

Overview of the ALA

The national biodiversity data infrastructure

Peter Brenton | Manager, Data Collection & Community Engagement

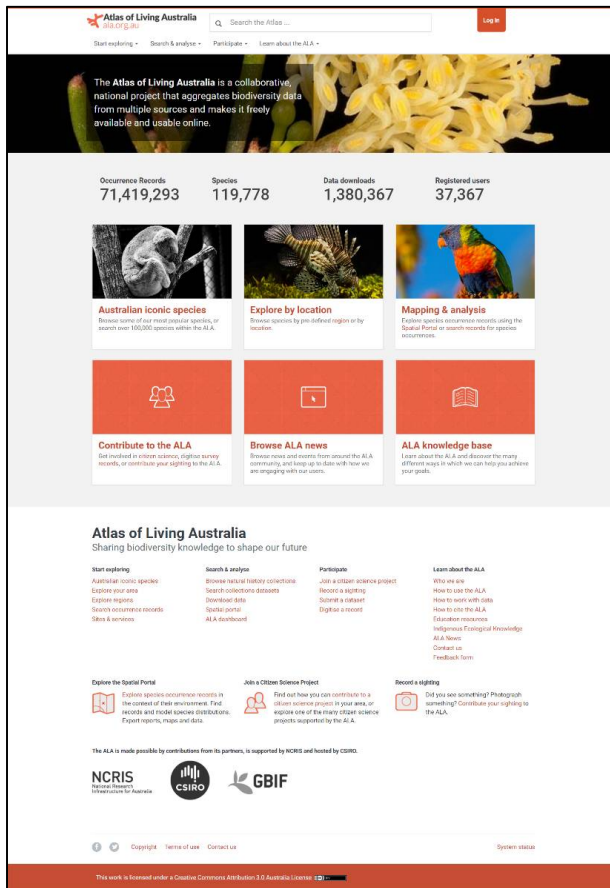
January 2019



The ALA is made possible by contributions from its many partners. It receives support through the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS) and is hosted by CSIRO.

What is the ALA?

Australia's aggregation & access facility for biodiversity data



The Atlas of Living Australia is a collaborative, national project that aggregates biodiversity data from multiple sources and makes it freely available and usable online.

Occurrence Records: 71,419,293
Species: 119,778
Data downloads: 1,380,367
Registered users: 37,367

Australian iconic species
Browse some of our most iconic species, or search over 10,000 species within the ALA.

Explore by location
Browse species by pre-defined regions or by location.

Mapping & analysis
Filter species occurrence records using the Spatial Portal or search records for species occurrences.

Contribute to the ALA
Get involved in citizen science, upload survey records, or contribute your sighting to the ALA.

Browse ALA news
Browse news and events from around the ALA community, and keep up to date with 'hear what we are engaging with our users'.

ALA knowledge base
Learn about the ALA and discover the many different ways in which we can help you achieve your goals.

Atlas of Living Australia
Sharing biodiversity knowledge to shape our future.

Start exploring
Australian iconic species
Explore regions
Search occurrence records
View & records

Search & analyse
Browse natural history collections
Search collections datasets
Download data
Spatial portal
ALA dashboard

Participate
Join a citizen science project
Record a sighting
Report a sighting
Digital a record




Learn about the ALA
Who we are
How to use the ALA
How to work with data
How to cite the ALA
Education resources
Indigenous Ecological Knowledge
ALA News
Contact us
Feedback form

Explore the Spatial Portal
Explore species occurrence records in the context of their environment. Find records and model species distributions. Export reports, maps and data.

Join a Citizen Science Project
Find out how you can contribute to a citizen science project in your area, or explore one of the many citizen science projects supported by the ALA.

Record a sighting
Did you see something? Photograph something? Contribute your sighting to the ALA.

The ALA is made possible by contributors from its partners, is supported by NCRIS and hosted by CSIRO.

Copyright Terms of Use Contact us System status

This work is licensed under a Creative Commons Attribution 3.0 Australia License

- One of several facilities funded by the Australian Government for **national research infrastructure**



