confidentiality, privacy, prior agreements	local
Dates	**transitory phenomena** – breeding times	external
	link to location – a collector (of sensitive species) was in a particular location on a particular date	external & local
	trade sensitive – recent events are more of a concern	external
Incomplete and unchecked data	**unpublished data** – premature results, intellectual property	local
	incorrect data – data integrity, quarantine-sensitive locations and misidentifications	local
Third party data	**licence restrictions** – many datasets are not to be transferred, or require attribution	local

Table 6-1 Summary of sensitive data types with examples. Some data types can be assessed for sensitivity using criteria that are decided externally, some using criteria that are decided locally, and some data types will require both. Support is needed with dealing with these data types.

7 Sensitive Threatened Species

Conservation status of flora and fauna is determined at both a State and federal level in Australia. The conservation agency of each State or Territory is responsible for creating lists of plants, animals and ecological communities that are threatened within their jurisdiction. The commonwealth conservation agency DEWHA, through a separate process, creates distinct national lists of endangered and threatened fauna, flora and ecological communities.

7.1 To restrict or not to restrict?

Why is information about threatened species sensitive?

"Making biodiversity data available should reduce the risk of damage to the environment. If it is likely to have the opposite effect, availability may need to be controlled." (NBN Trust, 2009).

Plants or animals that are confined to a fixed geographical area are easily located once sightings are recorded or specimens are collected.

Some threatened species of plants and animals are more at risk because they are:

- collectible, such as bird eggs and orchids;
- commercially valuable, such as ferns, cycads, reptiles and fish;
- at risk of disturbance, such as bird nests and delicate plants; or
- at risk of disease, such as frogs and some plants.

Reasons to restrict

Restricting access to the locations of threatened species is expected to reduce the likelihood of harm. However, it has been suggested that serious collectors know where to find their quarry anyway, and restrictions to relevant data will be futile. On the other hand, for the greater public it is generally considered that restricting location information makes the task of finding more difficult for *those who don't know but would like to know*.

Problems with reducing the precision of locations

- Any restriction acts as a barrier to the free flow of information
- Restrictions add complexity to data sharing
- Legitimate uses are hampered, such as for research, study, planning, and managing. This applies in particular to rare and threatened species as they are commonly the target of research, study, planning, and managing.
- Precise information is required for many studies, eg for scientific modelling
- Damage will occur when people don't know of threatened species sites or habitats. Fragile ecosystems are easily destroyed by mowing, road maintenance, fire, etc
- Common knowledge of precise locations of threatened species may be essential when tackling many key threatening processes[4] such as land clearing, and the impact of invasive pests such as weeds, foxes and rabbits

[4] See for example DEWHA, Listed Key Threatening Processes. Of the 17 processes on this list, none suggest that public access to information about the location of threatened species is a key threatening process. Two processes (dieback caused by *Phytophthora*, and amphibian infection caused by chytrid fungus) may have an indirect link to data restriction. Note that States and Territories also have separate lists of threatening processes.

Figure 7-1 How many trees does it take to make a forest? Pollard Park in Kirrawee, NSW, is one of the few locally remaining fragments of critically endangered Sydney Turpentine - Ironbark Forest. Common knowledge of the significance of this park, albeit less than half the size of a football ground, has averted its most recent threat – that of duplicating the adjacent railway line.

7.2 Determining sensitivity

Methods of determining the sensitivity of threatened species vary widely amongst the different State conservation agencies.

WA DEC has determined species as sensitive from the State threatened species and priority species lists. The WA threatened flora and fauna have legislative significance, whereas priority species have no legal protection but are considered *worthy of making a special case*. Locations of sensitive species are limited: users of *NatureMap* are prevented from zooming past a threshold scale; and spatial coordinates and other sensitive fields are hidden in locality species lists for public access.

NSW DECCW has adopted a formal approach to assessing sensitivity.[5] They have produced a sensitive data policy where specific threatened species are classified as sensitive and the accuracy of their location has been graded into three categories. There is no distinction based on the form of the data. For example, if the location of nesting sites of a particular bird is sensitive, then all sighting records of that bird are treated as similarly sensitive, whether at a nest or not.

GBIF worked with NSW DECCW to develop a best practices guide for generalising sensitive species occurrence data[6] in order to help determine the sensitivity of threatened species. To establish sensitivity, the guide suggests using a method based on three criteria: *Risk of harm*, *Impact of harm*, and *Sensitivity of data*. A decision is then made determining the *Category of sensitivity*, where the categories correspond to levels of generalisation. This guide is most relevant to conservation agencies familiar with the risks and possible impacts of harmful

[5] *NSW Department of Environment, Climate Change and Water sensitive species data policy* (2009).
[6] *Guide to Best Practices for Generalising Sensitive Species Occurrence Data*, Arthur Chapman and Oliver Grafton, (2008).

activity on threatened species. The guide is less relevant for data managers with limited knowledge of risks, impacts and sensitivities of the organisms in their datasets.

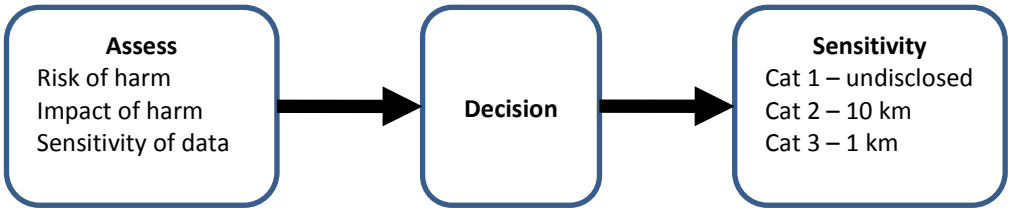

Figure 7-2 Assessment path for generalising the location of sensitive species. Based on Chapman and Grafton, 2008.

SA DEH does not treat information about threatened species as sensitive. They have determined that the advantages of publicly releasing information about threatened species exceed the risks of withholding information about selected species. They do however, manage third party datasets which have restrictions under their licence agreements.

VIC DSE has considered that the locations of all threatened species are sensitive and has reduced the precision of their occurrences on a map. They are currently reviewing their approach to determining sensitive species.

TAS DPIPWE engaged an expert panel to determine restricted species - the word 'sensitive' has connotations for fire and climate change. In general, public access to data about threatened species is best, and there are only three species that are restricted: spotted handfish, swan galaxias and *Lomatia tasmanica*. Access to authoritative information on Tasmania's wildlife is made through the *Natural Values Atlas*.

QLD EPA is responsible for the Queensland list of Rare and Threatened species. Of the species listed, the locations of nesting sites of some birds, and attractive plants such as orchids, ferns and cycads are considered sensitive. Outside the list, there are some species of invertebrates, spiders and scorpions, especially newly described taxa, which are attractive for international trade amongst collectors. Currently, *Wildlife Online*, the Queensland EPA public database, allows users to generate species lists to a limiting resolution of about 1 km.

NT NRETAS Herbarium has chosen to restrict location information for all taxa available through its public interface. Latitude and longitude are rounded to two decimal places, and locality is given as 'nearest named place'. There are good reasons for restricting only selected species, and although creating a filter is feasible for Northern Territory species, this would be a difficult task for a herbarium that has specimens from all Australian States.

ACT Parks, Conservation and Lands generalises point data for all threatened species. Locations are restricted to an accuracy of 250 m. They have used a risk assessment approach to sensitive species, and restrict virtually all public information about orchids treated as threatened species.

Local government and Catchment Management Authorities (CMAs) are commonly aware of communities that are not under threat at a State or federal level, but may be threatened and sensitive at a local level. An example of this would be the location of a penguin rookery in Sydney Harbour.

DEWHA has two facilities holding significant biodiversity data – SPRAT[7] and ANHAT[8]. Sensitivities extend to not including specific location information in the species profiles of threatened species, using maps with built-in fuzziness, and respecting wishes of third party providers.

Other agencies and specialists

Organisations such as museums, herbaria, research institutions, and National Parks Services, and specialists managing and studying plants and animals often understand closely the nature of the beast with which they are working. They can bring an experienced eye to species under threat, distributions of which cross political boundaries. Their fields of expertise may extend beyond that of State and Territory conservation agency listings. Timing may be critical, for example when releasing descriptions of new species, and legislative processes can be cumbersome and slow. Common examples of interest are invertebrates, shells, molluscs, fish and marine plants.

Figure 7-3 One of the largest spiders in Australia, *Selenotypus plumipes,* family Theraphosidae, is docile and popularly kept in captivity as a hobby. Few species in this family are widespread, and unrestricted access to data on accurate locations of newly described species will enable collectors to find them and substantially reduce their populations in the wild. Photo Robert Raven, Queensland Museum.

7.3 Other concerns

Ephemeral sensitivities

Much need for sensitivity in the scientific process may be short lived, basically until information in a more mature state is assessed and published. This may apply to information such as as-collected data, recovery plans in preparation, ongoing scientific work, and unpublished work such as range extensions or new host associations.

Duplicate records and reduction of precision of locations

Duplication of records will be a challenge for a service, such as the ALA, that aggregates data from many sources. The same data may arrive at the aggregator via different routes.

[7] SPRAT – Species Profile and Threats database provides information about species and ecological communities listed under the *Environment Protection and Biodiversity Conservation Act 1999.*

[8] ANHAT – Australian Natural Heritage Assessment Tool enables comparison of heritage values of different places. ANHAT is not publicly accessible.

Without some form of unique identifier applied at the source, obfuscation of locations or the idiosyncrasies of georeferencing could suggest erroneously that there are multiple occurrences, instead of a single observation. This has serious implications for records of rare species, as duplication could suggest that there are many instances of an organism, when in fact there may be just one.

Secondary effects
For survey data, sensitive data may be removed completely. It makes little sense to give accurate location information for some non-sensitive species, and also say this is the same location where particular sensitive species are found.

International implications
Some Australian agencies have data pertaining to records of species in other countries that may be sensitive to that country. These sensitivities need to be addressed.

7.4 Implications for data sharing

All conservation agencies have determined their own rules for sensitivity of the data that they release. Datasets held by third parties may hold data similar to that of any of those State or Territory agencies, and although the third party may make their data available publicly, it may not be released under the same set of sensitivity rules. For example, a State conservation agency may reduce the precision of the location of the nest site of a rare bird, whereas a local bird watching group may provide precise locality details of the same nest site on its website.

