ANNUAL BUSINESS PLAN 2009-2010

for

The National Collaborative Research Infrastructure Strategy's Research Capability

known as

5.2 Integrated Biological Systems: 5.2.3 Biological Collections –

The Atlas of Living Australia

July 2009

Contents

1. Executive Summary

ALA has made strong progress against the 2008-2009 business plan and looks forward to expansion and delivery in the remaining two years of funding.

During 2008-2009, the ALA concluded a detailed user needs analysis and performed a prioritisation of candidate projects to be addressed during 2009-2011.

Several core ALA infrastructure components were developed during the year, including a portal to index specimens and observations of Australian organisms, new interfaces to the key reference lists of species occurring in Australia, and tools to enable users to annotate information shared through the ALA.

The ALA has continued to collaborate closely with its partner capabilities within NCRIS Integrated Biological Systems, in particular through the Phenomics Ontology-Driven Database (PODD) project, which will provide a data management framework for the Australian Phenomics Network and the Australian Plant Phenomics Network.

The ALA has also been developing collaborations with other NCRIS capabilities, particularly with the Terrestrial Ecosystem Research Network (TERN) to develop integrated systems for managing biodiversity data from ecological research and natural history collections, with the Integrated Marine Observing System (IMOS) and TERN to ensure consistent approaches to handling geospatial data, and with the Australian Biosecurity Intelligence Network (ABIN) to ensure the availability of ALA data to support biosecurity planning.

On 12 May 2009 an additional $30M was announced as part of the Australian Federal Budget to extend the work of the ALA during 2009-2011. These funds have been provided through the Education Infrastructure Fund (EIF) and are to support the existing goals of the project.

This significant increase in funding has necessitated a major revision of the ALA delivery plan for 2009-2011, which is now organised around six major component areas. The Data Integration component is to be funded from NCRIS funds. The remaining component areas are to be funded from EIF funds and are labelled ALA-SS (for "Atlas of Living Australia – Super Science").

1. **Collection Data Management** (ALA-SS, EIF) – create and develop tools and services to optimise the data supply chain through Australia's natural history collections.
2. **Rich Data Stores** (ALA-SS, EIF) – create and develop shared infrastructure to manage and maintain biodiversity data sets on behalf of Australian institutions and projects.
3. **Australian National Species Lists** (ALA-SS, EIF) – create and develop enhanced tools and services to organise Australian biodiversity data according to a shared understanding of the species found in the country and of their classification.
4. **Geospatial Data Management** (ALA-SS, EIF) – create and develop shared models, tools and services to ensure interoperability of all geospatial data accessed through the ALA and compatibility with data shared through related NCRIS capabilities (particularly TERN and IMOS).
5. **Data Integration** (ALA, NCRIS) – create and develop core services to catalogue and index Australian biodiversity data resources to maximise their accessibility and usefulness to researchers, policy-makers and the public.
6. **Data Dissemination** (ALA-SS, EIF) – create and develop web portals and applications to deliver biodiversity data to end-user communities.

Additional staff have been recruited to assist with planning and project management and to lead activity within these component areas. The ALA is well prepared to expand the scale of its activity in accordance with the increased funding.

2. Project Contents

2.1 Status of ALA Project

During 2008-2009, the ALA concluded key planning activities:

1. ***ALA User Needs Analysis*** – The ALA commissioned a wide-ranging consultation with providers and users of biodiversity information around Australia, resulting in the identification of the most significant requirements for the ALA to address.
2. ***ALA Project Priority Setting*** – Following from the User Needs Analysis, the ALA engaged a panel of experts to review candidate projects for the ALA to address during 2009-2011. Their report provides a ranked set of priorities for ALA delivery.
3. ***Evaluation of repository software*** – The ALA evaluated three potential technologies for the development of the Biodiversity Information Explorer, the core component for cataloguing and indexing information resources relating to Australian biodiversity. This activity identified Fedora Commons as the preferred platform for this work.

Several core ALA infrastructure components were developed during the year:

1. ***Geospatial Data Cache/Portal*** – The ALA has developed a national portal for exploring data on the occurrence of Australian species (from specimens and observations). This portal is based on the GBIF Data Portal open source software and has been enhanced 1) to map data against significant divisions of the Australian continent (e.g. states/territories, ecoregions, water catchments), 2) to organise data using Australian species lists, and 3) to integrate ALA tools allowing users to annotate data records.
2. ***AFD/APC Service*** – The ALA has funded a team at the Australian National Botanic Gardens (ANBG) to develop web service interfaces to improve the accessibility and usefulness of the Australian Faunal Directory (AFD) and the Australian Plant Census (APC), the key national reference lists of species occurring in Australia. These tools will enable the ALA and other applications to access these data sets and use them to organise biodiversity data.
3. ***Annotation Tools*** – The ALA has funded a team at the University of Queensland School of Information Technology and Electrical Engineering (UQ ITEE) to develop tools to enable users to annotate web pages, images and data records. These tools will be used extensively in ALA user interfaces and have already been incorporated in the ALA Geospatial Data Portal (see example).
4. ***Citizen science data capture*** – The ALA has funded a team at Gaia Resources to develop software components for the capture of citizen science data. These are already in use within the ClimateWatch project (with which the ALA has a Memorandum of Cooperation) and will be incorporated in the ALA's own citizen science components.
5. ***IdentifyLife Descriptlet Store*** – The ALA and the Encyclopedia of Life (EOL) have been collaborating with the University of Queensland Centre for Biological Information Technology (CBIT) to develop tools to support species identification as part of the IdentifyLife project. The IdentifyLife team has been developing data stores to hold species descriptions.
6. ***Metadata Repository (Biodiversity Information Explorer)*** – The ALA has funded a team at the CSIRO ICT Centre to develop the ALA Metadata Repository using the open source Fedora Commons repository software. The team has developed software around Fedora to allow the ALA to catalogue information resources and to store additional data on the species, geographic regions, habitats, etc. to which these resources relate.

2.2 Project Outlook

2.3 Research Infrastructure

Note: this section describes only ALA components to be developed during 2009-2010 using NCRIS funds. *Attachment 1 – ALA Implementation Plan* includes a description of all components to be developed during this period, including those to be funded from EIF funds.

2.3.1 Data Integration (ALA, NCRIS)

This activity will be addressed using NCRIS funds.

Data integration activities represent most of the core tasks funded under the original ALA NCRIS funding. (Information on how other tasks from the original plan are being handled can be found in ***Error! Reference source not found. Error! Reference source not found.***). The focus is on the development of tools and services to index primary information sources and provide integrated views to enable users to select and access those resources which are most relevant to their concerns.

The following components are under development:

1. **Ontologies and Vocabularies** – data integration within the ALA and with other projects (e.g. AVH, OZCAM, GBIF, EOL, OBIS) depends on a shared understanding of the structure of biodiversity data and agreement about the data elements which can be shared. The ALA is working with TDWG and international projects to set up tools to engage the community in developing and maintaining the ontologies and vocabularies required for this purpose. These structures will be particularly important to the ALA Metadata Repository and will provide the models to be used within that tool for organising Australian biodiversity information.
2. **Metadata Repository** – The core component within the ALA's data integration programme will be the Metadata Repository. This is being constructed using the Fedora open-source content management system. It will serve as a catalogue of biodiversity information resources (databases, documents, images, etc.) with provider-supplied metadata describing the origins and nature of each resource, but will be extended to link these resources to the species to which they relate, the geographic regions which they cover, etc. and to model the relationships between species, regions, habitats, descriptive characters, etc. (using information from tools such as the ALA Geospatial Data Cache). This will allow the ALA to produce web pages giving overviews of the available information relating to each species, region, habitat, etc. The Metadata Repository will therefore act as the engine serving information links to the proposed Data Dissemination components (especially Biodiversity Information Explorer and Biosecurity Portal).
3. **User Authentication and Identity Management** – The ALA will require the ability to authenticate users for many different purposes: to allow data providers to manage the metadata for their resources; to allow users to identify themselves to make annotations or provide additional data; for taxonomists to contribute to the Australian national checklists; etc. Building an integrated understanding of the expertise of each individual will also allow the ALA to improve its use of the information supplied by each user. The ALA has already held discussions with the Australian Access Federation (AAF) on the use of its services as the identity management and authentication framework for those users belonging to AAF-enabled organisations.
4. **Annotation Services** – The ALA has received funding from the NCRIS Platforms for Collaboration capability's NeAT programme to develop annotation services to enhance the quality of data and to enable end users to contribute new information to the network. This work is being carried out at the University of Queensland School of Information Technology

and Electrical Engineering and early versions of some of the tools have been integrated into the GBIF Data Portal software at http://data.ala.org.au/. As the ALA proceeds, these tools will be used in many ways, including capturing user suggestions for corrected values within data records, free-text user comments, user tagging of species with descriptive terms, etc.

5. **Data Quality and Sensitive Data Tools** – The ALA contracted a review of concerns around potentially sensitive data within state conservation agencies, natural history collections and biosecurity activities. The goal is to develop best practice recommendations on the handling of occurrence records of conservation or biosecurity concern (e.g. reduction of precision of coordinates for records of species considered endangered in the state where they have been recorded, or diverting records of pest species with implications for Australian trade to the relevant authorities), and then to provide easy-to-use services to scan sets of records (e.g. as a spreadsheet) to evaluate any possible issues and report back to the data provider. This will be an important tool to help data providers become comfortable about sharing data and also to allow the broader community to develop consistent approaches to handling records for sensitive taxa. The tool will also support a wide range of additional data validation and other checks. In this form it will become a major component in the ALA's approach to improving data quality. Records with issues can be reported to the data holders and can automatically be annotated with notes or suggested corrections. End users will also be able to use annotation tools to contribute to data quality. Ultimately all such annotations should be handled through workflows which capture responses from the data providers.

Implementation

These activities are already in progress but reorganisation of the ALA to reflect the additional EIF funding will allow more resources to be applied to ensure delivery of a comprehensive set of tools and services.

The following resources have been assigned to this activity:

- Project lead, David Martin (ALA Technical Architect, full-time, ALA, NCRIS)
- Java developer, Nick dos Remedios (full-time, ALA, NCRIS)
- Metadata Curator, Lynette Woodburn (full-time, ALA, NCRIS)
- University of Queensland annotation services development team, Ron Chernich, Stephen Crawley (full-time, NeAT funding) – seeking to recruit one additional developer using ALA, NCRIS funds
- CSIRO ICT Centre metadata repository development team, Hon Hwang, Nerolie Oaks (total 1.5 EFT, NeAT/ALA, NCRIS funding), Carsten Friedrich (part-time team lead, ALA, NCRIS) – seeking to recruit two additional developers to contribute to this work and to liaise with the APPF/APN Phenomics Ontology-Driven Database NeAT project
- APPF bioinformatician (full-time, funded from ALA, NCRIS funds to support IBS data management)
- APN bioinformatician (full-time, funded from ALA, NCRIS funds to support IBS data management)
- Developers (including test, documentation, etc.) – up to 4 additional EFT

2.3.2 Hardware and Networking

The ALA is currently using the following servers provided by CSIRO IM&T:

Machine	Role	Services	Specification
alaprodweb1-cbr.vm.csiro.au	Production web server	Website (www.ala.org.au) GIS Portal (data.ala.org.au) Mapping Services (maps.ala.org.au) Annotation services (annotate.ala.org.au)	Apache/Tomcat 4GB, 2 CPU, 200GB of storage
Alaproddb1-cbr.vm.csiro.au	Production database	Database for production services	MySQL

			8GB, 2 CPU, 500GB of storage
alatstweb1-syd.nexus.csiro.au	Test web server	Development versions of GIS Portal and Mapping Services	Apache/Tomcat 4GB, 2 CPU, 200GB of storage
alatstdb1-syd.nexus.csiro.au	Test database	Database for development services	MySQL 4GB, 2 CPU, 500GB of storage
Alaslvweb2-cbr.vm.csiro.au	Indexing web server	Java indexing processes	Apache/Tomcat 4GB, 2 CPU, 200GB of storage
alaslvddb1-cbr.vm.csiro.au	Indexing database	Database for Java indexing processes	MySQL 4GB, 2 CPU, 500GB of storage
diasbdev1-cbr.vm.csiro.au	DIAS-B development server	Web server and database for development of Metadata Repository	Tomcat, MySQL, Fedora Commons 4GB, 2 CPU, 200GB of storage
diasbtest1-cbr.vm.csiro.au	DIAS-B test server	Web server and database for test of Metadata Repository	Tomcat, MySQL, Fedora Commons 4GB, 2 CPU, 200GB of storage

During 2009-2010, the ALA will work with ARCS to plan and deploy additional storage and servers located around Australia to reflect the locations of activities developing ALA components. The ALA has started discussions with ARCS about the possibility of leveraging ARCS Super Science funded hardware.

2.4 Governance

The existing ALA Management Committee has been working well and consists of representatives from CSIRO, the peak bodies representing Australian natural history collections, DEWHA and three related NCRIS capabilities (TERN, ABIN, APN), currently under the chairmanship of Frank Howarth (Council of Australian Museum Directors). This committee acts in an advisory role providing high level direction, strategy and guidance to CSIRO and the ALA Director.

CSIRO has responsibility to manage the ALA project and NCRIS funds and ensure that the project is carried out in accordance with the obligations under the DIISR Funding Agreement. Following the allocation of the new funding, a new ALA Oversight Committee is being established with representation from CSIRO, the ALA Management Committee, DIISR and DEWHA (as a key stakeholder with interest in the products from the ALA). The existing Management Committee will continue in its existing role.

2.5 Collaboration

2.5.1 NCRIS capabilities

The ALA is one of three capabilities included within NCRIS 5.2 Integrated Biological Systems (IBS) and holds funds to assist the other capabilities (Australian Phenomics Network, APN, and Australian Plant Phenomics Facility, APPF) with their data integration. During 2009-2010 these funds will support Kai Xu (APN) and James Eddes (APPF) working on data standards, ontologies and organisation of phenomics data sets. In addition the ALA is a partner in the NeAT-funded Phenomics Ontology-Driven Database (PODD) project to develop a data repository and metadata catalogue for APN and APPF. This project will use many of the same technologies as the ALA Biodiversity Information Explorer (in particular the Fedora Commons repository software). The ALA is seeking an additional Fedora developer to work on matters of interest both to PODD and the Biodiversity Information Explorer. All IBS capabilities will undergo an external review in the week of 8-12 March 2010.

The ALA has been working closely with the Terrestrial Ecosystem Research Network (TERN) Ecoinformatics and Rangelands Monitoring components. The two capabilities are committed to shared delivery of geospatial data infrastructure to ensure full compatibility between data and tools from ALA and TERN. The ALA will coordinate this relationship will be coordinated through the

ALA Geospatial Data Management component.

The ALA is in discussion with the Integrated Marine Observing Facility (IMOS) on reuse of the IMOS metadata portal software and shared development of future enhancements to this tool. The ALA will coordinate this activity through the ALA Geospatial Data Management component.

The ALA has discussed interactions with the Australian Biosecurity Intelligence Network (ABIN). During 2009-2010 ABIN and the ALA have agreed to establish a Memorandum of Understanding (MOU) where all points of collaboration will be identified and managed. Areas of collaboration already identified include management of Sensitive Data and use of the ALA Citizen Science Portal.