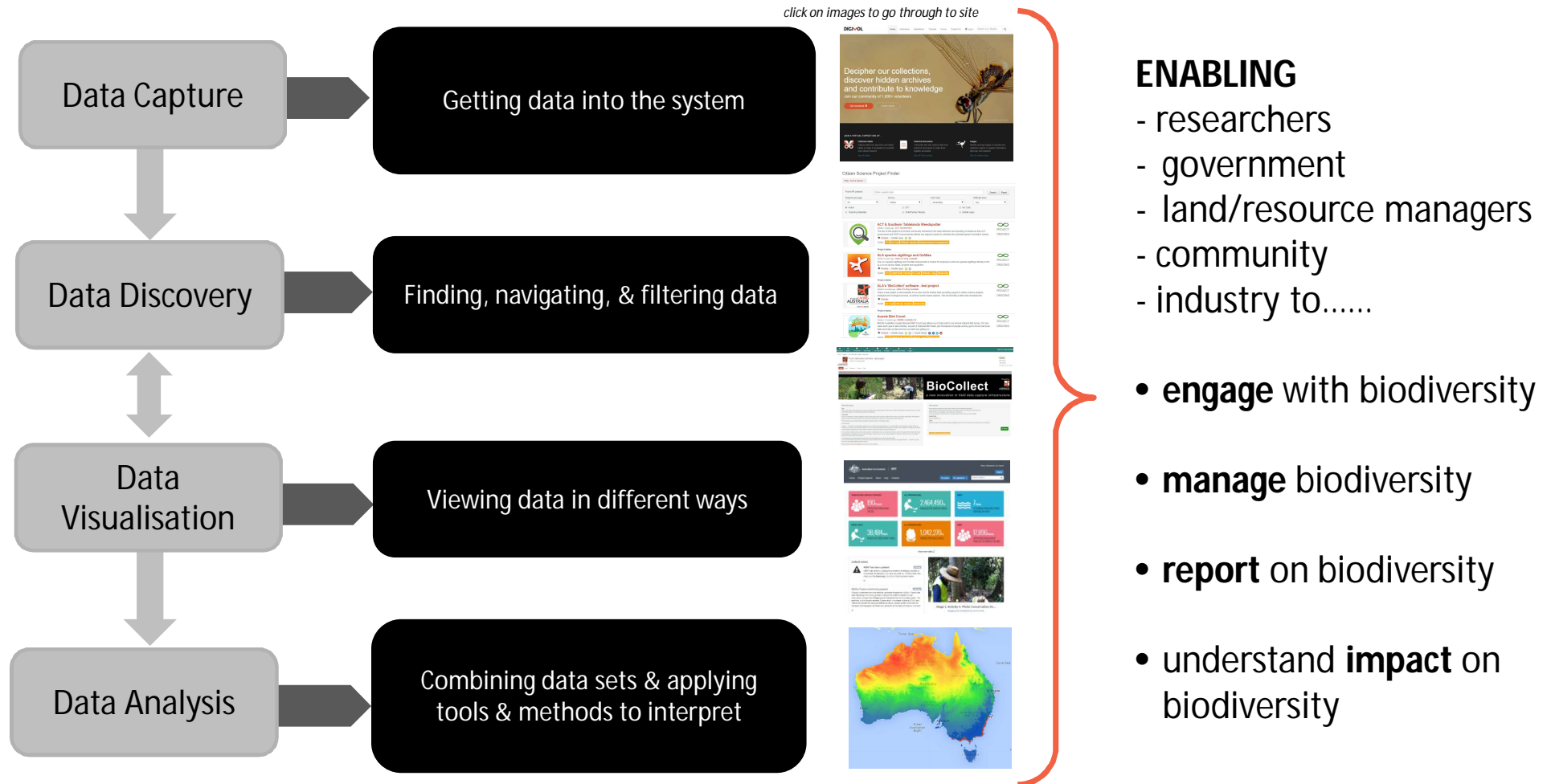
- Hosted by



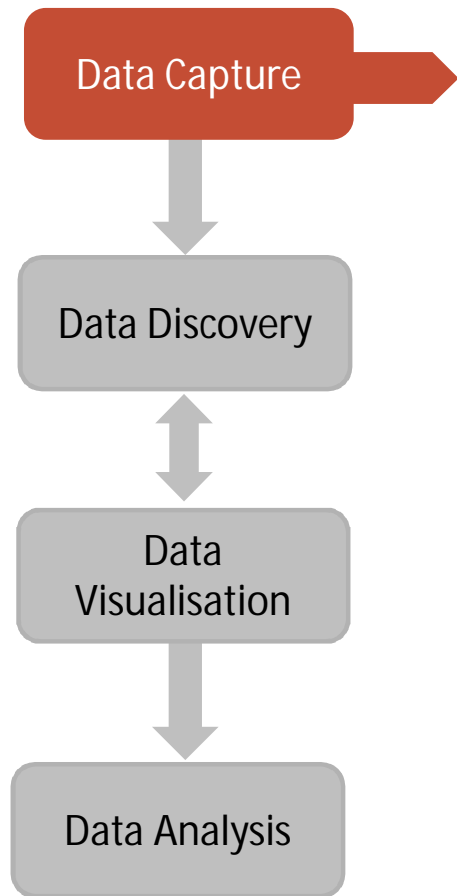
- Enables

- more **effective & efficient** scientific research
- expanded research **opportunities**
- informing **policy & management**
- **community** participation & connection
- **education** resource

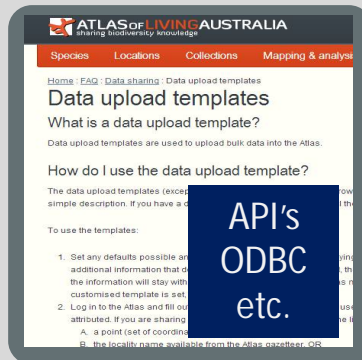
ALA - tools for biodiversity



ALA Tools – data capture



Direct Data Uploads



ATLAS of LIVING AUSTRALIA
Species Locations Collections Mapping & analysis

Home / FAQ / Data sharing / Data upload templates

Data upload templates

What is a data upload template?

Data upload templates are used to upload bulk data into the Atlas.

How do I use the data upload template?

The data upload templates (except simple description, if you have a...

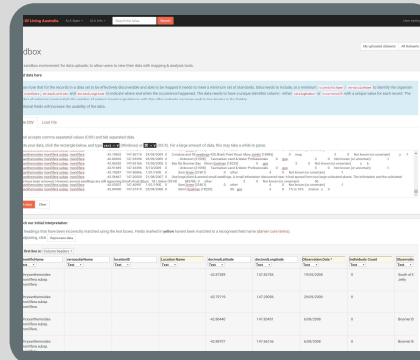
To use the templates:

1. Set any defaults possible and additional information that the information will stay with your customised template is set.
2. Log in to the Atlas and fill out the information.

A. a point (set of coordinates)
B. the locality name available from the Atlas gazetteer. OR

API's
ODBC
etc.

ALA Sandbox



ALA Sandbox

Species	Location	Collection	Mapping & analysis
...