This discrepancy occurs for several reasons:

- Sensitivity rules created within an agency may not be able to be enforced outside the agency
- Sensitivity rules are neither advertised, promoted nor recommended widely
- Sensitivity rules are not in a form that could be readily adapted and adopted for use with a third party dataset
- Sensitivity rules may change, for example, due to a change in conservation status
- There are many dataset owners, including institutions interstate and overseas, that may have data pertaining to a particular jurisdiction, and the conservation agency responsible for that jurisdiction is generally not in a position to contact them all
- There are eight State and Territory conservation agencies. Each dataset holder would find it a daunting task to contact all conservation agencies regularly, in order to obtain their rules for sensitivity.

Conservation status varies across the country. Situations are likely to arise in which some threatened species considered sensitive in one jurisdiction, may be common or even a weed in another.

A National Register of Sensitive Threatened Species?
Currently there is no register of sensitive threatened species applicable Australia-wide.
In order to create such a register that could assist managing of existing datasets, it will be important to engage the help and cooperation of the conservation agencies of each State and Territory, and the Commonwealth. The State and Territory conservation agency rules for sensitivity could be combined state-by-state to form the basis for establishing a national register of sensitive threatened species. State and Territory borders would form a sensible extent for each agency.

Specialists and other organisations working with biodiversity, can be in a good position to recommend sensitivity rules independent of, and in addition to, State conservation agencies.

Each sensitive listing will need accompanying documentation for annotating dataset records.

Names

Accurate and current names of species are important for people working with biodiversity.[9] This is especially important when it is the organism that needs to be treated specially, independently of any name used to describe it. It would be unacceptable for a sensitive species to be defined by one name, only to find that some records could not be restricted because they referred to that same organism by another name. Desirably, the service for names would be able to relate the current broadly accepted (valid) name to its synonyms, as well as its historical, common, international, national and regional names.

Annotations

Datasets and sensitivity criteria will occasionally contain errors and inconsistencies. Those working with rare and threatened species may be aware of additional reasons where the sensitivity of an organism may need to be changed. Provision of a facility for annotation of records and rules, and for clear channels for feedback from the users to the originators will allow for correction and improvements of the data and the service, as well as keep users engaged. Of course, for this service to work effectively, data providers will have a responsibility to respond promptly and visibly to these suggestions.

Figure 7-4 Without directly disclosing the names of rare plants in danger of collection, measures can be put in place to protect endangered species from some of their greater threats – accidental destruction by mowing, land clearing, and road maintenance. Photo Pam Smedley.

[9] See *Atlas of Living Australia user needs analysis*, John Tann, Lynda Kelly, Paul Flemons, 2008

7.5 Recommendations

These recommendations will assist those wishing to share their data while respecting the concerns of those responsible for threatened species.

- Develop and implement a rules-based national register of sensitive threatened species. The rules would be created from criteria for sensitive threatened species recommended by conservation agencies and specialists.
- Allow for the rules for sensitivity to differ depending on geographic area, time, species, and more
- Allow for the rules to hold true in combination
- As a foundation, encourage each State and Territory conservation agency to contribute to the sensitivity rules applicable to that State or Territory. For example, WA DEC could determine the sensitivity rules that would be used for data pertaining to Western Australia; SA DEH could determine the sensitivity rules that would be used for data pertaining to South Australia; and so on.
- Encourage contributions from museums, herbaria and other organisations working with biodiversity
- Allow for, and encourage, contributions from specialists and experts
- Adopt a standard set of metadata fields for documenting of sensitive data
- Provide annotation and feedback services to clarify sensitive criteria
- Link all names to a taxonomic names service
- Include alerts for data contributors to remove or minimise personal information
- Develop techniques to minimise duplicate records, especially for rare species

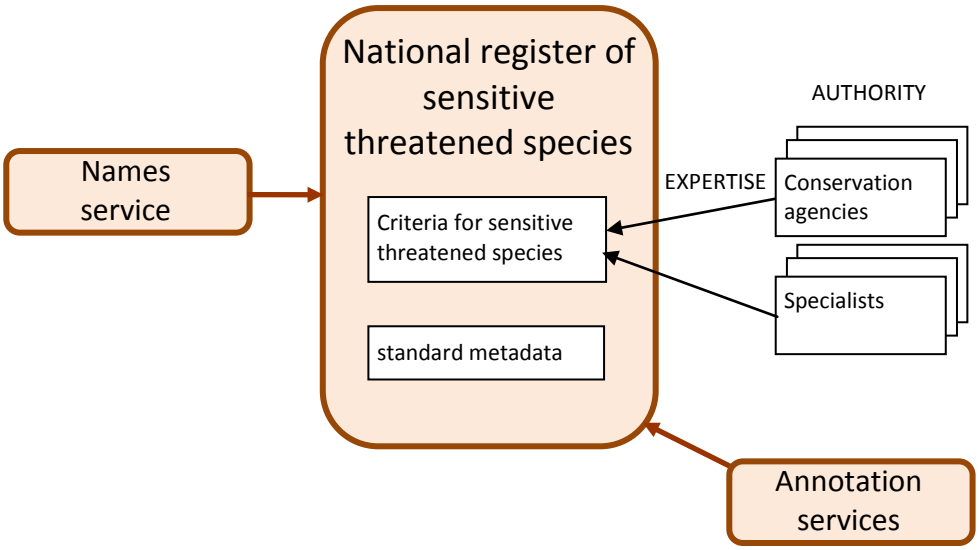

Figure 7-5 Diagram illustrating the operation of a proposed National register of sensitive threatened species. The criteria will be made available as a set of rules for assessing the sensitivity of biodiversity data, see Section 10 Managing sensitive data.

7.6 Participation

The following conservation agencies representing the Commonwealth and all the States and Territories will be encouraged to participate in helping establish criteria for a national register of sensitive threatened species

- ACT – Parks Conservation and Lands
- NSW DECCW – Department of Environment, Climate Change and Water
- NT NRETAS – Natural Resources, Environment, The Arts and Sport
- QLD EPA – Environmental Protection Agency
- SA DEH – Department of Environment and Heritage
- TAS DPIPWE – Department of Primary Industries, Parks, Water and Environment
- VIC DSE – Department of Sustainability and Environment
- WA DEC – Department of Environment and Conservation
- DEWHA – Department of the Environment, Water, Heritage and the Arts

Relevant organisations such as museums and herbaria, and specialists and experts throughout Australia, who have an understanding of threatened species sensitive data and how it would apply to their field of expertise, should also be invited to participate.

8 Sensitive plant pests and biosecurity

Quarantine procedures effect a barrier at Australia's borders to invasive and pest species. For plants, the pests of interest are insects, mites, nematodes, fungi, viruses and phytoplasma. For quarantine purposes, Australia's offshore territories such as Norfolk Island, Christmas Island, and some Torres Strait islands are regarded as 'artificial', ie outside the Australian quarantine border.

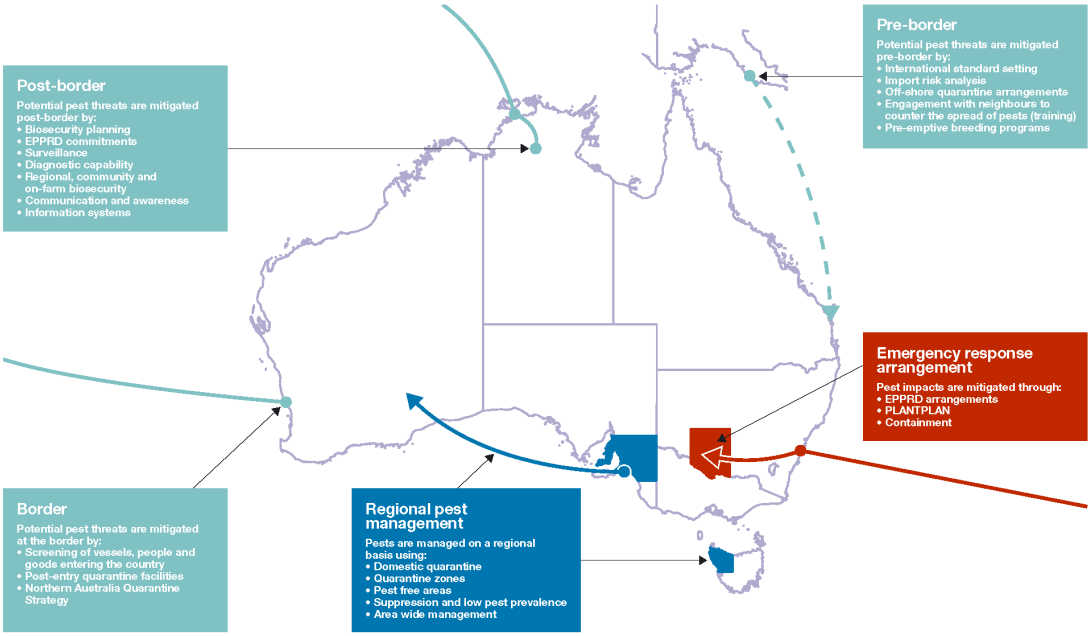

Figure 8-1 Key components of Australia's plant health system. Plant Health Australia (2009).

8.1 Quarantine-sensitive data

Quarantine data is highly sensitive. Australia's trade depends on the belief that certain pests are not present in the country. The perception of presence of a particular pest could trigger commercially difficult, international trade problems. Australia has obligations as a signatory to the International Plant Protection Convention (IPPC),[10] and as a member of the WTO, to provide information on its plant health status, and specimens held in reference collections provide verifiable evidence of that status.

For example, a record of a quarantine intercept, if not flagged as such, could suggest that it may be present in Australia. Depending on the species, under WTO agreements other countries may refuse to accept Australian produce if they know that it comes from an area where this pest has been found; or, from a different perspective, a foreign country may attempt to force Australia to accept imports at risk of harbouring a particular pest species, on the grounds that that pest already exists within the country.

[10] International Plant Protection Convention (IPPC). The IPPC is an international treaty to secure action to prevent the spread and introduction of pests of plants and plant products, and to promote appropriate measures for their control.

8.2 Types of data that are sensitive to plant biosecurity

Misidentifications – some species are easily misidentified. Modern genetic techniques can also produce misidentifications that are not readily obvious, for example sample mix-up, a misidentified vouchered specimen, or a sequencing error.

