Atlas of Living Australia Report to Partners

October 2011

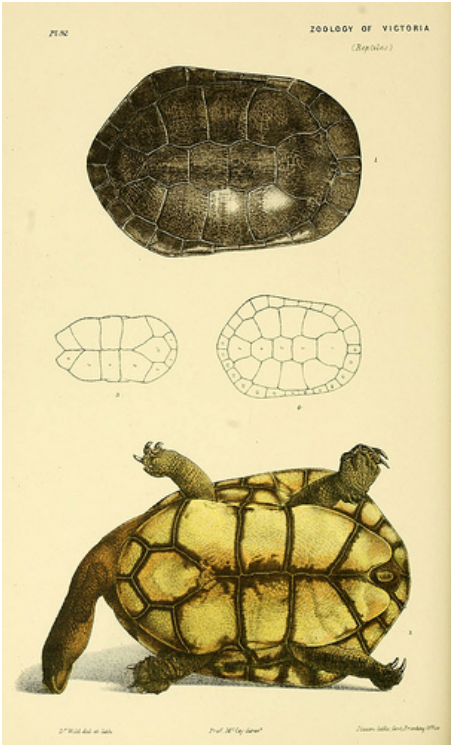
ATLAS OF LIVING AUSTRALIA
sharing biodiversity knowledge

Cover photo: Tatiana Gerus

Preface

The Atlas of Living Australia (ALA) is a significant and successful partnership between Australia's natural history collections and many other institutions, agencies and individuals committed to increasing our understanding of native species.

This report showcases the activity of ALA and its partners since 2007. The project has addressed many aspects of managing biodiversity data, but the combined result of so many partners working together is something much greater. Australia now has significant infrastructure both to generate and to manage its biodiversity data into the future as well as providing Australian researchers with world-class tools for exploring these data.



Chelodina longicollis, Natural history of Victoria, Frederick McCoy, 1881. Accessible through the Biodiversity Heritage Library—Australia bhl.ala.org.au/item/27160

The Atlas supports many different research and user communities. It brings together data, images and other resources from the national collections to support taxonomic research on Australia's unique biodiversity. It also serves as a critically-important spatial data set for all researchers and officers needing to explore and analyse the recorded distribution of native and introduced species in Australia. This same infrastructure also allows interested amateurs and the public to contribute to our knowledge of the fauna and flora and includes tools that allow these groups to assist with the databasing of historical specimens.

The Atlas has delivered foundational infrastructure allowing for national-scale collaboration. These foundations will ensure that future work in digitising Australian natural history collections and in conducting field surveys will deliver data of immediate benefit to taxonomy, ecological research, pest response, conservation and land-use planning.

The Atlas has now ensured that biodiversity data is readily accessible for all areas of research.

Donald Hobern
Director, Atlas of Living Australia

Contents

About the ALA project

The Atlas of Living Australia project is a national initiative focused on making Australia's biodiversity information more accessible and useable online: 'an online encyclopaedia of all living things in Australia'. The Atlas holds 25 million distribution records for Australia's fauna and flora, integrated with over 250 layers for mapping and analysis, as well as photos, maps, identification tools, reference species lists, literature and databases on biological collections.

As an organising hub for biodiversity data, ALA has natural linkages to the Australian National Data Service (ANDS) and other components of the Department of Innovation, Industry, Science and Research (DIISR) National Collaborative Research Infrastructure Strategy (NCRIS)/Education Investment Fund (EIF) e-research infrastructure investment. Related projects include the Terrestrial Ecosystem Research Network (TERN), the Integrated Marine Observing System (IMOS), the Australian Biosecurity Intelligence Network (ABIN), the Australian Phenomics Network (APN) and the Australian Plant Phenomics Facility (APPF). ALA has a significant partnership with the Great Eastern Ranges (GER) Initiative, delivering tools used by a range of stakeholders in monitoring and managing biodiversity.

ALA also serves as a gateway for data sharing between Australian and international biodiversity informatics programs, including the Global Biodiversity Information Facility (GBIF), the Encyclopedia of Life (EOL), the Biodiversity Heritage Library (BHL), IdentifyLife, Morphbank and the Barcode of Life Database (BOLD). It is an overseas partner of the European Union's Distributed Dynamic Diversity Databases for Life (4D4Life) and the Data Observation Network for Earth (DataONE).

CSIRO, Australian museums, herbaria and other biological collections, the Australian Government, and local communities are contributing to the development of ALA.

Funding

Funding for ALA until June 2012 comes from the Australian Government under NCRIS and the Super Science Initiative of the EIF. ALA partners are also contributing substantial in-kind resources, and are currently planning for the next stage of ALA.

For the period post 2012, ALA capabilities are already strongly acknowledged in the *2011 Strategic Roadmap for Australian Research Infrastructure* (www.innovation.gov.au/Science/ResearchInfrastructure/Pages/default.aspx). As such, this roadmap exercise is expected to result in additional funds for ALA from the 2013/14 financial year. In the meantime, CSIRO, as host agency for ALA, has indicated strong support for the project, and this is expected to result in additional supplementary funding through 2012/13. Furthermore, due to the value being delivered to a wide range of communities (including natural resource managers, Catchment Management Authorities, conservation agencies and amateur naturalists) in terms of biodiversity management—ALA is confident of attracting additional funds outside of the traditional research infrastructure funding base.

Governance

Two committees oversee the ALA project:

- a Management Committee represents ALA partners and provides strategic guidance on the development of the Atlas and its capabilities
- a CSIRO Steering Committee ensures that ALA operates within the governance principles of CSIRO and meets the project's obligations to the Australian Government.

Regular reports on ALA activities are provided to DIISR, which manages the NCRIS and EIF programs.

Communications

The delivery of core infrastructure and tools has made it possible for ALA to intensify its communication activities and raise its media profile substantially over the last twelve months.

ALA has actively strengthened its communications with partners through a series of roadmapping meetings, joint launches, symposiums, seminars, pilot projects and other activities, which have generated strong cross-agency communication and media coverage.

ALA presence at targeted events, coupled with the distribution of media releases and engagement with journalists, has generated significant interest from radio, print and web-based mass media. In July 2011 at the International Botanical Congress (IBC), a gathering of 2000 botanists from across the world, ALA staffed an information stand and display, held a public lecture, launched IdentifyLife (see page 20), published three media releases and conducted twelve interviews with media outlets.

As a result, the Atlas of Living Australia was featured in *The Age*, *Australian Geographic* online, *Gardening Australia* magazine, commercial Radio 6PR in Perth, ABC Radio National *Life Matters* show, ABC Radio National *Breakfast* show, ABC Melbourne *Breakfast* show, ABC Sydney *Drive time* show, and five ABC regional radio stations.

The intensified media coverage was closely followed by a surge in visits to the Atlas website, contacts from new potential sources of data and contributions to the Citizen Science portal.

As an example of a joint event, the ALA–Queensland Museum launch and symposium included launching a new Citizen Science portal, 'Wild Backyards' (www.qm.qld.gov.au/microsites/wild/) and a three month campaign of media promotion in fifteen Brisbane community newspapers (Quest). The joint symposium raised the awareness of over one hundred people from key departments and organisations, some of whom are now acting as ambassadors for ALA.

Status

ALA has now delivered its core infrastructure and a range of powerful new tools. New software has been developed for recording and managing biodiversity data and photos in the field. People and organisations around the country are using the Atlas as part of their work on research

projects, urban biodiversity surveys, museum outreach activities, science education, biosecurity monitoring, and natural resource management and reporting.

Partners and collaboration

While ALA has a number of formal partners it also engages with international organisations, in some cases with formal cooperative arrangements.

Australian partners

The Atlas of Living Australia was initiated by a group of 14 organisations:

- CSIRO
- Australian Museum
- Museum Victoria
- Queensland Museum
- Tasmanian Museum and Art Gallery
- The Council of Australasian Museum Directors (CAMD)
- The Council of Heads of Australasian Herbaria (CHAH)
- The Council of Heads of Australian Collections of Microorganisms (CHACM)
- The Council of Heads of Australian Entomological Collections (CHAEC)
- The Council of Heads of Australian Faunal Collections (CHAFC)
- Southern Cross University
- The University of Adelaide
- The Department of Agriculture, Fisheries and Forestry (DAFF)
- The Department of Sustainability, Environment, Water, Population and Communities (DSEWPaC).

In July 2010 ALA welcomed three new partners:

- Museum and Art Gallery of Northern Territory
- South Australian Museum
- Western Australian Museum.

International engagement

GBIF—Global Biodiversity Information Facility

ALA formally serves as the Australian GBIF node and is acting as the gateway for a number of Australian datasets to be shared with the GBIF network (approximately 14.5 million records from Birds Australia, NSW Office of Environment and Heritage's Atlas of NSW Wildlife, Eremaea, SA Department of Environment and Natural Resources and NT Department of Natural Resources, Environment, The Arts and Sport).

ALA has held a role as a mentor for GBIF India. A representative of the Wildlife Institute of India studied ALA data aggregation in Canberra and ALA participated in a training course for local data providers in Kolkata.