BioCollect + Mobile



BioCollect

Project discovery +
General field data
capture
cloud hosting service

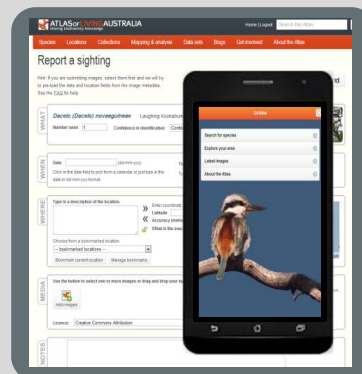
Seasonal Calendars



INDIGENOUS SEASONAL CALENDARS

Recording cultural & phenologic perspectives

ALA Sightings + Mobile



ATLAS of LIVING AUSTRALIA

Report a sighting

Search for species

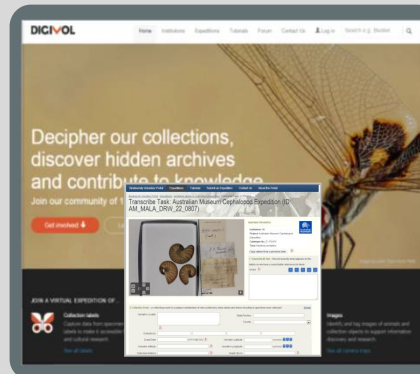
Location

When

How

What

DigiVol – Crowd Digitisation



DIGIVOL

Decipher our collections, discover hidden archives and contribute to knowledge

MERIT



Australian Government
Department of the Environment

MERIT

AUS govt sponsored
environmental
interventions

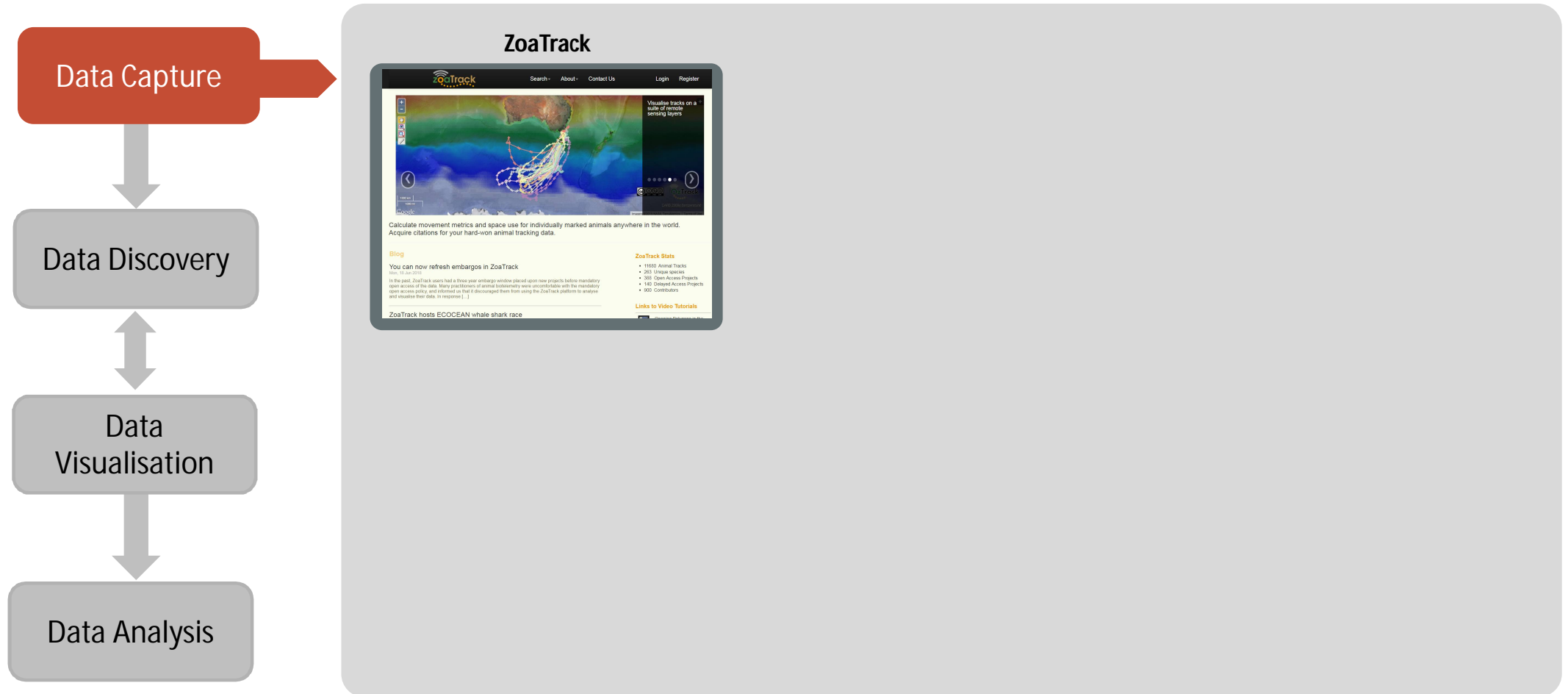
Profiles Tool



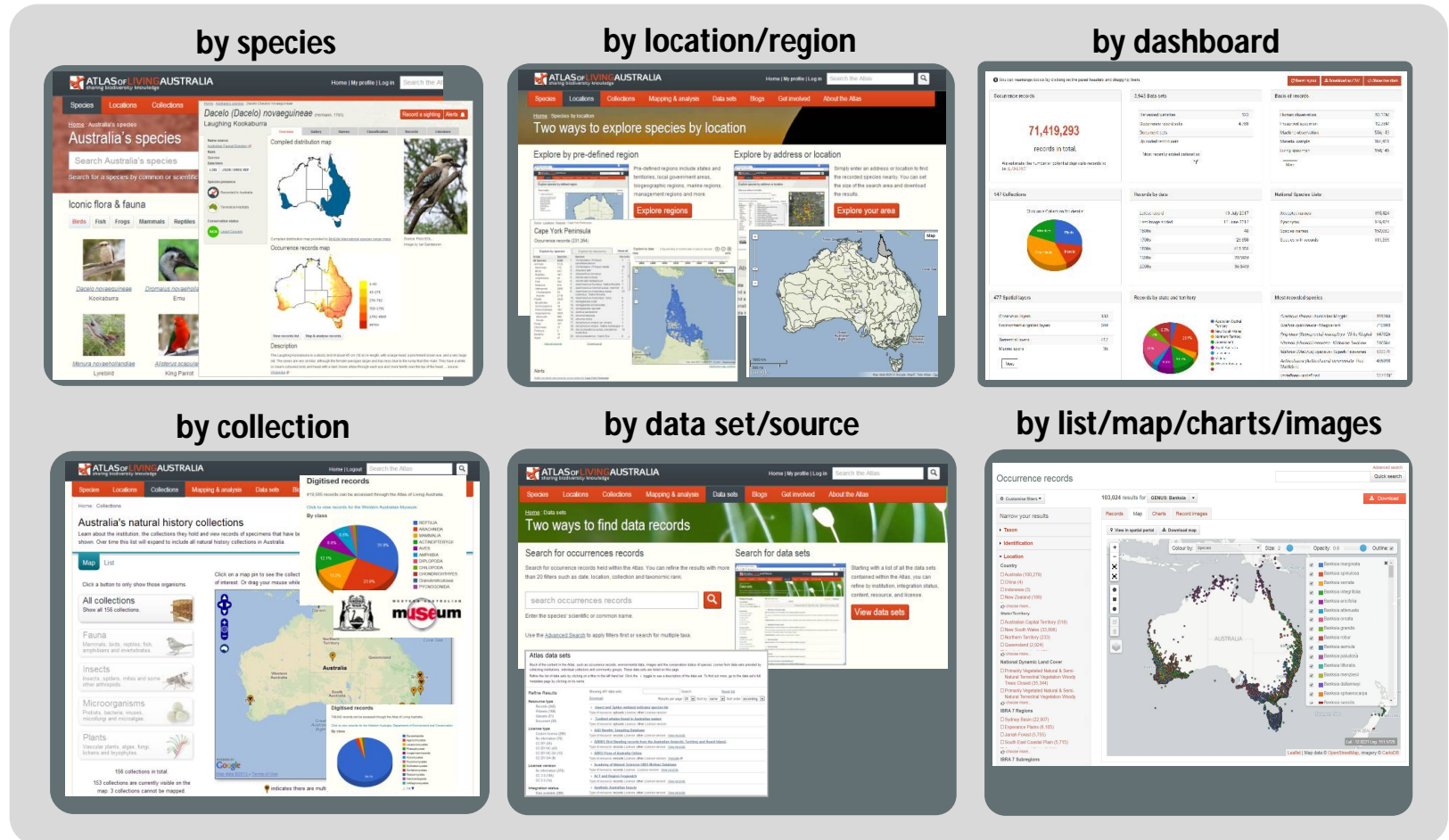
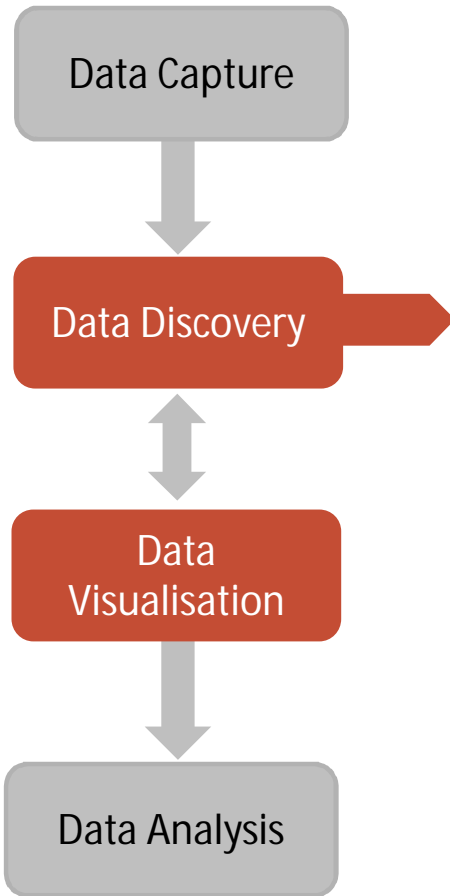
Profiles Tool