Incomplete identifications – Many quarantine intercepts may not have been identified to species. For some taxa, identification to genus or even family may help, but there are problems when those groups also have close relatives naturally occurring in Australia.

Inaccurate and imprecise locations – for example, there may be a fruit fly exclusion zone in the Riverina, yet a specimen marked NSW may be generalised to ALL OF NSW, and consequently fall erroneously inside the Riverina, implying that fruit fly could be found there.

Historical records – State-based departments of primary industries in the past were responsible for international quarantine, and so their databases have records of exotic species – many of these with no reference to their being quarantine intercepts. Having to determine quickly that a pest is not established in a particular place is expensive and disruptive.

Internal borders – Australia has long had a system to restrict transport of produce from areas affected by particular pests into unaffected areas. Interceptions associated with this internal quarantine thus occur along State borders and other regional boundaries. Again, data associated with these specimens may not describe adequately their source as quarantine interceptions. Misinterpretation of that data may be of serious concern for international trade.

Dates – Following eradication of a troublesome pest, a specimen and record may still remain in a collection and in a database. If this were marked on a map, it may lead to the interpretation that this pest is still to be found there. These records need to be marked SENSITIVE ERADICATED or a similar tag.

Unvalidated records – where identification is not confirmed, or locality is not confirmed. An example could be a report of a borer in a piece of furniture recently arrived from China.

Commercial confidentiality – An enterprise growing a particular crop may not want it to be known publicly known that they have a particular pest or even that the crop is grown on a particular parcel of land.

Weeds – Landholders consider their personal details and locations of weeds to be sensitive. There are legal and compliance issues surrounding weeds; court action may be pending, or in place; and there is a fear of prosecution.

Exotic specimens – Australian collections hold specimens collected overseas representing pests and diseases that do not occur in Australia. These specimens can be important reference material for diagnostic purposes.

Specific taxa – Certain groups of organisms, such as fruit flies (family Tephritidae)[11] are notably trade sensitive. Some areas of Australia are recognised by trading partners as free of particular fruit fly species. However, this status can be called into question if old records of eradicated species or quarantine interceptions are drawn unqualified into the limelight. There is uncertainty in the species-level taxonomy of many fruit flies, including some species of *Bactrocera*, a large and taxonomically difficult genus which includes some especially damaging pests. Unqualified, publication of records as "*Bactrocera sp.*" can provoke alarm that an unidentified specimen may represent one of the quarantine species.

Area Freedom concepts – Australia enjoys access to some international markets on the basis that some production areas are free of particular pests. For example, Queensland fruit fly (Q-fly), a species of *Bactrocera,* is a pest in many parts of Australia. However, citrus fruit coming out of the Murrumbidgee Irrigation Area (MIA) is accepted by other countries on the basis that it comes from an area where Q-fly is actively managed. Data suggesting that Q-fly is quite widespread in NSW and Victoria misrepresents the actual (and well documented) situation and could unnecessarily threaten trade.

Figure 8-2 An identified museum specimen of *Ceratitus capitata* (Wiedemann), commonly known as Mediterranean fruit fly, or medfly for short. This specimen is obviously very old, and labelled as being collected in Sydney. However, there is no record of a date. A serious pest of crops, medfly was eradicated in NSW more than 50 years ago. Publicly releasing an undated database record of this specimen is likely to have significant trade implications. Photo Lucy Tann.

[11] There is much confusion with the use of the term *fruit fly* in Australia. *Fruit fly* is the common name of members of two quite distinct families of flies:

1. Family **Tephritidae**, genus ***Bactrocera*** – a genus of fruit fly that causes rotting in fruit
2. Family **Drosophilidae**, genus ***Drosophila*** – a genus of fruit fly that is attracted to rotting fruit

It is ***Bactrocera*** species that cause serious problems to commercial fruit crops.
Museums across the country have records of *Bactrocera* species and other exotic insects. Some of them may be poorly labelled and lead to misinterpretation as to their origins.

8.3 Concerns with restrictions

Restricting access to data can bring its own difficulties. Timely access to knowledge of locations of pests and disease outbreaks is important for good plant health management. Delays in publicising of current work are of concern, as this can cause a quarantine breach to escalate.

Concerns arise with making available negative data, ie surveillance targeted to detect exotic pests. Restricted access to this data is likely to be counterproductive.

An example where information about certain pests should not be restricted is for Emergency Plant Pests (EPPs). For each of the major plant crops in Australia, a panel of experts is creating a list of the exotic pests which have mostly not yet arrived in Australia. There is a deed which gathers these EPPs and devises contingency plans (IBPs),[12] including funding for incursions. For example, there are 150 EPPs in Victoria.[13] Timely access to all data about EPPs must be a priority.

Reducing the precision of locations for pests can erroneously suggest that they are found where they are not.

8.4 Determining plant pest and disease status

Plant health and biosecurity is managed at both a federal and State level.

Commonwealth

The Office of the Chief Plant Protection Officer within the Commonwealth Department of Agriculture, Fisheries and Forestry is the national and international focal point for plant health in Australia.[14] It links with quarantine and market access efforts, and coordinates national plant health policy. The Department is Australia's "National Plant Protection Organisation" and has IPPC standards at its disposal to describe national plant health status and the status of particular species (eg "established", "transient", "no longer present").

Note that the federal Quarantine Act specifies "prohibited" imports. Schedule 6 is amended every 3 months[15] and lists pest plants. It has limited use for determining the quarantine sensitivity of database records, as it also includes species native to Australia, such as eucalypts.

States and Territories

In each State the chief plant health manager is responsible for lists of "regulated" pests and "proclaimed" pests.

---

[12] IBPs – Industry biosecurity plans

[13] PaDIL – Pests and Diseases Image Library has a Plant Biosecurity Toolbox with EPP lists.
See http://www.padil.gov.au/pbt/

[14] http://www.daff.gov.au/about/contactus/piaph

[15] Quarantine Proclamation 1998
The Quarantine Act creates a system by which things and activities that are likely to introduce pests or disease can be prevented from entering Australia. The basic way to impose these restrictions is by proclamation. Quarantine Proclamation 1998 contains these restrictions.
Schedule 5 - Permitted seeds
Schedule 6 - Kinds of plants that must not be imported (section 65)
Comment. Neither Schedule 5 nor Schedule 6 is very useful as a filter to trap potentially quarantine-sensitive species within datasets of existing collections, as they list species native to, and widespread in Australia – for example, Eucalyptus species.

Two lists are of interest:

1. Pests that are established in limited areas within the State
2. Pests that occur in other States, and are prohibited imports for the State[16]

APPD

The Australian Plant Pest Database is a nationally coordinated database of plant pests and diseases.[17] It provides a single point of access to existing electronic records of voucher specimens contained in databases across Australia. The system provides a powerful tool to assist bids for market access and to justify measures to exclude potentially harmful, exotic organisms.[18] The APPD provides access to most databases containing pest data, but does not provide access to databases of the Australian museums community.

Access to the APPD is restricted. Generally, only plant health or quarantine specialists and collection curators from contributing agencies have access to the database.[19]

Data held within the APPD is subject to ongoing review for integrity, quarantine intercepts and public availability.

8.5 Implications for data sharing

Agencies responsible for plant pests and biosecurity are concerned about the potential release of quarantine-sensitive data.

It is generally considered that all entomological and pathogenic data should be made available. However, due to the likelihood of misinterpretation of records, coupled with the expensive and disruptive complications arising from this, then for public use it will be necessary to filter and annotate records of organisms considered to be quarantine-sensitive.

Currently there is no definitive register of species that are considered quarantine-sensitive. In order to create such a register that could help manage existing datasets, it will be important to engage the help and cooperation of agencies currently working with plant pests and biosecurity.

Points to be considered in compiling a register of quarantine-sensitive species:

- The State and Territory lists of *plant pests not present in the State (Territory)* could form the basis for establishing a register of quarantine-sensitive species
- APPD will be an important resource to help determine quarantine-sensitive species
- Each quarantine-sensitive listing will need accompanying documentation for annotating data records
- A different set of rules may need to be created for each “problem” group. Take for example the *Bactrocera* group of fruit flies: some of these are native to Australia; some are native to particular states or regions; some are pests of agriculture and some are not. A set of rules governing the *Bactrocera* would need to take this information into account.
- Information associated with a species of interest may not always be resolvable

[16] An example of state/territory declared pest list (Victoria) is http://www.dpi.vic.gov.au/dpi/nrenfa.nsf/LinkView/0FA3479925C07065CA2573E7002348027A3C416170F25102CA2573E7007B22F5/$file/pests%20and%20disease.pdf

[17] http://www.planthealthaustralia.com.au/our_projects/display_project.asp?category=4&ID=1

[18] http://www.planthealthaustralia.com.au/APPD/help/index.asp

[19] http://www.planthealthaustralia.com.au/APPD/guidelines/index.asp

- Specimens that are of interest for quarantine purposes will need to be identified accurately
- It is inevitable that some minor things will be missed
- Annotation services where end users can provide feedback will help improve both the data quality and the criteria for determining sensitivity
- It will be important to relate the current broadly accepted name to its synonyms, as well as its historical, common, international, national and regional names

Other items to note in relation to data sharing

- Often informative notes accompany data records for exotics. Data associated with specimens of exotic origin can reveal possible pathways of introduction. Knowing that representative specimens of exotic species are held in particular collections can assist diagnosticians wishing to compare these specimens with unknowns.
- Privacy is no less important than for other biodiversity records
- Commercial confidentiality is important
- Location details of weeds on private property will need to be suppressed at the source

8.6 Recommendations

These recommendations will assist those wishing to share their data while respecting the concerns of those responsible for plant pests and biosecurity.