EOL—Encyclopedia of Life

ALA developers have been working with EOL to test processes for exposing Australian data resources on EOL pages. An initial trial is now in place with a number of ALA images resources shared.

The work of others

ALA is fortunate to be able to build on existing tools and the previous efforts of others bringing together biological data.

Data aggregators

ALA is working with data aggregators to develop compatible tools and data standards. ALA aims to contribute to the further development of these networks and to assist them in delivering information products targeted to their specific audiences. Data aggregators are:

- AVH—Australia's Virtual Herbarium (www.chah.gov.au/avh/)
- OZCAM—Online Zoological Collections of Australian Museums (www.ozcam.org.au/)
- PaDIL—Pest and Diseases Image Library (www.padil.gov.au/)
- OBIS—Ocean Biogeographic Information System (www.iobis.org/)
- AMRiN—Australian Microbial Resources Information Network (amrin.org/)
- APPD—Australian Plant Pest Database (www.planthealthaustralia.com.au/go/phau/capacity-and-capability/information-support-systems/appd).

Government and non-government organisations

ALA is assisting some organisations to better capture, mobilise and share their data while both government and non-government organisations are contributing invaluable data and ideas to ALA. These include:

- Birds Australia (www.birdsaustralia.com.au/)

- Earthwatch (www.earthwatch.org/australia)
- Bush Blitz (www.bushblitz.org.au/)
- Great Eastern Ranges Initiative (www.greateasternranges.org.au/).

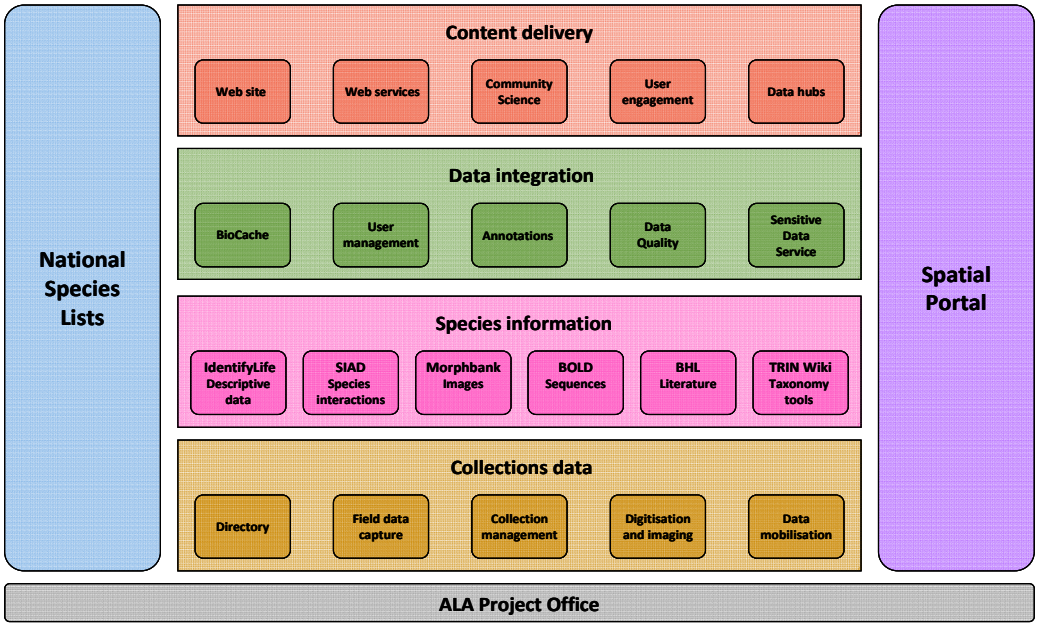
Global initiatives

ALA is contributing to global projects making use of existing work and adding local content, including:

- BHL—Biodiversity Heritage Library (www.biodiversitylibrary.org/)
- BOLD—Barcode of Life Database (www.boldsystems.org/)
- Morphbank (www.morphbank.net/).

ALA components

ALA is developing core systems, tools and services needed for supporting and managing biodiversity information. The diagram below gives an overview of the six main areas of work.



Key components to be delivered by the ALA project

National Species Lists. Tools are being developed to manage and share names and classification of all Australian organisms. Comprehensive lists are being built for all species found in Australia.

Spatial Portal. At the heart of the Atlas is the ability to map data from natural history collections, ecological field work and amateur observations. Mapping tools are interoperable with data caches, environmental layers and analysis frameworks.

Collections data. ALA is working with the whole collection community to enhance capacity for accession, curation and databasing of specimens, including novel techniques with volunteers.

Species Information. Repositories are provided to help with management of images, sequences, identification keys, species interactions and literature.

Data integration. This work is bringing together and managing data and information from all sources, addressing data quality issues and sensitivity of data for some species.

Content delivery. The Atlas web site delivers species pages, maps, occurrence records, and information. Access is also tailored to specific views, such as through AVH, OZCAM and other data hubs; for amateur naturalists; and in machine form.

Much of the software used to develop these components is Open Source and can be freely downloaded by anyone wishing to use it.

National Species Lists

Authoritative lists of scientifically defensible names are the backbone of biodiversity science. The National Species Lists (NSL) project provides this key infrastructure of reusable nomenclatural and taxonomic content, and systems and services to enable interoperability within biodiversity informatics. The NSL project is:

- an integrated, open-data repository
- owned and maintained by the taxonomic community
- based on published or citable sources
- based on web-based distributed data management solutions
- built on biodiversity standards-based web services
- using an architecture that accommodates both reusable open factual content and taxonomic overlays
- supportive of content evolution through incremental change
- able to ensure persistence of cited extracts
- able to integrate with global initiatives.

The core of NSL is provided by the Australian Faunal Directory (AFD), maintained by the Australian Biological Resources Study (ABRS), and the Australian Plant Census (APC) based on the Australian Plant Name Index (APNI) maintained by the Council of Heads of Australasian Herbaria (CHAH). These databases are the most complete on-line resources of information on the names and taxonomy of Australian plants and animals. They are of high quality, authoritative, actively maintained and widely accepted as the single point of truth for Australian nomenclature and taxonomy. Complementary additional content comes from Australian taxonomic compilations such as the Interactive Catalogue of Australian Fungi, the Australian Moss Catalogue, the Checklist of Lichens and the Australian Marine Algae Index.

A key component of the NSL project has been the direct ALA funding and engagement of the Australian taxonomic community in compiling, extending and enhancing content to build the essential NSL infrastructure. The importance of this community engagement cannot be underestimated; without it the NSL would not survive as a contemporary and on-going resource.

Existing services of the NSL include:

- persistent identifiers
- Life Science Identifiers (LSID) resolution
- OAI-PMH and Linked Data objects for names, taxa, publications, classifications and people using TDWG standard forms for RDF, XML, JSON and HTML
- name search incorporating fuzzy matching
- taxon name resolution services.

Names are now anchored to cited concepts or hypotheses of the precise taxon in question.

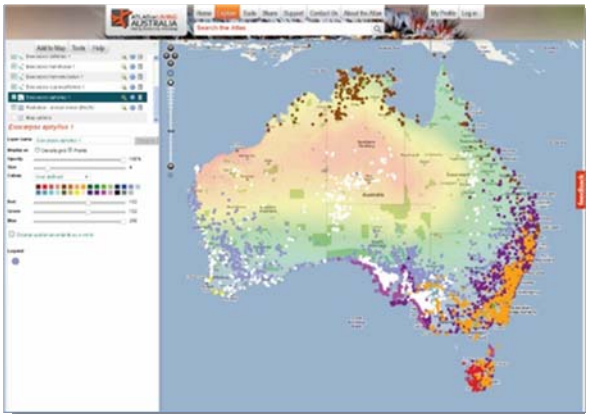
Spatial Portal

The ALA Spatial Portal (spatial.ala.org.au) aims to support the environmental management and research communities. The Spatial Portal brings together over 25 million (and growing) biological records, 250 environmental layers (eg mean annual temperature), 50 contextual layers (eg land cover) and over 800,000 named areas.

The interface of the Spatial Portal has been designed to easily map taxa, areas and layers. Taxa can be defined at any taxonomic level or be uploaded. Uploaded records use CSV-format and comprise an ID, decimal longitude and decimal latitude. LSIDs can also be uploaded (in CSV-format) to define species assemblages. In either case, records (imported or looked-up respectively) can be used in the Spatial Portal in the same way as any map produced from occurrence records.

Areas can be defined in 13 different ways:

- digitised or selected from the map
- searched via Google or our gazetteer
- uploaded (in three different formats)
- defined by an environmental envelope. This is a powerful tool to identify areas with specific environmental conditions based on any combination of available environmental layers (see spatial.ala.org.au/layers).