Species profile content authoring & management

ALA Tools – data capture



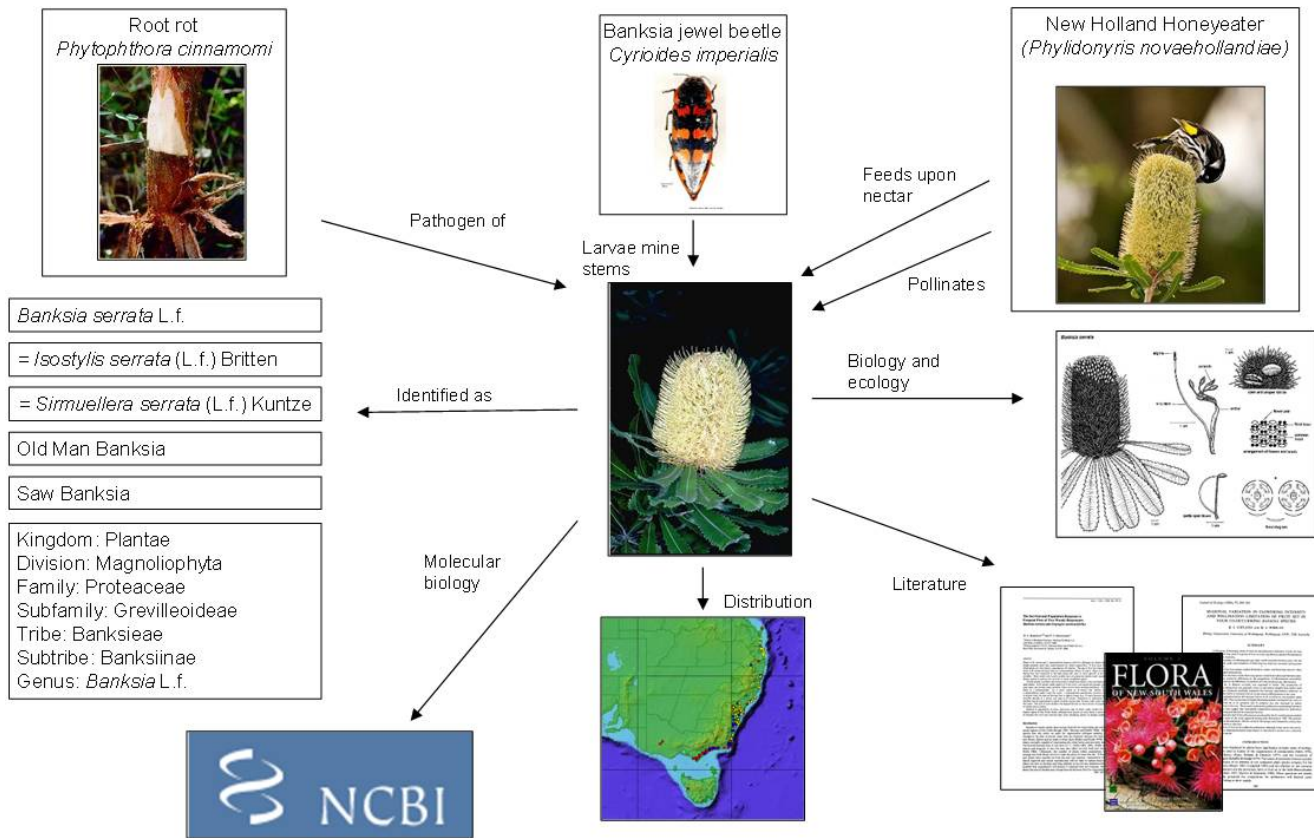
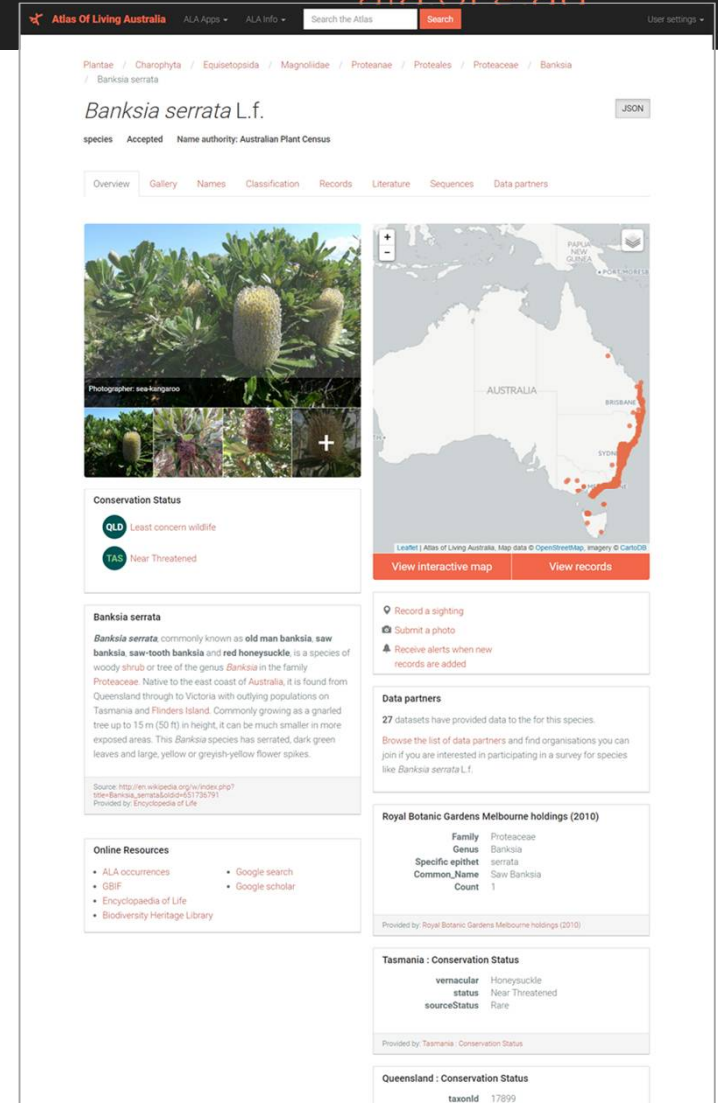
ALA Tools – navigating the data