- Develop and employ a rules-based register of quarantine-sensitive species, based on criteria recommended by agencies responsible for plant health and biosecurity.
- Ensure all publicly accessible records of quarantine intercepts are labelled clearly as such.
- Develop and employ a list of alert phrases to accompany quarantine-sensitive data records. Include also reference to special quarantine areas.
  For example:
  - TRADE SENSITIVE
  - QUARANTINE INTERCEPT
  - SPECIAL QUARANTINE AREA
  - NOT KNOWN TO OCCUR IN THIS AREA
  - NOT KNOWN TO OCCUR IN THIS STATE
  - NOT KNOWN TO BE PRESENT IN AUSTRALIA
  - ERADICATED
- Develop and employ a hierarchical scheme for categorising unresolved matters. For example:
  - highest priority taxa
  - list of possible taxa
  - intractable taxa, eg unidentified fruit fly *Bactrocera sp.*
- Provide annotation and feedback services for clarification of sensitive criteria.
- Link all names to a taxonomic names service.
- Include alert signals to data contributors to remove or minimise personal information

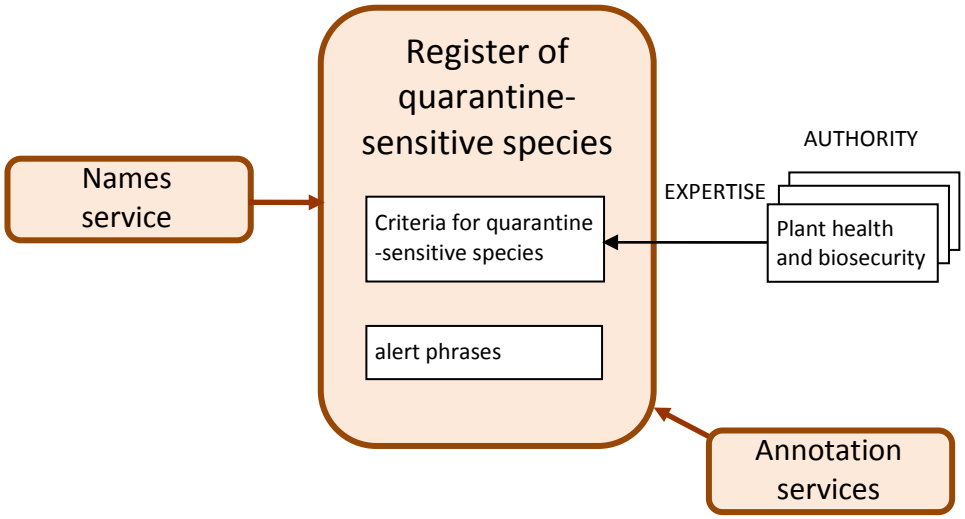

Figure 8-3 Diagram illustrating the operation of a proposed Register of quarantine-sensitive species. The criteria will be made available as a set of rules for assessing the sensitivity of biodiversity data, see Section 10 Managing sensitive data.

8.7 Participation

The following key groups will be encouraged to participate in helping establish criteria for plant pests and biosecurity sensitivity:

- **APPD – Australian Plant Pest Database** Steering Committee, comprising representatives of data providers, diagnostic specialists and users.
- **PHC – Plant Health Committee,** the overarching group responsible for plant health which comprises representatives of the Commonwealth, State and Territory Governments.[20]
- **PHA – Plant Health Australia,** another peak, national, coordinating body for plant health in Australia.[21] PHA works with the private sector and government to develop biosecurity strategies for major plant industries.

[20] See http://www.daff.gov.au/animal-plant-health/plant/committees/phc
[21] See http://www.planthealthaustralia.com.au/

9 Sensitive microorganisms, animal health and biosecurity


[22]

In Australia, animal health and biosecurity are subject to the same principles as plant health, and have similar implications for trade. However, there are some additional considerations, such as the possibility of some diseases affecting both livestock and humans, and the fact that some diseases are explicitly identified as threats to national security.

Testing of animals is well regulated. Testing must be carried out by a licensed person or accredited laboratory to ensure that accurate information is obtained. False positives must be avoided to reduce the possibility of unwarranted disruption to production and trade. False negatives are equally undesirable, since early and precise detection of disease outbreaks can be the key to successful eradication or management.

9.1 Types of data that are sensitive to animal biosecurity

Trade

Universities, hospitals, and animal testing and research laboratories hold isolates and collections of microorganisms. If an institution claimed (possibly erroneously) to hold a domestic isolate of an exotic pest or disease and this record were placed in the public domain, trade problems might ensue unnecessarily. An example of this would be if a university stated that it held a *Brucella abortus* isolate obtained from an Australian cow in 2008. Australia has implemented a major program to eliminate brucellosis from its herds. This sensitivity could apply to any pest or disease of which Australia is free; or to pests or diseases where certain areas are free.

Published lists are available for disease agents for both terrestrial and aquatic animals that could be potentially trade sensitive. Of interest would be:

Animal Health in Australia, report (Animal Health Australia, 2008)

- Table 2.1 Australia's status for OIE-listed diseases
- Table 2.2 Australia's status for other diseases of interest
- Table 3.1 Australia's status for OIE-listed diseases of aquatic animals
- Table 3.2 Australia's status for other diseases of aquatic animals

SSBA

For biosecurity reasons, the Department of Health and Ageing has created a list of Security Sensitive Biological Agents (SSBA)[23] and has legislated security controls on certain isolates,

[22] Biohazard symbol by Silsor. http://commons.wikimedia.org/wiki/File:Biohazard_symbol_(red).svg

bacteria and pathogens. Some SSBA affect only humans, some affect only animals, and some affect both. Based on risk analyses, those organisms that present a high risk are identified.

Some institutions holding these agents would prefer not to be identified and, at any rate legal obligations prevent the publication of where these organisms are held.

Care needs to be taken in publishing the source of these agents. References to historical sources, or where information has long been widely available, are likely to be acceptable.

Human donors
Human donor information is confidential and needs to be removed at the source.

9.2 Concerns with restrictions

Much information on disease has most value when timely. Restrictions can delay and frustrate actions.

Scientists and others who need samples of microorganisms including SSBA for diagnostic purposes need to know which institution to contact in order to obtain these. Rather than create a public facility for passing legitimate requests for SSBA and other restricted microorganisms it is hoped that increasingly, diagnosticians will be more effectively networked and able to share this kind of information directly.

9.3 Implications for data sharing

Agencies responsible for animal health and biosecurity are concerned about the potential release of sensitive data that may affect trade or security.

Documentation relating to potentially sensitive microorganisms is readily available. With appropriate advice and support from animal and human health agencies, lists of SSBA and diseases could be readily adapted to generate criteria for assessing sensitive microorganisms.

For managing disease outbreaks such as avian influenza, processed data such as daily statistics, trends, and broad-scale maps, can be more useful to managers than raw data. Processed data is less likely to have sensitivities attached – personal information is gone, locations are generalised, and individual reports are aggregated.

Providing the ability for end users to annotate data and provide feedback will help improve both the data quality and the criteria for determining sensitivity.

9.4 Recommendations

These recommendations will assist those wishing to share their data, while respecting the concerns of those responsible for microorganisms, animal health and biosecurity.

- Develop and employ a rules-based register of sensitive microorganisms from criteria for sensitive microorganisms recommended by agencies responsible for animal health and biosecurity.

[23] See *Health Emergency Preparedness and Response, Security Sensitive Biological Agents (SSBA)* and *SSBA guidelines*

- Incorporate the list of SSBA into the register of sensitive microorganisms
- Incorporate the published lists of disease agents for both terrestrial and aquatic animals into the register of sensitive microorganisms. These lists may need to be moderated to account for trade sensitivities.
- Develop and employ a list of alert phrases to accompany sensitive microorganism data records.
  For example:
  - SSBA
  - LAST OCCURRENCE ON <date>
  - HELD IN COLLECTION. MAY NOT EXIST OUTSIDE THIS INSTITUTION
  - ERADICATED FROM AUSTRALIA
- Provide annotation and feedback services to clarify sensitive criteria
- Link all names to a taxonomic names service
- Include alerts to warn to remove or minimise personal information

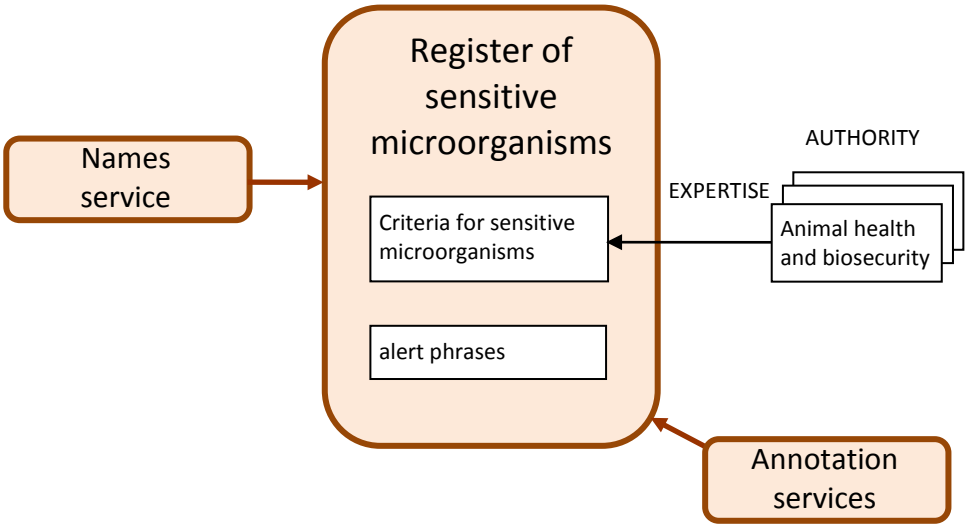

Figure 9-1 Diagram illustrating the operation of a proposed Register of sensitive microorganisms. The criteria will be made available as a set of rules for assessing the sensitivity of biodiversity data, see Section 10 Managing sensitive data.

9.5 Participation

The following key organisations will be encouraged to participate in helping establish criteria for animal health and biosecurity sensitivity:

- **Office of the Chief Veterinary Officer** provides scientific advice to minimise potential impacts of disease on animal health in Australia
- **Council of Heads of Australian Collections of Microorganisms** understands the diversity and use of microorganisms and culture collections

Other officials and organisations may be in a good position to contribute supplementary details. Such as:

- **Chief Veterinary Officers** in each State and Territory
- **Health Emergency Management**, Department of Health and Ageing

10 Managing sensitive data

Australia is signatory to principles and guidelines recommended by OECD to facilitate cost-effective access to digital research from public funding.[24] These recommendations set out collective and precise objectives which member countries are expected to implement. The Atlas of Living Australia will create open access to Australia's biodiversity data, and institutions will be able to share their research data through the ALA.

Information about Australia's biodiversity is held in many datasets. Support for management of these datasets varies widely across the country with some data managers having little in the way of extra resources. Many database managers and curators are in the process of preparing their datasets for sharing, however the complexity associated with sensitive data has made the task of sharing data intimidating.