2.5.2 Other Australian linkages

The following table highlights key consultations and linkages between the ALA and other agencies and projects in Australia:

Linkage	Areas discussed	Next steps
Department of Environment, Water, Heritage and the Arts (DEWHA)	Data management for Threat Abatement Plans for tramp ants and chytrid fungus	Candidate projects for ALA Citizen Science Portal
Department of Environment, Water, Heritage and the Arts (DEWHA)	Australian National Heritage Assessment Tool (ANHAT)	Explore use of ALA Geospatial Data Management to maintain ANHAT data caches
Department of the Environment, Water, Heritage and the Arts (DEWHA)	Environmental Resources Information Network (ERIN)	Explore use of ALA Geospatial Data Management to maintain ERIN data caches
Department of Agriculture, Fisheries and Forestry (DAFF)	Data management for recreational fisheries	Candidate project for ALA Citizen Science Portal
Earthwatch Australia	Climatewatch (http://www.climatewatch.org.au/)	Candidate project for ALA Citizen Science Portal
Australian Institute for Marine Science (AIMS)	Marine biodiversity data	Integrate data and tools
Australian Institute of Marine Science (Glenn De'ath)	e-Atlas	Continue discussions around code and data sharing
Australian Antarctic Data Centre	Antarctic biodiversity data	Integrate data and tools
CSIRO Marine and Atmospheric Research	Genus-level names and tools for matching scientific names	Evaluate linkages between ALA Australian National Species Lists work and IRMNG
Australian Museum	Software for Delta	Explore collaboration to provide Delta tools
University of New South Wales	Species interactions	Involve in Species Interactions work
Murray-Darling Freshwater Research Centre	Aquatic macroinvertebrate data and identification tool	Possible focus on aquatic macroinvertebrates as an indicator group
CSIRO Livestock Industries	Environmental genomics	Sequence data management and use of ecogenomics data in community analysis
University of Queensland	Data and tools to support ecological modelling	Continue consultation around ALA Geospatial Data Management
ANU Fenner School of Environment and Society	Bioclim and related environmental layers for portal	Identify environmental data requirements
Integrated Marine Observing System	Collaboration on spatial portal	Establish collaboration structure
CSIRO Marine and Atmospheric Research	Marine biodiversity analysis and CERF hub	Identify priority marine spatial analysis use cases and analysis
Terrestrial Ecosystems Research Network – Ecoinformatics	Spatial data requirements and analysis	Identify priority spatial analysis use cases and analysis
CSIRO Sustainable Ecosystems	Spatial data requirements and analysis	Identify environmental data requirements
Australian Tree Seed Centre	Database support for small collections	Potential to provide collection management tools
Australian National University (Archaeology and Natural History)	Australasian Pollen and Spore Atlas	Explore potential of sharing data through ALA
Australian National Algae Culture Collection	Management of a living collection	Potential to assist in exporting data
CSIRO Plant Industry	Dadswell Memorial Wood Collection Wood Inhabiting Fungi Collection	Continued consultation around small, under-supported collections
NSW Department of Environment,	State conservation agency Atlas of NSW	Potential to provide taxonomic,

Climate Change and Water (DECCW)	Wildlife	validation tools and services
Biodiversity Informatics Working Group represents: NSW DPI Fisheries NSW DPI Forests NSW DECCW NSW Herbarium Australian Museum	BioNet NSW	Potential to assist with mapping of BioNet NSW public data, data sharing
Museum Victoria	Imaging	Continue consultation about standards for specimen imaging
ANU Fenner School of Environment and Society	Conservation of wood library and collection	Potential to assist with methods for digitising and managing data of small collections
NSW DPI Forests	West Pennant Hills wood collection	Candidate project for small collection revival
NSW DPI	Small collections	Continued collaboration with species-species interactions
Tasmania Department of Primary Industries, Parks, Water and Environment (DPIPWE)	State conservation agency, Natural Values Atlas	Continued close collaboration, with taxonomy, sensitive data, data sharing
SA Department for Environment and Heritage (DEH)	State conservation agency data management	ALA to facilitate collaboration between all state conservation agencies. Potentially there are many important tools and services for ALA to deliver. Potential sharing of data through ALA
QLD Department of Environment and Resource Management (EPA)	State conservation agency, Wildlife Online	Potential to provide taxonomic, synchronization validation and mapping tools and services. Potential sharing of data through ALA.
Northern Territory Department of Natural Resources, Environment, The Arts and Sport (NRETAS) Biodiversity Conservation Unit	State conservation agency, Fauna Atlas	Potential to provide tools and services for names, public access to specific records, and assistance with delivery over narrow bandwidth
NT NRETAS Herbarium	State conservation agency, flora data management	Potential to provide tools and services for bibliography, mapping, field capture of site data. Potential sharing of data through ALA.
Western Australia Department of Environment and Conservation	State conservation agency, NatureMap	Linkages
DAFF	Australian Plant Pest Database (APPD)	Continued collaboration with data and its management
University of Sydney	Management of university-based collections	Continued consultation with data management
Museum and Art Gallery of the Northern Territory	Data management	ALA support for MAGNT, further engagement in ALA

2.5.3 International Linkages

The ALA continues to manage the payments for Australia's participation in the Global Biodiversity Information Facility (GBIF). The ALA has developed its geospatial data cache using GBIF software and will use this tool to manage the interface between Australian biodiversity data sets and the global cache maintained by GBIF.

The ALA has also been in discussions with the Consortium for the Barcode of Life, and specifically with the Barcode of Life Database project, to establish tools here for Australian projects to manage their barcode sequence data as part of the international barcoding programme.

The ALA continues its relationship with the Encyclopedia of Life and will begin integrating some relevant EOL data sets into the ALA during 2009-2010. A particular linkage will be with the Biodiversity Heritage Library (BHL). The ALA intends to establish a BHL node during 2009-2010.

The ALA has been identified as an Australian partner for two large international projects:

- 4D4Life – a European Union-funded project to accelerate the development of a global checklist of all species (through the work of the Catalogue of Life) – the ALA Australian National Checklists component will provide content and probably tools for this work.
- DataONE – a United States NSF-funded project to address long term data management for ecological data – the ALA (and TERN) will track this activity to identify tools and standards which can be adopted in Australia.

2.6 Promotion

The ALA's web site at http://www.ala.org.au/ (hosted on CSIRO infrastructure) is successfully being used for general communication about the project's goals and progress. Web traffic has been increasing (Jan-Apr 2009: 2,018 visits; May-Aug 2009: 3,872 visits). During 2008/09 the website was updated to incorporate support for user registration with associated wiki-based discussion facility, site feedback and newsletter mailing list

For 2009-2010 the ALA is engaging with the CSIRO communications team to prepare and co-ordinate additional communications and promotions - reflecting the broadened project scope and visibility to the public and other stakeholders.

2.7 Access and Pricing

There are no access and pricing issues associated with the Atlas of Living Australia. The vision for the Atlas is to the greatest extent possible to provide free and open access to information. As the Atlas proceeds, it is likely that the same infrastructure will also serve for more restrictive point-to-point sharing of restricted data between data owners and authorised users. This scenario is however likely to result in reduced functionality and interoperability for the data resources so secured.

An Intellectual Property Rights Policy has been developed and was included as Attachment 5 to the 2007-2008 Business Plan.

2.8 Long-term Viability

There are a number of complementary strategies being undertaken to ensure the sustainability of the ALA and its attractiveness for future external funding and support. Some of those strategies include:

- **High value systems***:* Successful delivery of systems by ALA to provide key services to state and federal agencies and others for handling species names and to support conservation/land-use planning and biosecurity activity will be the most appropriate way to justify seeking future support for operating costs from DEWHA, DAFF, etc.
- **Less expensive shared infrastructure***:* The systems to be developed will reduce the internal costs for each ALA partner to manage biodiversity data, making it feasible for the ALA partners to possibly contribute ongoing support costs beyond June 2011.
- **Integrated and project based funding**: The development of an integrated end-to-end model for managing biodiversity information will position the ALA and its partners to bid efficiently for project-based funds to create additional content and to enhance the systems and their interfaces.
- **Use of open source:** Where possible the ALA is reusing and contributing to open source components developed by the international biodiversity informatics community, with the prospect that maintenance for these systems will be shared with others.
- **Simple low cost mode**: The ALA is designing the infrastructure so that it is as cheap as possible to maintain it into the future. All the systems should operate in a simple low-cost mode during periods where additional funds are not available. The vast majority of data resources will be hosted by the

provider organisations. ALA hosting will be limited to central indexes and mirrors of a few international databases.

- **Hosting and development by others**: The implementation plan will closely involve different ALA partners in the development of major components based on the institution's interest in the areas concerned. In the extreme case of there being no further external funding, some of these components could be adopted by these institutions.

2.9 Financial and Human Resources

Balance forward	$3,427,912
NCRIS fourth payment	$1,548,000
Interest to June 2009	$143,600
Total NCRIS contribution (including interest)	**$5,109,512**
CSIRO cash	$1,000,000
Australian Museum cash	$100,000
Museum Victoria cash	$100,000
Queensland Museum cash	$100,000
Tasmanian Museum and Art Gallery cash	$70,000
University of Adelaide cash	$30,000
Southern Cross University cash	$50,000
Total participant cash contribution	**$1,450,000**
NeAT DIAS-B project	$400,000
EIF funds	$10,000,000
Total other cash contribution	**$10,400,000**
Total income all sources	**$16,959,512**

CSIRO will credit the ALA with interest for the NCRIS funds held during the period. Interest will be credited for 2009-2010 during end-of-year accounting and will be reported in the 2009-2010 progress report. Interest on EIF 2009-2010 funds will be credited in 2010-2011.

The following table summarises **in-kind co-investment** by ALA participants:

CSIRO	$500,000
Australian Museum	$400,000
Museum Victoria	$850,000
Queensland Museum	$78,000
Tasmanian Museum and Art Gallery	$80,000
University of Adelaide	$18,000
Southern Cross University	$145,000
Australian Virtual Herbarium	$1,000,000
DEWHR/ABRS	$321,000
Total participant co-investment	**$3,392,000**

The following table presents the budget for 2009-2011.

Note: this table details only budget for activities during 2009-2011 using NCRIS funds. *Attachment 2 - ALA Budget* includes budget for all components to be developed during this period, including those to be funded from EIF funds.

	2009-2010	2010-2011	Total	Notes

EXPENDITURE				
Output 1 Project Office				
1.1 Director salary	175,959	184,757	**360,716**	
1.2 Director Overheads (CSIRO in kind)	*140,767*	*147,806*	***288,573***	
1.3 Programme Manager	213,000	213,000	**426,000**	
1.4 Project Officer salary	107,163	112,521	**219,684**	
1.5 Project Officer Overheads (CSIRO) in kind	*85,730*	*90,017*	***175,747***	
1.6 CSIRO Line Management (10%) cash	*18,721*	*20,097*	***38,818***	
1.7 Operating / travel etc	40,000	40,000	**80,000**	
1.8 Hardware (staff and development)	5,000	5,000	**10,000**	
1.9 Informatics support for 5.2.1, 5.2.2	350,000	450,000	**800,000**	
1.10 Recruitment and relocation costs	5000	5000	**10,000**	
1.11 Publicity and outreach	150000	150000	**300,000**	
Total Spend	**1,291,340**	**1,418,198**	**2,709,538**	
Cash and in kind	***245,218***	***257,920***	***503,138***	
NCRIS Spend	**1,046,122**	**1,160,278**	**2,206,400**	
Output 6 - Data Integration				
6.1 Technical Architect	220,760	231,797	**452,557**	
6.2 Java Developer	186,733	196,070	**382,803**	
6.3 Metadata Curator	186,733	196,070	**382,803**	
6.4 DIAS-B project (NeAT in-kind)	*400,000*	*200,000*	***600,000***	
6.5 DIAS-B Annotation Services developers	130,000	260000	**390,000**	
6.6 DIAS-B Metadata Repository developers	70,000	180,000	**250,000**	
6.7 Developers	379,200	632,000	**1,011,200**	**Year 1: 4*60%, Year 2:4*100%**
Total Spend	**1,573,426**	**1,895,937**	**3,469,363**	
Cash and in kind	***400,000***	***200,000***	***600,000***	
NCRIS Spend	**1,173,426**	**1,695,937**	**2,869,363**	
Output 8 - International Engagement				
8.1 GBIF Membership	165,200	165,200	**330,400**	
8.2 TDWG Membership	500	600	**1,100**	
Total Spend	**165,700**	**165,800**	**331,500**	
Cash and in kind	***0***	***0***	***0***	
NCRIS Spend	**165,700**	**165,800**	**331,500**	
Output 9 - Governance				
9.1 External review of usability and function	25,000	25,000	**50,000**	
9.2 Operating / travel etc	30,000	30,000	**60,000**	
Total Spend	**55,000**	**55,000**	**110,000**	
Cash and in kind	***0***	***0***	***0***	
NCRIS Spend	**55,000**	**55,000**	**110,000**	
Output 10 - Network Infrastructure				
10.1 Hardware	50,000	200,000	**250,000**	
Total Spend	**50,000**	**200,000**	**250,000**	
Cash and in kind	***0***	***0***	***0***	
EIF Spend	**50,000**	**200,000**	**250,000**	
Output 11 – Populating the Atlas				
11.1 Australian Virtual Herbarium in-kind	*1,000,000*	*1,000,000*	***2,000,000***	
11.2 CSIRO cash	*500,000*	*500,000*	***1,000,000***	
11.3 CSIRO in kind	*500,000*	*500,000*	***1,000,000***	
11.4 Australian Museum cash	*100,000*	*100,000*	***200,000***	
11.5 Australian Museum in kind	*400,000*	*400,000*	***800,000***	
11.6 Museum Victoria cash	*100,000*	*100,000*	***200,000***	
11.7 Museum Victoria in kind	*850,000*	*850,000*	***1,700,000***	
11.8 Queensland Museum cash	*100,000*	*100,000*	***200,000***	

11.9 Queensland Museum in kind	*78,000*	*78,000*	***156,000***	
11.10 Tasmanian Museum & Art Gallery cash	*70,000*	*70,000*	***140,000***	
11.11 Tasmanian Museum & Art Gallery in kind	*80,000*	*80,000*	***160,000***	
11.12 University of Adelaide cash	*30,000*	*30,000*	***60,000***	
11.13 University of Adelaide in kind	*18,000*	*18,000*	***36,000***	
11.14 Southern Cross University cash	*50,000*	*12,500*	***62,500***	
11.15 Southern Cross University in kind	*145,000*	*38,000*	***183,000***	
11.16 ABRS in kind	*321,000*	*321,000*	***642,000***	
11.17 CSIRO External	*500,000*	*0*	***500,000***	
Total Spend	**4,842,000**	**4,197,500**	**9,039,500**	
Cash and in kind	***4,842,000***	***4,197,500***	***9,039,500***	
NCRIS Spend	**0**	**0**	**0**	
ALL OUTPUTS				
Total Spend	**7,927,466**	**7,732,435**	**15,659,901**	
Cash and in kind	***5,487,218***	***4,655,420***	***10,142,638***	
NCRIS Spend	**2,440,248**	**3,077,015**	**5,517,263**	

The *ALA Business Plan 2008-2009* provided a budget for 2006-2011, organised according to five Outputs:

1. Building the Atlas
2. Populating the Atlas
3. Tools for Data Discovery
4. International Engagement
5. Governance & Management

The increases to ALA funds have necessitated a revision of the entire budget and planned outputs. The project has now been reorganised to target eleven Outputs, five of which will be funded from NCRIS funds and six from EIF funds:

1. Project Office (ALA, NCRIS)
2. Collection Data Management (ALA-SS, EIF)
3. Rich Data Stores (ALA-SS, EIF)
4. Australian National Species Lists (ALA-SS, EIF)
5. Geospatial Data Management (ALA-SS, EIF)
6. Data Integration (ALA, NCRIS)
7. Data Dissemination (ALA-SS, EIF)
8. International Engagement (ALA, NCRIS)
9. Governance (ALA, NCRIS)
10. Network Infrastructure (ALA-SS, EIF)
11. Populating the Atlas (ALA, NCRIS in-kind)

The following table reproduces the detail from the budget presented in the *ALA Business Plan 2008-2009* for the final two years of the project, as well as the five-year total. For each item for which funds were allocated for 2009-2011, the table identifies the reorganised Output to which these activities are now assigned. Activities to be funded from NCRIS funds in 2009-2011 have been highlighted in gold and activities to be funded from EIF funds have been highlighted in green.

	Expenditure			New Output
	2009-2010	**2010-2011**	**5-year Total**	
TOTAL INVEST (original budget)			**34,786,255**	
NCRIS by year	**1,548,000**	**1,443,000**	**8,233,000**	

Cash and in kind	***5,146,539***	***4,510,198***	***26,553,255***		
EXPENDITURE (revised budget)					
Output 1 Building the Atlas					
1.1 Recruitment and relocation costs			**30,000**		
1.2 Project Leader salary	175,959	184,757	**614,946**	**1**	**Project Office**
1.3 Project Leader Overheads (CSIRO in kind)	*140,767*	*147,806*	***491,957***	***1***	***Project Office***
1.4 Operating / travel etc 5 yrs	30,000	30,000	**95,000**	**1**	**Project Office**
1.5 User needs analysis			**72,853**		
1.6 Technical Architect	220,760	231,797	**610,242**	**6**	**Data Integration**
1.7 Java Developer	186,733	196,070	**560,643**	**7**	**Data Dissemination**
1.8 External review of usability and function		25,000	**50,000**	**1**	**Project Office**
1.9 Hardware (staff and development)	10,000	10,000	**65,000**	**1**	**Project Office**
1.10 Taxonomic Names Service	100,000	100,000	**400,000**	**4**	**Australian National Species Lists**
1.11 Data hosting	70,000	70,000	**210,000**	**10**	**Network Infrastructure**
1.12 Data Provision Services	200,000	200,000	**700,000**	**2**	**Collection Data Management**
1.13 Informatics support for 5.2.1, 5.2.2	350,000	350,000	**1,000,000**	**1**	**Project Office**
1.14 Contingency	276,000	276,000	**828,000**		
Total Spend	**1,760,219**	**1,821,430**	**5,728,641**		
Cash and in kind	***140,767***	***147,806***	***491,957***		
NCRIS Spend	**1,619,452**	**1,673,624**	**5,236,684**		
Output 2 - Populating the Atlas					
2.1 Australian Virtual Herbarium cash			***741,195***	***11***	***Populating the Atlas***
2.2 Australian Virtual Herbarium in-kind	*1,000,000*	*1,000,000*	***7,758,388***	***11***	***Populating the Atlas***
2.3 CSIRO cash	*500,000*	*500,000*	***2,000,000***	***11***	***Populating the Atlas***
2.4 CSIRO in kind	*500,000*	*500,000*	***2,788,458***	***11***	***Populating the Atlas***
2.5 Australian Museum cash	*100,000*	*100,000*	***523,602***	***11***	***Populating the Atlas***
2.6 Australian Museum in kind	*400,000*	*400,000*	***2,448,238***	***11***	***Populating the Atlas***
2.7 Museum Victoria cash	*100,000*	*100,000*	***627,746***	***11***	***Populating the Atlas***
2.8 Museum Victoria in kind	*850,000*	*850,000*	***4,413,883***	***11***	***Populating the Atlas***
2.9 Queensland Museum cash	*100,000*	*100,000*	***674,078***	***11***	***Populating the Atlas***
2.10 Queensland Museum in kind	*78,000*	*78,000*	***392,000***	***11***	***Populating the Atlas***
2.11 Tasmanian Museum & Art Gallery cash	*70,000*	*70,000*	***350,000***	***11***	***Populating the Atlas***
2.12 Tasmanian Museum & Art Gallery in kind	*80,000*	*80,000*	***400,000***	***11***	***Populating the Atlas***
2.13 University of Adelaide cash	*30,000*	*30,000*	***154,518***	***11***	***Populating the Atlas***
2.14 University of Adelaide in kind	*18,000*	*18,000*	***77,490***	***11***	***Populating the Atlas***
2.15 DAFF (APPD) cash			***378,589***		
2.16 Southern Cross University cash	*50,000*	*12,500*	***212,500***	***11***	***Populating the Atlas***
2.17 Southern Cross University in kind	*145,000*	*38,000*	***618,000***	***11***	***Populating the Atlas***
2.18 ABRS in kind	*321,000*	*321,000*	***1,605,000***	***11***	***Populating the Atlas***
2.19 CSIRO External	*500,000*		***1,500,000***	***11***	***Populating the Atlas***
Total Spend	**4,842,000**	**4,197,500**	**27,663,685**		
Cash and in kind	***4,842,000***	***4,197,500***	***27,663,685***		
NCRIS Spend	**0**	**0**	**0**		
Output 3 – Tools for Data Discovery					
3.1 Metadata Curator	186,733	196,070	**560,643**	**6**	**Data Integration**
3.2 Tools survey			**35,000**		
3.3 DIAS-B project (NeAT in-kind)	*400,000*	*200,000*	***1,000,000***	***6***	***Data Integration***
3.4 DIAS-B Annotation Services developer	120,000		**240,000**	**6**	**Data Integration**
3.5 DIAS-B project facilitation	10,000	10,000	**50,000**	**6**	**Data Integration**
3.6 Online Identification Services	120,000	120,000	**360,000**	**3**	**Rich Data Stores**
3.7 GIS Services	120,000	120,000	**360,000**	**5**	**Geospatial Data Management**
Total Spend	**956,733**	**646,070**	**2,605,643**		
Cash and in kind	***400,000***	***200,000***	***1,000,000***		