The screenshots demonstrate various navigation tools:

- by species:** Shows a search for *Dacelo (Dacelo) novaeguineae* with a distribution map and species details.
- by location/region:** Shows options to explore by pre-defined region (e.g., Cape York Peninsula) or by address/location.
- by dashboard:** Provides a high-level overview with 71,419,293 records in total, including charts for 148 Collections, 477 Spatial Layers, and 181 Occurrences.
- by collection:** Displays 'Australia's natural history collections' with a pie chart showing the distribution of digitised records across various institutions like the Australian Museum and Queensland Museum.
- by data set/source:** Offers ways to find data records, including searching for occurrences and data sets, with filters for species and collection.
- by list/map/charts/images:** Shows a detailed view of occurrence records on a map of Australia, with a legend for different species and a list of records on the right.

ALA Tools - biodiversity information

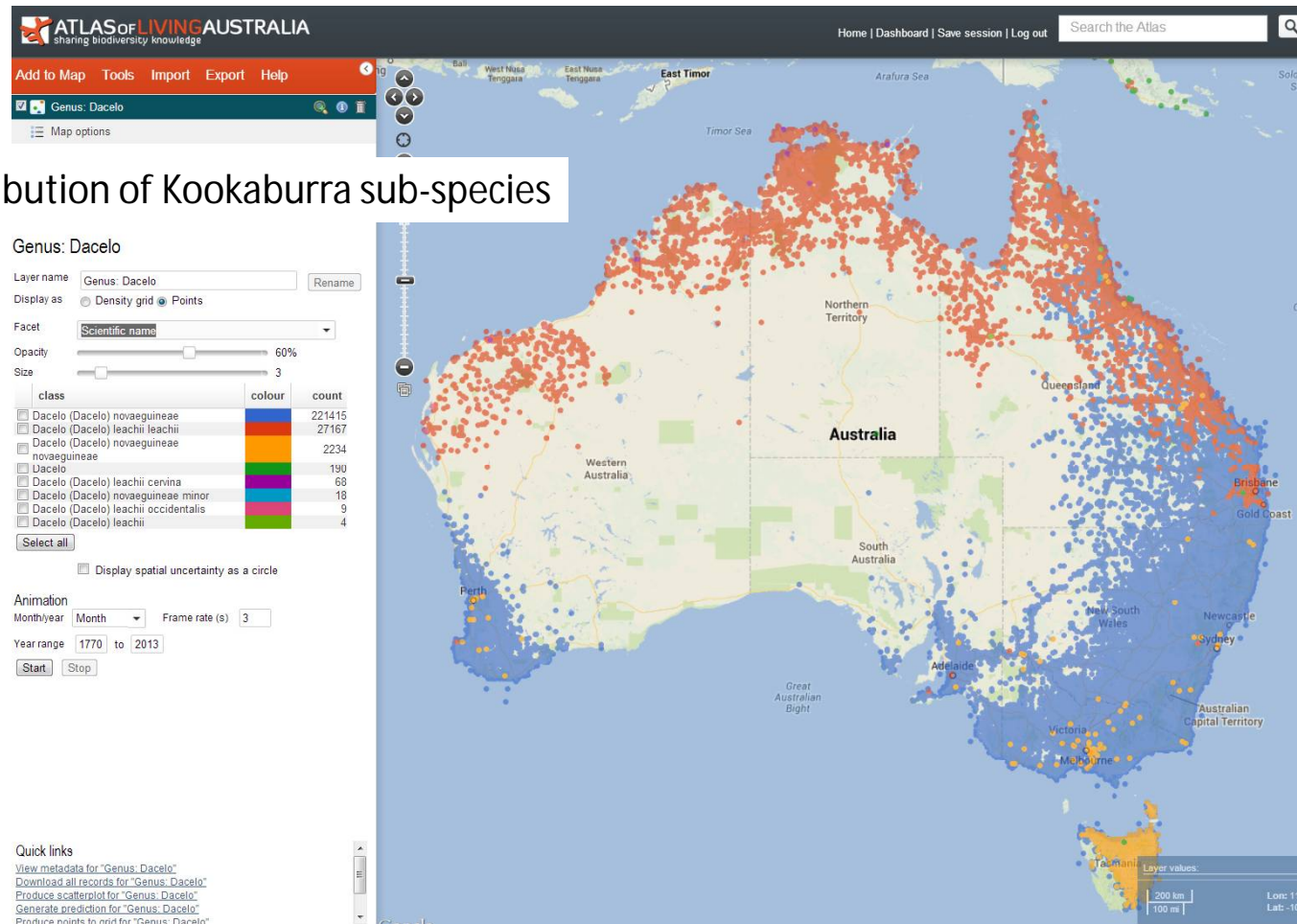



The screenshot shows the species page for *Banksia serrata* L.f. on the Atlas of Living Australia website. The page includes:

- Navigation:** Home, ALA Apps, ALA Info, Search the Atlas, Search, User settings.
- Taxonomy:** Plantae / Charophyta / Equisetopsida / Magnoliidae / Proteales / Proteaceae / Banksia / Banksia serrata.
- Species Information:** *Banksia serrata* L.f., Accepted, Name authority: Australian Plant Census.
- Overview:** Overview, Gallery, Names, Classification, Records, Literature, Sequences, Data partners.
- Media:** A gallery of photos showing the plant in its natural habitat, with a map of Australia highlighting its distribution.
- Conservation Status:**
 - QLD: Least concern wildlife
 - TAS: Near Threatened
- Description:** *Banksia serrata* commonly known as old man banksia, saw banksia, saw-tooth banksia and red honeysuckle, is a species of woody shrub or tree of the genus *Banksia* in the family Proteaceae. Native to the east coast of Australia, it is found from Queensland through to Victoria with outlying populations on Tasmania and Flinders Island. Commonly growing as a gnarled tree up to 15 m (50 ft) in height, it can be much smaller in more exposed areas. This *Banksia* species has serrated, dark green leaves and large, yellow or greyish-yellow flower spikes.
- Online Resources:**
 - ALA occurrences
 - Google search
 - GBIF
 - Google scholar
 - Encyclopaedia of Life
 - Biodiversity Heritage Library
- Other Data:** Royal Botanic Gardens Melbourne holdings (2010), Tasmania Conservation Status, Queensland Conservation Status.