Spatial Portal

Basic tools include the generation of species checklists over any defined area and sampling of any combination of environmental and contextual layers at occurrence locations. All results can be exported as CSV-formatted files. More advanced tools that use the environmental layers include:

- scatterplots of taxa (eg www.ala.org.au/explore/themes/case-studies/scatterplot-case-study/)
- the classification of environments (eg www.ala.org.au/explore/themes/case-studies/classification-case-study/)
- spatial prediction of taxa (eg www.ala.org.au/explore/themes/case-studies/prediction-case-study/).

The Spatial Portal database structure now permits a great deal of flexibility to facet and filter occurrence records. Species occurrence records can also be uploaded with up to 256 fields that can be used for faceting and filtering. Future plans include the ability to generalise the scatterplot function to include the tabulation of any combination of environmental and contextual layers. Simon Ferrier's Generalized Dissimilarity Modelling will also be added to the tools before the end of 2011.

Collections data

ALA has been developing tools to help effectively manage data over the entire collection process—from capturing information directly in the field, through accession and curation of specimens, adding new records to a database, mobilising data out of institutions, to sharing that data.

The Directory

A repository for information about Australia's natural history collections, the Directory is a showcase for each collection, its governing institution, key people and datasets. See page 11.

Field data capture

Part of the Citizen Science toolkit and extended for professional science, Gaia Resources has been developing applications for smart phones and other mobile devices to enable people in the field to capture information about plants and animals. See Community Science page 34.

Collection management

ALA has re-engineered BioLink, a database used by entomology collections; supported the roll-out out of BioloMICS into microorganism collections; and conducted workshops to introduce Specify 6 as an Open Source database for managing collections. See pages 12–14.

Digitisation support

Museums and herbaria are capturing high resolution photographs of specimens using imaging equipment provided by ALA. Images are being shared publicly and are readily accessible to researchers world-wide. Guidance materials have been produced to help institutions with their imaging and digitisation strategies. See Digitisation support, page 15.

Rapid digitisation

Two museums are trialling projects with volunteers to rapidly digitise their collections. At the South Australian Museum, volunteers are capturing high quality images of original type specimens. At the Australian Museum, volunteers are photographing specimen labels and registers. See Rapid digitisation, page 16. Volunteers working from home will be able to assist with transcribing labels. See Volunteer Portal page 17.

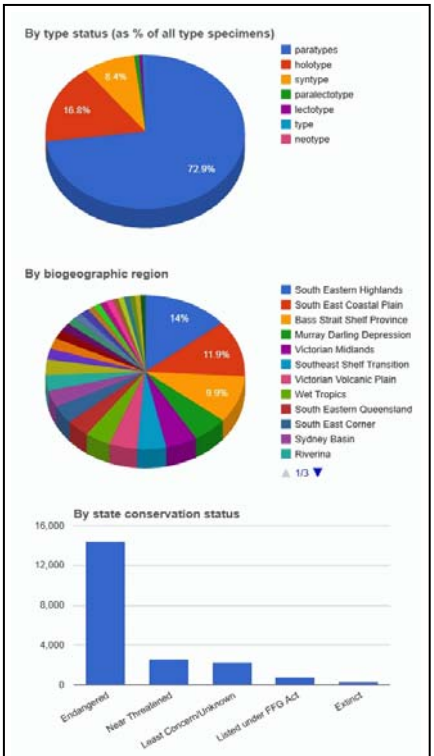
Data mobilisation

Moving data out of an institutional database is surprisingly tough. ALA has run workshops for technical staff, assisted with developing standards for data sharing, and built and implemented tools to export data regularly. See page 18.

The Directory

The role of the Directory (formally the Atlas Registry) has expanded from its initial conception as a catalogue of biological collections, to a major repository of metadata for the Atlas.

In addition to describing biological collections and associated institutions, the Directory now holds information about the datasets used by the Atlas as well as identifying other useful biodiversity resources. Datasets can be browsed, sorted, filtered and searched from the datasets page (collections.ala.org.au/datasets).



Graphic view of dataset

The Directory is highly integrated with other Atlas components. It provides support, and in some cases coordination, for automated processes including data mobilisation, ongoing harvesting, record processing and indexing. It also supplies other components with contextual information such as citations and rights, organisation names and relationships, and groupings of occurrence records to support a hub view of data in the ALA BioCache.

In keeping with the Atlas modular, re-usable architecture, almost all metadata and functionality of the Directory is available via web services. Work is underway to actively publish metadata via initiatives such as the Australian National Data Service (ANDS).

The Directory provides alternative entry points for exploring occurrence data including using visualisations such as charts and graphs. Increasingly these visualisations are being made available to users who wish to put similar charts on their own web pages.

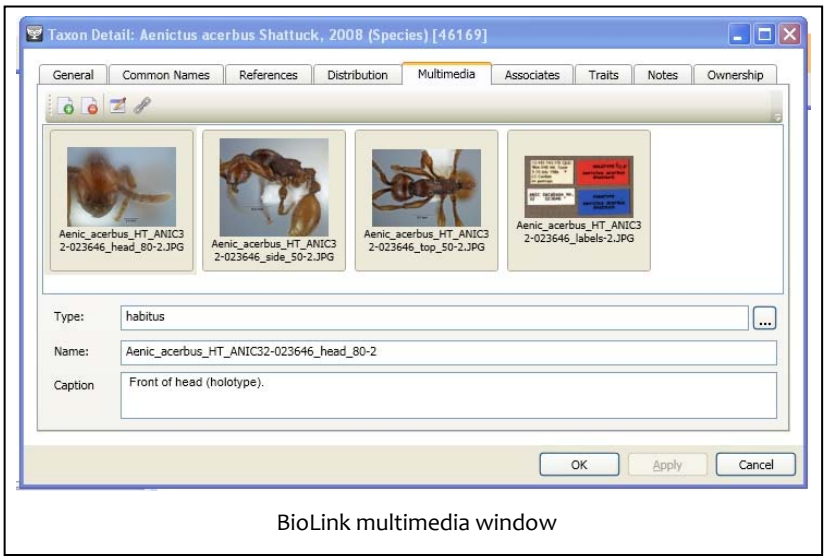
It was always intended that collection managers and data managers should have the ability to maintain their own metadata. This 'self-service' facility is currently under development and will be delivered in 2012.

In the same timeframe, it is expected that collection metadata will be expanded to include a more detailed account of holdings. Institutional data currently collected by the Australian Biological Resources Study (ABRS) and the Council of Heads of Australasian Herbaria (CHAH) will be collected through the ALA Directory infrastructure including a self-service facility. CHAH plans to deliver its data on its own website which will be branded as 'powered by ALA'. ABRS will be able to extract from the Directory the data needed for institutional collections reporting. ALA will also provide infrastructure for the ABRS Taxonomic Capacity Surveys.

BioLink

It's been just over 10 years since BioLink Version 1.0 made its first public appearance. Since then there have been a number of releases, the most recent being several years ago when Version 2.5 was made available. Now, with support from ALA, BioLink has undergone a complete redevelopment and Version 3.0 is now ready to carry on the long BioLink tradition.

While BioLink 3.0 will be familiar to existing users, the new user interface offers many features that simplify data entry, collection management, report generation and image handling, and provides a rich environment to customise the interface to meet individual user needs. Feedback from early users and tests has been very positive.



BioLink multimedia window

The cornerstone of BioLink 3.0 is its dockable windows. The Explorers and Tools windows can now be docked to the edges of the main window so they remain in the same location throughout a session, always ready for use. Windows can be set to 'auto-hide', sliding out of the way when not needed but popping open with a quick slide of the mouse pointer. This is an especially useful feature for the new Pinboard, a place to collect frequently used items such as taxa, locations and references.

BioLink 3.0 has been made much more user friendly, often in small but important ways. For example, previously when the Taxon and Site Explorers were closed they 'forgot' where the user was working. Now, when they are opened they return to the exact spot the user was working previously. This can save considerable time by eliminating the need to perform a search or to click down the classification to a favourite family or genus. The Multimedia Manager has been upgraded to show thumbnails of all images attached to a taxon or other object.

And a number of new reports have been added. For example, do you need to know what specimens have been collected within the Wet Tropics World Heritage Area? With the help of the Mapping tool and an appropriate Shape file the answer is only a couple of mouse clicks away. And in another couple of clicks the results can be in Word or Excel, ready for reporting or further analysis, or transferred to a data partner such as OZCAM or ALA.

The new BioLink can be downloaded from Google Code at code.google.com/p/biolink/. And the complete source code will soon be available at this same site. If early indications are correct BioLink 3.0 will prove as useful and popular as its predecessors and the new user interface will made it even easier to use.

BioloMICS

Following consultation with the Council of Heads of Australian Collections of Microorganisms (CHACM), ALA is providing licences for BioloMICS software (see www.bio-aware.com) to assist in managing and sharing data on microorganisms.

A 100-license bundle has been purchased and implemented in 7 of the 12 sites that have agreed to adopt BioloMICS. The remaining sites will be completed by mid-November. Each implementation involves installing software on servers and workstations, face-to-face training, and significant assistance migrating data and customising the product.

As a consequence of this activity, about 20 Australian microorganism collections will be digitally recorded, managed and analysed through the same product. This will, in turn, allow for much closer levels of collaboration both within and externally to CHACM. An extended maintenance period for BioloMICS has been purchased: to the end June 2013.