NCRIS Spend	**556,733**	**446,070**	**1,605,643**		
Output 4 - International Engagement					
4.1 GBIF Membership	165,200	165,200	**826,000**	**8**	**International Engagement**
4.2 TDWG Membership	500	600	**1,900**	**8**	**International Engagement**
Total Spend	**165,700**	**165,800**	**827,900**		
Cash and in kind	***0***	***0***	***0***		
NCRIS Spend	**165,700**	**165,800**	**827,900**		
Output 5 - Governance & Management					
5.1 Operating / travel etc 5 yrs	50,000	50,000	**184,000**	**9**	**Governance**
5.2 Branding					
5.3 Project Officer salary	107,163	112,521	**370,344**	**9**	**Governance**
5.4 Project Officer Overheads (CSIRO) in kind	*85,730*	*90,017*	***296,275***	***1***	***Project Office***
5.5 CSIRO Line Management (10%) cash	*18,721*	*20,097*	***72,673***	***1***	***Project Office***
Total Spend	**261,614**	**272,635**	**931,792**		
Cash and in kind	***104,451***	***110,114***	***368,948***		
NCRIS Spend	**157,163**	**162,521**	**562,844**		
ALL OUTPUTS					
Total Spend	**7,986,267**	**7,103,435**	**37,757,661**		
Cash and in kind	***5,487,219***	***4,655,419***	***29,524,590***		
NCRIS Spend	**2,499,048**	**2,448,015**	**8,233,071**		

See *Attachment 7 - Details of participant contributions by organisation* 2009-2010 for details of planned activities by each participant during the current period.

During 2009-2010 The following staffing positions will be funded by the ALA using NCRIS funds:

Director (Donald Hobern)	CSIRO Entomology
Programme Manager (Peter Doherty)	CSIRO Entomology
Executive Officer (Wolf Wanjura)	CSIRO Entomology
Technical Architect (David Martin)	CSIRO Entomology
Metadata Curator (Lynette Woodburn)	CSIRO Entomology
Java Developer (Nick dos Remedios)	CSIRO Entomology
Business Analyst (Peter Brenton)	CSIRO Entomology
Mouse Phenomics Bioinformatician (Kai Xu)	ANU
Plant Phenomics Bioinformatician (James Eddes)	University of Adelaide
Metadata Repository lead (Carsten Friedrich, 25%)	CSIRO ICT Centre
Annotation Services developer (TBA)	UQ ITEE
Metadata Repository developer (TBA)	CSIRO Entomology
Metadata Repository developer (TBA)	CSIRO Entomology
Communications Officer (Julie Carter, 20%)	CSIRO Entomology

2.10 Milestones

Note: this section lists only milestones relating to activities during 2009-2010 using NCRIS funds. *Attachment 3 - Activities and Milestones for 2009-2010* includes milestones for all components to be developed during this period, including those to be funded from EIF funds.

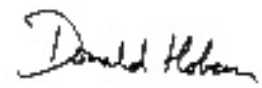
Activities and Milestones for 2009-2010	Due Date	Status
Output 1. Project Office		
Programme Officer appointed	1 Jul 2009	Complete
ALA all-hands coordination workshop	31 Oct 2009	ALA-SS, EIF

		Initial Milestone
Output 6. Data Integration		
Harvesting Demonstrator	15 Aug 2009	Complete
Revised project plan agreed by ALA MC	15 Oct 2009	
First public release of ALA components	28 Feb 2010	
Second public release of ALA components	31 May 2010	
Third public release of ALA components	30 Aug 2010	
Fourth public release of ALA components	30 Nov 2010	
Fifth public release of ALA components	28 Feb 2011	
Sixth public release of ALA components	31 May 2011	
Output 8. International Engagement		
Renewal of GBIF membership	31 Dec 2009	
ALA workshops (project publicity and user needs) at TDWG conference	9 Nov 2009	
Output 9. Governance		
09/10 Management Committee Meeting 1	9 Oct 2009	
09/10 Management Committee Meeting 2	31 Dec 2009	
09/10 Management Committee Meeting 3	31 Mar 2010	
09/10 Management Committee Meeting 4	30 Jun 2010	
09/10 CSIRO Oversight Committee Meeting 1	22 Oct 2009	
09/10 CSIRO Oversight Committee Meeting 2	31 Mar 2010	
IBS Review	12 Mar 2010	
Output 10. Network Infrastructure		
Establishment of agreement with ARCS	15 Mar 2010	
Output 11. Populating the Atlas		
Report on progress against activities proposed by ALA participants for 2008/2009	30 Sep 2009	
Report on progress against activities proposed by ALA participants for 2009/2010	30 Sep 2010	

2.11 Attachments

There is no confidential information attached, although several supplementary attachments are provided in appendices.

Donald Hobern
Project Director



29 September 2009

3. Attachments

3.1 Attachment 1 – ALA Implementation Plan

Note: this section includes all components planned for ALA development in 2009-2011, including those to be funded from NCRIS funds and those to be funded from EIF funds.

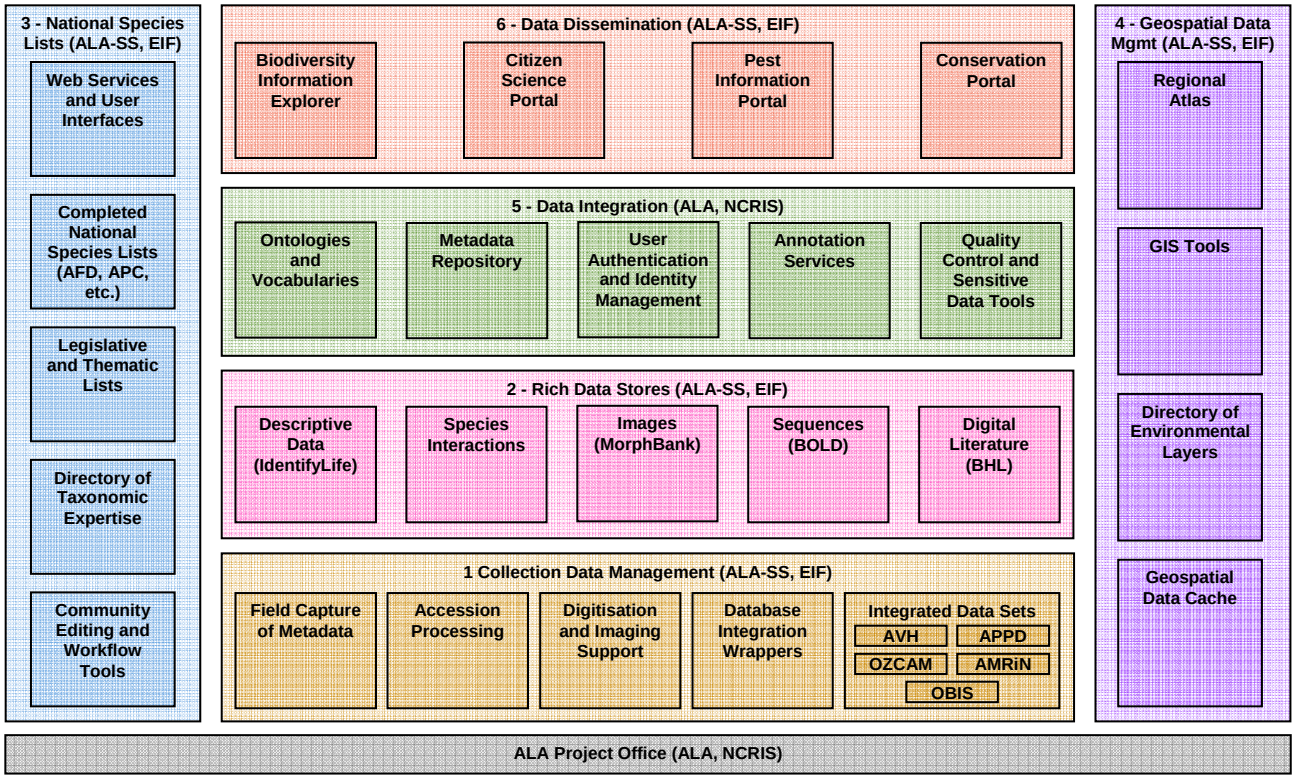


Figure 1: ALA project components

The ALA has revised its delivery plans since the 2008-2009 Business Plan to reflect the additional EIF funding made available for 2009-2011. This section provides an overview of all planned components to reflect the high degree of interconnectedness between components. Those components to be funded from ALA NCRIS funds are marked "ALA, NCRIS". Those components to be funded from ALA EIF funds are marked "ALA-SS, EIF".

The project will deliver research infrastructure organised into six major component areas:

1. **Collection Data Management** (ALA-SS, EIF) – create and develop tools and services to optimise the data supply chain through Australia's natural history collections.
2. **Rich Data Stores** (ALA-SS, EIF) – create and develop shared infrastructure to manage and maintain biodiversity data sets on behalf of Australian institutions and projects.
3. **Australian National Species Lists** (ALA-SS, EIF) – create and develop enhanced tools and services to organise Australian biodiversity data according to a shared understanding of the species found in the country and of their classification.
4. **Geospatial Data Management** (ALA-SS, EIF) – create and develop shared models, tools and services to ensure interoperability of all geospatial data accessed through the ALA and compatibility with data shared through related NCRIS capabilities (particularly TERN and IMOS).
5. **Data Integration** (ALA, NCRIS) – create and develop core services to catalogue and index

Australian biodiversity data resources to maximise their accessibility and usefulness to researchers, policy-makers and the public.

6. **Data Dissemination** (ALA-SS, EIF) – create and develop web portals and applications to deliver biodiversity data to end-user communities.

3.1.1 Collection Data Management (ALA-SS, EIF)

This activity will be addressed using EIF funds.

The goal of this project component is to optimise the supply of data through Australia's natural history collections. Collections occupy a central role in understanding the state of biodiversity in the country and in responding intelligently to changes in biodiversity. The herbaria and faunal and microbial collections house the reference materials and expertise required for identification of Australian organisms and for carrying out taxonomic research to understand the range of species found on the continent. The collections also have significant holdings of literature, images and other data of relevance to the ALA. Ecological monitoring activities use the collections for identification of materials and for long-term storage of voucher specimens.

The ALA has a two-fold interest in supporting data management within collections. First, the collections need to be able to process new specimens from field activities (e.g. long-term ecological monitoring work, environmental impact assessments) as efficiently and rapidly as possible. Data should flow smoothly from initial documentation of field work, through accessioning, curation, imaging and digitisation of specimens, and into applications to support monitoring, conservation, land-use planning, etc. Secondly, the collections still hold millions of historical specimens for which no data are available in a digital format. Tools and processes need to be in place to enable the collections to bring as much as possible of this information online. For many species such records are the richest available source of data on their historical distribution. For most species this data will help us to develop a clearer understanding of their environmental and habitat requirements.

The ALA has therefore established a working group with representatives from CHAH, CHAFC, CHAEC and CHACM, with John Tann contracted to coordinate their discussions. The working group is consulting with collections of all types and sizes and has been tasked with identifying a set of infrastructure components which the ALA can fund or develop to improve the capacity of collections to derive reliably-managed data, images and other information resources from specimens. The group will provide its recommendations in October 2009.

The ALA Collection Data Management component will include activities in the following areas:

1. **Field capture of metadata** – create and develop tools and processes to ensure that basic information is captured when specimens are collected (locality, coordinates, date, collecting team, methods, taxonomic groups collected, etc.). Centralised management of such metadata can simplify subsequent digitisation activities, allow preliminary information to be represented through GIS tools, advertise the existence of materials to researchers, and improve subsequent integration of specimen records as part of a single sample.
2. **Accession processing** – create and develop tools and processes to maximise the capture of information (data and images) as new materials are added to collections.
3. **Digitisation and imaging support** – create and develop tools and processes to improve the efficiency of data and image capture for all collections. These activities are expected to include development of tools to enable researchers and other users to annotate images and to assist in remote curation of materials.
4. **Database integration wrappers** – improve availability and reliability of software packages to assist collections and other holders of biodiversity data to connect their data to the ALA in well-structured formats. This work will also support the ongoing development of the AVH and

OZCAM as integrated caches of specimen records from herbaria and faunal collections respectively.

5. **Integrated data sets** – create and develop tools to support the ongoing operation of the key national integrated caches of collection-based biodiversity data (Australia's Virtual Herbarium, AVH; Online Zoological Collections of Australian Museums, OZCAM; Australian Plant Pest Database, APPD; Australian Microbial Resources Information Network, AMRIN; and Ocean Biogeographic Information System, OBIS). These data sets provide efficient community-driven hubs to organise and integrate data to feed into the ALA. Consequently the ALA will partner with these activities to ensure their robustness, compatibility and long-term viability.

Implementation
The ALA has contracted John Tann as project lead to coordinate activity around Collection Data Management and to convene an expert panel to make recommendations on the optimal use of ALA funds in this area. The panel comprises representatives from each of the peak bodies representing Australian natural history collections:

- Penny Berents (CHAFC)
- Jim Croft (CHAH)
- John Jennings (CHAEC)
- Lindsay Sly (CHACM)

The panel will report with its recommendations in October 2009. A working draft overview of the areas under investigation is included as *Attachment 5 - Candidate services and processes for Collection Data* Management.

3.1.2 Rich Data Stores (ALA-SS, EIF)

This activity will be addressed using EIF funds.
Many classes of biodiversity data are complex and difficult to manage in a consistent way between different institutions. In particular, many institutions lack sufficiently robust and flexible web infrastructure to publish identification keys, images, sequence data and digital literature online. The ALA, therefore intends to develop a set of shared national repositories to handle these more complex data types. Each of these repositories will be enhanced to simplify upload of data by partner institutions and to facilitate the presentation of this data both from within ALA portals and also within institutional web sites. These repositories will manage data as a set of projects from contributing institutions, each such project being branded to identify the source institution and associated with project metadata.

Five major sub-components have been identified as part of the ALA Rich Data Stores activity:

1. **Descriptive Data** - The ALA is working with the Encyclopedia of Life and the University of Queensland Centre for Biological Information Technology (CBIT) on IdentifyLife (http://www.identifylife.org/), a project to develop shared stores for species descriptive data and identification keys and to facilitate the population of these stores and the reuse of the data as a core element within the ALA infrastructure. The ALA will also provide support to ensure that IdentifyLife can support descriptive data in the Delta format.
2. **Images** – The ALA plans to establish a national mirror of the Morphbank (http://www.morphbank.net/) biological image repository (currently around 226,000 images) and to enhance upload/download interfaces for Australian projects. Morphbank provides a platform for disparate projects to share their images and to manage a wide range of metadata for each image, including placement within a taxonomic hierarchy, geospatial data and morphological tags.