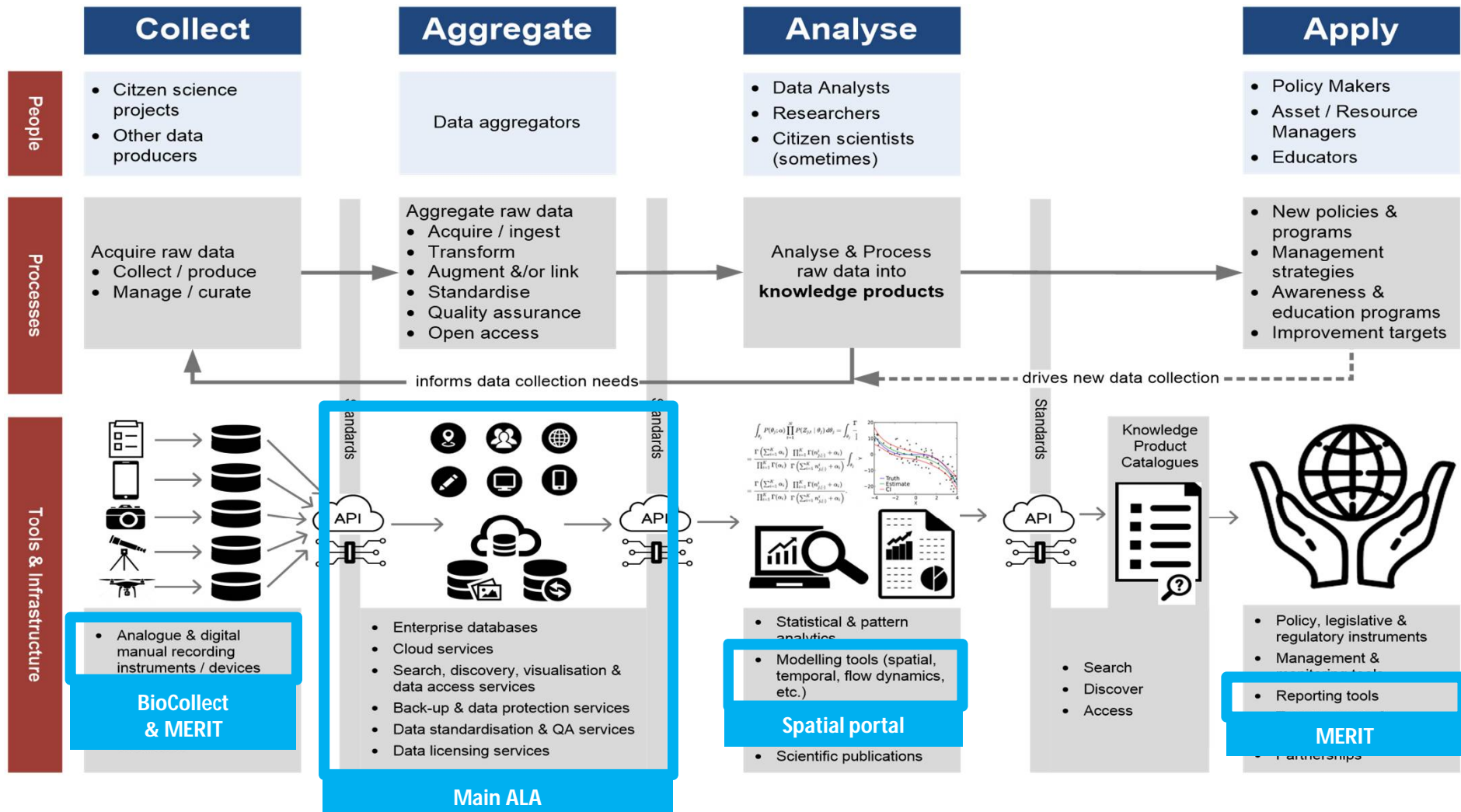
'Big data' analytics

Distribution of Kookaburra sub-species



Before the ALA a map like this was not possible!

ALA Tools – the information supply chain



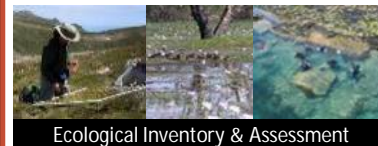
Citizen Science

Public Local & State Govt. NRM's Researchers NGOs Schools



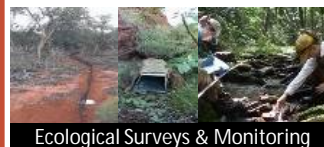
Applied Ecology - Industry

Planning & Development NRM's Consultants Mining & Extractives



Ecological Research

Ecological Consultants Government Researchers Universities



Natural Resource Management

Land holders/managers NGOs Govt. NRM's Developers

ACSA Public



Collaborative Species Distribution Modelling Program



Aims to produce scientifically robust and reproducible Species Distribution Models for transparent decision making across and between governments (States and Commonwealth).

