



## ARDC Location Infrastructure

### 1 Background

ANDS has been established to catalyse the emergence of the Australian Research Data Commons.

The term ‘commons’ traditionally refers to resources that are made available for community use and re-use. ANDS is creating a digital realisation of this concept for research data to enable greater data sharing and re-use.

Research data is all data used by and created by researchers. The Australian Research Data Commons is a combination of several related attributes:

- the set of shareable Australian research collections;
- the descriptions of those collections including the information required to support their re-use;
- the relationships between the various elements involved (the data, the researchers who produced it the instruments that collected it and the institutions where they work), and
- the infrastructure needed to enable, populate and support the commons

ANDS does not hold the actual data, but provides electronic pointers to the location where the data can be accessed. ANDS may sponsor projects that establish strategic infrastructure at research organisations and government agencies, using a distributed model where infrastructure is typically operated on an ongoing basis by the organisations and agencies themselves.

### 2 Introduction

Inter-disciplinary and eResearchers increasingly require the integration of data and information from spatial and non-spatial sources. An example might be overlaying real-time disease outbreak by common location name onto a transport surface or locating human experience by region within a spatial representation of a climate change scenario.

A recent review conducted by ANDS has confirmed that there is an increasingly urgent need for Australian eResearchers to have the capacity to link both current and past place names to their location in space and to other geographies—such as traditional or administrative boundaries or regions—and to other attributes, spatial and non-spatial. Such a service needs to be web-based, machine readable and readily searchable.

### 3 A location infrastructure for the data commons

An important goal of the Australian Research Data Commons is to enable cross-disciplinary discovery of related research data, and spatial location is a vital linkage mechanism in this process. The value of the data commons will be increased if the dataset descriptions include spatial coverage data encoded as geographical points or polygons rather than just text.

ANDS vision for a data commons would see non-GIS-experts from arts, humanities, and science able to enrich their discipline specific data with standardised spatial information.



Achieving this goal requires the establishment of a robust national infrastructure that would allow place names to be validated by both individuals and software systems against an Australian gazetteer service in an efficient manner. There will need to be distributed sources of gazetteer data, depending on jurisdiction, feature types, temporal coverage and language. A comprehensive national service will need to provide inter-operable query services across these sources, irrespective of their differing construction.

This infrastructure is intended to increase the amount and quality of spatially-marked-up research data. This will enable new kinds of research and innovation based on new data linkage and data merging opportunities. The infrastructure aims to unlock significant innovation and productivity. It will bring benefits well beyond the research and innovations sector.

ANDS will partner with long-lived government agencies and research institutions to establish the various components of location infrastructure which the agencies and institutions will operate into the future.

#### **4 An enabling infrastructure to benefit research**

The research and innovations sector is the primary beneficiary of the ANDS initiative. The location infrastructure projects should benefit researchers and innovators by enabling more data to be spatially enriched. This includes both:

- Data *used by researchers* (i.e., data from various agencies and research organizations); and
- Data *created by researchers* themselves

The infrastructure should enable agencies and research organizations to add spatial dimensions to their data holdings. This will benefit the end users of that data in the research and innovations community and beyond.

#### **5 A spectrum of needs**

The infrastructure will need to enable and support a range of government agencies, research organizations, research groups and individual researchers. For example, in the research sector, humanities researchers are often interested in the previous name of a location; ornithologists may want to plot a precise location and may only be secondarily interested in its name; oral historians might be interested in eliciting commonly held (perhaps conflicting) views on the name or extent of a location; social scientists and health researchers typically map from a location name to a larger location (suburb, post code, census block) in order to maintain the privacy of their subjects; natural resource managers may be interested in mapping the agricultural activity only within a location defined by its rainfall; observational scientists may want to refer to a creek rather than a town with the same name; marine scientists are interested in named and unnamed locations in the ocean and below.

With such diverse groupings there will necessarily be a spectrum of needs, technical expertise and expectations and thus the ANDS approach needs to recognise and address this logically. The approach also needs to recognize that the infrastructure needs to be scalable in terms of coverage and function as and when appropriate, but commencing with a significant first step.

The following principles have been adopted to help guide the development of ARDC Location Infrastructure:

- A phased development approach will be adopted, where the earlier phases will focus on addressing the most commonly articulated needs for which a ready solution seems extant



- For example, the first phase is likely to be focused on web-enabling existing, maintained, location sources (databases)
- Equal emphasis to be placed on supporting individual researchers (e.g. via a web interface) and automated (machine-readable) services
- Infrastructure development to use existing national and international formats, standards and protocols
- Design will be scalable in coverage and function

## 6 Phased Approach

In consultation with the research community, the Committee for Geographical Names of Australasia (CGNA), the Spatial Information Council (ANZLIC) and Geoscience Australia (GA), ANDS will prepare a separate options paper on future directions (post phase 1) of the ARDC Location Infrastructure. In the first phase, the focus will be on producing an electronic gazetteer which is essentially about point locations; some digital boundaries may be included, where these help provide context for the locations. In the second phase the focus may be on the inclusion of marine locations and well as more digital boundaries. Future phases may include historical information such as former names of features, the capacity to deal with informal names or names used by indigenous communities, the capacity to deal with fuzzy boundaries (e.g. regions which are described anecdotally) or the capacity to accept information (such as revisions) from outside of the system (i.e. *crowd-sourcing*).

## 7 The broader ANDS location infrastructure agenda

ANDS is also committed to promoting the use of this infrastructure. ANDS supports the building of capability within the research sector to make use of the location infrastructure. The broader location infrastructure agenda includes:

- developing resource materials and best-practice guides
- hosting training and workshops
- promoting the use of open international standards
- furthering the policy and governance frameworks for location infrastructure
- working with research organisations to help data repositories, research instruments, and cataloguing tools to be more “spatially” capable.

For more information on the ANDS location infrastructure, please contact:

Adrian Burton

Director, Services, Australian National Data Service

[adrian.burton@ands.org.au](mailto:adrian.burton@ands.org.au)

July 2010