

What ANDS might look like

Components of ANDS

ANDS covers four broad areas:

1. physical storage of data;
2. development of a system for overall national data management;
3. operations of ANDS;
4. providing end user support to researchers.

Physical storage of data

ANDS should not care where data is physically stored, as the overarching management system will be aware of where data is.

There are quite a few sizeable silos for research data around the country at present (ACT, Vic, QLD), but none in NSW, unfortunately. It is not clear whether ANDS will fund physical data storage, as there are competing priorities, but if funding is available through ANDS, this would provide an opportunity for the acquisition of a sizeable data storage facility, which would be located at **ac3**, adjacent to the supercomputers that would invariably be used to analyse the data.

Development of the ANDS Management framework

This is a project — i.e. it has a start point and an end point, although there will invariably be ongoing enhancements, like any product. The specifications are yet to be developed, but there will be two main functions:

1. **Depositing data:** When a research project is being wound up, or when researchers wish their results to be available globally, they submit their results to ANDS via an intuitive web-form. This web form captures all the necessary meta data, which is used to create the directory entry for the data. If not already done so, the data is then ingested to one of the ANDS approved data repositories. This activity happens once only.
2. **Retrieving Data:** Researchers use this facility to source and retrieve data. Whilst the depositing of data occurs once, data retrieval may occur many times. A search engine is required, along with appropriate permissions engine (sourced from AAF). The CSIRO developed *FunnelBack* is a candidate search engine.

Operations of ANDS

Having completed product development and beta testing, then the operational system needs to reside somewhere where there is 24x7 support. It is proposed that **ac3** host the ANDS infrastructure (which would be funded by NCRIS), providing first level support through the existing Support Desk. Second level support would be provided by Melcoe. The hardware required would not be significant — possibly a web server and a business server, both duplicated for redundancy — but they do need to be managed like any other operational system.

Supporting end users

A research team has, to their delight, discovered several sources of data that relate to their current work, stored physically in different parts of the country, and perhaps overseas. ANDS informs them that they are entitled to use the data, perhaps with some restrictions. But the databases are all stored in different formats. So they approach one of the state-based e-Research Support centres for assistance. In due course the IT experts in the eResearch Support Centre provide a framework to the research team that make the disparate data bases look like a single logical database. The research team is able to carry out their research without the need to hire an IT specialist, or to become part time computer scientists themselves.

In the case of NSW, this end user support would be carried out by the NSW eResearch Support Centre.

P. McCrea