Nathalie van de Wiele (in black) delivering training

In supporting BioloMICS implementation, ALA has:

- documented requirements for sharing data
- developed a logical data model to be used in BioloMICS
- developed a Darwin Core implementation of the data model to be used to mobilise data into the Atlas
- established protocols for securely transferring data to the Atlas
- developed a range of guidance notes and software scripts that will be published on the Atlas.

Future development of BioloMICS is expected to allow the software to automatically access species names via web services, such as those provided by the National Species Lists (see page 8).

Specify 6

Specify 6 (specifysoftware.org) is Open Source software designed to manage natural history collections. ALA has been actively exploring with those institutions in need of new or better collection management software how best to help—including determining whether Specify 6 would be suitable for their collections.

To date, Specify 6 has been implemented in one Australian institution (National Herbarium of Victoria, noting that this implementation was completed prior to ALA involvement) while a number of others have indicated an intention to adopt the software. Several people expressed interest in further training in Specify 6, but the numbers were too small for ALA to host specialist courses.

Potential users at ALA-hosted workshops identified features not in Specify 6 that could benefit Australian institutions. These features are documented and discussed at bit.ly/otjNme and have been referred to the Specify development team for consideration in future releases.

ALA would like to be able to provide additional assistance to those interested in Specify 6, including provision of resources for data analysis, prototyping, project management and implementation/migration. Unfortunately, this is one area where ALA funding has not been able to stretch during this current funding period. Instead, ALA will make available documents that summarise the level of interest in Specify 6, issues that have arisen, and lessons learnt from existing implementations—information that will be ready to be used when additional funding is available.

Digitisation support

To help biological collections expose their specimen information, ALA has engaged in a number of activities intended to support the digitisation of collections, especially creating images of specimens. Notable activities include:

- visiting partners to gain an understanding of their digitisation activities and the constraints on digitisation
- funding the purchase of a range of imaging hardware and software, eg cameras, microscopes, x-ray machine, Visionary Digital systems, hard disk storage, digital asset management software, image manipulation software, computers
- arranging for an imaging expert to assist partners in optimising their use of the new hardware and software, as well as improving existing imaging practices
- developing and publishing a range of imaging case studies and guidance, including how to establish a digitisation strategy (www.ala.org.au/tools-services/imaging/)

Tachinid fly
Photo by Geoff Thompson, Queensland Museum

- assisting partners mobilise their images into the ALA instance of Morphbank (see page 22).

ALA is currently working with several partners to develop digitisation strategies, and these will be used as models for other partners to develop their strategies. ALA is also expecting to publish more case studies to demonstrate the value of digitisation and how the new hardware is being used as a research tool.

Rapid digitisation

ALA has funded an experimental project at the Australian Museum in Sydney and the South Australian Museum in Adelaide to rapidly digitise their collections using volunteers (see also the next section on the Volunteer Portal).

At the Australian Museum, a laboratory was established with imaging equipment and six digitising workstations operated by a team of 50 volunteers who keep the lab working Tuesday through to Saturday each week. The lab is staffed by two part-time digitisation officers coordinating the volunteers with technical assistance for photographic and database management.

The South Australian Museum Digitisation Suite is located in the Artlab building. Previously used for model-making it now houses the project manager and three volunteer digitisation workstations: one designed for macrophotography, one for microphotography and one for post-processing and databasing. 15 volunteers contribute about 85 hours per week.

A moth and its label imaged at the Australian Museum

What the project is about

At the Australian Museum, the project aims to establishing best practice for using volunteers for digitising Museum specimens and registers. Volunteers take photos of specimens and their labels, and register pages, and add them into the Museum's databases. The images of specimen labels are submitted to the ALA Volunteer Portal. Volunteers working online then help to transcribe the label information—location, date, collector, identification, collection method and other details.

At the South Australian Museum, 8500 holotype specimens and their associated labels are being photographed and processed by volunteers. Undocumented holotype specimens (approximately 83% of those being photographed) are being databased.

Both projects produce images and collection data that can be used not only within each museum, but shared more widely through ALA. Researchers world-wide and the broader community will have access to high resolution images of unique and fragile museum specimens.

The processes and practices of each project are being documented in detail for other museum departments as well as other museums. This will promote consistent practices and digitisation standards, and encourage the use of volunteers for digitisation.

Volunteer Portal

The Volunteer Portal (volunteer.ala.org.au/) has been created to test crowd-sourcing, also known as online volunteering, as a mechanism for transcribing specimen labels, herbarium sheet labels and field notes.

Images of hand-written labels and field notes, such as those generated through rapid digitisation activities as detailed in the previous section, are uploaded to a website (the Volunteer Portal) where the text in the images can be transcribed into specific database fields. The latitude and longitude of given locations are also determined, an activity known as georeferencing. Transcription and georeferencing tasks are carried out by volunteers working on their computers from home, or elsewhere.

After each label is transcribed the text is validated online, also by volunteers, and then imported into the database of the custodial institution. The images and metadata can then be shared more widely through ALA, OZCAM, AVH or GBIF.

Volunteer Portal website

Images

Images for transcription by volunteers are currently generated through the Rapid Digitisation Project being run at the Australian Museum (see page 16). Other images are planned for the coming months. Projects at the CSIRO Australian National Insect Collection and the John Ray Herbarium at Sydney University will soon also be providing images for label transcription. A specific template for each category of label may need to be created.

The first project

The Australian Museum Cicada Expedition trial used 47 online volunteers and transcribed 1183 records.

Data mobilisation

Support for data mobilisation into the Atlas continues on two fronts, in close association with our partners and data providers.

System-specific mobilisation software suites are being developed for representative collection management systems, eg KE EMu and Vernon. These suites provide standard schemas and software for easy export of data to the Atlas. Source code will be issued as Open Source with the expectation that a community of users will grow around each suite. Development activities have focussed on:

- Texpress (EMu) on Linux—a near-complete prototype is available for use in three institutions
- SQL Server on Windows—a prototype is being developed for use in one institution
- Vernon—a plan for sharing data within these systems is under way
- MySQL (Specify 6) on Windows—first steps have been taken in understanding the mobilisation challenge.

Mobilisation solutions for secondary data are also being devised. Secondary data represent, in a standard form, collection items or artefacts that are derived from the more traditional 'specimen brought in from the field', eg sound files, image and other multimedia files, gene sequences and specimens shared amongst collections.

ALA is working with five collecting institutions to mobilise their image collections into Morphbank, including very high resolution images that have not been attempted before. These partnerships will eventually load many thousands of images, but work to date has largely been in bulk digitisation, metadata capture and image set preparation. Loading is expected to happen in October and November 2011.

ALA has also streamlined data mobilisation processes based on experience to date. For instance, registration and description of resources (data providers and datasets), including the data mobilisation settings, can now be managed entirely within the Directory (see page 11).

For ongoing developments see:

- code.google.com/p/ala-datamob/
- Darwin Core completeness model code.google.com/p/ala-dataquality/wiki/CompletenessModelDwC

Species information

Information on species is central to the Atlas, and ALA has been actively supporting development of tools to facilitate the capture, creation and management of species information. Integration of the information from these tools into the Atlas is continuing.

IdentifyLife

A suite of tools to develop, manage and share species identification keys, IdentifyLife is an international project aimed at producing a next generation keys manager. See page 20.

Species Interactions of Australia Database

ALA is building a public domain, Open Source information system —Species Interactions of Australia Database (SIAD)—to allow users to capture and share information on species interactions and associations. See page 21.

Morphbank

Morphbank is an Open Source web application and repository for storing and sharing specimen-based images, principally for scientific use. See page 22.

Molecular data

ALA is working with the Barcode of Life Database project and European Molecular Biology Laboratory Australia to provide a repository for molecular data and to gain access to molecular data on Australian species held in overseas repositories. See page 23.

Biodiversity Heritage Library—Australia

An Australian node of the Biodiversity Heritage Library is being established as a repository for biodiversity literature. See page 24.

TRIN Wiki

An Open Source, collaborative tool, TRIN Wiki provides contributors with simple online interfaces to create, collate and compile well-structured biodiversity knowledge. See page 25.

Open DELTA

Open DELTA is an Open Source, updated version of the DELTA software used to encode taxonomic descriptions for machine processing. See page 26.

IdentifyLife

The IdentifyLife project (www.identifylife.org/), supported by ALA, the Moore Foundation, the Encyclopedia of Life and the University of Queensland, is building a global, collaborative web application and services to provide identification and descriptive data management tools for organisms. IdentifyLife has three main parts:

- Keys Central—a crowd-sourced, searchable index of web-based identification resources, allowing users of IdentifyLife to find keys in a wide variety of formats to a wide variety of organisms anywhere in the world
- My IdentifyLife—a collaboration space that allows groups of researchers and other experts to record detailed information about the morphology of a group of organisms and the characters needed to describe and identify them
- Key to All Life—an ambitious project to build a key that, at least potentially, could allow identification of any organism anywhere in the world.