Builds on:

- Advanced national cloud infrastructure
- National expertise
- State & Commonwealth Partnerships
- The Biodiversity and Climate Change Virtual Laboratory (BCCVL)
- The Atlas of Living Australia (ALA)

Program Partners:

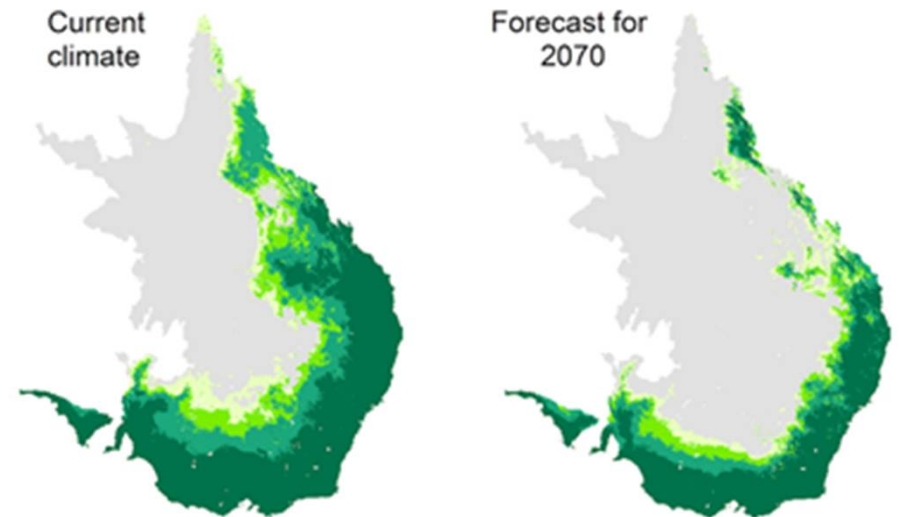
- Three government departments
 - NSW OEH
 - QLD DES
 - DoEE
- Atlas of Living Australia (CSIRO)
- Biodiversity and Climate Change Virtual Laboratory (Griffith University)

Species Distribution Models (SDMs)?

SDMs quantify the response of species to environmental and habitat variables including climate in terms of their distribution, abundance and functional traits.

SDMs can be used to identify locations and regions with potentially suitable environmental and habitat conditions for a species to survive.

They can also be used to assess how species may respond to projected future climate or habitat loss, guide new survey efforts, and assess overlap with threatening processes.



The 9-month pilot program is structured in three complimentary pillars:

Pillar 1: Expert Committee

- Establish a panel of interdisciplinary scientific and operational experts
- Develop a set of best-practice methods for SDMs for particular use cases (e.g. where few or no observation data are available)
- Establish a method to update standardised method based on new evidence
- Provide direction and advice to the Technology Platform to integrate standards

Pillar 2: Technical Platform

- 3 key products:
 - CSDM Spatial Portal
 - CSDM Modelling Platform
 - CSDM Model Explorer
- Building on existing national infrastructure capabilities

Pillar 3: Training, Engagement and Sustainability

- Support site for the CSDM system with user guides
- Engagement and sustainability:
 - Roadmap for Phase 2
 - Outreach to other jurisdictions and departments for Phase 2
 - Business Model for ongoing funding and operations

Aspects of data quality

Important considerations in data collection & curation

Peter Brenton | Manager, Data Collection & Community Engagement

January 2019

Concepts of Data Quality

What does “data quality” mean to you?

- **Accuracy** How closely a data value for a property represents the real/true value of the property being measured
- **Precision** The range of uncertainty in the measured resolution of a property value.
- **Completeness** Whether all of the required data is provided for all records.
- **Consistency** Whether methods have been applied in the same way for all records.
- **Validity** Whether records are factually correct.
- **Timeliness** Whether records are temporally relevant to the intended use (eg. Previously present, now absent).

Data Quality (spatial example)

Precision vs. Accuracy

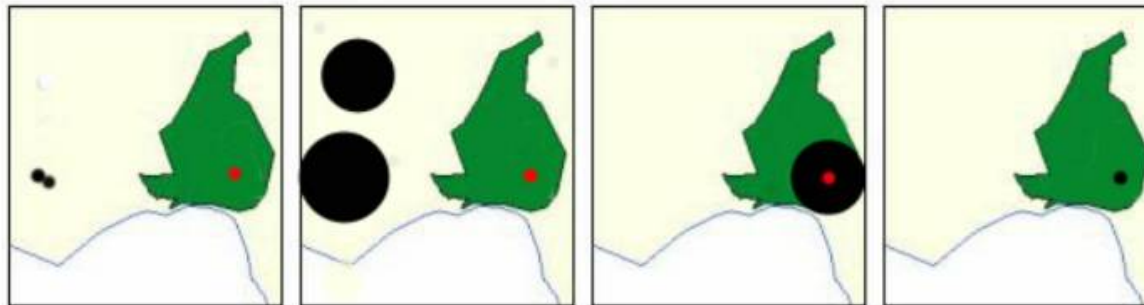


Fig. 1. Shows the differences between accuracy and precision in a spatial context. The red spots shows the true location, the black spots, represent the locations as reported by a collector.

- a. High precision, low accuracy.
- b. Low precision, low accuracy showing random error.
- c. Low precision, high accuracy.
- d. High precision and high accuracy.

Principal also applies to:

temporal; **taxonomic** and **non-taxonomic** parameters

Concepts of Data Quality

No data is perfect

- Accuracy
- Precision
- Completeness
- Consistency
- Validity
- Timeliness

Data quality is about “**fitness for use**”, not personalities or reputations

Is the ‘quality’ of the data adequate for the intended use?

How can I know whether the data is valid, complete and/or consistently collected/treated?

Tips:

Metadata is critical for assessing fitness for use.

PROJECT

Description of Project

PPSR-Core PMM

SURVEY

DATASET

DATA (records)

SURVEY (dataset metadata)

Description of Dataset

PPSR-Core DMM

+

DATA (records)

DATA AGGREGATORS

Occurrence Records	Species	Data downloads	Registered users
84,447,583	124,758	1,726,165	46,701

GBIF | Global Biodiversity Information Facility

www.gbif.org

Impact:

Used by ...

- researchers
- government
- land/resource managers
- community
- industry to ...

- engage with biodiversity
- manage biodiversity
- report on biodiversity
- understand impacts on biodiversity
- International monitoring
- Pan-national conservation & biosecurity management

The credibility conundrum

What's needed to improve data credibility?

Data Set Level

- Trust
 - Where has it come from?
 - How was it collected?
(methodology, equipment, etc.)
 - Who collected it?
 - How accurate is it?
 - QA, curation, validation & management processes?
- Known biases in the data and conditions on use?
- How to cite/reference?