3. **Sequences** – The ALA plans to establish a national mirror of the Barcode of Life Database (BOLD, http://www.barcodinglife.org/views/taxbrowser_root.php, currently around 688,000 records) and to enhance upload/download interfaces for Australian projects to store molecular sequences (barcode sequences and others). Many Australian projects are already contributing to thematic barcode networks (e.g. AllLeps, FishBOL, TreeBOL). An Australian node would give the opportunity to provide an integrated national view of all of this data and of data from other Australian sequencing projects. It would also provide a focus for integrating sequence data into the ALA's GIS capabilities.
4. **Digital Literature** – The ALA plans to establish a national mirror of the Biodiversity Heritage Library (BHL, http://www.biodiversitylibrary.org/) digital biodiversity literature archive (currently around 15,800,000 pages) and to enhance upload/download interfaces for Australian projects. The BHL platform allows publications to be accessed in a range of formats. New BHL projects are under way or starting in China, Europe and Japan. All of these will contribute to the global pool of accessible digital literature. BHL is keen to establish a replica node in Australia and to assist the ALA and its partners in planning and executing a scanning strategy here. The existence of such infrastructure could serve as a focus for project-based contributions of relevant literature and to explore collaborations with Australian libraries and publishers. The ALA will also explore options for infrastructure to accelerate the scanning of relevant published materials within Australia and for automated text mining of literature for key terms and concepts.
5. **Species Interactions** – The ALA has identified a create tools to simplify management and access of simple data records identifying interactions between species (predator/prey, pollinator/plant, host/parasite, etc.) derived from specimen records, literature and field observations.

The ALA has already held preliminary discussions with the MorphBank, BOLD and BHL projects on the use of their software and replication of their data stores. All three of these projects already hold some data relating to Australian biodiversity. Adopting their solutions will help the ALA to minimise initial costs for developing equivalent function and will allow future costs in maintaining these open-source components to be shared between the ALA and other partners around the world. Other mirrors are also being established for these projects in other countries under similar arrangements. The ALA would then work with these projects to enhance the software to address national requirements and to integrate cleanly with processes and web infrastructure within Australian institutions.

Implementation

The ALA is identifying expert project leads for each of the sub-components and will subcontract or second these experts in a part-time capacity to coordinate work in their area. The following experts have already been identified:

- Kevin Thiele (Western Australian Herbarium) – Descriptive Data
- Ely Wallis (Museum Victoria) – Digital Literature
- Gerry Cassis (University of New South Wales) – Species Interactions

The ALA is already funding one developer working at CBIT on the IdentifyLife components and has received initial requirements definitions for work to support descriptive data in the Delta format (from Jim Lowry, Australian Museum) and for the Species Interactions component (from Gerry Cassis).

The ALA expects to allocate the following additional resources to the development of the Rich Data Stores components:

- Business analyst – 50% EFT

- Developers (including test, documentation, etc.) – up to 10 EFT

3.1.3 Australian National Species Lists (ALA-SS, EIF)

This activity will be addressed using EIF funds.

All of the ALA's activities depend on understanding how to combine information on Australian biodiversity from different sources. These sources vary in their use of scientific and common names for the organisms in question. Some of these variations reflect earlier and later understanding of the taxonomy for a given group. Others reflect differing scientific opinion on the most appropriate way to interpret biological variation. Others reflect the number of species in Australia which have not yet formally been described and named or the difficulties of workers in different times and places in discovering previously published names for a species.

Information is required both on the published names for organisms and on the application of these different names to refer to currently accepted species concepts. There is also a need for global data sets (particularly for published names) and also for a well-curated set of national species lists which identify all the native and non-native species found in Australia and the various names that have been applied to these species.

These national species lists provide the framework for the ALA to integrate its data and in effect serve as the primary keys for all data records. The ALA consequently depends on these lists as a critical part of the national biodiversity informatics infrastructure. These lists are also of great importance to state and federal agencies and other organisations in structuring their own data sets and in supporting the merging of data from multiple sources.

This information exists in a number of national and international data sets, in particular:

- Australian Plant Name Index (APNI) – published names for Australian plants
- Australian Plant Census (APC) – national consensus view of the species of plants found in Australia and the various published names relating to each of these species (partial checklist – major families to be added)
- Australian Faunal Directory (AFD) – national consensus view of the species of animals found in Australia and the various published names relating to each of these species (partial checklist – major families to be added)
- AusMoss – published names
- Census of Freshwater Algae of Australia – published names
- Australian Marine Algal Name Index – published names
- Checklist of Australian Liverworts and Hornworts – national consensus view of liverwort/hornwort species found in Australia and the various published names relating to these species
- Interactive Catalogue of Australian Fungi – published names
- International Plant Name Index (IPNI) – published names of world plants
- Index Fungorum (IF) – published names of world fungi
- ZooBank – published names of world animals (new resource, only partially populated)
- Catalogue of Life (CoL) – global consensus view of species from all groups and the various published names relating to each of these species (partial checklist – major groups to be added)
- World Register of Marine Species (WoRMS) – global consensus view of species from all marine groups and the various published names relating to each of these species (partial checklist, will feed into CoL)
- Interim Register of Marine and Non-marine Genera (IRMNG) – global list of genus-level names

from all groups (maintained by CSIRO Marine, 90-95% complete)

As indicated, most of the existing resources are currently incomplete, reflecting the significant history of taxonomic work preceding the Internet age. However there is significant activity globally in improving and integrating these data sets. The ALA has been identified as a partner in the latest phase of activity for the Catalogue of Life and has been in discussion with Landcare New Zealand about synergies between the ALA Australian National Species Lists component and the New Zealand Organism Register, which has similar goals for New Zealand.

The ALA Australian National Species Lists activity will create and develop infrastructure around the core data sets on Australian published names and species. It will also address the use of these lists as data management tools in government and other agencies. The goals for the activity are to:

- Engage taxonomists to complete the national species lists (APNI/APC, AFD, etc.) for all groups to the fullest extent possible.
- Build a distributed editing framework to support this activity and to facilitate workflows for approval and effort-logging.
- Develop tools to allow other classifications to be related to these species lists (CoL, red lists, CITES, existing classifications from databases, etc.)
- Integrate TAXAMATCH (http://www.cmar.csiro.au/datacentre/taxamatch.htm) and known synonyms at all stages to guide users and tools to the appropriate species.
- Use CoL and IRMNG to provide at least working answers for the appropriate placement for names not currently included on national species lists.
- Support informal names of various kinds (common names, unpublished names for species recorded in field work, etc,).
- Manage a stable system of unique identifiers for species concepts sufficient to allow a user or tool to track future taxonomic changes for concepts for which they have stored such an identifier.
- Provide a range of web service and browser interfaces to access and explore the data.
- Provide interfaces for government agencies and others to simplify associating their own database records with the species lists.
- Provide interfaces for approved agencies to request handling for names not yet included in the checklists.
- Provide workflow processing to assign unhandled names to the appropriate taxonomic expert and update the checklists to place the name as a valid species name, synonym, misspelling, etc.
- Integrate data from Australian species lists into the global data sets such as CoL and WoRMS.

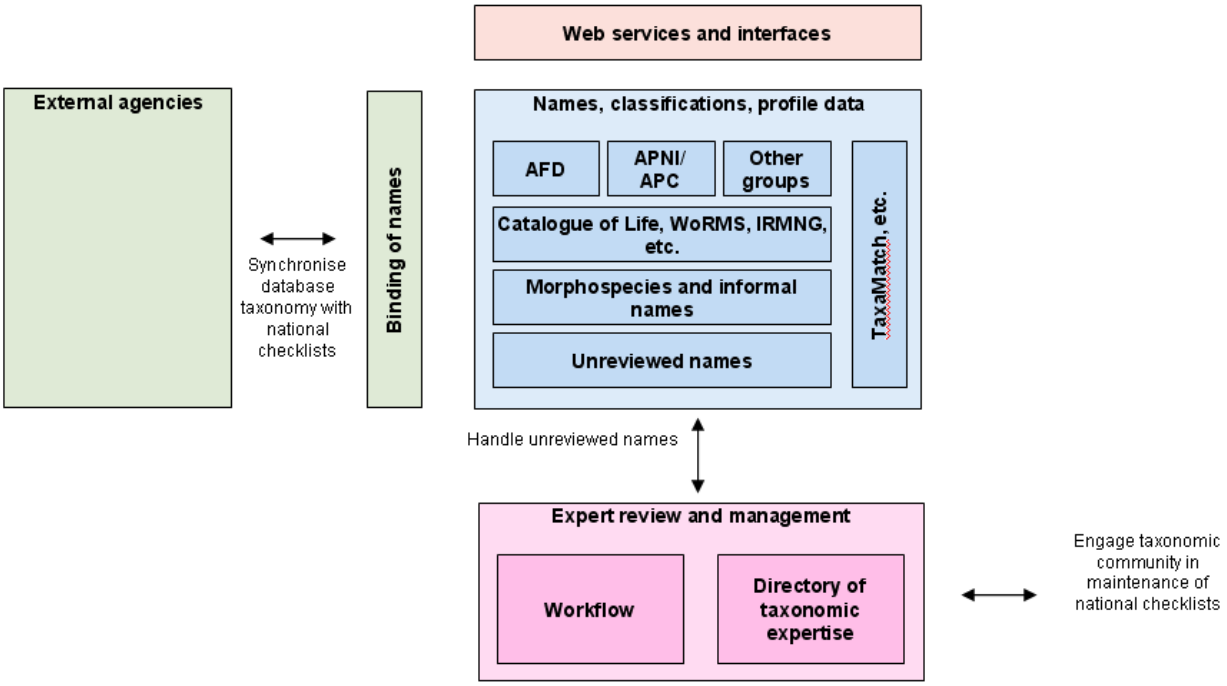


Figure 2: High-level model for Australian National Species Lists component

Five major sub-components have been identified for this activity:

1. **Community Editing and Workflow Tools** – create and develop software and processes to support collaborative editing and maintenance of these lists.
2. **Directory of Taxonomic Expertise** – create a registry of taxonomists and others with expertise to maintain components within these lists (and to support other information needs).
3. **Completed National Species Lists** – develop national species lists to the fullest extent possible.
4. **Legislative and Thematic Lists** – create and develop tools to support mapping other species lists (lists of species of legislative significance, state lists, alternative taxonomic views, etc.) against the national species lists.
5. **Web Services and User Interfaces** – create and develop tools to support the use of data from these lists by the ALA and other user groups.

Implementation
The ALA has received preliminary proposals from the Australian Biological Resources Study (ABRS) and from the Council of Heads of Australasian Herbaria (CHAH) for projects to complete the national species lists (ABRS for the fauna, CHAH for the flora and fungi). A part-time project lead will be identified to coordinate these activities.

The ALA has secured the continued involvement of the development team based at ANBG:

- Greg Whitbread (Architect)
- Paul Murray (ANBG developer)
- Brendan Douglas (ANBG developer)
- Christy Geromboux (ABRS analyst)

In addition, the ALA has allocated the following resources to advance the development of the software components:

- Peter Brenton (ALA business analyst, 50% allocation)
- Haydn Lowe (Wollemi Systems, lead developer, 80% allocation)
- Bruce Hyslop (Wollemi Systems, developer, 80% allocation)

Up to two additional developers may be allocated to this task as requirements are further specified.

3.1.4 Geospatial Data Management (ALA-SS, EIF)

This activity will be addressed using EIF funds.
All project elements dealing with the organisation, integration, analysis and visualisation of geospatial data are being coordinated as a single project area, to simplify coordination with TERN and IMOS and collaboration with other groups in Australia developing solutions in the same area (especially AVH and OZCAM). The goal is to deliver a well-integrated reusable set of services and to ensure that any project developing GIS outputs using Australian data can benefit from tools and visualisations developed by other projects.

The following components are already under development with the existing ALA work programme:

1. **Geospatial Data Cache** – the ALA is deploying an instance of the GBIF Data Portal software which has been customised to provide richer information management within the national context (mapping all records against smaller divisions of the continent such as local government areas, water catchments, nature reserves and ecoregions, use of Australian national checklists). This will serve as a central cache of geospatial biodiversity data records from all sources (AVH, OZCAM, other collections, ecological field work, amateur observations, etc.) and will support many other tools and services. The test version of this portal is online at http://data.ala.org.au/.
2. **GIS Tools** – The GBIF Data Portal software at http://data.ala.org.au/ provides a number of valuable GIS outputs (WMS, WFS and KML layers, map images) but the ALA will collaborate closely with IMOS, TERN and other relevant initiatives to develop an effective spatial portal. This portal will access contextual (land use, border etc), environmental layers (geology, soil, vegetation classification, climate, etc.) and biological data and provide tools for visualisation and analysis.
3. **Regional Atlas** – The ALA aims to provide a general-purpose portal for mapping of species distributions and for species reports by region. The initial version will be based on the map interfaces offered by the GBIF Data Portal software, but the ALA aims to build a rich platform for visualising and analysing geospatial data, including use of a wide range of base layers (geology, soils, vegetation, climate, land-use, etc.) and to integrate the outputs from this tool into the Biodiversity Information Explorer, Conservation Portal and other web portals.

These activities will be enhanced to support site-based data collection (i.e. batches of specimens or observations sharing the same collecting/recording metadata) and to form components into integrated and interoperable cross-capability services. Additionally the ALA, TERN and IMOS will collaborate to ensure that suitable catalogues of relevant environmental data layers are accessible and exposed to allow optimal reuse of such data resources (**Directory of Environmental Layers**).

Implementation
The ALA has engaged Lee Belbin (Hobart) as project lead for the ALA Geospatial Data Management component and has initiated discussions with TERN, IMOS, BioMaps (Australian Museum), CSIRO Entomology, the Australian Institute of Marine Science and others to define technical requirements. The ALA is exploring the reuse of the IMOS geospatial data portal software and contribution of ALA effort to further its development. Two preparatory workshops are being organised to address the

following areas:

- Requirements for a spatial data analysis framework
- Identifying and managing key environmental layers

The following additional resources will be allocated to Geospatial Data Management:

- Peter Brenton (ALA business analyst, 50% allocation)
- Paul Flemons (Workshop facilitation, Australian Museum, short-term contract)
- Developers (including test, documentation, etc.) – up to 8 EFT

3.1.5 Data Integration (ALA, NCRIS)

This activity will be addressed using NCRIS funds.
Data integration activities represent most the core tasks funded under the original ALA NCRIS funding. The focus is on the development of tools and services to index primary information sources and provide integrated views to enable users to select and access those resources which are most relevant to their concerns.

The following components are under development:

1. **Ontologies and Vocabularies** – data integration within the ALA and with other projects (e.g. AVH, OZCAM, GBIF, EOL, OBIS) depends on a shared understanding of the structure of biodiversity data and agreement about the data elements which can be shared. The ALA is working with TDWG and international projects to set up tools to engage the community in developing and maintaining the ontologies and vocabularies required for this purpose. These structures will be particularly important to the ALA Metadata Repository and will provide the models to be used within that tool for organising Australian biodiversity information.
2. **Metadata Repository** – The core component within the ALA's data integration programme will be the Metadata Repository. This is being constructed using the Fedora open-source content management system. It will serve as a catalogue of biodiversity information resources (databases, documents, images, etc.) with provider-supplied metadata describing the origins and nature of each resource, but will be extended to link these resources to the species to which they relate, the geographic regions which they cover, etc. and to model the relationships between species, regions, habitats, descriptive characters, etc. (using information from tools such as the ALA Geospatial Data Cache). This will allow the ALA to produce web pages giving overviews of the available information relating to each species, region, habitat, etc. The Metadata Repository will therefore act as the engine serving information links to the proposed Data Dissemination components (especially Biodiversity Information Explorer and Biosecurity Portal).
3. **User Authentication and Identity Management** – The ALA will require the ability to authenticate users for many different purposes: to allow data providers to manage the metadata for their resources; to allow users to identify themselves to make annotations or provide additional data; for taxonomists to contribute to the Australian national checklists; etc. Building an integrated concept of the expertise of each individual will also allow the ALA to improve its use of the information supplied by each user. The ALA has already held discussions with the Australian Access Federation (AAF) on the use of its services as the identity management and authentication framework for those users belonging to AAF-enabled organisations.
4. **Annotation Services** – The ALA has received funding from the NCRIS Platforms for Collaboration capability's NeAT programme to develop annotation services to enhance the quality of data and to enable end users to contribute new information to the network. This work is being carried out at the University of Queensland School of Information Technology

and Electrical Engineering and early versions of some of the tools have been integrated into the GBIF Data Portal software at http://data.ala.org.au/. As the ALA proceeds, these tools will be used in many ways, including capturing user suggestions for corrected values within data records, free-text user comments, user tagging of species with descriptive terms, etc.

5. **Data Quality and Sensitive Data Tools** – The ALA contracted a review of concerns around potentially sensitive data within state conservation agencies, natural history collections and biosecurity activities. The goal is to develop best practice recommendations on the handling of occurrence records of conservation or biosecurity concern (e.g. reduction of precision of coordinates for records of species considered endangered in the state where they have been recorded, or diverting records of pest species with implications for Australian trade to the relevant authorities), and then to provide easy-to-use services to scan sets of records (e.g. as a spreadsheet) to evaluate any possible issues and report back to the data provider. This will be an important tool to help data providers become comfortable about sharing data and also to allow the broader community to develop consistent approaches to handling records for sensitive taxa. The tool will also support a wide range of additional data validation and other checks. In this form it will become a major component in the ALA's approach to improving data quality. Records with issues can be reported to the data holders and can automatically be annotated with notes or suggested corrections. End users will also be able to use annotation tools to contribute to data quality. Ultimately all such annotations should be handled through workflows which capture responses from the data providers.