Material in this section derives from interviews and surveys conducted with managers of biodiversity data at institutions throughout Australia.

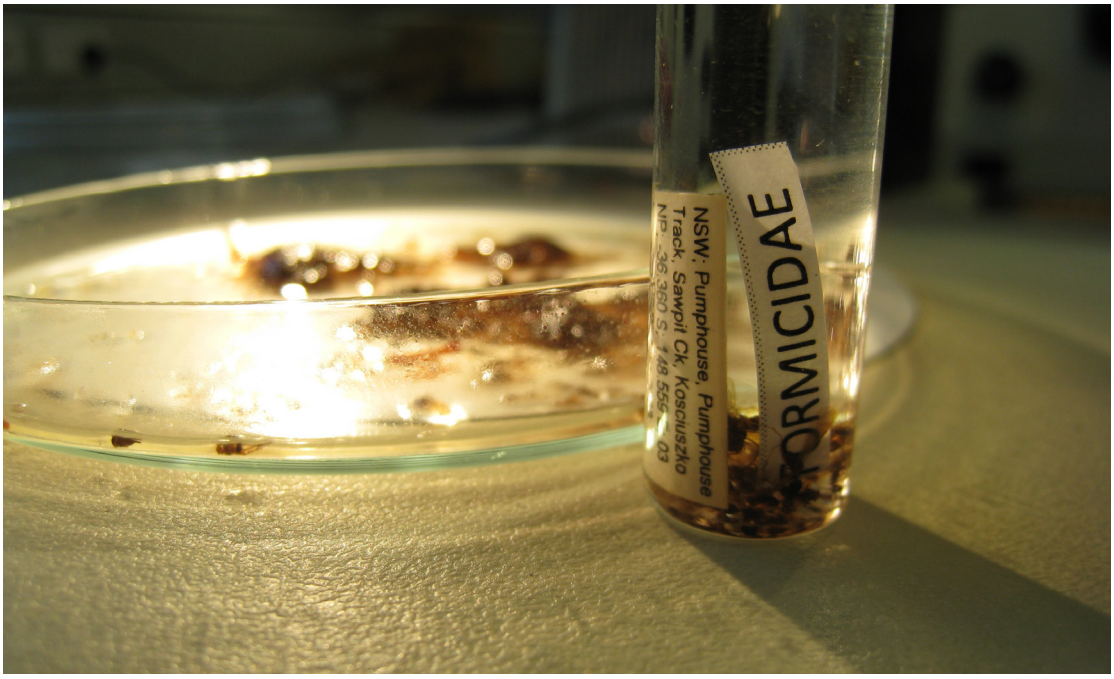

Figure 10-1 Work-in-progress. Sorting the contents of a malaise trap is one of the many tasks involved in assessment of species richness. Occurrence data from this ongoing research will not be available publicly until data is analysed and results are published.

10.1 Types of data that are sensitive for a data custodian

Unvalidated and uncertain data. The presence of unchecked or unvalidated data is one of the common explanations for non-sharing of data. Data managers are reluctant to release data that has not been checked thoroughly. Reputations are at risk. Grading of unvalidated and uncertain data may help. Annotation and validation services may help.

Poor quality and incorrect data. For similar reasons, public release of poor quality or erroneous data can reflect on the reputation of a researcher or institution. However,

[24] See *OECD Principles and Guidelines for Access to Research Data from Public Funding*, 2007
Also see *OECD Declaration on Access to Research Data from Public Funding*, January 2004

unconfirmed sightings, minimal labelling of historical specimens, incomplete identifications, and even data containing mistakes, can all be used in some contexts, especially if clearly graded. Points for consideration are: how data could be graded; and whether it is advisable that grading be the task of the data custodian. Annotation services may help.

Incomplete and insufficient data. Institutions have limited resources to apply to queries about incomplete identifications, distributions, unknown dates, etc. Filters have traditionally been put in place to minimise workload.

Data that is unpublished or subject to ongoing research. A scientific research project involves many steps – from applying for funding and collaborating, through collection of data, analysis, and publication of results. For the natural sciences this process may be protracted over several years. The ability to attract future support relies much on the quality of the output of previous research. Premature release of data could prejudice these efforts. Although occurrence data and other measurements may be valuable to others, scientists require some time to exclusively analyse and present their results.

Data categorised as intellectual property. This might also apply to data on or related to indigenous knowledge, such as the traditional uses of particular plants. There may be uncertainties, release restrictions such as time delays, or access permissions. Photographs, images and published materials have copyright concerns.

Names of donors. Agreements exist with some donors that their names are not to be released. Privacy laws generally prohibit unsolicited disclosure of information that may identify a person.

Names of collectors. Release of personal names appears to be mostly a concern of collectors of vertebrates such as reptiles, who do not want their identity revealed for fear of harassment by avid collectors wanting specimens for commerce, trade or for personal collections. This appears not to be a problem for collectors of plants, or other faunal groups. Although privacy laws may be relevant, collectors' names are routinely included in publications referring to their specimens.

Questionably obtained specimens. Some institutions refuse any specimens that have not been obtained legally. Others acknowledge that a museum may make the best use of questionably obtained fauna. An example might be a donation from a deceased estate of a well-documented collection of bird eggs. Collection for scientific purposes generally requires a licence; historical records may have no attached associated licence information. Information of a sensitive nature might include names of donors and collectors, dates, and locations; restrictions may have time limits.

Manuscript names. Premature publication of manuscript names creates taxonomic and nomenclatural problems. The rules of nomenclature (and several different Codes are involved) are evolving to cope with digital publication in its various forms and the appearance of manuscript names in databases may cause problems into the future. A means of dealing with manuscript names needs to be considered as well as whether working names, field names, phrase names, and convenience names like *Genus sp*. and *Genus sp. A*, are acceptable or useful.

Gazumping and pre-publication names. Release of even sketchy details of new names that are in-press, raises fears that the names will be unscrupulously usurped; some isolated gazumping attempts have occurred in the past.

Third party data. Data owned by others may be subject to restrictive licence agreements that prevent sharing further.

Data that is "MINE". Even when an institution has agreed to make its collective, scientific data public, data managers may be reluctant or unable to allow access to data subsets "controlled" by individuals within the institution. These individuals may be fearful of becoming redundant, of losing relevance or academic edge, of loss of income, of loss of control, of what someone might do with the data, or of making hard decisions. These perceived data sensitivities could stem from deeper administrative or rivalry concerns. They may have little to do with science or rational policy but can be a serious deterrent to data sharing. Revised reward structures and support may make a difference.

In-confidence. Some data are provided to custodians on an "in-confidence" basis. For example, observations of commercial fishing activities commonly are recorded on the agreement or understanding that the data is confidential and will be used for a specific scientific or regulatory purpose. Fishing is a competitive trade; catch sizes, or deep-sea locations are sensitive information and, not unreasonably, operators prefer to keep this kind of intelligence from competitors. Other examples of in-confidence concerns are: environmental surveys carried out under contract, where one party, or the contract does not permit open release of the data; or surveys that might have commercial significance, such as mining surveys. Much of the nature of environmental monitoring is carried out in-trust and built on mutual goodwill. Breaking that trust may jeopardise other ventures. Care needs to be taken to ensure both formal and informal agreements are honoured.

Private property. Landholders in particular, have refused access for surveys on their land unless the results were only to be used under specific conditions. There are instances where these arrangements were negotiated individually and been found unwieldy to manage, defining different and restrictive conditions of data use. Landholders could require data to be generalised to a level where their property could not be identified. Data obtained from these expensive surveys is restricted and virtually unusable elsewhere.

10.2 Concerns about managing sensitive data

Concerns raised by data managers in the course of this study are below. Some of these questions have been addressed in other parts of this report, such as "What data is sensitive?" They are reiterated here to emphasise the difficulties faced by those wishing to share their datasets.

Concerns about preparing data for public release include:

- What data is sensitive?
- Who knows what data is sensitive?
- Is there a checklist of potentially sensitive data types or fields?
- Are there any tools to help flag sensitive data?
- What is sensitive about this particular taxon?
- What secondary information also needs to be restricted to ensure primary sensitive data is protected?

- Different jurisdictions have different criteria for sensitivity. How is data from many places best managed?
- What is the best way to check a dataset of records that use old names, against lists with modern names?

Concerns about releasing data
- When data is withheld from the public due to sensitivity, how secure is it? Who can see it or use it?
- Once data is released, it is generally gone. Copies of that data may perpetuate irrespective of correctness or sensitivity. Care will be required to ensure there are no accidental releases of sensitive information.
- What are the consequences of public release of data that someone else considers to be sensitive?

Concerns about maintaining sensitive data

Review
- How to keep track of changing sensitivities?
- How to incorporate an irregularly changing set of sensitivities?
- How often need a dataset be reviewed for sensitivities?
- How important is it to be up-to-date with any changes to sensitive data?

Feedback
- If publicly accessible data is to be subject to annotation services and feedback, what are the implications for a data manager?
- What is the best way to deal with feedback?
- How to accept/reject changes?
- How to incorporate changes to original data?
- How to re-issue corrected data?
- How to apply feedback to third party data?

10.3 An external Sensitive Data Service to help manage sensitive data

In the course of this study, the concept of an external Sensitive Data Service was suggested to data managers as a possible practical way to help classify their sensitive records. This hypothetical service was described as an online site where data managers could freely submit their datasets. Each record in their dataset would then be compared against nationally agreed criteria for sensitivity, and records regarded as sensitive would be tagged and, potentially, graded. The results of this service would be returned to the data manager in an appropriate form for review.

Concerns raised by data managers about a Sensitive Data Service:
- What are the sensitivity rules?
- Can the sensitivity rules be used locally?
- Can an external service be trusted with sensitive data? Can it be trusted with any data?
- Who takes responsibility if the Sensitive Data Service is wrong?
- What data formats are needed?
- What is the process for extracting records from a database and providing those records to an external service?
- How would alerts from an external service be incorporated into a database?

- The overhead of sending data to a separate service to determine sensitivity may be too great for those with limited support. Can the Sensitive Data Service be incorporated into existing data sharing facilities such as OZCAM or AVH?

The next chapter discusses in more detail implications of a Sensitive Data Service.

