

## 4th International Digital Curation Conference

December 2008

### Design and Implementation of the Australian National Data Service

Andrew Treloar,  
Director, Australian National Data Service Establishment Project,  
Monash University, Australia

Dec 2008

#### Abstract

This paper will describe the genesis and realisation of the Australian National Data Service (ANDS). It will commence by outlining the context within ANDS was conceived, both in the international research and Australian research support domains. It will then describe the process that brought about the ANDS vision and the principles that informed the realisation of that vision. The paper will then outline each of the four ANDS programs (Developing Frameworks, Providing Utilities, Seeding the Commons, and Building Capabilities) while also discussing particular items of note about the approach ANDS is taking. The paper concludes by briefly examining related work in the UK and US.

## Introduction

There is a rich research literature dealing with the changes to the scholarly communications system brought about by the rise of the Internet (Borgman, 2007). These changes have been underway since the late 1980's, but have accelerated rapidly in the last decade. As a result, a majority of all the scholarly publications produced each year are available online, and in time this will be true for all scholarly publications produced ever. Some of these publications are available in toll-access publications, some in open-access publications. The latter category, augmented by pre-prints, are often captured and made available through various institutional repositories (Lynch, 2003).

The rise of data-driven research (Hey and Trefethen, 2003) has meant that publication of the research publication alone as a largely textual document will increasingly not be seen as sufficient. The reader will need access to the data, and maybe the associated computational models, used to derive the conclusions being argued for. As well, the data will need to be managed for the long term as part of the scholarly record. Consequently, a number of significant reports (Association of Research Libraries, 2006), (National Science Board, 2005), (ACLSCCHSS, 2006) have argued in recent years for a more systematic set of solutions to long-term curation of research data outputs. It was in this context that Australia started to look at a national approach.

## ANDS Context

The Australian National Collaborative Research Infrastructure Strategy was funded in FY 2004/5, and the programs commenced in FY 2006/7. As well as up to fifteen areas of identified research capability, the NCRIS plan recognised the need for a sixteenth area: Platforms for Collaboration. This area is made up of a number of discrete services, addressing the need for ongoing network investment (Australian Research and Education Network – AREN), a co-ordinated approach to authentication (Australian Access Federation – AAF), collaboration and middleware (Australian Research Collaboration Services – ARCS) high-performance computing (National Computational Infrastructure – NCI), and data-management and sharing (Australian National Data Service – ANDS). This paper will focus on ANDS.

ANDS arose from a series of consultations undertaken during 2006 and early 2007 by Dr Rhys Francis (now Executive Director of the Australian eResearch Infrastructure Committee, the advisory committee for Platforms for Collaboration). A consistent theme arising from the consultation meetings was the need to improve data management and availability. In April of 2007, a forum with wide representation met and agreed to set up an ANDS Technical Working Group (ANDS TWG). The TWG produced in October 2007 a report (Australian National Data Service Technical Working Group, 2007) that laid out a vision for how ANDS might operate. This document envisaged four programs of activity within ANDS: Frameworks, Utilities, Repositories, and Researcher Practice.

In late 2007, the Australian Commonwealth Department of Education, Science and Training (DEST) asked Monash University to be the lead agent (working in close collaboration with ANU and CSIRO) on a project to establish ANDS. This project

commenced in January 2008 and will conclude in late 2008 with the submission of its final report. This project conclusion will overlap with the commencement of ANDS itself, which came into existence in September 2008. The author was a member of the ANDS TWG, was subsequently appointed as the Director of the ANDS Establishment Project, and will likely be one of the ANDS Deputy Directors.

## **Establishing ANDS**

The Australian National Data Service (ANDS) is being established to ensure that Australian research data is well managed, made available for access, and discoverable so that researchers can find and access any relevant data in the Australian ‘research data commons’, Australian researchers are able to discover, exchange, reuse and combine data from other researchers and other domains within their own research in new ways, and Australia is able to share data easily and seamlessly to support international and nationally distributed multidisciplinary research teams. The overall goal for ANDS is to “deliver greater access to Australia’s research data assets in forms that support easier and more effective data use and reuse” (ANDS TWG 2007, p. 18). This will be achieved by the creation of the Australian Research Data Commons (ARDC), a virtual space which aggregates information about available datasets (ANDS TWG 2007, pp. 10-17). The centrality of the ARDC can be demonstrated by the decision to refer to ‘ANDS: Building the Australian Research Data Commons’ in the marketing materials.

Consistent with these goals, over the course of 2008 the ANDS Organisational Network (ANDS ON) developed an Interim Business Plan (Australian National Data Service 2008). It will be updated in March 2009. At the same time ANDS Project Management Committee focussed on the Agreement for the collaboration that will be responsible for ANDS, the contract between this collaboration and the Department of Innovation, and the search for the ANDS Executive Director.

The ten-year goals for ANDS (bearing in mind that it will initially only be funded for three years) are that:

- A. A national data management environment exists in which Australia’s research data reside in a cohesive network of research repositories within an Australian ‘data commons’.
- B. Australian researchers and research data managers are ‘best of breed’ in creating, managing, and sharing research data under well formed and maintained data management policies.
- C. Significantly more Australian research data is routinely deposited into stable, accessible and sustainable data management and preservation environments.
- D. Significantly more people have relevant expertise in data management across research communities and research managing institutions.
- E. Researchers can find and access any relevant data in the Australian ‘data commons’.
- F. Australian researchers are able to discover, exchange, reuse and combine data from other researchers and other domains within their own research in new ways.
- G. Australia is able to share data easily and seamlessly to support international and nationally distributed multidisciplinary research teams. (ANDS

TWG 2007, p. 6)

The funding for ANDS will be A\$21M over three years, commencing in the second half of 2008 and concluding in the middle of 2011.