IdentifyLife was launched at the International Botanical Congress in Melbourne in August 2011. The three core modules are complete and operational at least in part. IdentifyLife is open to users. Core functionality is built and is being tested and extended through real-world use cases. While much remains to be done, IdentifyLife is being used to commence significant projects such as standardised character lists (ontologies) and keys to the Asteraceae, Euphorbia and Solanum species of the world; and a key to the genera of flowering plants of Australia. Discussions are underway to index a large body of keys published in the online journals Zootaxa and Phytotaxa into the Keys Central index.

While the IdentifyLife framework is built, much proofing, testing and refining needs to be done. New developments in the next 12 months will include a mechanism for uploading, archiving and integrating into IdentifyLife legacy keys and other datasets in a variety of formats, and mechanisms for handling dichotomous keys. Other work, time permitting, will include support for more data types, more work on the collaboration interface and tools, and support for the generation of natural-language descriptions from IdentifyLife coded data.

Species Interactions of Australia Database

ALA is building an Open Source information system —Species Interactions of Australia Database (SIAD)—that will enable users to gather and share information on species interactions and associations.

Components of the system include:

- a species interaction data schema and associated controlled vocabulary
- mirrored data warehouses of publicly accessible files with a searchable index and associated metadata
- software tools to upload, download and check the integrity of data records.

Coridromius monotocopsis Tatarnic and Cassis 2008 (Insecta: Heteroptera: Miridae) eating *Monotoca elliptica* (Sm.)

The system will accommodate:

- literature based data, eg the Australian Faunal Directory (AFD www.environment.gov.au/biodiversity/abrs/online-resources/fauna/afd/home) and the Australian Plant Name Index (APNI www.cpbr.gov.au/apni/)
- specimen-based data, eg Plant Bug Inventory Database (research.amnh.org/pbi/)
- observational data from various sources, including photographs submitted by the public.

To test and refine the system's functions, the system will initially be populated with datasets that include information on Australian species, with the true bugs of Australia (Heteroptera) and their host plants as the first test case.

Activities to date have focused on developing the data schema; acquiring and processing test data; and enhancing the true bug test case dataset. The data schema and controlled vocabulary is being developed in association with the Encyclopedia of Life (www.eol.org) and BioInfo (UK) (www.bioinfo.org.uk). Version 1 of the data schema was discussed at a workshop at the 2011 TDWG meeting in New Orleans (abstract at www.discoverlife.org/siad/tdwg2011).

See:

- www.discoverlife.org/siad/questions for more information on SIAD
- www.discoverlife.org/siad/interactions for typical interactions to be included in SIAD

Morphbank

Morphbank is an Open Source web application and repository for storing and sharing specimen-based images, principally for scientific use.

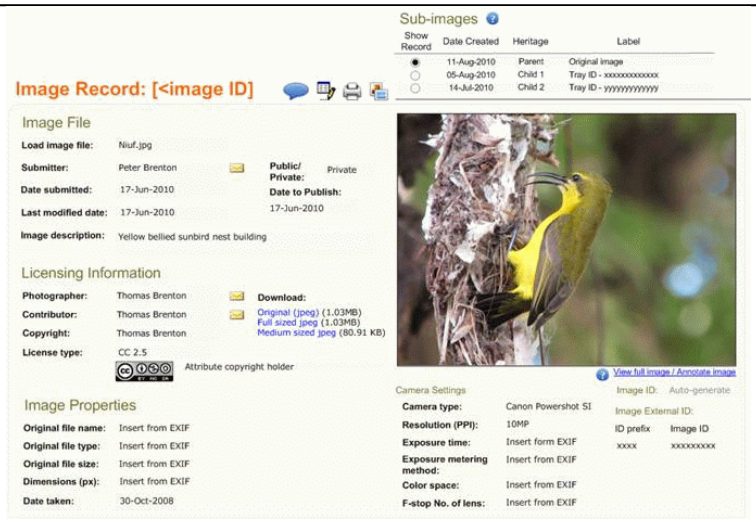
ALA, in collaboration with the US Morphbank development team, is working to:

- implement the current version of Morphbank on Atlas infrastructure; this is now in place
- refresh and extend the software to take advantage of new technologies to better serve the requirements of Australian users.

Use of Morphbank has been actively promoted to all major collections and a number have decided to adopt it as their own image management system.

Images from Morphbank have been indexed in, and are available from, the Atlas.

Development of the new Drupal-based version of Morphbank is about three months behind schedule, but is proceeding steadily.



The new version of Morphbank has an improved user interface

By the end of June 2012 it is expected that all of the new application design and planned features will have been built, tested and deployed for production use. More robust image loading methods will also be in place to enable independent bulk loading capability by institutions.

Finally, the new version of Morphbank is expected to be installed in at least three major institutions and be the primary image repository for at least another two major institutions. By June 2012, about 10,000 additional images should have been loaded.

See also:

- morphbank.ala.org.au/?id=2000569h for a sample image in current version of Morphbank
- bie.ala.org.au/species/Platygyra daedalea for a sample of an ALA link to a Morphbank image

Molecular data

BOLD

ALA has been working with the BOLD (Barcode of Life Database) development team at the University of Guelph in Canada (www.boldsystems.org/) to integrate BOLD data into the Atlas and to plan an Australian node of BOLD. The Australian node will enhance upload, download and storage for Australian barcoding projects.

ALA hosted an international meeting to progress the development of BOLD nodes in several countries. Existing BOLD mirroring software developed by the Chinese Academy of Sciences (www.boldmirror.net/) is likely to provide a rapid mirroring solution for the Atlas. A software stack is now with ALA for assessment and likely implementation.

Many Australian projects already contribute significantly to thematic barcode networks with over 33,000 sequences generated.

Australian molecular data

ALA is collaborating with the European Molecular Biology Laboratory Australia (EMBL Australia) mirror (www.emblaustralia.org) as the source for molecular sequence data on Australian species. EMBL Australia was established at the University of Queensland by the Australian National Data Service (ands.org.au).

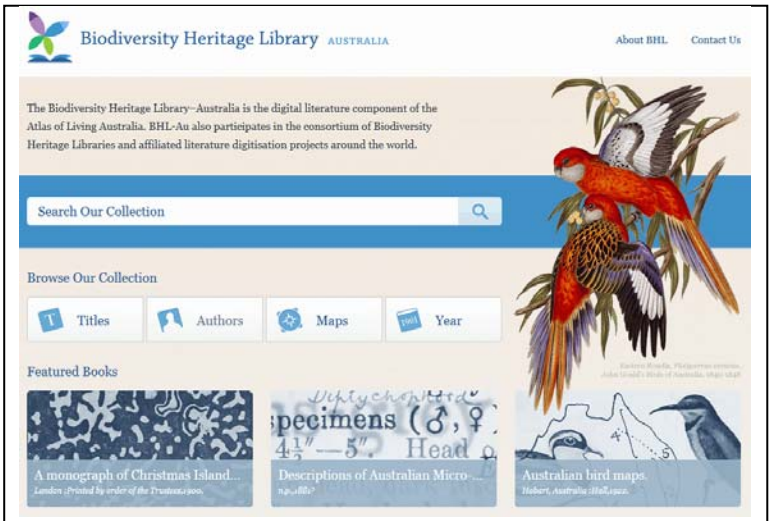
There are about 500,000 Australian sequence records that will be indexed by the Atlas providing a view of all existing sequence data for Australian species.

Around 70,000 sequence records also include geospatial coordinates and would be suitable for display through the ALA Spatial Portal. This resource will provide an important new view of the extent of Australian sequence data especially when viewed with combinations of the 25 million existing records and overlaid with environmental and contextual layers from the Spatial Portal.

Biodiversity Heritage Library—Australia

Scientific literature used in the Atlas is provided by the Australian node of the Biodiversity Heritage Library (BHL). BHL is a global consortium of museum, herbarium, university and research institute libraries whose aim is to provide free and open access to biological literature. Originally started by 12 libraries in the US and the UK, BHL now has partners in Europe (through the Europeana project), China, Brazil and Egypt. The overall aim is to provide extensive, open and global access to digitised literature.

The Australian node of BHL (BHL–Au bhl.ala.org.au) is being developed by Museum Victoria on behalf of ALA in close consultation with the US global technical managers.