Implementation
These activities are already in progress but reorganisation of the ALA to reflect the additional EIF funding will allow more resources to be applied to ensure delivery of a comprehensive set of tools and services.

The following resources have been assigned to this activity:

- Project lead, David Martin (ALA Technical Architect, full-time, ALA, NCRIS)
- Java developer, Nick dos Remedios (full-time, ALA, NCRIS)
- Metadata Curator, Lynette Woodburn (full-time, ALA, NCRIS)
- University of Queensland annotation services development team, Ron Chernich, Stephen Crawley (full-time, NeAT funding) – seeking to recruit one additional developer using ALA, NCRIS funds
- CSIRO ICT Centre metadata repository development team, Hon Hwang, Nerolie Oaks (total 1.5 EFT, NeAT/ALA, NCRIS funding), Carsten Friedrich (part-time team lead, ALA, NCRIS) – seeking to recruit two additional developers to contribute to this work and to liaise with the APPF/APN Phenomics Ontology-Driven Database NeAT project
- APPF bioinformatician (full-time, funded from ALA, NCRIS funds to support IBS data management)
- APN bioinformatician (full-time, funded from ALA, NCRIS funds to support IBS data management)
- Developers (including test, documentation, etc.) – up to 4 additional EFT

3.1.6 Data Dissemination (ALA-SS, EIF)

This activity will be addressed using EIF funds.
The ALA aims to organise Australian biodiversity data in forms which meet user needs, particularly in contexts in which it can be support decision-making processes.

The ALA is already developing a general-purpose data access tool to support any users of biodiversity information:

1. **Biodiversity Information Explorer** – An interconnected overview of all available data resources for any species, habitat, geographic region, etc. This will be based on the data index created in the Metadata Repository component.

In addition, the ALA has identified three focus areas for more targeted use of biodiversity data:

2. **Conservation Portal** – This will be a specific application of the data managed through the ALA Geospatial Data Management component and will focus on delivery of key indicator data for the reserves in the Australian national reserve system and more generally for the state of Australian habitats. Site-based data, particularly from survey work by DEWHA in the National Reserve System (NRS) and from the TERN Rangelands monitoring activity, and collection data will be integrated to support clustering and comparison of sites by their community composition and to evaluate the health and comprehensiveness of the NRS. The final deliverable should serve as a toolbox and suite of standard reports for each reserve or habitat.
3. **Pest Information Portal** – This is intended to provide a focus for organising information on species of biosecurity concern (including distribution modelling, species fact sheets, literature, etc.) to address needs within AQIS. ABIN, etc. This will build on the work of the Pests and Diseases Image Library (PaDIL, http://www.padil.gov.au/).
4. **Citizen Science Portal** – This is a two-way portal, designed to engage members of the public in data recording projects. Amateur naturalists will be given the opportunity to upload and manage observational data and will be placed in contact with organisations, projects and researchers with activities to which they can contribute. This work will build on software components from the ClimateWatch (http://www.climatewatch.org.au/) citizen science project developed by Gaia Resources using ALA, NCRIS funds. The ALA has identified a number of existing projects in DEWHA, DAFF, ABIN, state museums and elsewhere which will benefit from this shared infrastructure.

Implementation
The ALA has placed a contract with Gaia Resources (Piers Higgs, with Nigel George from Redskink) initially to document requirements for the Citizen Science Portal (report due mid-November 2009).

The ALA is also in discussion with PaDIL (based at Museum Victoria) about developing the PaDIL site to become a more comprehensive ALA Pest Information Portal.

The ALA expects to recruit the following additional resources to deliver the Data Dissemination components:

- User interface analyst
- Developers (including test, documentation, etc.) – up to 8 EFT

3.1.7 Hardware and Networking

The ALA is currently using the following servers provided by CSIRO IM&T:

Machine	Role	Services	Specification
alaprodweb1-cbr.vm.csiro.au	Production web server	Website (www.ala.org.au) GIS Portal (data.ala.org.au) Mapping Services (maps.ala.org.au) Annotation services (annotate.ala.org.au)	Apache/Tomcat 4GB, 2 CPU, 200GB of storage
Alaproddb1-cbr.vm.csiro.au	Production database	Database for production services	MySQL 8GB, 2 CPU, 500GB of storage
alatstweb1-syd.nexus.csiro.au	Test web server	Development versions of GIS Portal and Mapping Services	Apache/Tomcat 4GB, 2 CPU, 200GB of storage

alatstdb1-syd.nexus.csiro.au	Test database	Database for development services	MySQL 4GB, 2 CPU, 500GB of storage
Alaslvweb2-cbr.vm.csiro.au	Indexing web server	Java indexing processes	Apache/Tomcat 4GB, 2 CPU, 200GB of storage
alaslvddb1-cbr.vm.csiro.au	Indexing database	Database for Java indexing processes	MySQL 4GB, 2 CPU, 500GB of storage
diasbdev1-cbr.vm.csiro.au	DIAS-B development server	Web server and database for development of Metadata Repository	Tomcat, MySQL, Fedora Commons 4GB, 2 CPU, 200GB of storage
diasbtest1-cbr.vm.csiro.au	DIAS-B test server	Web server and database for test of Metadata Repository	Tomcat, MySQL, Fedora Commons 4GB, 2 CPU, 200GB of storage

During 2009-2010, the ALA will work with ARCS to plan and deploy additional storage and servers located around Australia to reflect the locations of activities developing ALA components. The ALA has started discussions with ARCS about the possibility of leveraging ARCS Super Science funded hardware.

3.2 Attachment 2 - ALA Budget

Note: this section includes all components planned for ALA development in 2009-2011, including those to be funded from NCRIS funds and those to be funded from EIF funds.

3.2.1 Revised budget 2009-2011

The following table shows the projected expenditure of both NCRIS and EIF funds during 2009-2011. Activities to be funded from NCRIS funds in 2009-2011 have been highlighted in gold and activities to be funded from EIF funds have been highlighted in green.

	2009-2010	2010-2011	Total	Notes
EXPENDITURE				
Output 1 Project Office (ALA, NCRIS)				
1.1 Director salary	175,959	184,757	**360,716**	
1.2 Director Overheads (CSIRO in kind)	*140,767*	*147,806*	***288,573***	
1.3 Programme Manager	213,000	213,000	**426,000**	
1.4 Project Officer salary	107,163	112,521	**219,684**	
1.5 Project Officer Overheads (CSIRO) in kind	*85,730*	*90,017*	***175,747***	
1.6 CSIRO Line Management (10%) cash	*18,721*	*20,097*	***38,818***	
1.7 Operating / travel etc	40,000	40,000	**80,000**	
1.8 Hardware (staff and development)	5,000	5,000	**10,000**	
1.9 Informatics support for 5.2.1, 5.2.2	350,000	450,000	**800,000**	
1.10 Recruitment and relocation costs	5000	5000	**10,000**	
1.11 Publicity and outreach	150000	150000	**300,000**	
Total Spend	**1,291,340**	**1,418,198**	**2,709,538**	
Cash and in kind	***245,218***	***257,920***	***503,138***	
NCRIS Spend	**1,046,122**	**1,160,278**	**2,206,400**	
Output 2 - Collection Data Management (ALA-SS, EIF)				
2.1 Project lead	213,000	213,000	**426,000**	
2.2 Travel and workshops	20,000	20,000	**40,000**	
2.3 Implementation	4,000,000	6,000,000	**10,000,000**	
Total Spend	**4,233,000**	**6,233,000**	**10,466,000**	
Cash and in kind	***0***	***0***	***0***	
EIF Spend	**4,233,000**	**6,233,000**	**10,466,000**	
Output 3 - Rich Data Stores (ALA-SS, EIF)				
3.1 Project leads	372,750	532,500	**905,250**	**5*50% expert leads**
3.2 Business analyst	77,600	97,000	**174,600**	**50%**
3.3 Developers	884,800	1,580,000	**2,464,800**	**Year 1: 8*70%, Year 2:10*100%**
3.4 Travel and workshops	20,000	20,000	**40,000**	
3.5 Test/commissioning projects	200,000	800,000	**1,000,000**	
Total Spend	**1,555,150**	**3,029,500**	**4,584,650**	
Cash and in kind	***0***	***0***	***0***	
EIF Spend	**1,555,150**	**3,029,500**	**4,584,650**	
Output 4 - Australian National Checklists (ALA-SS, EIF)				
4.1 Project leads	149,100	213,000	**362,100**	**2*50% (one taxonomy, one IT)**
4.2 Business analyst	77,600	97,000	**174,600**	**50%**
4.3 Developers	758,400	948,000	**1,706,400**	**Year 1: 6*80%, Year 2: 6*100%**
4.4 Travel and workshops	20,000	20,000	**40,000**	
4.5 ABRS contract	750,000	1,750,000	**2,500,000**	
4.6 CHAH contract	700,000	1,300,000	**2,000,000**	
Total Spend	**2,455,100**	**4,328,000**	**6,783,100**	
Cash and in kind	***0***	***0***	***0***	

EIF Spend	**2,455,100**	**4,328,000**	**6,783,100**	
Output 5 - Geospatial Data Management (ALA-SS, EIF)				
5.1 Project lead	159,750	127,800	**287,550**	**Year 1: 75%, Year 2: 60%**
5.2 Business analyst (50%)	77,600	97,000	**174,600**	**50%**
5.3 Developers	884,800	1,264,000	**2,148,800**	**Year 1: 8*70%, Year 2:8*100%**
5.4 Expert modellers	70,000	100,000	**170,000**	
5.5 Travel and workshops	20,000	20,000	**40,000**	
Total Spend	**1,212,150**	**1,608,800**	**2,820,950**	
Cash and in kind	***0***	***0***	***0***	
EIF Spend	**1,212,150**	**1,608,800**	**2,820,950**	
Output 6 - Data Integration (ALA, NCRIS)				
6.1 Technical Architect	220,760	231,797	**452,557**	
6.2 Java Developer	186,733	196,070	**382,803**	
6.3 Metadata Curator	186,733	196,070	**382,803**	
6.4 DIAS-B project (NeAT in-kind)	*400,000*	*200,000*	***600,000***	
6.5 DIAS-B Annotation Services developers	130,000	260000	**390,000**	
6.6 DIAS-B Metadata Repository developers	70,000	180,000	**250,000**	
6.7 Developers	379,200	632,000	**1,011,200**	**Year 1: 4*60%, Year 2:4*100%**
Total Spend	**1,573,426**	**1,895,937**	**3,469,363**	
Cash and in kind	***400,000***	***200,000***	***600,000***	
NCRIS Spend	**1,173,426**	**1,695,937**	**2,869,363**	
Output 7 - Data Dissemination (ALA-SS, EIF)				
7.1 Project leads	127,800	213,000	**340,800**	**2*50%**
7.2 Business analyst	77,600	97,000	**174,600**	**50%**
7.3 Developers	316,000	1,264,000	**1,580,000**	**Year 1: 4*50%, year 2: 8*100%**
7.4 Travel and workshops	20,000	20,000	**40,000**	
7.5 Test/commissioning projects	200,000	800,000	**1,000,000**	
Total Spend	**741,400**	**2,394,000**	**3,135,400**	
Cash and in kind	***0***	***0***	***0***	
EIF Spend	**741,400**	**2,394,000**	**3,135,400**	
Output 8 - International Engagement (ALA, NCRIS)				
8.1 GBIF Membership	165,200	165,200	**330,400**	
8.2 TDWG Membership	500	600	**1,100**	
Total Spend	**165,700**	**165,800**	**331,500**	
Cash and in kind	***0***	***0***	***0***	
NCRIS Spend	**165,700**	**165,800**	**331,500**	
Output 9 - Governance (ALA, NCRIS)				
9.1 External review of usability and function	25,000	25,000	**50,000**	
9.2 Operating / travel etc	30,000	30,000	**60,000**	
Total Spend	**55,000**	**55,000**	**110,000**	
Cash and in kind	***0***	***0***	***0***	
NCRIS Spend	**55,000**	**55,000**	**110,000**	
Output 10 - Network Infrastructure (ALA, NCRIS)				
10.1 Hardware	50,000	200,000	**250,000**	
Total Spend	**50,000**	**200,000**	**250,000**	
Cash and in kind	***0***	***0***	***0***	
EIF Spend	**50,000**	**200,000**	**250,000**	
Output 11 – Populating the Atlas (ALA, NCRIS)				
11.1 Australian Virtual Herbarium in-kind	*1,000,000*	*1,000,000*	***2,000,000***	
11.2 CSIRO cash	*500,000*	*500,000*	***1,000,000***	
11.3 CSIRO in kind	*500,000*	*500,000*	***1,000,000***	
11.4 Australian Museum cash	*100,000*	*100,000*	***200,000***	
11.5 Australian Museum in kind	*400,000*	*400,000*	***800,000***	

11.6 Museum Victoria cash	*100,000*	*100,000*	***200,000***	
11.7 Museum Victoria in kind	*850,000*	*850,000*	***1,700,000***	
11.8 Queensland Museum cash	*100,000*	*100,000*	***200,000***	
11.9 Queensland Museum in kind	*78,000*	*78,000*	***156,000***	
11.10 Tasmanian Museum & Art Gallery cash	*70,000*	*70,000*	***140,000***	
11.11 Tasmanian Museum & Art Gallery in kind	*80,000*	*80,000*	***160,000***	
11.12 University of Adelaide cash	*30,000*	*30,000*	***60,000***	
11.13 University of Adelaide in kind	*18,000*	*18,000*	***36,000***	
11.14 Southern Cross University cash	*50,000*	*12,500*	***62,500***	
11.15 Southern Cross University in kind	*145,000*	*38,000*	***183,000***	
11.16 ABRS in kind	*321,000*	*321,000*	***642,000***	
11.17 CSIRO External	*500,000*	*0*	***500,000***	
Total Spend	**4,842,000**	**4,197,500**	**9,039,500**	
Cash and in kind	***4,842,000***	***4,197,500***	***9,039,500***	
NCRIS Spend	**0**	**0**	**0**	
ALL OUTPUTS				
Total Spend	**17,903,566**	**25,104,935**	**43,008,501**	
Cash and in kind	***5,487,218***	***4,655,420***	***10,142,638***	
NCRIS Spend	**2,440,248**	**3,077,015**	**5,517,263**	
EIF Spend	**10,246,800**	**17,793,300**	**28,040,100**	

3.2.2 Relationship to budget from ALA Business Plan 2008-2009

The *ALA Business Plan 2008-2009* provided a revised budget for 2006-2011, organised according to five Outputs:

1. Building the Atlas
2. Populating the Atlas
3. Tools for Data Discovery
4. International Engagement
5. Governance & Management

The increases to ALA funds have necessitated a revision of the entire budget and planned outputs. The project has now been reorganised to target eleven Outputs, five of which will be funded from NCRIS funds and six from EIF funds:

1. Project Office (ALA, NCRIS)
2. Collection Data Management (ALA-SS, EIF)
3. Rich Data Stores (ALA-SS, EIF)
4. Australian National Species Lists (ALA-SS, EIF)
5. Geospatial Data Management (ALA-SS, EIF)
6. Data Integration (ALA, NCRIS)
7. Data Dissemination (ALA-SS, EIF)
8. International Engagement (ALA, NCRIS)
9. Governance (ALA, NCRIS)
10. Network Infrastructure (ALA-SS, EIF)
11. Populating the Atlas (ALA, NCRIS in-kind)

The following table reproduces the detail from the budget presented in the *ALA Business Plan 2008-2009* for the final two years of the project, as well as the five-year total. For each item for which funds were allocated for 2009-2011, the table identifies the reorganised Output to which these activities are now assigned. Activities to be funded from NCRIS funds in 2009-2011 have been highlighted in gold and activities to be funded from EIF funds have been highlighted in green.