10.4 Implications for data sharing

Data managers and custodians wishing to share their data are concerned with the processes involved in extraction of data from their institution in a manner that ensures that any data considered sensitive will be treated appropriately. They would benefit from clear guidelines and advice as to what types of data their institution might consider to be sensitive. They also need a means to alert them to the data records that they manage, but which others consider to be sensitive.

Tools and techniques need to be developed that will help with the data sharing process. In general, data managers have limited resources; whatever is developed needs to minimise the complexity and size of the task.

Register of sensitive species

Currently there is no register of sensitive taxa that is applicable nationally. Having a single, authoritative, up-to-date register would improve the process of determining sensitivity for all data managers and custodians.

- A single register minimises the task of discovery
- An authoritative register gives confidence to data managers and custodians that they are instituting the sensitivity requirements that will obviate or minimise the harm that could be caused by releasing some forms of data
- An up-to-date register allows for re-checking as priorities change

This national register could track sensitivity locally without generalising it. It could, for example, accept and apply sensitive criteria that apply to an organism only in one part of the country and not elsewhere.

Sensitive Data Service

If a national register of sensitive species were to be created, a facility for testing datasets against the sensitivity rules would be needed. Creation of an online tool that would be a service for data managers and custodians, would be their interface to the register of sensitive species. This service would need to address the concerns of data managers and custodians such as security, quality of service, reliability, and ease of use. An instance of a Sensitive Data Service that would function locally yet refer remotely to the national register of sensitive species may be appropriate for some institutions.

Support for data managers

Data managers need support. Having access to checklists, FAQs, how-to's, standards, schemas, and guides relating to data sharing in a single one-stop-shop would help. A data management wiki to help implement this would make it possible for data managers to help other data managers working with similar difficulties.

Quality rating

A rating system may be a useful way to discriminate between, for example, an unconfirmed observation by a novice and a vouchered specimen in a well-curated collection, identified by a respected expert. Known errors need to be annotated clearly. A rating system would need

to be defined. Whether data managers and custodians grade their own data for quality would need to be considered.

Taxonomic names
Managing scientific names for a biodiversity dataset is currently difficult as there is no standard reference that keeps track of the constantly changing and updating taxonomy. It is the name of an organism that allows it to be referenced to others. When sharing data, names associated with records will either need to be current, or easily cross-referenced to a current name. A well maintained taxonomic names service may be the best way to address this.

Annotation services and feedback
Creating the ability for end users to annotate records and datasets, and enable feedback back to the data provider will be an ongoing way to help with:

- poor quality data
- correcting and adding locations, identifications, names
- re-grading data
- manual tasks with a complexity beyond the capability of current computing

Consideration will need to be given to the best way to manage the results of data annotation, and how to apply that to original data. A workflow may need to be established that will enable engagement of specialists, acceptance or rejection of annotations, an auditable set of changes, appropriate documentation, duplication removal, some automation, and a means to obviate the need to re-visit the same changes.

Data security
There are strong concerns about data security and access to restricted information which will need to be dealt with by the ALA. They are noted here, but further investigation into the means of restricting and enabling access is beyond the scope of this study.

10.5 Recommendations

These recommendations will assist those wishing to share their data, while respecting the concerns of data managers and custodians with regard to sensitive data.

National Register of Sensitive Species

- Develop and implement a rules-based National Register of Sensitive Species by including three separate registers:
    1. the national register of sensitive threatened species
    2. the register of quarantine-sensitive species
    3. the register of sensitive microorganisms

Sensitive Data Service

- Develop and implement a Sensitive Data Service that references the National Register of Sensitive Species and is readily accessible for managers of datasets to use to determine records referring to sensitive species
- Include in the service for assessing sensitivity, a facility for grading sensitivity, and apply appropriate labelling
- Ensure that the service can readily accommodate changes to rules
- Ensure that the service is readily usable by a data manager who has little support
- Ensure that the service can be re-used both irregularly and frequently, without the need to manually re-visit the same data records multiple times
- Take steps to ensure transparency, integrity, accuracy, security, currency and care

Sensitive Data Toolbox

- Develop and implement for data managers, a support Toolbox which contains standards, schemas and checklists. Consider a wiki format to facilitate data manager interaction.
- Adopt a standard set of metadata fields to document sensitive data
- Provide annotation and feedback services to clarify and correct suspect data, as well as to clarify sensitive criteria
- Develop tools for managing feedback at the data source
- Develop and implement a pre-release checklist of potentially sensitive data types and fields, to help detect other, locally sensitive information not discovered with a rules-based register
- Include alerts to remove or minimise personal information
- Link all names to a taxonomic names service
- Devise a rating scheme to accommodate data of varying quality

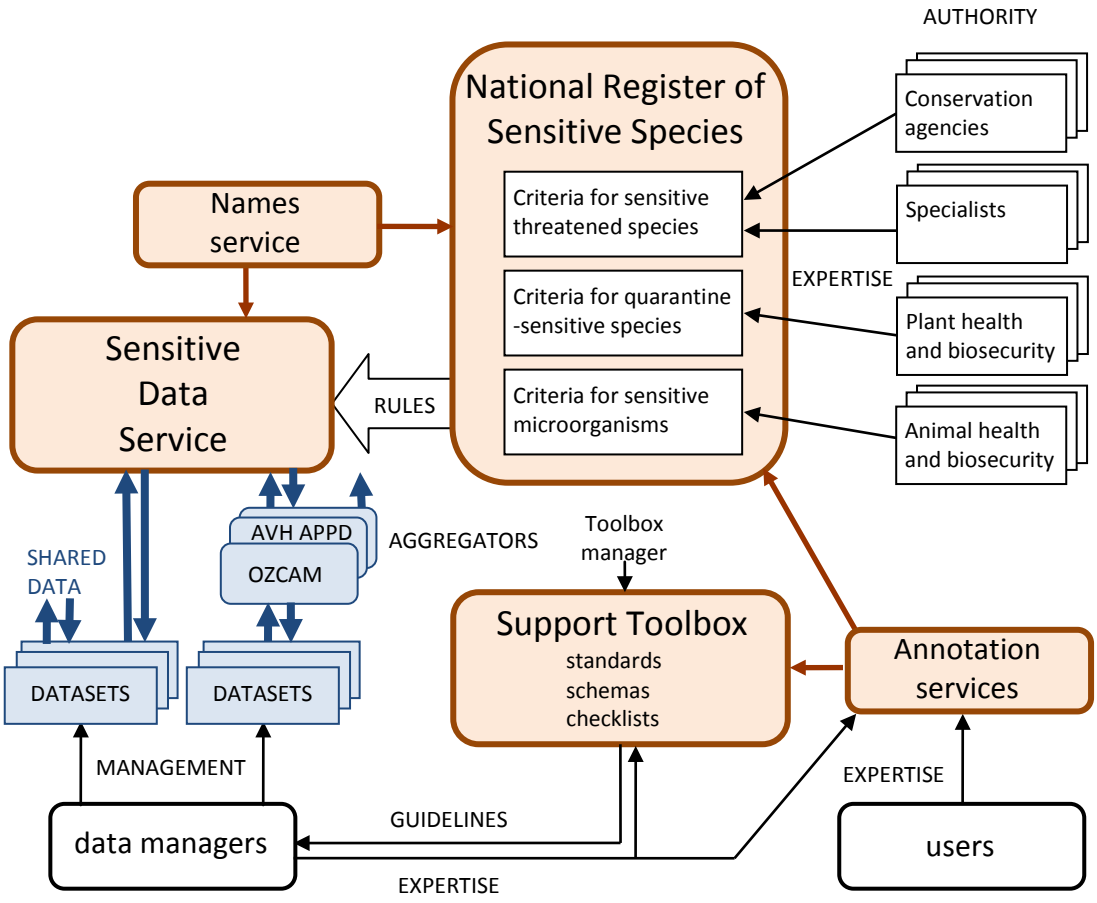

Figure 10-2 Diagram illustrating the operation of a proposed Sensitive Data Service to assist data managers and custodians with assessment of their datasets for sensitivity. The Sensitive Data Service references criteria for sensitivity based on rules held within the National Register of Sensitive Species. Data managers can also receive guidance on sensitivities that may be known only locally by using a peer-supported service, shown here as a Support Toolbox.

10.6 Participation

Managers of datasets interested in sharing their biodiversity data will be encouraged to participate in defining, developing and testing a service for assessing sensitivity.

11 Implications of a Sensitive Data Service

11.1 A model for an Atlas of Living Australia Sensitive Data Service

One model of a Sensitive Data Service could be an online service to which data managers and custodians would be able to submit their datasets, and through which each record would be checked against a National Register of Sensitive Species. The results of the check would be returned to the data manager in a form suitable for review, with each sensitive record flagged and graded. In order to address concerns about data security this would be a stateless system, and neither the data nor the results would be stored.

The National Register of Sensitive Species would be a rules-based filter comprising:

1. a national register of sensitive threatened species
2. a register of quarantine-sensitive species
3. a register of sensitive microorganisms

It is anticipated that these registers would contain criteria for sensitivity determined by interested agencies and experts concerned with the consequences of releasing sensitive data.

Associated with the Sensitive Data Service would be a wiki-based Sensitive Data Toolbox with standards, schemas and checklists. These tools and checklists could assist managers with datasets that may hold confidential or locally defined sensitive information.

Local implementation of the Service may be feasible. If a local sensitive data tool were available, it could allow the checking process to be iterated several times before data leaves the originating site. The tool would not only refer to the latest set of rules which would sit externally, but also allow for locally defined rules to be invoked as well. This tool would make it easier for providers to slot the processes into their own workflows at whatever point makes most sense and conceivably could be built to handle many in-house needs.

The Sensitive Data Service would be freely available to be used by data managers and custodians in the process of preparing their datasets for sharing.