BHL-Au home page

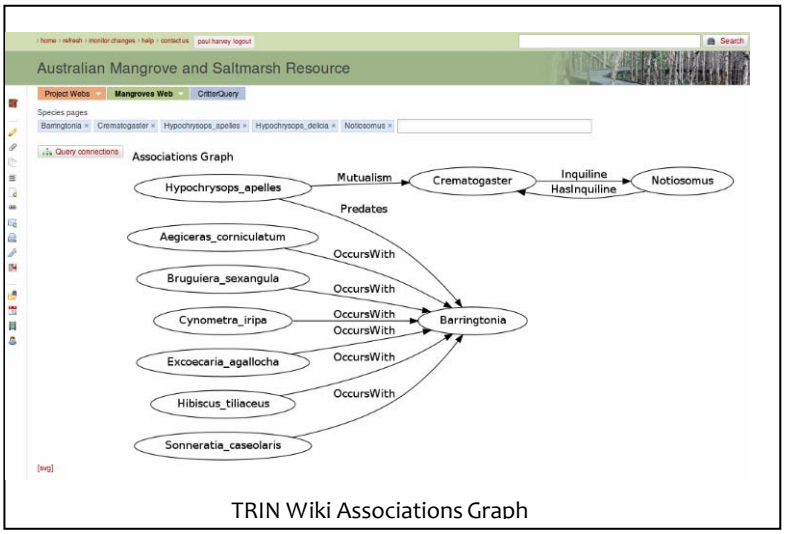
Progress on the project so far has seen the BHL-Au node go live with a brand new design and streamlined functionality. Content available in the US node, more than 33 million pages of scanned literature, can all be searched through the Australian node. New scanning operations, focusing on Australian published literature, are getting underway. An ATIZ Book Pro scanner has been purchased and commissioned at Museum Victoria. A 'bid list' is in development. This will be presented as a website that will allow users to nominate titles they wish to see scanned and made available through BHL. Development of workflows to allow the Australian node to upload literature and metadata to the Internet Archive and then publish to BHL is progressing.

Plans for BHL-Au include adding an annotations function to the website, exposing BHL content through the National Library of Australia Trove service and making OCR correction tools available. BHL-Au will also work with the global team to provide better access direct to articles in scanned journals, in order to better service the scientists who are the principal users.

TRIN Wiki

TRIN Wiki (wiki.trin.org.au/Main) provides tools for collaboration in biodiversity research and documentation. Open Source and collaborative, it provides contributors with simple online interfaces for the creation, collation and compilation of well-structured biodiversity knowledge.

TRIN Wiki has been developed through close collaboration with biodiversity research teams and is organised around specific projects which vary in their maturity and target audience. Some projects contain information put together for public consumption; others are moderated or closed spaces, in which taxonomic projects collaborate on their own business prior to publication. These projects have provided templates for:



TRIN Wiki Associations Graph

- primary pre-publication research in taxonomy and systematics
- enlisting public contributions to document species diversity for research
- managing molecular research into species biodiversity, linked to online resources
- documenting species and their interactions in a given ecosystem
- using phylogenies to explore biodiversity.

Atlas web services are used to provide a broader context for specimens, samples and species through online resources which provide access to national species lists, synonymies, maps that include useful biologically relevant layers and other online biodiversity resources.

Examples of TRIN Wiki projects are:

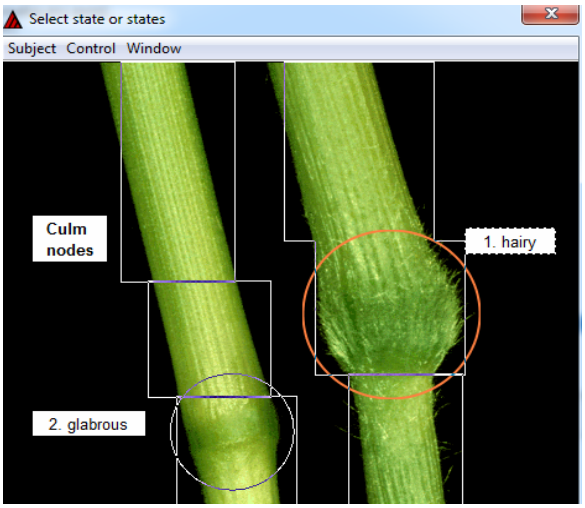
- Mangroves (wiki.trin.org.au/Mangroves/)
- Acacia (wiki.trin.org.au/Acacia/WebHome)

Open DELTA

DELTA (DEscriptive Language for TAxonomy) software was developed by CSIRO between 1971 and 2000 (see delta-intkey.com/) and is a flexible tool for encoding taxonomic descriptions for machine processing.

DELTA-format data can be used to produce natural-language descriptions, conventional or interactive keys and phylogenetic and phenetic classifications. By capturing data in DELTA, taxonomists ensure uniformity and completeness in their descriptions. They can more efficiently update descriptions as new knowledge comes to hand and re-use the same data for multiple purposes.

Since 2000 there has been no ongoing maintenance of DELTA resulting in some of its core components not operating under Windows Vista and Windows 7.



A pop-up Intkey screen in Open DELTA

Given the widespread use of DELTA, in February 2011 ALA announced that it would re-develop DELTA. This project—Open DELTA (code.google.com/p/open-delta/)—has the following objectives:

- Develop a consolidated and extensible DELTA code base using modern programming tools and release the code as Open Source software. This will allow ongoing development and maintenance of DELTA beyond the current term of ALA.
- Replicate at least the functionality of the current software.
- Deliver a cross platform solution, to allow DELTA to run on all current operating systems.
- Collaborate with IdentifyLife (see page 20) to facilitate interoperability between the two tools for creating taxonomic keys.

To inform the redevelopment and better assess the work required to meet the project objectives, ALA established an advisory working group and conducted a review, outlined in *DELTA Redevelopment Options Paper* (www.ala.org.au/wp-content/uploads/2011/02/DELTA-Redevelopment-Options-Paper.pdf). In response to feedback, ALA undertook to focus work on re-implementing existing DELTA functionality rather than enhancing or extending DELTA functionality.

Timeframes for developing the new software are outlined in the project roadmap (code.google.com/p/open-delta/wiki/RoadMap). Consistent with the roadmap, Milestone 7 was released on 14 September this year and development is on track to meet its objectives by 30 June 2012.

Data integration

More than 25 million occurrence records are shared via the Atlas, with over 30 million records expected by June 2012.

Over 250 environmental, geospatial and contextual layers are available in the Spatial Portal (spatial.ala.org.au/layers, page 9). Significant layers include:

- 3 second (~100m) terrestrial Digital Elevation Model (DEM)
- estuary habitat classifications
- (100m) Great Barrier Reef DEM
- Great Eastern Ranges Initiative regions.

The Atlas holds descriptions of over 300 data providers (or resources), each of which is listed in the Directory (collections.ala.org.au/public/datasets, see page 11). Providers can be found by their content types, eg the list can be filtered on whether the resources contains habitat information. In this way, users can find relevant resources even where the Atlas itself does not (or does not yet) display that information.

Current data providers comprise:

- 160 web sites
- 138 record providers
- 17 documents (species lists or sensitive data rules)
- data shared through the Atlas 'Share your data' capability (various individuals).

Data from 186 of these providers is already in the Atlas while the remainder are working through the process of sharing information and links.

Existing information gaps, such as reef life and coral data, or customary medicinal knowledge, are being filled through a number of collaborations.

There are three main components of data integration:

The BioCache

Specimen and occurrence data is held and integrated in the BioCache, which is also responsible for harvesting data from providers. See page 28.

Data quality assurance

Before data is added to the BioCache it is subject to a number of data quality checks to improve its fitness-for-use. See page 29.

Data sensitive check via the Sensitive Data Service

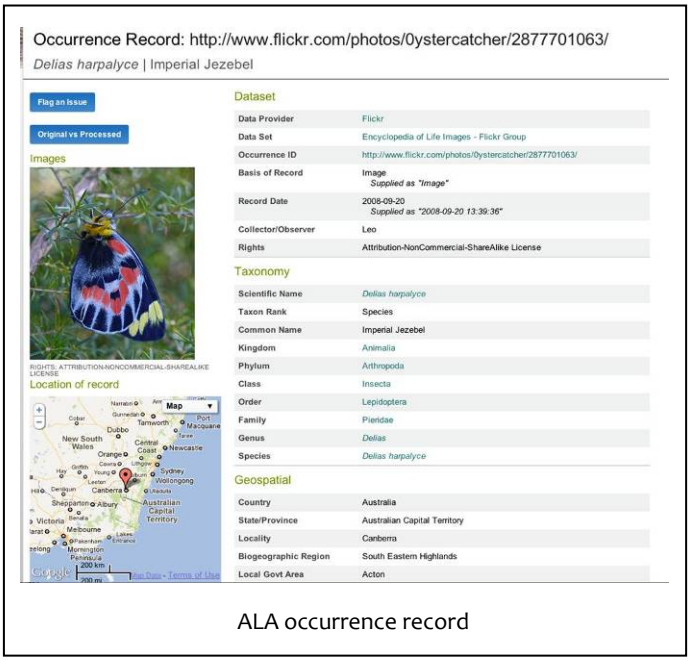
The Sensitive Data Service identifies information that might cause harm if it were made public. It then applies rules to limit or modify that data made visible through the Atlas. See page 30.

The BioCache

The BioCache (biocache.ala.org.au) is the cache of specimen and occurrence records that the Atlas keeps up to date by harvesting data from institutions. This data helps drive mapping applications for species and any application describing the distribution of Australian taxa within the Atlas or externally.

