

## Aus-e-Lit - started

### Project Committee

Dr Elizabeth McMahon, ASAL President (Chair)  
 Prof Tony Williams, ARCS E/D  
 Dr Ross Wilkinson, ANDS E/D nominee  
 Kerry Kilner, AustLit Executive Manager  
 Prof Jane Hunter, UQ eResearch Lab

### Reference Group

Dr Philip Mead, ASAL; Keith Webster, UQ Librarian; Ross Coleman, SETIS; Prof Robert Dixon, USyd; Prof Paul Eggert, ADFA; Margie Burn or Warwick Cathro, NLA; Gavan McCarthy, eScholarship Research centre; Willard McCarty, Kings College; Carl Lagoze, OAI-ORE, Sheila Anderson, Director AHDS

### Project Manager: Roger Osborne, UQ

### Milestones

- |        |  |
|--------|--|
| Year 1 | Federated Search Portal for the community that supports: Searching across metadata terms as well as full-text searching across corpuses; Search, retrieval and presentation of records and documents (text, images, audio);<br>Empirical reporting services that support quantitative analysis of historical patterns in Australian literature and spatio-temporal visualizations for browsing AustLit collections   |
| Year 2 | Shared annotation spaces based on secure annotation creation, editing and attachment services.<br>A range of annotation presentation, browse and search services and GUIs that enable users to search and browse across the content of annotations or for annotations by particular individuals (tag clouds will be provided for keywords);<br>RSS notification services – that notify subscribers of new annotations on particular texts or by particular individuals |
| Year 3 | Metadata attachment and publishing service for OAI-ORE objects – enabling their upload to a Fedora repository<br>Search, retrieval and presentation interface for OAI-ORE objects based on both metadata terms plus component objects.   |

### **Service Outcomes**

- Data integration and search services across multiple national databases relevant to Australian literature, supporting metadata and content based access as well as timeline and spatial search representations
- Development of interest specific community knowledge bases built using shared annotation services over the linked data bases
- Publication tools allowing researchers to create compound digital objects capturing and explaining relationships in the corpus

### Resources

NeAT will fund 2.5 EFTs at UQ – 2 software developers and 0.5 EFT for project management and liaising with data providers and AustLit users

UQ will provide a cash contribution of \$105k, service hosting and 0.5 EFT pa towards project direction

The project will build on outputs from an AustLit ARC LIEF grant; DART, ARCHER, PILIN, RUBRIC projects

QCIF will provide some system support

Project size is about 4 EFT pa

## Data Integration and Annotation Service in Biodiversity (DIAS-B) – Started

### **Project Committee**

(Chair) Donald Hobern, ALA Director  
 Dr Paul Coddington, ARCS E/D nominee  
 Dr Ross Wilkinson, ANDS E/D nominee  
 Prof Jane Hunter, University of Queensland  
 Prof Hugh Possingham (Fed Fellow), user representative, UQ

### **Reference Group – to be drawn from:**

ALA Scoping Group, Australian Phenomics Network, Australian Plant Phenomics Network, CSIRO ICT Centre, CSIRO IM&T, ABIN, TERN, IMOS eMII, Global Biodiversity Information Facility, the Encyclopedia of Life, IdentifyLife and the Ocean Biogeographic Information System

**Project Manager: Lynette Woodburn, ALA/CSIRO**

### **Milestones**

Year 1	Core metadata repository implemented – basic repository available as basis for development, including basic registration of metadata Annotation store and metadata schema repository implemented providing the basis for development of other components
Year 2	Metadata tagging components, providing interfaces (including web services) for tagging metadata documents with ontology terms Metadata search interface implemented. Annotation retrieval services implemented, providing web services and user interfaces for search and retrieval of annotations
Year 3	Data retrieval interfaces implemented providing intelligent discovery and access of structured data from resources catalogued in repository Priority annotation user interfaces implemented and interfaces (widgets) for all priority document structures released

### **Service Outcomes**

Data Integration Services, including:

- Operational metadata repository for biodiversity data resources, including registration and discovery services

Annotation Services, including:

- Operational annotation repository for annotations relating to biodiversity data (but potentially open for use by users in any domain), including services to create, recover and harvest annotations
- Tools to use annotations for flagging data quality problems and possibly proposing corrections, and allowing data providers to use the annotations to correct the problem

Reusable software implementations and associated expertise for use by other NCRIS capabilities, etc.

### **Resources**

NeAT will fund 3.5 EFTs - 2 developers for Annotation Services at UQ and 1.5 developers for Data Integration Services at CSIRO.