11.2 Advantages

For **conservation and biosecurity agencies**, a National Register of Sensitive Species:

- would provide an outlet from which they could promote their policies with regards to sensitive species
- would provide a service through which they could readily introduce changes to new and existing sensitivities
- would provide the means for others to restrict their information about the species that an agency considers to be sensitive

For **data managers and custodians**, a Sensitive Data Service:

- would provide a single place to discover the sensitivities of many jurisdictions, species, agencies, etc
- would be an interface to an authoritative and current National Register of Sensitive Species
- would be freely accessible and could be re-used frequently
- would be a form of good-practice – it would be an up-to-date service that other data managers use

- would be consistent. Although it would be expected to change from time-to-time, the one National Register of Sensitive Species would be used to assess all datasets submitted to the Sensitive Data Service. The same National Register of Sensitive Species would also be expected to be used by the ALA and other data aggregators when they release data publicly.
- would be an obligation-free service. Using this service would not mean that data need be shared through the ALA.
- would be available for review of data before release, giving the opportunity to verify or withhold suspect records

For **data users**, a Sensitive Data Service that refers to a National Register of Sensitive Species

- would ensure that data accessed publicly through the Atlas of Living Australia could be re-used, re-purposed and re-published without the concerns that some of it may be sensitive

11.3 Weaknesses

A Sensitive Data Service that refers to a National Register of Sensitive Species:

- would need the support and cooperation of many agencies across the country – all national, State and Territory conservation agencies, plus agencies responsible for plant and animal health and biosecurity. If one or more agencies declines to contribute, then the effectiveness of the Sensitive Data Service would be reduced.
- would best work in a spirit of cooperation. Consensus regarding sensitivity criteria could be achieved through a Steering or Expert Advisory Committee. Principles for submitting sensitivity criteria may need to be clearly defined.
- would require data managers to do "one more thing" in order to enable them to share their data
- would not be perfect. For example, spelling problems, use of incorrect or out-of-date names, and information in unexpected data fields, may allow some sensitive records to slip through the filter.

11.4 Management

The Atlas of Living Australia would appropriately take on the task of managing the development and operation of a Sensitive Data Service. Its role would be to deliver a workable solution. The ALA would need to work with, and help, data managers and custodians with their task of sharing data while caring for their sensitive data.

In order to use the best organisation for the task, a separate agency with appropriate expertise will need to be selected to take responsible for establishment and management of the registers of sensitive species. Included in its role would be liaison with the agencies and specialists responsible for defining sensitive criteria.
This agency should:

- be interested primarily in science
- have infrastructure and capabilities to support biodiversity informatics
- have credibility as an authority on plants, animals and microorganisms
- have the respect of experts and specialists, museums, herbaria, State and federal conservation agencies, plant and animal health and biosecurity agencies
- be independent of close vested interests
- have appropriate governance and, possibly, security arrangements in place

A group or agency would need to be selected to moderate the Sensitive Data Toolbox. Its role would be to facilitate interaction between data managers involved in the task of data sharing.

This group would be required to:

- have an understanding of data standards and schemas
- understand management of biodiversity data
- have the respect of data managers

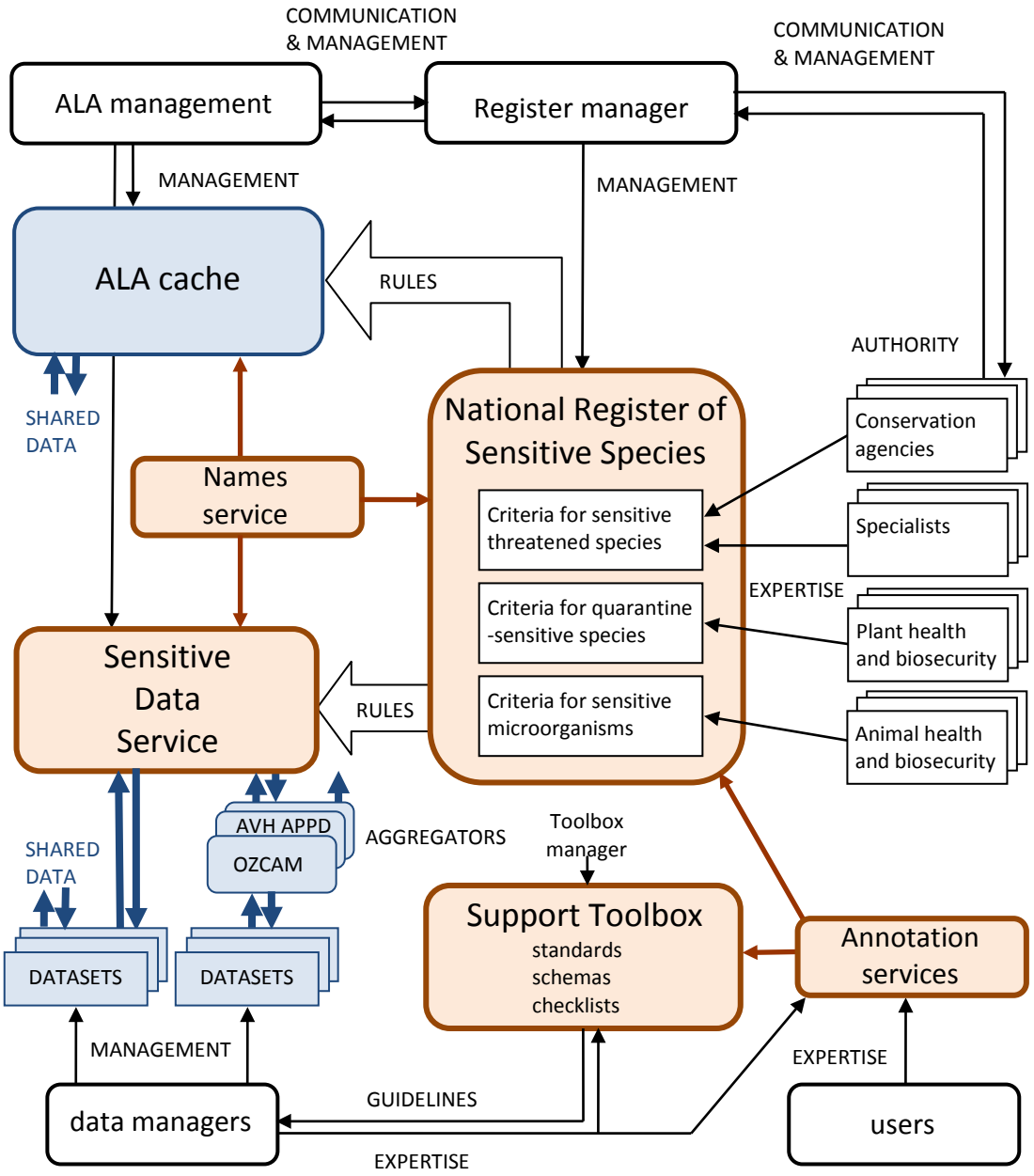

Figure 11-1 Diagram of a potential model of a Sensitive Data Service including its management structure. At its core is the National Register of Sensitive Species, which would develop and implement rules based on sensitive criteria. These rules would be used by the Sensitive Data Service to check the sensitivity of data records held by data custodians as well as data aggregators. The rules would be also used by the ALA itself before public release of data from its cache, independent of the source of the data. An independent agency would manage the National Register, and liaise with specialists and other agencies that have the expertise to define appropriate sensitivity criteria.

12 Recommendations

These recommendations will assist those wishing to share their data while respecting not only the concerns of data managers and custodians with regard to information that is sensitive, but also the concerns of those responsible for threatened species, plant pests, microorganisms, animal health and biosecurity.

The recommendations below have been collated from previous sections.

12.1 Key recommendation 1 – National Register of Sensitive Species

Develop and implement a rules-based National Register of Sensitive Species, an amalgam of three newly developed registers:

- a national register of sensitive threatened species
- a register of quarantine-sensitive species
- a register of sensitive microorganisms

1. Allow for the rules for sensitivity to differ depending on geographic area, time, species, authority, and more
2. Allow for the rules to hold true in combination

12.2 Key recommendation 2 – Sensitive Data Service

Develop and implement a facility which data managers and custodians can use to determine which of their records refer to sensitive species. This service will use as its reference standard, the National Register of Sensitive Species.

1. Include in this facility for assessment of sensitivity, the ability to rank the sensitivity, with appropriate labelling
2. Include alerts to remove or minimise personal information
3. Take steps to ensure transparency, integrity, accuracy, security, currency and care
4. Ensure that the service can readily accommodate changes to the sensitivity rules
5. Ensure that the service is readily usable by a data manager who has little support
6. Ensure that the service can be re-used both irregularly and frequently, without the need to manually re-visit the same data records multiple times

12.3 Key recommendation 3 – Sensitive Data Toolbox

Develop and implement a support service for data managers and custodians involved in sharing biodiversity data.

1. Include a pre-release checklist of potentially sensitive data types and fields, to help detect other, locally sensitive information not detected with a rules-based register
2. Include guidelines, standards and schemas for managing and sharing data
3. Include tools for managing feedback at the data source
4. Consider using a wiki-format through which data managers can assist other data managers

12.4 Supplementary recommendations

Recommendation 4 – Information infrastructure

1. Adopt a standard set of metadata fields to document sensitive data
2. Provide annotation and feedback services to clarify and correct suspect data
3. Link all names to a taxonomic names service
4. Develop and implement a hierarchical scheme to categorise possibilities when information associated with a species of interest cannot be resolved.
5. Devise a rating scheme to accommodate data of varying quality

Recommendation 5 – Threatened Species

These recommendations will assist those wishing to share their data, whilst respecting the concerns of those responsible for threatened species.

1. As a foundation, encourage each State and Territory conservation agency to contribute to the sensitivity rules applicable to that State or Territory
2. Encourage contributions from museums, herbaria and other organisations working with biodiversity
3. Allow for, and encourage, contributions from specialists and experts
4. Develop techniques to minimise duplicate records, especially for rare species

Recommendation 6 – Plant pests

These recommendations will assist those wishing to share their data, while respecting the concerns of those responsible for plant pests and biosecurity.

1. Encourage key groups associated with plant pests and biosecurity to contribute to the register of quarantine-sensitive species
2. Ensure all publicly accessible records of quarantine intercepts are labelled clearly as such
3. Create a list of alert phrases to accompany quarantine-sensitive data records

Recommendation 7 – Microorganisms

These recommendations will assist those wishing to share their data, while respecting the concerns of those responsible for microorganisms, animal health and biosecurity.