The BioCache supports:

- an extensible range of data fields, supporting any field required by the botanic, zoological and micro-organism communities
- a platform for multiple web applications, internal and external to the Atlas. This includes support for OZCAM, AVH, APPD and other applications requiring search capabilities on top of the occurrence data such as the Spatial Portal
- integration with the Sensitive Data Service (see page 30) so that locations for sensitive species are either generalised or removed from public view



ALA occurrence record

- a range of data quality tests (see page 29) and the ability to search for records by specific issues—allowing researchers to select data sets according to their immediate 'fitness for purpose' requirements
- simple annotation integration allowing users to assert there are issues with records and for data custodians to be able to provide a verification of the records (over-ruling any previous assertions by users)
- an extensive range of methods for searching for records including by conservation status.

Demonstration versions of OZCAM and AVH using the same BioCache backend have been created.

ALA has also set up a sandbox environment (sandbox.ala.org.au) of the BioCache that allows users to upload their data, run it through data quality tests and visualise the data on a map. This is currently considered an early release version that will eventually be a key tool of the Atlas.

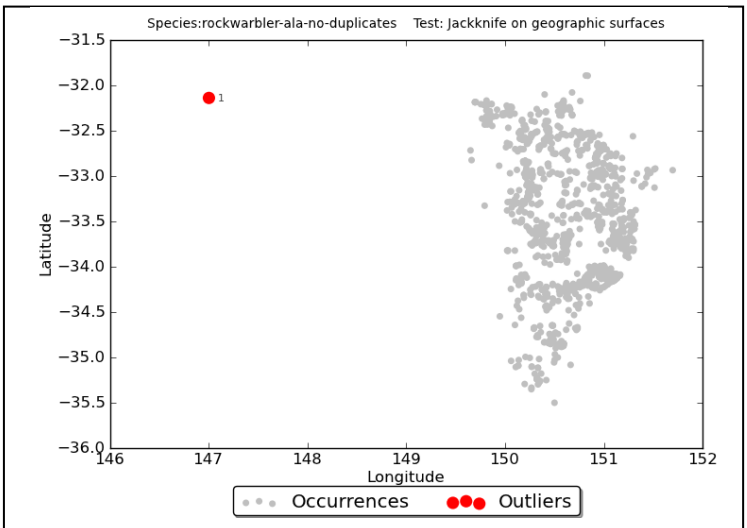
Data quality assurance

To assure the quality of its data, ALA has developed and is implementing a data quality framework.

A Data Quality Wiki has been established (code.google.com/p/ala-dataquality/) to facilitate discussion of quality issues with our data providers.

Initiatives are under way to enhance 'quality at source'. A completeness model for Darwin Core-based datasets has been devised and is being trialled in a number of faunal collections. Consideration is being given to developing a similar model for the herbarium standard, HISPID. To support this model, tools are being developed for use by our data providers to improve the completeness and standards-compliance of the data they provide to the Atlas based on the completeness model. Increased fitness-for-use of data is expected in due course.

A suite of data quality checks and metrics is also being developed and implemented. The suite will flag occurrence records that may have particular quality issues and advise the issues to both the data custodian—for review and remediation if appropriate—and users—to alert them to potential data quality shortcomings. Details of the quality checks already implemented and those planned to be rolled out are listed in a summary sheet on the Data Quality Wiki. Where necessary, detailed wiki pages are being developed to fully describe the checks and associated algorithms.



An example of a spatial outlier detected during trials of the reverse Jackknife method on occurrences of the Rock Warbler

The range of data quality tests will be expanded over the coming months.

ALA is also looking at the issue of spatial precision in our data—ie the uncertainty associated with the latitude/longitude of an occurrence record—and its consequent implications for fitness-for-use. Emphasis is being placed on the detection of potential spatial outliers. A reverse jackknife method has been documented (code.google.com/p/ala-dataquality/wiki/DETECTED_OUTLIER_JACKKNIFE) and is being implemented in the Atlas. Advice has also been received from statisticians at CSIRO Mathematics, Informatics and Statistics and the Australian National University on additional methods that could be employed. Several of these are presently being investigated and trialled.

Sensitive Data Service

Some species are of particular concern to researchers, resource managers and biosecurity officers because they are endangered or threatened, are pests, or because they are potential threats to Australian agriculture or other industries. Inappropriate release of data related to such species may cause harm to those species, people or Australia's trade relations.

The Sensitive Data Service (SDS) is being implemented in stages. In the first stage, the SDS lists species with a conservation status or conservation sensitivity aggregated from state and territory agencies (collections.ala.org.au/public/dataSets#filters=contentTypes%3Asensitive+species+lists). These lists allow records with conservation related sensitivity to be identified. Associated rules allow records to be modified or withheld as appropriate. Users are informed when data has been generalised or withheld. This functionality allows datasets to be made public while protecting sensitive species. For example, Loveridge's Frog is considered a sensitive species in NSW and as a result coordinate information is generalised.

> Due to sensitivity concerns, the coordinates of this record have been generalised: "Location in NSW generalised to 0.1 degrees Sensitive in NSW [Endangered, NSW DECCW]".

An SDS notice of geospatial generalisation for Loveridge's Frog (biocache.ala.org.au/occurrences/f489e55c-424e-4800-83b5-14ca91b4fb1f)

The SDS currently provides a prototype online service (sds.ala.org.au/) where data providers can check individual records or upload spreadsheets and have each record checked for sensitivity against publicly available criteria. ALA has also set up a sandbox environment (sandbox.ala.org.au) of the BioCache that allows users to upload their data, run it through sensitive data processing and data quality tests, and then visualise the data on a map. This is considered an early release version that will eventually be a key tool of the Atlas.

In addition to addressing conservation related concerns, ALA has been working with the Commonwealth Department of Agriculture, Fisheries and Forestry (DAFF) to collate rules to deal with data that has plant biosecurity-related sensitivities. An exposure document outlining these rules (www.ala.org.au/plant-biosecurity-sensitive-data-public-consultation-draft) is open for public comment until 21 October 2011. An interactive flowchart (www.ala.org.au/wp-content/uploads/2010/09/Draft_plant_biosecurity_rules_flowchart_30August2011.pdf) is also provided that gives an overview of the rules.

In the coming months ALA will work to incorporate plant biosecurity rules into the SDS and collate rules on animal health-related sensitivities.

Content delivery

Information in the Atlas is provided to the public, researchers, educators, natural resource managers and policy makers through a number of channels. In addition, ALA is providing tools to facilitate capture and management of biodiversity information.

Atlas website

The Atlas website is the front face of the Atlas, from which a user can access all information and analytical tools and functionality. See page 32.

Web services

Many features of the Atlas can be integrated with other sites and information systems using Atlas web services. See page 33.

Community science

ALA is producing an Open Source Citizen Science software toolkit to record and manage sightings of Australia's biodiversity by community groups as well as by researchers conducting field work. See page 34.

NRM engagement

A close working relationship has been forged with the Great Eastern Ranges (GER) Initiative to provide functionality specifically to meet the needs of Natural Resource Management groups. See page 35.

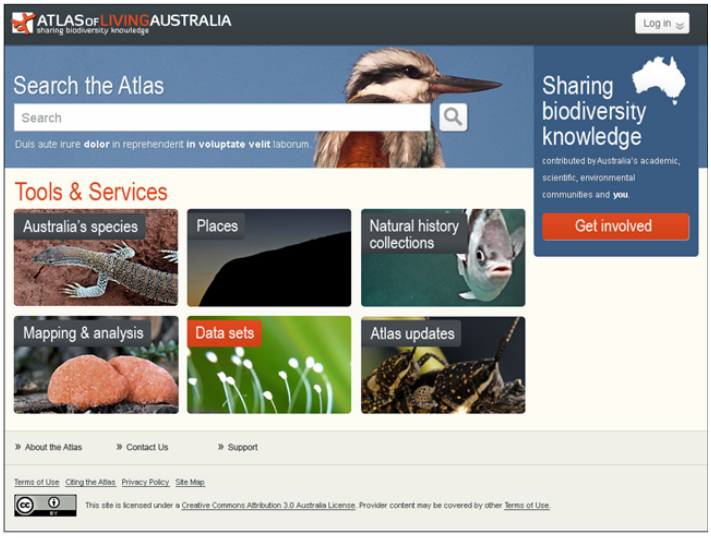
Hubs

ALA is working with botanical, zoological and other collection communities to develop their next virtual collections based on a specialist view of the Atlas. See page 36.

Atlas website

The Atlas website (www.ala.org.au) is where users access all aspects of the Atlas. Since the site was launched in October 2010 it has been constantly upgraded with additional datasets; new and improved functionality; better data presentation and discoverability (eg enhanced search interface from OZCAM); improved performance; and closer integration between modules.

Recently the BioCache (see page 28) was updated with a new database capable of storing a much broader range of data and supporting greater flexibility in analytical and searching tools. The Citizen Science toolkit (see page 34) has now been integrated into the ALA main site, allowing people to submit their field observations directly to the BioCache.