The ALA will provide 0.5 EFT for Data Integration Services and 0.5 EFT for the project manager, as well as significant input from 4 ALA staff, particularly the ALA Architect and Director

The CSIRO ICT Centre will provide in-kind supervision and support for the developers

CSIRO IM&T will provide effort to host ALA and DIAS-B services, and eSIM will provide 3-4 EFTs to support ALA including DIAS-B  
 University of Queensland will supervise the Data Annotation Services developers

Project size is around 8 EFT pa

## Marine and Climate Data Discovery and Access Project (MACDDAP) - started

### Project Committee

(Chair) Prof Gary Meyers, IMOS E/D  
 Prof Anthony Williams, ARCS E/D  
 Dr Ross Wilkinson, ANDS E/D nominee  
 Prof Roger Proctor, eMii Director  
 Prof Andy Pitman, ARC NESS Convenor  
 Prof Nathan Bindoff, Uni of Tasmania and TPAC,  
 representing the user community and data providers

### Reference Group (tbc)

Prof Andrew Rohl, iVEC CEO; Rob Woodcock, AuScope, Dr Jon Blower, University of Reading; Dr Andrew Woolf, CCLRC; Tim Pugh, OpenDAP Inc.; Kim Finney, AAD; Kate Roberts, BlueNET; Peter Oke, SIRO/CAWCR, Amanda Lynch, Monash; Alex Sen Gupta, UNSW; Ian Suthers, IMOS

### Project Manager: Dr Ray Williams, TPAC

### Milestones

Discovery Services	MEST enhancements (GEONETWORK, OPeNDAP) TPAC Digital Library Portal enhancements (data registration, geospatial coordinate aware, gridftp, WCS and WMS service aware, simple visualisation) Enhanced OPeNDAP harvester (geospatially aware, gridftp, WCS and WMS service aware, simple visualisation) Aggregation Service for remote sensing
Access Services	OPeNDAP data delivered as WMS, WCS and trialled with AAF OPeNDAP enhancements (authentication, administration, remote management, and data handler enhancements, server-side functions)
Support Services	Translation Service (web based workflow for OPeNDAP to ISO19115 standards) The enhancements of the software, community profiles, translation tools and related workflows in the proposed services will be delivered back to the international community, through their respective mechanisms or provided on Sourceforge where appropriate

### **Service Outcomes**

General access to a wide variety of marine and climate data in a range of standard protocols, including Open Geospatial Consortium (OGC) standards, through OPeNDAP, the international standard access method

Support Services including workflow tools that process existing data sets in order to generate standard metadata and enable the data to be accessed via the OGC services.

Discovery Services to improve the conformity of data holdings, including a Metadata Entry Search Tool, an OPeNDAP digital library metadata harvester, a Catalogue Exchange Service and an Aggregation Service.

### Resources

NeAT will provide \$1M over three years to fund approx 3 EFTs pa for software and service development, at TPAC and CSIRO and some contracted commercial software developers

Key members of the user community (comprising BoM, CSIRO and TPAC) will contribute 2.45 EFTs per year for three years

IMOS community will contribute 3.2 EFT per year for three years (and \$30k for hardware)

ARCS and ANDS will each provide approx 1 EFT per year from operational services to support the project and the required service deployment

Project size is about 10 EFT pa

## Spatial Information Services Stack (SISS) - started

### **Project Committee**

(Chair) Scott McTaggart, CEO of Auscope  
 Dr Paul Coddington, ARCS E/D nominee  
 Dr Ross Wilkinson, ANDS E/D nominee  
 David Gray, CSIRO Minerals Down Under Flagship  
 Ben Searle, General Manager, Office of Spatial Data Management, Geosciences Australia

### **Reference Group**

Alan Willocks , Geological Survey of Victoria;  
 David Lemon, CSIRO Land and Water; Donald Hobern, Atlas of Living Australia; Nathan Bindoff, IMOS; Andrew Rohl, iVEC.

### **Project Manager: Robert Woodcock, CSIRO**

### **Service Outcomes**

An increase in the number and variety of spatial data sets made available through common access mechanisms.

A library of maintained common software components that can be deployed with spatial data holdings to make the holdings accessible within the spatial data commons.

A sustained level of expertise needed to maintain and deploy those components and assist holders of spatial data to deploy and operate relevant data servers and OGC compliant services

An OGC Catalog Service – supporting the registration of spatial data service meta-data.

A discovery portal providing access to federated registry/catalogue services.

### Resources

NeAT will fund 3 software developers hosted at iVEC and 0.4 EFT for the project manager at CSIRO

AuScope will provide 2.2 EFT pa (5 people) as a contribution to the development team

AuScope will provide the servers for testing, development, deployment and the project management and QA infrastructure

The user community as indicated in the milestones will resource the spatial data service deployments

Project size is about 6 development EFT pa and approx 2 EFT pa deployment effort

### **Milestones**

- July 2008-Dec 2008 OGC Catalog Services and Discovery Portal, hosted by iVEC for ARCS, and CSIRO for AuScope  
 Auscope GPS data WFS service deployed; Geological Survey of Victoria GeoSciML testbed collaboration deployed
- Jan 2009- Dec 2009 ARCS SISS support service fully available providing deployment and user assistance  
 Auscope National Virtual Core Library WFS service deployed  
 CSIRO Minerals Down Under Flagship deploys laterite geochemical data, airborne hyperspectral data service; thermodynamic data service; Northern Yilgarn hydro-geochemistry
- Jan 2010 – Dec 2010 Auscope Virtual Rock Laboratory and Tsunami workflows utilise SISS; Auscope deploys WCS for geophysics imagery with large data set support  
 Report on the advantages and disadvantages of SISS approach to data access and costs benefits analysis
- Jan 2011-June 2011 Auscope Earth Model and portal service infrastructure fully established and using SISS