Data Record Level

- Accuracy & precision
 - Taxonomic
 - Spatial
 - Temporal
 - Non-taxonomic attributes
- Consistency of method application
- Standardised formats
- Comprehensive & complete

Building trust in data

Solutions

Metadata describing the dataset

- Guided
- Standardised
- Comprehensive
- Accessible
- Simple
- Methods (measuring & processing)
- QA/QC/curation/validation processes

Survey Info Visibility Alert Survey Form Species Locations Publish

Published

Step 1 of 7 - Describe the survey

Name: * Macadamia Descendants

Description: * Citizen scientist data collection on macadamia trees across south east Queensland and northern New South Wales

Survey attribution/citation: * Healthy Land & Water
Macadamia Conservation Trust
University of Queensland
Southern Cross University

Survey method type: * Systematic method-based survey

Survey/sampling method used: * This is the name of the published systematic survey methodology (eg. 2ha 20 minute bird survey)

Description of survey/sampling method: *
Community engagement
Tree identification and recording
Collection of test samples

Web site of survey/sampling method: *
<https://biocollect-test.ala.org.au/acsa/project/index/035af93e-2fa6-4b00-8152-5b85ec7a2af0?hub=acsa>

Support document of survey/sampling method: *
+ Attach Document


Start Date: * 06-11-2017

End Date: *

Allow public users to enter data: *

Comments on records allowed: *

Data Sharing License: * Creative Commons Attribution-NonCommercial (Australia)

Attach logo:  CC BY-NC 3.0 AU

Legal Custodian of the Dataset: * Healthy Land and Water

Dataset reliability tags: *
Spatial accuracy confidence *
 High
 Moderate
 Low
Temporal accuracy confidence *
 High
 Moderate
 Low
Species identification confidence *
 High
 Moderate
 Low
Non-tax *
 High
 Moderate
 Low

Data quality assurance method: *
 Data owner curated
 Crowd-sourced record verification
 System supported data attribute configuration
 Not applicable
 Subject matter expert record verification
 Record verification
 No Data

Data access method: * Limited structured raw data access in this system - via request (subject to embargo)

Data access external URL: * <https://biocollect-test.ala.org.au/acsa/project/index/035af93e-2fa6-4b00-8152-5b85ec7a2af0?hub=acsa>

Data quality assurance description: * This section describes the methods used to assure data quality, as part of processes and raw data curation processes.

Data biases and constraints: * This section describes any biases in the raw data which may affect interpretation, spatial issues, temporal factors, methodology driven biases, etc. The aim of the raw data is any factors which may affect data quality and/or data interpretation, constraints or assessment of fitness for purpose.

Is data management policy documented? * Yes

Description of data management policy: * This section describes how the raw data is managed, curated and made available for use, including the terms of use, licensing conditions, etc. This section is not required if a document artifact is uploaded for this purpose.

Link to data management policy description: * <https://biocollect-test.ala.org.au/acsa/project/index/035af93e-2fa6-4b00-8152-5b85ec7a2af0?hub=acsa>

Document describing data management policy: * <https://ecodata-test.ala.org.au/uploads/2018-08/TEST%20DOCUMENT.docx>

+ Remove Document

Note: Changes made to this tab will be reflected in search results after 24 hours.

Description

Methods

Temporal Range

License & Ownership

QA Tags

QA Methods


Data Access Methods

Data Constraints & Biases

Data Mngt Plan

Building trust in data

About Blog Resources **Surveys** Data Admin




Individual sighting

General template for individuals to log single sightings of flora, fauna or fungi into the Atlas of Living Australia

Status: Active, In progress Show more

Spatial accuracy confidence - Moderate **Species identification confidence - Moderate** **Temporal accuracy confidence - High** **Non-taxonomic data accuracy - High**

Number of data recording events in the dataset (count of activity records)	91,686
Number of species records in the dataset (count of species records)	91,588
Last Updated:	24 February, 2019
Public access to data	True
Survey attribution/citation:	Atlas of Living Australia. (2019) Individual sighting dataset download. Retrieved from https://biocollect.ala.org.au/bioActivity/projectRecords/f813c99c-1a1d-4096-8eeb-cbc40e321101 , 24/02/2019, 04:51.
Survey method type:	Opportunistic/ad-hoc observation recording
Survey/sampling method used:	Opportunistic/ad-hoc observation recording
Data Sharing License:	 Creative Commons Attribution-NonCommercial (Australia)
Legal Custodian of the Dataset:	Atlas of Living Australia
Data quality assurance description	This dataset comprises opportunistic ad-hoc sightings contributed by members of the public via both web app and mobile apps. Both the web and mobile apps employ in-built software mechanisms to maximise recording accuracy, but there are still opportunities for human decisions to cause data inaccuracies, particularly in respect to species identifications. These are primary data which are not directly validated or curated in this application before being harvested into the Atlas of Living Australia (ALA). Therefore, whilst most of the data is likely to be reliably accurate and of high quality, they do contain some variability in the accuracy of species identifications and also potentially temporal and spatial accuracy. Once these records are harvested into the ALA, they can have accuracy issues flagged via the Atlas annotation service, through which the record owner is notified and can correct any errors in this application. Corrected records are then re-harvested to update the Atlas record.
Data access external URL	https://biocache.ala.org.au/occurrences/search?q=%3A*&fq=%2Bdata_resource_uid%3A364%29
Data biases and constraints:	Opportunistic observations are presence only occurrence records. Being opportunistic, these records are also temporally and spatially sporadic in nature and are not made as a result of deliberate survey effort in a particular location. Owing to the unvalidated nature of these data it is recommended that users pre-qualify the data and remove records which are unsuitable for the required purpose from the downloaded dataset before using it.
Is the data management policy documented?	No

Important to capture survey (dataset) metadata and make accessible to users.

- Builds confidence
- Informs 'fitness for use' decisions
- Authoritative
- DQ accuracy tags
- Citation/attribution
- Data accessibility & methods
- Licensing & ownership
- QA treatments
- Survey methodology
- Biases, constraints & usage requirements

The credibility conundrum

What's needed to improve data credibility?

- Meta-attributes describing the data
 - Accuracy/precision tags
- Attribute level standards (*PPSR-Core ODM*)
- Database managed rules, validations, algorithms, data types, etc.
- Configuration constraints to manage applicable range/scope – spatial, temporal, taxonomic, non-taxonomic.
- Data curation & validation mechanisms

BDBSA Fauna Observations Data Sheet

Location

Select a location

Site Visit Details

09-03-2019

Reliability of the entered date: Accuracy to day

Visit comments

Observer name: Peter Drenton

Data release authorisation: approved

Latitude: -34.24355311941496

Longitude: 140.28364852327338

Source of coordinates: Google maps

Location notes

Saved locations: -- saved locations --


Matched locality: Wachtells Lagoon, Sheward Road, Coodloga, The Berrri Barmers Council, South Australia, 5248, Australia

Location accuracy: Please select

Reserve name: Please select

Location confidence: U_unverified

Species Observations

Species type	Species name	Identification confidence	Number of individuals seen	Observation / collection method	Field voucher #	Photo	Species comments	Verified by (office use only)
R_	Hand on ballam (grey)	U_unver	present	Seen		 Or, drop images here		

+ Add a row

Upload data for this table

Thank you

Peter.Brenton@csiro.au



Sharing biodiversity knowledge to shape our future



<https://biocollect.ala.org.au>



Solving the greatest challenges through
innovative science and technology