1. Encourage key organisations associated with animal health and biosecurity to contribute to the register of sensitive microorganisms
2. Incorporate the list of SSBA into the register of sensitive microorganisms
3. Incorporate the published lists of disease agents for both terrestrial and aquatic animals into the register of sensitive microorganisms. These lists may need to be moderated to account for trade sensitivities.
4. Create a list of alert phrases to accompany sensitive microorganism data records

Recommendation 8 – Executive

1. In order for the most appropriate agency to make best use of its expertise, the ALA should create a management structure under which the administration and compilation of the National Register of Sensitive Species is distinct from the administration and operation of the Sensitive Data Service

13 References

Animal Health in Australia Report, Animal Health Australia, 2008.
http://www.animalhealthaustralia.com.au/status/ahia.cfm

Atlas of Living Australia user needs analysis, John Tann, Lynda Kelly, Paul Flemons, 2008.
http://www.ala.org.au/wp-content/uploads/ALAUserNeedsAnalysisReportExtract.pdf

Guide to Best Practices for Generalising Sensitive Species Occurrence Data, Arthur Chapman and Oliver Grafton, 2008, published by GBIF. See
http://circa.gbif.net/Public/irc/gbif/pr/library?l=/webfiles/bpsensitivedata/_EN_1.0_&a=d

Health Emergency Preparedness and Response, Security Sensitive Biological Agents (SSBA), Australian Government Department of Health and Ageing. See
http://www.health.gov.au/internet/main/publishing.nsf/Content/ssba.htm
SSBA guidelines are available at:
http://www.health.gov.au/internet/main/publishing.nsf/Content/ssba-guidelines.htm
(Accessed June 2009).

Herbarium Networks Part III: A prototype for exchange of botanical specimen data to reduce duplicative effort and improve quality using a 'filtered push', James A. Macklin, Paul Morris, Robert Morris, Zhimin Wang & Maureen Kelly, SPNHC conference July 2009, Leiden, The Netherlands.
http://www.spnhc2009.org/

Listed Key Threatening Processes, Australian Government, Department of the Environment, Water, Heritage and the Arts,
http://www.environment.gov.au/cgi-bin/sprat/public/publicgetkeythreats.pl
(Accessed June 2009).

National Plant Health Status Report 07/08, Plant Health Australia, 2009.
http://www.planthealthaustralia.com.au/corporate_documents/display_document.asp?category=7&ID=654

Natural Values Atlas, Tasmania Department of Primary Industries and Water. See
http://www.dpiw.tas.gov.au/inter.nsf/WebPages/LJEM-6TV6TV?open

NatureMap, Western Australia Department of Conservation
http://naturemap.dec.wa.gov.au

NBN data exchange principles and their rationale, National Biodiversity Network Trust, UK. See http://www.nbn.org.uk/About/The-Organisation/Data-exchange-principles.aspx
(Accessed June 2009).

NSW DECCW Sensitive species data policy, 2009. See
http://www.environment.nsw.gov.au/policiesandguidelines/SensitiveSpeciesPolicy.htm

OECD Principles and Guidelines for Access to Research Data from Public Funding, OECD, 2007. http://www.oecd.org/dataoecd/9/61/38500813.pdf

OECD Declaration on Access to Research Data from Public Funding, Science, Technology and Innovation for the 21st Century. Meeting of the OECD Committee for Scientific and Technological Policy at Ministerial Level, 29-30 January 2004 - Final Communique. http://www.oecd.org/document/0,2340,en_2649_34487_25998799_1_1_1_1,00.html (Accessed June 2009).

PaDIL – Pests and Diseases Image Library. See http://www.padil.gov.au/

Quarantine Proclamation 1998, Australian Government http://www.comlaw.gov.au/comlaw/Legislation/LegislativeInstrumentCompilation1.nsf/0/E881571745BA7082CA2575AC000ED232?OpenDocument

SPRAT – Species Profile and Threats database, Australian Government Department of the Environment, Water, Heritage and the Arts. See http://www.environment.gov.au/cgi-bin/sprat/public/sprat.pl

Wildlife Online, Queensland Environment Protection Agency. See http://www.epa.qld.gov.au/nature_conservation/wildlife/wildlife_online/

14 Acknowledgements

This report has only been made possible by those who were generous in contributing their ideas and time. Through the meetings and conversations over the past months have we been able to understand the many intricacies of the sensitivity of biodiversity data.

We would like to thank those people from the herbaria and museums in every State and Territory in Australia, as well as those that work with microorganisms. They were able to convey the details of their world to us in terms of anecdotes and examples from laboratories and collections, and had a deep understanding of specific subjects that went well beyond the walls of their institutions.

We would like to thank those hospitable entomologists in the departments of primary industries, who conveyed the importance and difficulties of working with current and historical information about plant pests and diseases.

We would like to thank those people working with biodiversity data and its management in each of the Commonwealth, State and Territory conservation agencies. Interestingly all groups had different approaches to the subject, yet all had a strong sense of what was important in the interests of the wildlife with which they worked.

We would like to thank those members of the Office of Chief Plant Protection Officer, Plant Health Australia, Department of Health and Ageing and the Office of the Chief Veterinary Officer, who were able to explain so clearly the complexities of a world of pests and diseases and biosecurity.

We would like to thank those data managers at many institutions across the country who were able to tell us clearly about the difficulties they faced when working with sensitive data.

We would like to give a special thanks to Margaret Cawsey of ANWC, Piers Higgs of Gaia Resources, Robert Morris of SAM, Ian Naumann of APPD, and Alice Wells of ABRS who made many helpful suggestions at crucial points in this project.

And, of course, a big appreciation to Donald Hobern for his smart direction, good humour and support.

John Tann
Paul Flemons

15 Appendix A – List of agencies and organisations consulted

Representatives of these agencies contributed to the content of this report through personal and telephone interviews and discussions.

AAD – Australian Antarctic Division
AAHL – Australian Animal Health Laboratory
ABIN – Australian Biosecurity Intelligence Network
ABRS – Australian Biological Resources Study
ACAP – Agreement on the Conservation of Albatrosses and Petrels
Australian Collection of Microorganisms
ACT Parks Conservation and Lands
AMRiN – Australian Microbial Resources Information Network
ANBG – Australian National Botanic Gardens
ANFC – Australian National Fish Collection
ANIC – Australian National Insect Collection
ANWC – Australian National Wildlife Collection
APPD – Australian Plant Pest Database
AQIS – Australian Quarantine Inspection Service
Australian Biosecurity CRC for Emerging Infectious Disease
Australian Museum
AVH – Australia's Virtual Herbarium
CCAMLR – Commission for the Conservation of Antarctic Marine Living Resources
CHACM – Council of Heads of Australian Collections of Microorganisms
CHAEC – Council of Heads of Australian Entomological Collections
CHAFC – Council of Heads of Australian Faunal Collections
CHAH – Council of Heads of Australian Herbaria
CMA – Catchment Management Authority
CSIRO – Commonwealth Scientific and Industrial Research Organisation
DAFF – Department of Agriculture, Fisheries and Forestry
DAFF Australian Wildlife Health and Environment
DAFF Invasive Marine Species
DAFF Office of the Chief Plant Protection Officer
DAFF Office of the Chief Veterinary Officer
DAF-WA – Department of Agriculture and Food, Western Australia
Department of Health and Ageing
Desert Knowledge CRC
DEWHA – Department of the Environment, Water, Heritage and the Arts
DEWHA ANHAT – Australian Natural Heritage Assessment Tool
DEWHA ERIN – Environmental Resources Information Network
DEWHA Genetic Resource Management
DEWHA SPRAT – Species Profile and Threats Database
DEWHA Threatened Species
FCIG – Faunal Collections Informatics Group
Forestry Tasmania
Gaia Resources
Hunter - Central Rivers CMA, Catchment Management Authority
MAGNT – Museum and Art Gallery of the Northern Territory
Museum Victoria
NAQS – Northern Australia Quarantine Strategy
National Herbarium of New South Wales

National Herbarium of Victoria
NSW DECCW – Department of Environment, Climate Change and Water
NSW DECCW Wildlife data
NSW DPI – Department of Primary Industries
NSW DPI Fisheries
NSW DPI Entomology
NSW DPI Plant Biosecurity
NT DRDPIFR – Department of Regional Development, Primary Industry, Fisheries and Resources
NT DRDPIFR Biosecurity and Product Integrity
NT DRDPIFR Entomology
NT Herbarium
NT NRETAS – NT Natural Resources, Environment, The Arts and Sport
NT NRETAS Biodiversity Conservation
OZCAM – Online Zoological Collections of Australian Museums
PHA – Plant Health Australia
QLD DPI – Queensland Primary Industries and Fisheries
QLD DPI Weeds
QLD EPA – Environmental Protection Agency
Queensland Museum
QVMAG – Queen Victoria Museum and Art Gallery
Royal Botanic Gardens, Sydney
SA DEH – South Australia Department of Environment and Heritage
South Australian Museum
Sutherland Shire Council
Tasmania DPIPWE – Department of Primary Industries, Parks, Water and Environment
Tasmanian Herbarium
Tasmanian Museum and Art Gallery
TRIN – Taxonomy Research & Information Network
University of Adelaide
University of Queensland
VIC DPI – Department of Primary Industries
VIC DSE – Department of Sustainability and Environment
WA DEC – Department of Environment and Conservation
WA DEC NatureMap
Western Australian Museum

16 Colophon

Front cover design by Alison Murphy, agraphique.

Front cover photographs

Donated eggs
Ill-gotten gains. Museums may be the best place to hold specimens of rare or extinct species that may have been collected under unknown or questionable legality. Donations are made on the understanding that personal details will not be disclosed.

Ground Parrot
With occasional reports of being heard in urban areas, the vulnerable Ground Parrot, *Pezoporus wallicus wallicus,* has locations of sightings reduced in precision to minimise disturbance.

Grey Falcon
Sensitivities to the Grey Falcon *Falco hypoleucos,* a threatened species, vary across its range. In New South Wales, the location of nesting sites of the Grey Falcon is reduced in precision, while in South Australia there are no restrictions on public access to information about where the Grey Falcon has been found. Photo Tony Morris. http://www.flickr.com/photos/tonymorris/2098388135/

Wollemi Pine
With fewer than 100 trees growing in the wild and those susceptible to exploitation and disease, the location of the biologically significant Wollemi Pine *Wollemia nobilis,* is not disclosed.

Morning Glory
Landholders with weeds consider the location of their properties to be confidential. There are legal obligations and compliance issues surrounding weeds, and fear of prosecution.

Stuttering Frog
The Stuttering Frog, *Mixophyes balbus,* one of the larger of the Australian ground frogs, is now believed to be extinct in Victoria. The precision of its location is reduced to 1 km in NSW. Photo GA Hoye.