	Expenditure			New Output	
	2009-2010	2010-2011	5-year Total		
TOTAL INVEST (original budget)			**34,786,255**		
NCRIS by year	**1,548,000**	**1,443,000**	**8,233,000**		
Cash and in kind	***5,146,539***	***4,510,198***	***26,553,255***		
EXPENDITURE (revised budget)					
Output 1 Building the Atlas					
1.1 Recruitment and relocation costs			**30,000**		
1.2 Project Leader salary	175,959	184,757	**614,946**	**1**	**Project Office**
1.3 Project Leader Overheads (CSIRO in kind)	*140,767*	*147,806*	***491,957***	***1***	***Project Office***
1.4 Operating / travel etc 5 yrs	30,000	30,000	**95,000**	**1**	**Project Office**
1.5 User needs analysis			**72,853**		
1.6 Technical Architect	220,760	231,797	**610,242**	**6**	**Data Integration**
1.7 Java Developer	186,733	196,070	**560,643**	**7**	**Data Dissemination**
1.8 External review of usability and function		25,000	**50,000**	**1**	**Project Office**
1.9 Hardware (staff and development)	10,000	10,000	**65,000**	**1**	**Project Office**
1.10 Taxonomic Names Service	100,000	100,000	**400,000**	**4**	**Australian National Species Lists**
1.11 Data hosting	70,000	70,000	**210,000**	**10**	**Network Infrastructure**
1.12 Data Provision Services	200,000	200,000	**700,000**	**2**	**Collection Data Management**
1.13 Informatics support for 5.2.1, 5.2.2	350,000	350,000	**1,000,000**	**1**	**Project Office**
1.14 Contingency	276,000	276,000	**828,000**		
Total Spend	**1,760,219**	**1,821,430**	5,728,641		
Cash and in kind	***140,767***	***147,806***	***491,957***		
NCRIS Spend	**1,619,452**	**1,673,624**	5,236,684		
Output 2 - Populating the Atlas					
2.1 Australian Virtual Herbarium cash			***741,195***	***11***	***Populating the Atlas***
2.2 Australian Virtual Herbarium in-kind	*1,000,000*	*1,000,000*	***7,758,388***	***11***	***Populating the Atlas***
2.3 CSIRO cash	*500,000*	*500,000*	***2,000,000***	***11***	***Populating the Atlas***
2.4 CSIRO in kind	*500,000*	*500,000*	***2,788,458***	***11***	***Populating the Atlas***
2.5 Australian Museum cash	*100,000*	*100,000*	***523,602***	***11***	***Populating the Atlas***
2.6 Australian Museum in kind	*400,000*	*400,000*	***2,448,238***	***11***	***Populating the Atlas***
2.7 Museum Victoria cash	*100,000*	*100,000*	***627,746***	***11***	***Populating the Atlas***
2.8 Museum Victoria in kind	*850,000*	*850,000*	***4,413,883***	***11***	***Populating the Atlas***
2.9 Queensland Museum cash	*100,000*	*100,000*	***674,078***	***11***	***Populating the Atlas***
2.10 Queensland Museum in kind	*78,000*	*78,000*	***392,000***	***11***	***Populating the Atlas***
2.11 Tasmanian Museum & Art Gallery cash	*70,000*	*70,000*	***350,000***	***11***	***Populating the Atlas***
2.12 Tasmanian Museum & Art Gallery in kind	*80,000*	*80,000*	***400,000***	***11***	***Populating the Atlas***
2.13 University of Adelaide cash	*30,000*	*30,000*	***154,518***	***11***	***Populating the Atlas***
2.14 University of Adelaide in kind	*18,000*	*18,000*	***77,490***	***11***	***Populating the Atlas***
2.15 DAFF (APPD) cash			***378,589***		
2.16 Southern Cross University cash	*50,000*	*12,500*	***212,500***	***11***	***Populating the Atlas***
2.17 Southern Cross University in kind	*145,000*	*38,000*	***618,000***	***11***	***Populating the Atlas***
2.18 ABRS in kind	*321,000*	*321,000*	***1,605,000***	***11***	***Populating the Atlas***
2.19 CSIRO External	*500,000*		***1,500,000***	***11***	***Populating the Atlas***
Total Spend	**4,842,000**	**4,197,500**	27,663,685		
Cash and in kind	***4,842,000***	***4,197,500***	***27,663,685***		
NCRIS Spend	**0**	**0**	0		
Output 3 – Tools for Data Discovery					
3.1 Metadata Curator	186,733	196,070	**560,643**	**6**	**Data Integration**
3.2 Tools survey			**35,000**		
3.3 DIAS-B project (NeAT in-kind)	*400,000*	*200,000*	***1,000,000***	***6***	***Data Integration***
3.4 DIAS-B Annotation Services developer	120,000		**240,000**	**6**	**Data Integration**

3.5 DIAS-B project facilitation	10,000	10,000	**50,000**	**6**	**Data Integration**
3.6 Online Identification Services	120,000	120,000	**360,000**	**3**	**Rich Data Stores**
3.7 GIS Services	120,000	120,000	**360,000**	**5**	**Geospatial Data Management**
Total Spend	**956,733**	**646,070**	**2,605,643**		
Cash and in kind	***400,000***	***200,000***	***1,000,000***		
NCRIS Spend	**556,733**	**446,070**	**1,605,643**		
Output 4 - International Engagement					
4.1 GBIF Membership	165,200	165,200	**826,000**	**8**	**International Engagement**
4.2 TDWG Membership	500	600	**1,900**	**8**	**International Engagement**
Total Spend	**165,700**	**165,800**	**827,900**		
Cash and in kind	***0***	***0***	***0***		
NCRIS Spend	**165,700**	**165,800**	**827,900**		
Output 5 - Governance & Management					
5.1 Operating / travel etc 5 yrs	50,000	50,000	**184,000**	**9**	**Governance**
5.2 Branding					
5.3 Project Officer salary	107,163	112,521	**370,344**	**9**	**Governance**
5.4 Project Officer Overheads (CSIRO) in kind	*85,730*	*90,017*	***296,275***	***1***	***Project Office***
5.5 CSIRO Line Management (10%) cash	*18,721*	*20,097*	***72,673***	***1***	***Project Office***
Total Spend	**261,614**	**272,635**	**931,792**		
Cash and in kind	***104,451***	***110,114***	***368,948***		
NCRIS Spend	**157,163**	**162,521**	**562,844**		
ALL OUTPUTS					
Total Spend	**7,986,267**	**7,103,435**	**37,757,661**		
Cash and in kind	***5,487,219***	***4,655,419***	***29,524,590***		
NCRIS Spend	**2,499,048**	**2,448,015**	**8,233,071**		

3.3 Attachment 3 - Activities and Milestones for 2009-2010

Note: this section includes all components planned for ALA development in 2009-2011, including those to be funded from NCRIS funds and those to be funded from EIF funds.

Activities and Milestones for 2009-2010	Achievement Date	Status
Output 1. Project Office (ALA, NCRIS)		
Programme Officer appointed	1 July 2009	Complete
ALA all-hands coordination workshop	31 Oct 2009	ALA-SS, EIF Initial Milestone
Output 2. Collection Data Management (ALA-SS, EIF)		
Collection Data Management project lead appointed	15 Aug 2009	Complete
Expert panel recommendations for implementation activities	15 Oct 2009	
Project plan with milestones agreed by ALA MC	15 Nov 2009	
Output 3. Rich Data Stores (ALA-SS, EIF)		
Descriptive Data project lead appointed	31 Oct 2009	
Project plan for supporting Delta format agreed by ALA MC	30 Nov 2009	
Digital Literature project lead appointed	31 Oct 2009	
Project plan for Digital Literature agreed by ALA MC	30 Nov 2009	
Output 4. Australian National Species Lists (ALA-SS, EIF)		
Contract for fauna species lists signed (ABRS)	15 Nov 2009	
Contract for flora/fungi species lists signed (CHAH)	15 Nov 2009	
Project plan for species list software agreed by ALA MC	30 Nov 2009	
Implement revised common architectural framework for species names	15 Mar 2010	
Implement ALA web services for access to cache data	15 Mar 2010	
Implement streamlined data review and validation processes	30 Jun 2010	
Deploy common format bulk data entry tools to taxonomy community	30 Jun 2010	
Interfaces for names information & databases operational	30 Jun 2010	
Output 5. Geospatial Data Management (ALA-SS, EIF)		
Geospatial Data Management project lead appointed	1 Sep 2009	Complete
Load IMOS portal code base	15 Oct 2009	
Spatial analysis toolkit workshop	15 Nov 2009	
Environmental layers workshop	15 Dec 2009	
Contract/appointment of GIS Manager	15 Dec 2009	
Contract for portal design signed by ALA MC	20 Dec 2009	
Contract/appointment of Spatial Analysis Toolkit developers	15 Jan 2010	
Draft implementation of contextual data within portal	28 Feb 2010	
Implementation of portal design	1 Apr 2010	
Draft implementation of environmental data within portal	31 May 2010	
Draft implementation of Spatial Analysis Toolkit within portal	30 Jun 2010	
User assessment of portal	31 Jul 2010	
Revised portal design implemented	30 Sep 2010	
Commission contextual data in portal	30 Oct 2010	
Commission environmental data in portal	30 Nov 2010	
Commission Spatial Analysis Toolkit	15 Dec 2010	

Output 6. Data Integration (ALA, NCRIS)		
Harvesting Demonstrator	15 Aug 2009	Complete
Revised project plan agreed by ALA MC	15 Oct 2009	
First public release of ALA components	28 Feb 2010	
Second public release of ALA components	31 May 2010	
Third public release of ALA components	30 Aug 2010	
Fourth public release of ALA components	30 Nov 2010	
Fifth public release of ALA components	28 Feb 2011	
Sixth public release of ALA components	31 May 2011	
Output 7. Data Dissemination (ALA-SS, EIF)		
User interface analysis delivered	15 Feb 2010	
Prototype interfaces delivered	30 Jun 2010	
Output 8. International Engagement (ALA, NCRIS)		
Renewal of GBIF membership	31 Dec 2009	
ALA workshops (project publicity and user needs) at TDWG conference	9 Nov 2009	
Output 9. Governance (ALA, NCRIS)		
09/10 Management Committee Meeting 1	9 Oct 2009	
09/10 Management Committee Meeting 2	31 Dec 2009	
09/10 Management Committee Meeting 3	31 Mar 2010	
09/10 Management Committee Meeting 4	30 Jun 2010	
09/10 CSIRO Oversight Committee Meeting 1	22 Oct 2009	
09/10 CSIRO Oversight Committee Meeting 2	31 Mar 2010	
IBS Review	12 Mar 2010	
Output 10. Network Infrastructure (ALA, NCRIS)		
Establishment of agreement with ARCS	15 Mar 2010	
Output 11. Populating the Atlas (ALA, NCRIS)		
Report on progress against activities proposed by ALA participants for 2008/2009	30 Sep 2009	
Report on progress against activities proposed by ALA participants for 2009/2010	30 Sep 2010	

Milestones in **bold** relate to delivery of public data services.

3.4 Attachment 4 - Risk Management Strategy

This section updates the ALA Risk Management Strategy focussing particularly on the risks of particular relevance in 2009-2010. Changes and additions from the 2008-2009 Risk Management Strategy are in **bold**.

3.4.1 Specific risks

NCRIS Investment Plan 5.2.3 Risk/hazard Identification and Management Strategy		
Specific risks of particular relevance to the NCRIS Investment Plan in the 2009-10 period		
Area	**Specific risk/hazard**	**Management Strategy**
Informatics technology	Lack of the Informatics capacity and vision to create a world class version of the ALA.	Development of the Atlas of Living Australia is not a trivial exercise, and no-one has done it in an acceptable way yet, and certainly not on an inclusive, comprehensive national scale. It will require a stringent iterative process around testing and standards, collaboration on a national and international scale, and some clear direction and vision. **The** steering committee **has been created** with the appropriate talent and backgrounds, proactive contact with colleagues in biodiversity informatics **must continue in 09/10**.
Recruitment	Inability to recruit staff with key skills, particularly in the biological informatics domain.	In some cases we know of suitable candidates in Australia – either to hire or to serve as collaborators. In other cases we will probably need to recruit from a broader pool of candidates. We will use well defined job descriptions and selection criteria to recruit appropriate personnel. **Core ALA staff positions and two bioinformatician positions have been filled. Part of the new ALA Programme Manager's role is to ensure appropriate staff levels are built and maintained.**
Managing client/stakeholder relationships	Clients/stakeholders not having ownership of outcomes leading to lack of uptake of outputs	Uptake of outputs by stakeholders is critical to achieve project outcomes. Explicit relationship building with key stakeholders will be established to achieve agreed outputs. Stakeholders will be engaged throughout the life of the project to ensure from the start that end user needs are considered in all stages of project planning. **The ALA has engaged resources from the CSIRO Entomology Communications team to assist with the update and delivery of a comprehensive communications**

		plan as well as generation of the existing newsletter. **The communications plan identifies all stakeholders and the type and frequency of communications required.** **Furthermore, each of the ALA-SS, EIF components are engaging directly with their key stakeholders and communities as part of scoping the work to be delivered over the funding period.** **The newsletter** is intended to serve as a key information tool to build stakeholder involvement. The ALA user needs assessment are also intended to strengthen communication between the ALA and its stakeholders. Good communication and realistic setting of expectations will be essential since there are so many possible areas in which the ALA could contribute**.**
Project selection	Projects for creating tools or populating the ALA will need to be chosen to ensure meaningful outputs are achieved.	Governance and steering committees **have been** formed to make decisions on the projects to be included in the business plan. Project selection criteria will include feasibility and impact. The ALA user needs analysis is **being used** as a key management strategy to give the ALA Management Committee the information needed to select priorities, and to provide criteria for measuring success in meeting user requirements**.**
Technology Failure	Failure of technology to provide products essential to the delivery of outputs in the appropriate electronic environment.	Monitoring of existing technology, and working in conjunction with technology suppliers to ensure that suitable platforms are chosen for project delivery. Working with technology suppliers and informatics specialists to guarantee that where possible specific requirements can be built into existing application and products without needing large investment in programming costs.

3.4.2 General Project Management Risks

General Project Management Risks		
Area	**Specific risk/hazard**	**Management Strategy**
Resource management	Ineffective management resulting in slippage, non delivery, poor resource allocation	Effective project management, with realistic budgeting, milestones and workforce planning **In order to manage the ALA-SS, EIF scope enhancements, the ALA has**

		hired a Programme Manager as well as team leads where required.
Quality standards	Outputs not meeting agreed standards will jeopardize uptake	Project monitoring to ensure compliance with agreed global standards and stakeholder recommendations
Performance management; project team communication	Disenfranchised unhappy staff, resulting in poor performance	Clear and regular feedback. Performance agreements with clear , achievable objectives and appropriate training opportunities for staff at all levels **With ALA-SS, EIF funding, the number of contractors rather than staff will be increasing dramatically. Contractors are not under formal performance agreements so this risk will be managed directly by the ALA Programme Manager and team leads**
Managing team members in other institutions/ locations	Poor performance due to lack of communication, and proper controls to ensure delivery	Established framework for communication between all project members, with certain line management duties relegated to appropriate staff in other institutions.
Dependence on key staff	Loss of expertise essential to project delivery	Succession planning for life of project; ensuring that at least 2 ALA staff have knowledge of all relationships and plans; internal project web site and document store to include briefing notes on all relationships and agreements.
Skills & training	Lack of appropriate competencies can affect project delivery	Funding for necessary training and development built into project
Financial planning, control and reporting	Poor financial management can lead to cost overruns and insufficient budget for key tasks	Good project management, with regular review of budgets and expenditure
Data loss/backup/ storage	Data loss jeopardises delivery	Implement proper IT policy and procedures
Obsolete technology	Poor equipment and software will affect outcomes and quality	Sector knowledge and budget forecasting to ensure proper equipment and software
Technology transfer (publishing knowledge, website)	Failure to deliver outputs in appropriate format can lead to failed outcomes	Proper planning and monitoring to ensure outputs are to appropriate standards, which have been agreed between all parties.
Effective communication	Poor marketing of achievements would be detrimental to uptake of project results	Appropriate communication strategy, developed through engagement with stakeholders. This can include press releases, product launches, presentations at national and international conferences.

3.5 Attachment 5 - Candidate services and processes for Collection Data Management

The goal of the Collection Data Management component is to improve information flow and support processes in herbal, faunal and microbial collections and to provide tools, systems and procedures that will enable Australia to respond better to its biodiversity challenges.

The following sections provide an overview of the candidate activities under consideration for this area.

3.5.1 Principles

The services and processes proposed on the following pages are under consideration by the collections data management panel and their communities. The following principles are used to guide the selection of appropriate suggestions:

- Achievable – in the timeframe, 12 months / 24 months
- Applicable to others – other researchers, other institutions, other taxa, other disciplines, other states, other countries
- Open Source – owned by the community and will continue into the future. Creative Commons (Non Commercial) may leave a possible revenue stream into the future
- Badged and attributed – ideas, websites, and data, from people and institutions are important for promoting a presence and acknowledging efforts. The ALA does not intend to swallow up contributors.
- Connectible – able to interact with, and make use of other projects, eg TERN, ABIN, PfC, EoL, MorphBank, BHL, etc
- Interoperable – use of common standards, protocols, schemas
- Low future cost – should be maintainable into the future
- Low marginal cost – eg digital photographs, sound recordings, scanned images
- Preventing further backlog – methods to include digitising as part of acquisition
- Reducing existing backlog – methods to include digitisation of legacy items as part of the workflow

3.5.2 Suggestions for services and processes for collection data management

The following is a summary of suggestions under consideration by the Collection Data Management panel.

Imaging centres

Several high-throughput imaging centres to be established around the country, each with an obligation to make each centre and its resources available for ALA-related institutions in the region to use for their imaging needs. This Imaging infrastructure should not be tied to any institution, although it is highly appropriate for it to be hosted at a partner institution. This work would include provision of software to manage images.

These resources would have a major aim of generating images of primary types. Prioritisation, would of course be a potential issue.

ALA could pay the costs to develop the facilities and to ensure that the software linkages exist. During 2009-2011 explore use NCRIS funds to assist with operating costs. Beyond 2011, unless the successor to NCRIS has operational funds available, costs would need to come from institutional or project funds.