## Data in Microscopy Imaging Neutron X-ray (Data-MINX) – Project Plan being developed

### **Project Committee**

Prof Cameron Keppert, convenor MMSN ARC Research Network, Fed Fellow, Uni Sydney (chair)

Prof Jill Trewella, Fed Fellow, Uni of Sydney, tbc

Prof Anthony Williams, ARCS E/D or nominee  
ANDS Executive Director, or nominee

Dr Allan Jones, Chair of AMMRF eResearch Committee

Nick Hauser, ANSTO

Richard Farnsworth, Australian Synchrotron  
Director of VeRSI, or nominee

**Project Manager: Dr Peter Turner, MMSN co-convenor, Uni of Sydney (tbc)**

### **Service Outcomes**

Easy and reliable transfer of experimental data from all major Australian characterisation facilities to all significant remote data storage or data repositories as determined by principle investigators.

The automated conversion of data to standard formats.

The automated capture, storage and transmission of associated metadata.

The authenticated sharing of data with colleagues as determined by principle investigators.

The means to publish data in a way that it is easily discoverable and accessible by any researcher.

### **Resources**

NeAT will provide \$1.5M over three years to fund additional software and service development support, about 5 EFT pa

The immediate user community (comprising ANSTO, AS, AMMRF and MMSN) will contribute 5.5 EFTs per year for three years

VeRSI and INTERSECT will contribute an EFT each per year for three years

ARCS will provide 1-2 EFT per year from operational services to support the project and the required service deployments

Project size is about 15 EFT pa

### **Milestones**

#### Year 1

- Development of a design solution for the federated data repositories, metadata catalog and associated web portal, based and incorporating outputs from work done by STFC, ARCHER, VeRSI and GRANI project
- Development of prototype AAA services based on AAF for the necessary data sharing, authorization and authentication mechanisms
- Deployment of data transfer service from ANSTO, some Australian Synchrotron beamlines, AMMRF facilities and X-ray labs to MARCS and other institutions that have appropriate storage support; and provision of data repositories for ANSTO and some AMMRF facilities.

#### Year 2

- Authentication and authorization mechanisms integrated with operational AAF and ARCS services.
- Development of documentation and researcher training material, some researcher training programs run
- Metadata extraction and interchange format services developed
- Some data processing workflow services implemented
- Widening the deployment of the system to additional appropriate facilities, instruments, beamlines and data types not addressed in Year 1

## ASSDA Services for e-Social-Sciences (ASeSS) – Project Plan being developed

### Project Committee

Deborah Mitchell, ASSDA (Chair of S/C)

Tony Williams, ARCS E/D

Paul Bonnington, nominee of ANDS E/D

Ben Evans, ANUSF

Mark Western, UQ

Peter Nicholson, DIISR

**Project Manager: Mr Stuart Hungerford,  
ANUSF**

### **Service Outcomes**

The foundation for an e-Social Science Virtual Organisation (SSVO) including:

- Data curation software that supports reliable data ingestion, and sets up the appropriate access controls.
- Search tools that support authenticated discovery across archives
- Integration of a suite of analytic tools that enable easy use by social scientists (the tools themselves will be developed outside this project)
- Integration of visualisation tools, particularly spatially oriented, and temporally oriented

### Resources

NeAT will provide \$1M over three years to fund additional software and service development support

The immediate user community based around ASSDA will provide 6 EFT per year

Some ARCS operational support will be provided to deploy ASeSS services at MARCS and other institutions hosting social sciences data

Project size is about 10 EFT pa

### Milestones

Dec 2008	Demonstration versions of GIS data visualisation web tool and longitudinal data analysis web tool
June 2009	Q/A component of curation service for unit record (UR) data; Establishment of e-Social Science Virtual Organisation web interface. Demonstration of cross-archive data search between two major archives; Historical Census and Colonial Data Archive (HCCDA)
Dec 2009	Time Series & Panel Archive, and Qualitative Data Archive available; Text based analysis tool demonstration for Qualitative Data Archive; Search available on UR data archive
June 2010	SSVO web site upgraded to include search for Qualitative Data and HCCDA; longitudinal data analysis web service; data exchange service for Time Series & Panel Data and Qualitative Data; Q/A component of curation software for Qualitative data Production version of web-based UR archive curation service on VO web site
Dec 2010	SSVO web site upgraded to include Time Series & Panel Archive services; Q/A component of Indigenous data archive curation service; search on Electoral Database
June 2011	Production version of web-based Qualitative data archive curation service and an Indigenous data archive curation service on VO web site. Search available over Indigenous archive; Generalised version of GIS service on UR data available on VO web site