An example of a possible new Atlas home page

ALA is currently working on revamping the complete web site—not just with a new look and feel. Plans include substantial enhancements to usability, information accessibility and presentation, as well as new and enhanced tools and functionality. Many of these changes reflect extensive user feedback.

The revised look and feel will be implemented across the site before end of 2011.

Subsequent updates and additions to functionality may drive a further round of enhancements before project completion.

Web services

Many features of the Atlas can be integrated with other sites and information systems. This is primarily done through exposing Atlas features as 'web services', in addition to web pages. This approach also supports other ALA components, such as specialised OZCAM and AVH websites (hubs).

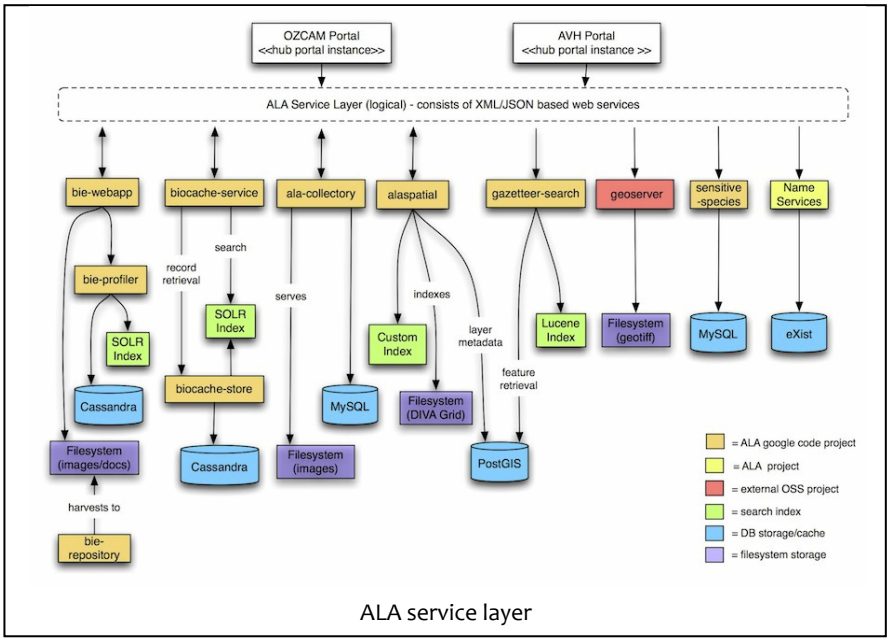
Achievements

The Atlas supports REST style web services offering JSON and XML. Between May and September 2011, several existing components were re-factored to better support exposing features as web services.

The web services have been proven in a production setting. Both the Atlas web site (www.ala.org.au/) and the prototype OZCAM site (ozcam-demo.ala.org.au:8080/) use the same services.

Development teams have produced some great technical documentation, which is invaluable for other developers. Examples are:

- Spatial Portal at code.google.com/p/alageospatialportal/wiki/GazetteerApi
- Biodiversity Heritage Library at bhl.ala.org.au/resources#api.



ALA service layer

Future

Over the coming months ALA will:

- continue to expose any new Atlas features as web services
- improve technical documentation where required
- add general documentation for business and management audiences. This will assist other parties to integrate particular features by describing the services that are involved.

A summary list of web services is at www.ala.org.au/tools-services/ala-web-services-list/.

Community science

ALA is producing an Open Source Citizen Science software toolkit to record and manage sightings of Australia's biodiversity (www.ala.org.au/tools-services/citizen-science/). Initially intended for the public to use, this software is now also in use by a range of scientific institutions for supporting field work and data collection on scientific surveys.

The Citizen Science toolkit software is designed to record and manage a diverse range of biodiversity (and non-biodiversity) data in standardised and structured ways. The software is highly customisable in both its look and feel and in the templates it uses to capture data. It can be used to generate stand alone web sites or be embedded in existing web sites. There is also a mobile component, which is currently being field tested by several institutions.

Web sites using ALA developed software

Currently the Citizen Science toolkit is used in a number of web sites including the Atlas itself (www.ala.org.au), ClimateWatch (www.climatewatch.org.au), Biodiversity Snapshots (www.biodiversitysnapshots.net.au), Birds Australia (birdsaustralia.ala.org.au) and Wild Backyards (www.wildbackyards.net.au). ALA is also working with a range of Natural Resource Management groups (see page 35) to install and configure the software. Several of these sites have won awards.

NRM engagement

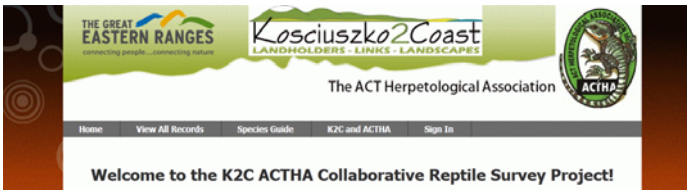
Following on from the early ALA user needs analysis and subsequent engagement across a broad range of stakeholder groups, ALA identified Natural Resource Management (NRM) and Catchment Management Authorities (CMA) as potential primary users of Atlas capability.

Accordingly, in 2011 ALA and the Great Eastern Ranges (GER) Initiative established a collaboration around NRM activities to explore requirements and provide functionality. These activities are proving very beneficial to both partners. The GER Initiative is trialling new ALA-developed tools that build on the existing Citizen Science toolkit as well as core Atlas and mapping analysis capabilities, to improve reporting and evaluation of biodiversity information and environmental outcomes.

These tools support broader engagement with an extended cross-section of NRM and related communities. The GER Initiative is helping ALA to improve its software, tools and services to meet the needs of regional NRM groups and projects across Australia.

The Great Eastern Ranges Initiative

The GER Initiative is establishing a 3400 km conservation corridor from the Victorian Alps to the Atherton Tablelands and beyond, along the Great Dividing Range and eastern escarpment.



Reptile survey project

ALA is working with the GER Initiative partners on five priority projects for trial and implementation:

- extending the Citizen Science toolkit to include on-ground activity planning, management, monitoring, evaluation and reporting
- increase the quality, reliability and frequency of monitoring and reporting on birds as indicators of landscape health
- developing new ways to use ALA components and communicate their value to users
- producing educational resources to be used when demonstrating the new ALA tools to other community groups
- sharing of partners' information in a web accessible space across NRM communities.

Each of the projects is taking paper- and people-based systems and implementing them in ALA citizen science tools while other enhancements are being made to the Atlas core to support NRM requirements.

For more information about the GER Initiative, see www.greateasternranges.org.au/.

Hubs

A data hub is a specific view into data held by ALA. ALA aggregates data into its BioCache (see page 28) from many different sources and a facet of that data related to a particular subject can be viewed through a hub. Because the ALA BioCache delivers content and tools through web services, an external website can make use of the ALA components best suited for the job.

The prototype OZCAM hub illustrates this (see ozcam-demo.ala.org.au:8080/). The OZCAM prototype sits inside a mock-up of the OZCAM site with its information about CHAFC and the Faunal Collections Informatics Group (FCIG), contacts, news and so on. An OZCAM search page has been created with search categories of interest to fauna people. A search returns results from the ALA BioCache, restricted to data provided by OZCAM member collections.

The ALA priority for development is to create an OZCAM and an AVH hub. This will be followed by other hubs—APPD for plant pests, OBIS for marine biology, AMRiN for microorganisms and hubs for botanic gardens living collections and seed banks.

ALA: The future

The Atlas of Living Australia has achieved a great deal during the first few years of its existence. Just as importantly, the ALA partners, along with ongoing support from CSIRO, will continue to develop this national resource. Now that these IT tools and services are in place, Australia is well-positioned to handle a massive expansion in biodiversity data. One major success has been the investment in developing accelerated systems and processes for digitising Australia's natural history collections. From this success, a future focus for ALA will be on supporting the collections community in creating open and accessible digital data on the millions of specimens currently undigitised.

The Atlas will continue to integrate collections data, along with significant data from groups such as State conservation agencies, Natural Resource Management authorities, other non-government organisations, amateur naturalists and the community. This work will, in turn, continue to build upon a data set of national significance—a data set which itself serves as fundamental infrastructure to support research. Noting that these data are already being applied in areas of environmental and ecological research, the future vision of ALA includes the continuing engagement of these groups to further their research efforts and to save them significant time previously spent in identifying, gathering (including negotiating licences), vetting and visualising data. The Atlas has been particularly strongly welcomed by NRM organisations, community groups and the public, and we look forward to enhancing their work to produce positive outcomes for conservation land management in Australia.

As it expands, the Atlas will also increasingly provide a platform for web-enabled taxonomic activity. Tools such as the TRIN Wiki can serve as taxonomic workbenches powered by the aggregated data sourced through the Atlas.

There are many other potential areas of value that should be mentioned—particularly in regards to areas such as metagenomics, name services, phylogenetics, literature, species identification, species interactions and so on—but these will have to be detailed in other papers. In closing, ALA partners are actively planning the next stages for the ALA project and will establish a stable and robust model for long-term governance and finance.

Atlas of Living Australia Report to Partners

October 2011