The following conditions would need to be met:

1. Sets of hardware can be identified to offer significant value in bringing images online but to exceed the needs/budget of individual institutions
2. There is sufficient demand for such hardware to ensure its use more or less solidly for the next few years.
3. It would be acceptable for institutions to ship/convey materials to the centres for imaging - if this is not workable, shared infrastructure of this type will not work.
4. There are resources to fund necessary staffing for such centres – including a sensible longer term model.

Note that this model might be an opportunity to go for high-end hardware rather than cheaper solutions, if going for a more high-end solution is likely to translate into higher throughput.

Replace BioLink
BioLink has been widely used, especially for entomological and microorganism collections. It is no longer well supported and has been long due for replacement. An open source alternative will be sought, enabling continued and future community support.

BioloMICS
BioloMICS has been identified as preferred database management software for culture collections. BioloMICS addresses specific requirements of collections of microorganisms, such as morphological, physiological, biochemical, chemical, chromatographical, DNA, RNA and protein sequences. Support will be given for conversion of existing databases to appropriate standards.

Support for OZCAM, AVH, AMRiN, APPD and OBIS hubs
OZCAM, AVH, AMRiN, APPD and OBIS are important data hubs for their specialist communities. Tools and services will be developed to provide support for managing these aggregated databases: ALA is in a position to broker cheap storage and distributed delivery through ARCS; to make available services for taxonomy, data validation and annotation; to assist with authentication; to create a generic and configurable public interface which includes specialised mapping; and to provide a sensitive data service for public data.

Support for wrappers for OZCAM, AVH, AMRiN, APPD and OBIS
Ensures that data served by the aggregating data hubs is in a form compatible with, and acceptable by other systems.

Gazetteers and names lists
When entering metadata associated with specimens and observations, accuracy and use of common terms and standard phrases are important. Gazetteers with locally-relevant place names that are available as drop-down lists for form entry, ease repetitive tasks, and assist with consistency and accuracy. Similarly for the names of species, or collectors, etc, tools are needed to create and deliver drop down lists for relevant or likely possibilities for say, a particular institution, or taxonomic group, or geographic area.

Hosted data
Many smaller collections and researchers have little or no IT support. Assistance is needed with the tasks of managing data associated with their collections and sharing that data with others. In order to increase the utility and functionality of this data, the ALA will assist with:

- database hosting for smaller collections, such as university herbaria and TMAG
- smaller datasets, such as those of a single researcher
- spreadsheets, as a means to deal with generic data
- staging data, as a buffer between simple data and a complex institutional database

Experiment with rapid digitisation via images
Novel attempts to reduce the backlog of legacy material. Possibilities include transcription from photographs of labels, and scanning of paper documents, such as field notes and log books. Potential methods also involve the community, both professional and public, in contributing to attempts to speed up digitising processes.

Embed metadata in images
Automatic and assisted ways of embedding geocodes, collection events, names of collectors, specimen references, etc, into standard image metadata formats. This helps relevant information stay with images as they are transferred, and simplifies data entry tasks.

Field capture of metadata
Capturing data at the source allows the information associated with collection events to be tied to observations and specimens as metadata, even when those observations or specimens may be unidentified or unsorted. Subsequent work on those specimens, such as classifying or naming, would create extra metadata to be associated with those specimens, and would keep the same collection event metadata with each component specimen, even if they were to be separated.

Once data is captured, it allows that information to be readily available to others if appropriate.

The capture of data in the field may involve barcodes, photographs, and other technologies. ALA will provide generalised applications for uploading data and images, and readily attaching metadata.

3.6 Attachment 6 - Acronyms

AAF	Australian Access Federation
ABCD	Access to Biological Collections Data
ABIF	Australian Biodiversity Information Facility
ABIN	Australian Biosecurity Intelligence Network
ABRS	Australian Biological Resources Study
ACPFG	Australian Centre for Plant Functional Genomics
AFD	Australian Faunal Directory
ALA	Atlas of Living Australia
ALA MC	Atlas of Living Australia Management Committee
AM	Australian Museum
AMRIN	Australian Microbial Resources Information Network
AMRRN	Australian Microbial Resources Research Network
ANBG	Australian National Botanic Gardens
ANDS	Australian National Data Service
ANHAT	Australian Natural Heritage Assessment Tool
ANU	Australian National University
APC	Australian Plant Census
APF	Australian Phenomics Facility
APN	Australian Phenomics Network
APNI	Australian Plant Name Index
APPD	Australian Plant Pest Database
APPF	Australian Plant Phenomics Facility
AQIS	Australian Quarantine and Inspection Service
ARCS	Australian Research Collaboration Service
AVH	Australia's Virtual Herbarium
BA	Birds Australia
BHL	Biodiversity Heritage Library
BOLD	Barcode of Life Database
CAMD	Council of Australian Museum Directors
CBIT	Centre for Biological Information Technology
CERF	Commonwealth Environmental Research Facilities
CHACM	Council of Heads of Australian Collections of Microorganisms
CHAEC	Council of Heads of Australian Entomological Collections
CHAFC	Council of Heads of Australian Faunal Collections
CHAH	Council of Heads of Australasian Herbaria
CITES	Convention on International Trade in Endangered Species
CoL	Catalogue of Life
CSIRO	Commonwealth Scientific, Industrial and Research Organisation
CSIRO ICT	CSIRO Information and Communication Technologies
CSIRO IM&T	CSIRO Information Management & Technology
CSV	Comma Separated Value
DAFF	Department of Agriculture, Fisheries and Forestry
DELTA	DEscription Language for TAxonomy
DEWHA	Department of the Environment, Water, Heritage and the Arts
DIAS-B	Data Integration and Annotation Services for Biodiversity
DIISR	Department of Innovation, Industry, Science and Research
EFT	Equivalent Full Time
EIF	Education Infrastructure Fund
EoL	Encyclopaedia of Life

ERIN	Environmental Resources Information Network
eSIM	eScience Information Management
FTE	Full-time Equivalent
GBIF	Global Biodiversity Information Facility
GIS	Geographic Information System
GPS	Global Positioning System
GUID	Globally Unique Identifier
HTML	HyperText Markup Language
IBS	Integrated Biological Systems
IF	Index Fungorum
IMOS	Integrated Marine Observing System
IPNI	International Plant Name Index
IPR	Intellectual Property Rights
IRMNG	I Interim Register of Marine and Non-marine Genera
KML	Keyhole Markup Language
LSID	Life Science Identifier
LTER	Long Term Ecological Research
MDBC	Murray-Darling Basin Commission
MoC	Memorandum of Cooperation
NATA	National Association of Testing Authorities
NCRIS	National Collaborative Research Infrastructure Strategy
NeAT	National eResearch Architecture Taskforce
NGOs	Non-government Organisations
NRS	National Reserve System
NSF	National Science Foundation
OAI-PMH	Open Access Initiative Protocol for Metadata Harvesting
OBIS	Ocean Biogeographic Information System
OBO	Open Biomedical Ontologies
OCR	Optical Character Recognition
OECD	Organisation for Economic Co-operation and Development
OGC	Open GIS Consortium
OWL	Web Ontology Language
OZCAM	Online Zoological Collections of Australian Museums
PaDIL	Pests and Diseases Image Library
PBI	Planetary Biodiversity Inventory
PDA	Personal Digital Assistant
PDFs	Portable Document Format
PfC	Platforms for Collaboration
PKI	Public Key Infrastructure
PODD	Phenomics Ontology-Driven Database
QM	Queensland Museum
RDF	Resource Description Framework
REST	Representational state transfer
RFID	Radio-frequency identification
SAM	South Australian Museum
SEM	Scanning Electron Microscope
SOAP	Simple Object Access Protocol
SOP	Standard Operating Procedure
TAPIR	TDWG Access Protocol for Information Retrieval
TCS	Taxon Concept Schema
TDWG	Taxonomic Databases Working Group

TERN	Terrestrial Ecosystem Research Network
TMAG	Tasmanian Museum and Art Gallery
TRIN	Taxonomic Research Information Network
uBio	Universal Biological Indexer and Organizer
UQ	University of Queensland
UQ CBIT	UQ Centre for Biological Information Technology
UQ ITEE	UQ School of Information Technology & Electrical Engineering
URL	Uniform Resource Locator
WAM	Western Australian Museum
WCS	Web Coverage Service
WFCC	World Federation for Culture Collections
WFS	Web Feature Service
WMS	Web Map Service
WoRMS	World Register of Marine Species
XMP	Extensible Metadata Platform

3.7 Attachment 7 - Details of participant contributions by organisation 2009-2010

This attachment provides plans for use of cash funds committed by ALA participants during 2009-2010.

3.7.1 CSIRO Contributions to ALA 2009-2010

Participant	Contributions	Projected $	Reported $
CSIRO	Cash	500,000	
	Cash (CERF)	500,000	
	In-kind	722,194	
	Total	***1,722,194***	

Project Title: Digital Liaison (ANIC)

Description	This will be a very important role within the new CSIRO Biodiversity Theme. The position will provide the linkages within the Status and Trends Stream between collections/taxonomy, the Monitoring and Modelling Group, and the Atlas of Living Australia. It will help to provide a path to impact for collection data and ensure we are providing pertinent biodiversity data to inform policy. CSIRO Entomology is investing in developing new routes for bringing biodiversity information to users – this goal is dependent on establishing a team with the skills in the intersection between computing and biodiversity. This appointment is a key element in developing this capability. We are hampered in the delivery of these projects in not having a staff member to oversee the delivery on all these projects, and to support increased digital content development and a migration to web-based delivery. We propose to use CSIRO ALA cash funding to hire a Digital Liaison Officer to manage the team generating digital content and ensure that they provide data at the appropriate standards through liaison with the ALA team. The Digital Liaison Officer will adopt an exciting role in the new Biodiversity Theme, making the connection between the CSIRO National Biological Collections, the Atlas of Living Australia, and ecologists and conservation biologists who will use biodiversity information. The appointee will be expected to: liaise with stakeholders within the BRABA Theme on matters of digital content (particularly in terms of monitoring, modeling and surveillance); have oversight on contracted delivery of digitized data from the CSIRO biological collections to the ALA; liaise with the biodiversity informatics group to make recommendations about the appropriate quality of digital information; liaise with IT support to ensure that we have the necessary hardware and software support for our activities; encourage and promote collaboration between key stakeholders within the Status and Trends Stream to enable key outputs.
Contact(s)	John La Salle
Taxa or biome	Australian Insects (plus others as appropriate from other CSIRO collections)
Deliverables	A range of high quality digital records and information, including database records, images, taxonomic products to be made available to ALA at appropriate standards.
Cost	Salary $169,269 (including overheads)
Staffing	None in house for this specific role. It will be advertised and filled externally. Guidance, management and supervision will be supplied through existing ANIC management.
Risks	Database unavailable (low)
ALA linkages	This person would be available to offer assistance/advice to other ALA projects delivering digital data. This need not be confined to CSIRO based projects.
Other linkages	This is a capability position. It links to all ANIC projects which deliver digital data, and could link to any other projects as well.
Data access	This person will be put in place specifically to ensure that the data is shared at the appropriate standards, protocols, etc., and will work directly with ALA staff to ensure this.
Users	Users of the ALA/digital data.

Project Title: Family Planning Project (ANH, as part of Australian Vascular Plants in APNI)

Description	Provision of an agreed family-level classification for Australian vascular plants (including ferns) for the *Australian Plant Census* project. Currently the APC uses Cronquists' 1981 family classification (that

	adopted by the *Flora of Australia*), a system now widely acknowledged as having been superseded. Many Australian herbaria are using, or have indicated a desire to use, the family classification developed by the Angiosperm Phylogeny Group (APG) (http://www.mobot.org/MOBOT/research/APweb/). This project will provide an agreed family-level classification for Australian vascular plants, using the most recent APG classification as a starting point, with additional taxonomic ranks (e.g. order to kingdom) also included. Synonymies and constituent taxa will be provided for families and all higher ranks. Generic-level synonymy will not be considered (this process takes place at the Genus and Species component of the APC, separate to this project), but assignment of constituent genera to families (and families to orders, orders to classes, etc) will be determined as part of this ALA-funded project. This project will extend the APC beyond its initial genus-and-species focus to higher classifications, improve congruence across the different systems currently used by State and Territory herbaria, and provide users with a nationally agreed classification for Australian vascular plants across the taxonomic hierarchy.
Contact(s)	Brendan Lepschi, Anna Monro
Taxa or biome	Australian vascular plants
Deliverables	An agreed higher-level (family to kingdom) classification for Australian vascular plants (including ferns), available electronically via the APC website, and including all relevant synonymy and protologue information. Selected important alternative family-level classifications will also be captured and entered into APNI, and will be available via the APC or APNI interfaces. All agreed concepts are dynamic and flexible, endorsed by CHAH, and will be revised and amended via the APC Working Group consensus model as new data is published. Program for 2009-2010 is as follows: • Agreed family-level classifications for Pteridophyta endorsed by CHAH (by end September 2009) • Data entry of remaining agreed family-level classifications (remainder of vascular plants, Pteridophytes) into APNI (approximately 3000 names, including constituent genera[1]) • APNI data entry of additional higher-level classifications (e.g. remainder of Kubitzki series, Moxie, Takhtajan), including protologues and relevant synonymies. This is a lower-priority component and will be undertaken as time allows. Completion of data entry of the Kubitzki series is the highest priority. • Agreed higher-level classification (ranks from order to kingdom) circulated to APC WG and subsequently endorsed by CHAH • Data entry of agreed higher-level classification into APNI (*c.*300 names) [1] Constituent genera are already entered into APNI, but need to be linked to agreed family-level classification concepts
Cost	Salary – CSOF 3.5
Staffing	Staff available to oversee project, working space and equipment (PC terminal) available. Staff available to manage and disseminate data deliverables derived from this work Staff available to undertake the project
Risks	Database unavailable (low)
ALA linkages	Data available to other ALA participant activities (e.g. *Australia's Virtual Herbarium* (AVH), other specific ALA projects) immediately.
Other linkages	Automatic linkage of data between APNI and APC, as well as other databases such as the *Australian Plant Image Index* (APII). Data available to DEWHA applications (ERIN, SPRAT, etc.) Data available to AVH and on-line national, state and territory floras.
Data access	Access to data is the same as that for other elements of the ANBG's IBIS databases (e.g. the *Australia's Virtual Herbarium,* the *Australian Plant Name Index* and the *Australian Plant Image Index*) and follows the web service guidelines provided by the ALA and the international biodiversity information standards and protocols of GBIF and TDWG.
Users	Both the APNI and APC have significant numbers of users, both nationally and internationally, with approximately 6000 hits per week. Users include all State and Territory herbaria, overseas herbaria, ABRS, DEWHA and other government departments, NGOs (e.g. Greening Australia) and the general public. Users access APNI and (particularly) APC to obtain nomenclatural and taxonomic information on the Australian vascular flora, including recommended (nationally agreed) scientific names. APNI and APC also provide the

	nomenclatural framework for the AVH. The number of potential users of this data can be expected to grow as more data is added and the product is refined and developed.

Project Title – Australian Mangrove and Saltmarsh Species (ANH)

Description	The aim of this collaborative project is to demonstrate species level content for projects such as the Atlas of Living Australia and for research, education and public information. It will build on existing biodiversity information management activities, resources and expertise around the nation and beyond. Information on the plant and animal species found in Australian mangroves and saltmarsh will be collated, with emphasis on taxonomy, appearance, identification, biology, distribution and ecology. The project will focus on developing active and inclusive partnerships to provide freely available on-line access to information using current biodiversity information management technologies. This particular project is being planned to be the nucleus of a larger and self sustaining national project documenting the species of Australian mangrove and saltmarsh communities and to act as a model for possible future initiatives in other Australian ecosystems. Links will be developed with mangrove and saltmarsh researchers, experts and other collections and institutions to facilitate the development of species lists and profiles for plant and animal taxa occurring in Australian mangrove and saltmarsh habitats. The approach is to begin with the obligate core mangrove and saltmarsh species in order to develop the collaborations, methodologies and technology, and the content format. As the information builds and the system is established, the project will expand to incorporate the facultative, transient and marginal species of the mangrove and saltmarsh community.
Contact(s)	Emma Clifton, Brendan Lepschi
Taxa or biome	Australian Mangrove and Saltmarsh Species
Deliverables	Liaise with other collections, institutions and experts. For areas of interest and expertise: • Seek species lists of organisms occurring in mangrove and saltmarsh communities (flag high, medium, low fidelity) • Compile / contribute information to species profiles • Compile distributional information • Identify associations and co-evolutionary relationships • Source images, multimedia and other relevant information Source additional support and funding to assist with obtaining species information, images, multimedia and expanding the network of mangrove and salt marsh experts. Continue compiling and refining species lists – focus now on Medium and Low Fidelity plant species and High fidelity groups of other organisms. Complete species profiles for all high fidelity plant species. Check CANB specimens of high and medium fidelity plant species. Verify current taxonomy in APNI/APC. Ensure determinations and geocodes are accurate and all specimens fully databased. Continue to source contributions of species lists, species profiles, information on associations and co-evolutionary relationships and images, multimedia and related data from other collections and institutions.
Cost	Salary CSOF 3.3 ½ time for 12 months to June 2010.
Staffing	Staff available to undertake the plant aspects of the project Staff available to oversee project, working space and equipment (PC terminal) available.
Risks	Risks include: • A lack of or limited data for some species descriptions • Availability of suitable images for all taxa • Lack of contributions from other CSIRO collections to include organisms other than plants • Inability to engage national mangrove community biologists A significant challenge for this project is to persuade others to voluntarily contribute their data, information, time and expertise. Experts that are self-employed or employed by other institutions are fully occupied with their own or institutional priorities. Their capacity to contribute to this project may be limited. The project staff will need to be very persuasive and to develop appropriate incentives to secure the necessary cooperation.
ALA linkages	This project will provide species level content on a specific biome to the ALA.
Other linkages	This project will link directly to data from the Australian Plant Name Index (APNI), Australia's Virtual Herbarium (AVH) and the Australian Plant Image Index (APII).
Data access	Data from this project will be provided to the ALA and other biodiversity information projects using web services and the standards and protocols of TDWG and GBIF as outlined on the ALA website.

Users	This project will provide information for research, education and public information. It will collate data from current literature and available on-line resources, providing a convenient and up-to-date source of information for Australian mangrove and saltmarsh plant species. The species profiles and electronic key produced for this project will be of interest to local, state and federal government departments, researchers and organisations such as Landcare, Coast Care and Greening Australia as well as the general public.

Project Title - Biodiversity of Australian Marine Fishes

Description	Continuation of projects within the Australian National Fish Collection (ANFC) to further the identification, description and management of Australia's marine fishes.
Contact(s)	Alastair Graham: 03 62325351, alastair.graham@csiro.au
Taxa or biome	Australian marine fishes
Deliverables	**Part 1 – Ongoing digitisation, databasing and enhancement of the PIAF image collection** ANFC's Photographic Index of Australian Fishes (PIAF) has been compiled over more than 30 years of ichthyological research. Staff continue to expand this database (already the most comprehensive collection of images of Australian fishes), with the goal of depicting every species of Australian fish. The resource is used extensively for species identification, and for the illustration of taxonomic publications, websites and identification books for industry and the public. This project will continue, with the capture of additional (digital) images, and the selective archiving and enhancement of the existing photographic slide collection. **Part 2 – Digital archiving of unique collections of Australian fish images** ANFC staff have selectively digitised 35 mm slides from two of Australia's most important regional image collections: Barry Hutchins' images of species from the southern coast, and Ken Graham's images from NSW. Photographic slides deteriorate with time, and this work will continue to ensure the conservation of key collections of Australian fish images. Previously non-incorporated images from Australian museums will be targeted, and possibly a unique collection of Australian tropical reef fishes held at the Bishop Museum, Hawaii. **Part 3 – Replacement of obsolete photographic-facility equipment** Some equipment in the ANFC photo lab is obsolete and requires upgrading to ensure that images captured continue to be of publication standard. **Part 4 – Selective processing of frozen GBR/TS specimens** Frozen fish specimens collected during habitat mapping projects in the Great Barrier Reef Marine Park and the Torres Strait will be selectively processed (identified, photographed, muscle sample collected, and the specimen catalogued and preserved). Some specimens were collected from sites that are now closed to extractive monitoring. These specimens are therefore crucial for obtaining photographs and muscle samples, and will contribute to the ongoing delineation of cryptic species in this region.
Cost	**Part 1 –** Louise Conboy salary @ ~1 day per week (CSOF 4) ~$44,000 (including divisional overheads) **Part 2 –** Digital archiving of key photographic collections ~$10,000 (depending on the libraries accessed) **Part 3 –** Replacement of obsolete photographic-facility equipment ~$10,000 **Part 4 –** Selective processing of frozen GBR/TS specimens ~$16,000 (salary, consumables etc)
Staffing	Louise Conboy and Daniel Gledhill are available to contribute to this project, and additional casual staff may also be utilised for the processing of material. The ANFC will contribute support from two staff, and various resources to the project.
Risks	Granting of access to images at the Bishop Museum is still to be finalised. If this is not possible, resources will be directed to other aspects of this project.
ALA linkages	All aspects of this project continue to build on ANFC's existing contribution to the ALA through the PIAF database, and possibly toward a demonstration project for the ALA.
Other linkages	Images from the PIAF database have been used in a number of books, guides, taxonomic papers and websites, including CAAB. The ANFC houses one of the most important, and largest, collections of fish tissues in the world, it has the most diverse coverage for marine fishes and is the largest and most diverse collection of shark and ray tissues. The majority of samples are linked to a registered and curated specimen, many of which have also been photographed. The ANFC has strong connections with the Barcode of Life project through FishBOL, and will continue to contribute tissues to this project.
Data access	Metadata for specimens, tissues and images can be accessed from the ANFC for scientific purposes (subject to commercial-in-confidence, and other standard restrictions).
Users	Images, tissues and specimens have been utilised in numerous projects such as species descriptions, guidebooks, FishBOL and websites for industry, public and researchers.

3.7.2 Australian Museum contribution to ALA 2009-2010

Participant	Contributions	Projected $	Reported $
Australian Museum	Cash	100,000	
	In-kind	400,000	
	Total	***500,000***	

Project Title: Marine and Terrestrial specimen data capture

Description	Marine and Terrestrial specimen data capture
Contact(s)	Dr Penny Berents (Head of Natural Science Collections) penny.berents@austmus.gov.au
Taxa or biome	Australian land & freshwater snails, Australian Arachnids and Myriapods, NSW Marine Fishes, Lizard Island Marine Fishes, Australian Lepidoptera (various moth families associated with rangelands survey projects), Australian nepomorph water bugs, Various insect taxa submitted to BOLD from the Townsville region collected by G. Cocks, Australian and Pacific Mammals including Grey Kangaroo and Swamp Wallabies, Australian Birds, Australian Polychaetes/Echinoderms/Crustacea
Deliverables	Approximately 20,000 records
Cost	2.6 FTE Technical Officers = around $168000
Staffing	20% of 13 Technical Officers time, computers, database and collections already available
Risks	None. Staff, equipment and collections available and AM has given priority to the project for 2009-010.
ALA linkages	Data will be available through OZCAM portal and will complement data provided by other natural history museums.
Other linkages	Data available to OZCAM, GBIF, OBIS
Data access	Data will be made available through the OZCAM portal
Users	Scientists, natural resource managers, government agencies (eg. DEWHA, DECC, AQIS, DAFF), special interest groups and other users of biodiversity information

3.7.3 Museum Victoria contribution to ALA 2009-2010

Participant	Contributions	Projected $	Reported $
Museum Victoria	Cash	100,000	
	In-kind	850,000	
	Total	***950,000***	

Project Title: Type Specimen Image Capture

Description	Image capture of Australian primary insect types (holotypes, lecotypes and neotypes)
Contact(s)	Ken Walker / Dermot Henry
Taxa or biome	Insecta
Deliverables	Approximately 500 types at 3 images per specimen = approx 1500 images
Cost	Staff costs 1.0 FTE = approx. $25,000
Staffing	Equipment and database already available
Risks	None projected as funding provided until Dec 2009
ALA linkages	Previously GBIF funded all museums to image capture of Australian vertebrate and mollusca types
Other linkages	Previously GBIF funded all museums to image capture of Australian vertebrate and mollusca types
Data access	Data and images will be available through the OZCAM portal and through Museum Victoria's database online project and hopefully in the near future from either or both Tapir and OAI wrapper on our EMu database.
Users	Rapid species level identification or confirmation through ability to view the type specimen online.

Project Title: Barrow Island Image capture

Description	Image capture of 2000 invertebrate species on Barrow Island, WA
Contact(s)	Ken Walker
Taxa or biome	Insecta
Deliverables	Approximately 1000 species at 3 images per specimen = approx 3000 images
Cost	Staff costs 2.0 FTE = approx. $100,000
Staffing	Equipment and database already available
Risks	None projected as funding provided until June 2010
ALA linkages	None identified
Other linkages	None identified
Data access	All images and survey data available through PaDIL
Users	Rapid species level identification or confirmation through ability to view the invertebrate fauna online.

Project Title: PaDIL upgrade

Description	Software upgrade to PaDIL – Pests and Diseases Image Library
Contact(s)	Ken Walker
Taxa or biome	Biosecurity Invertebrates
Deliverables	New platform delivering many social network attributes to a scientific website
Cost	Staff costs 1.0 FTE = approx. $100,000
Staffing	Equipment and database already available
Risks	None projected as funding provided until Dec 2009
ALA linkages	Biosecurity Portal
Other linkages	IPPC, SPHDS, OCPPO, Bugwood Network, Forest Images.org
Data access	All images and data available through PaDIL
Users	Rapid species level identification or confirmation through ability to view the biosecurity specimens online.

3.7.4 Queensland Museum contribution to ALA 2009-2010

Participant	Contributions	Projected $	Reported $
Queensland Museum	Cash	100,000	
	In-kind	78,000	
	Total	***178,000***	

Project Title: Data capture and delivery.

Description	Digital capture and delivery of biological, palaeontological and geological collection data, images and phenotypic data related to the permanent collections of the Queensland Museum
Contact(s)	Dr John Hooper (collections and research products), Ms Cecelia Ryan & Mr Paul Avern (database infrastructure, data digitisation and dissemination) john.hooper@qm.qld.gov.au, cecelia.ryan@qm.qld.gov.au, paul.avern@qm.qld.gov.au
Taxa or biome	All zoological taxa (living & fossil), fossil plants & mineralogical specimens
Deliverables	Delivery of (currently) approx. 800,000 datapoints and associated data to the ALA via OZCAM. [The exact number of final datapoints that will be delivered is still uncertain due to these data containing a mixture of point-data ranging from individual specimens of single taxa from a single locality to multiple specimens of multiple taxa (specimen lots) from single localities]. In 09-10 a GIS module will be incorporated into Vernon CMS for internal use in data validation, verification and reporting.
Cost	Approximately $500,000 over 5 years (cash expenditure on implementation of Vernon CMS), and $390,000 (in kind contribution, such as salaries, overheads, IT infrastructure, towards ongoing data capture and verification by curatorial, collection management and IM/IT staff) Note: These financial targets (cash expenditure and in kind contribution) were met by the end of the 07/08 financial year.
Staffing	IM/IT have 1.5 FTE staff responsible for the collection database development and management and liaising with scientific and technical staff. Biodiversity & Geosciences Programs have approximately 10 FTE staff whose duty is (partially) to continue digital capture of collection data, and to verify accuracy of this data on an ongoing basis.
Risks	Data quality variable amongst collections due to e.g. age of collections, accuracy of locality data, the taxonomic hierarchy used, taxonomic authority of identification etc. Phenotypic data exists for some but not all phyla, and availability dependent on concurrent projects and funding to assemble species descriptions (e.g. Barcoding of Life project funding to QM from Alfred P. Sloan Foundation for some marine collections)
ALA linkages	Data contribute to ALA
Other linkages	OZCAM; Barcoding of Life (Census of Life/ Census of Marine Life), Environment Australia (e.g. ANHAT) data sets
Data access	Data conforms to Darwin Core, and access by ALA will be guided by protocols developed by OZCAM. QM will have its own public access portal and web tools but contribution of particular datasets to ALA will require guidance from and development of appropriated software (e.g. wrappers) by ALA
Users	Internal QM users for collection management, specimen loans and tracking, GIS, ecological modelling, inventories for EIS, etc. Direct public access to QM data aimed at local communities and groups (e.g. schools) for inventories, mapping of local biotic communities, illustration of type specimen holdings, etc. Scientific users would probably use the OZCAM/ ALA portal for GIS, Bioclim modelling etc

3.7.5 Tasmanian Museum & Art Gallery contribution to ALA 2009-2010

Participant	Contributions	Projected $	Reported $
Tasmanian Museum & Art Gallery	Cash	70,000	
	In-kind	80,000	
	Total	***150,000***	

Project Title: Curation and digitisation

Description	Curation and digitisation of existing collections and development of collection management systems at TMAG
Contact(s)	Dr Catherine Young Senior Curator of Zoology Tasmanian Museum and Art Gallery GPO BOX 1164 HOBART TASMANIA 7001 Phone: (03) 62 11 4122 Fax: (03) 62 11 4112 Mobile: 0437 102 543 Web: www.tmag.tas.gov.au
Taxa or biome	1. Continue echinoderm and mollusc collections. 2. Pursue repatriation of all invertebrate loan collections including types from interstate and overseas.
Deliverables	1. Complete capture of all record information from cards and registers for echinoderm and mollusc collections. 2. Reintegrate returned loan collections and ensure that databases are updated with current taxonomic data. 3. Continue to update our interface to improve data capture to OZCAM cache. Continue to a review the CMS, with the intent of standardising database fields and content management in Biodiversity with the aim of implementing the ABCD and SDD standards.
Cost	1. $25K 2. $15K 3. $30K
Staffing	Dependent upon a range of museum commitments and possible impact of TMAG redevelopment.
Risks	The major risks are lack of staff and resources, competing projects and lack of in–house taxonomic expertise. Limited high level IT support for either project, similarly for projects from the Tasmanian Herbarium.
ALA linkages	Not known
Other linkages	The zoology and botany collections link into national projects such as AVH, Ozcam, etc. Cetacean study links informally into State, inter museum and National projects. Echinoderm/Mollusc data – informal linkages to government and private individuals.
Data access	Through existing AVH and OZCAM services.
Users	All users of biodiversity information.

3.7.6 University of Adelaide contribution to ALA 2009-2010

Participant	Contributions	Projected $	Reported $
University of Adelaide	Cash	30,000	
	In-kind	18,000	
	Total	***48,000***	

Project Title: LucID key to Australian Invertebrates

Description	Although production of this key was initially funded by ABRS (to Austin & Harvey), the additional funds will allow the project to be completed by mid 2010.
Contact(s)	John Jennings, Andy Austin (U of A) and Mark Harvey (WA Museum)
Taxa or biome	Invertebrates
Deliverables	LucID key to Australian Invertebrates
Cost	~ $25,000
Staffing	Staff available in WA Museum and University of Adelaide, and via subcontracting, for imaging, key construction, and testing, and writing text boxes. We have employed Dr Claire Stevens and Kate Muirhead on a casual basis to provide many of the check boxes and images. See 2008/09 report on progress.
Risks	low
ALA linkages	Key will be publically available through CBIT in the first instance
Other linkages	
Data access	See above
Users	Potential users include State and Territory agencies, government departments, school, undergraduate and postgraduate students, and the general public who are interested in invertebrates.

Project Title: Databasing of Hymenoptera

Description	Databasing of Australian Ichneumonoidea (Hymenoptera). This is part of a long-term goal to provide to the entire described Australian bee, wasp, ant and sawfly fauna to the Australian Faunal Directory. It is also probable that some of the remaining superfamilies and families of Hymenoptera not yet databased will also be started in 2009-2010. See Australian Faunal Directory website for missing groups.
Contact(s)	John Jennings
Taxa or biome	Ichneumonoidea (Hymenoptera)
Deliverables	The database will be downloaded to the Australian Faunal Directory
Cost	~ $7,500
Staffing	Staff available to undertake databasing and final product delivery
Risks	Very low
ALA linkages	Database also available to other ALA participants
Other linkages	Also links in with GBIF
Data access	Data will be on-line through Australian Faunal Directory
Users	Potential users include State and Territory agencies, overseas researchers, ABRS, and other government departments and the general public who are interested in systematic of the superfamily.

3.7.7 Southern Cross University contribution to ALA 2009-2010

Participant	Contributions	Projected $	Reported $
Southern Cross University	Cash	50,000	
	In-kind	145,000	
	Total	***195,000***	

Project Title: Australian Plant DNA Bank

Description	Australian Plant DNA Bank
Contact(s)	Prof Robert Henry, Nicole Rice
Taxa or biome	The focus for 2009-2010 is to ensure that there is at least one species from each of the Australian plant families in the collection. In addition to this the Australian Plant DNA Bank will hold the DNA reference samples for numerous whole genome sequencing projects including wheat, sugarcane, Australian Oryza species and taxa from the Eucalypts.
Deliverables	Genomic DNA samples and associated data for distribution including DNA quantity and quality.
Cost	
Staffing	1 x 60%, 1 x 40% and 1 x 60%
Risks	
ALA linkages	Australian Herbaria and botanic gardens. Particularly important that the DNA samples link to a vouchered Herbarium specimen.
Other linkages	Linkages to genetic resource collections, other DNA banks, herbaria and botanic gardens. Current linkages include collaborative projects, for example the collection of Australian species of Oryza with Japanese colleagues and Australian Tropical Herbarium. International plant genome sequencing projects.
Data access	Data available online at www.dnabank.com.au. Additional data available on request of DNA samples.
Users	Molecular biologists

3.7.8 Australia's Virtual Herbarium contribution to ALA 2009-2010

Participant	Contributions	Projected $	Reported $
AVH	Cash	0	
	In-kind	1,000,000	
	Total	***1,000,000***	

No cash contribution committed in 2009-2010.

3.7.9 Australian Biological Resources Study contribution to ALA 2009-2010

Participant	Contributions	Projected $	Reported $
DEWR/ ABRS	Cash	0	
	In-kind	321,000	
	Total	***321,000***	

No cash contribution committed in 2009-2010.

3.7.10 Australian Plant Pest database contribution to ALA 2009-2010

Participant	Contributions	Projected $	Reported $
DAFF/ APPD	Cash	0	
	In-kind	0	
	Total	***0***	

No cash contribution committed in 2009-2010.