### **ANDS Principles**

In seeking to turn Australian National Data Service Technical Working Group (2007) into a business plan and work towards these goals, the ANDS Establishment Project identified a number of principles to be followed. These are essentially the theoretical underpinnings for the implementation approaches that will be followed.

#### ***Commons Framework***

ANDS will start in a way that anticipates the need to scale up and adapt over time via an extensible framework of data stores, federations and services that enable better data creation, capture, management and sharing.

#### ***Focus***

ANDS will identify and work with those who are ready, willing, and able to contribute significantly to the Australian Research Data Commons vision, and who provide the most strategic return to the Australian Research Data Commons for the effort expended.

#### ***Content***

ANDS will initially focus on content recruitment into stores and federation across stores so as to achieve a wide coverage of data quickly at an agreed level of quality; in later years the emphasis will shift towards quality improvement.

#### ***Service Provision***

ANDS is focussed on service provision, not research and exploration; its programs will develop, integrate, and continually improve production-level systems in support of well-understood services. A separate process will fund the development of more innovative and exploratory domain-focused initiatives that may become ANDS services in later years.

#### ***Strategic Partners***

ANDS recognises the need to be open to, and engage appropriately with, innovations and external institutions relevant to the nascent Australian Research Data Commons, including the Access Australia Federation (AAF – a national authentication system based on Shibboleth and PKI) and the Australian Research Collaboration Services (ARCS – providing collaboration tools and an SRB-based data fabric).

#### ***Stores***

ANDS assumes an environment where storage and long-term curation primarily occur in institutionally-supported stores, either existing or brought into being over the life of ANDS. ANDS will facilitate public/restricted access and re-use across these institutional stores. These stores will preferably hold objects described by various discipline-specific and documented metadata schemas. Note that ANDS is not funding storage. Nor is it setting up discipline-aligned data centres (the lack of these is one of

the key differences between the Australian and UK environments).

### *Sustainability*

Research data management requires a long-term commitment. ANDS is developing its three-year plan on the assumption that this does not represent a one-off investment in data. The enduring changes forecast in this document within each program are also intended to be sustainable beyond the end of the ANDS planning period.

## **ANDS Programs**

To enable researchers to work in the new world of data-intensive research outlined in the introduction to this paper, they will need:

- Policies that support a new way of working
- A technical data fabric that enables storing and moving data
- A repository to store their data well
- A referencing mechanism that enables input data, modelling outputs (such as visualisations), software code and documents to be cross referenced
- The ability to search across all the collections that have been registered
- The training and training materials that enable the infrastructure to be used well

The research infrastructure to be provided by ANDS to enable this will be delivered under four programs of activity (Developing Frameworks, Providing Utilities, Seeding the Commons, Building Capabilities). These are different to the programs originally envisaged in ANDS TWG (2007). The need to focus what ANDS does has been driven by the complexity of the environment and the limited funding available. This meant that it was not possible to achieve critical mass in the planned Repositories and Researcher Practice programs. As a result, the original programs envisaged in Australian National Data Service Technical Working Group (2007) have morphed into Developing Frameworks, Providing Utilities, Seeding the Commons (the Research Data Commons, that is), and Developing Capabilities. Some elements of the underlying data fabric will also be delivered through the Australian Research Collaboration Services (ARCS). ANDS will work with the National e-Research Architecture Taskforce (NeAT) in defining projects that will primarily contribute towards the Seeding the Commons and Providing Utilities programs. Some of the outputs of these activities will appear in ANDS programs in later years. So, what exactly are the ANDS Programs going to achieve, and how are they approaching their task?

## **Developing Frameworks**

### *Overview*

The Frameworks program aims to influence and simplify the overall policy framework within which the ANDS goal is to be achieved, as well as defining how activities by researchers (in order to comply with their grant conditions) and institutions (in order to comply with funding requirements) can contribute to a national research data commons. The Frameworks program will progressively work to simplify and reduce the number of licenses under which data is created and shared.

The primary collaborators in the Frameworks Program are institutional data holders, national initiatives such as the National Committee for Data for Science, Cross-governmental groups, research funding government departments, research funding schemes, discipline leaders within institutions, and research office staff at institutions. The focus of Frameworks will, where possible be on the higher-level aggregations of data-managing entities rather than the individual entities themselves.

### *Discussion*

ANDS is able to benefit in the Frameworks program from an environment that is somewhat unique. Australia is a small country (population of 21M), with a fairly homogenous university sector as the result of reforms that took place in the 1980/1990s. All Australian universities cover a broad range of disciplines, and all span the range from undergraduate through doctoral degrees. This is quite different from the very heterogenous US situation, for instance. Because of the structure of the Australian Constitution, the Commonwealth Government is responsible for funding the higher education sector (although the proportion that this funding makes up of the total funding required has steadily fallen to an average of 40% across the sector). This has led to a tradition of government central direction and policy setting, particularly with respect to research funding. ANDS plans to take advantage of this to encourage desired data management behaviours. A positive development was the release last year of the Australian Code for the Responsible Conduct of Research (NHMRC/ARC/UA 2007). This laid out obligations on researchers for data management and on institutions for data management support. Recent statements from research-funding bodies also indicate that they are keen to move towards progressively requiring greater data deposit and sharing.

## **Providing Utilities**

### *Overview*

This program will provide a range of utility data services at a sector-wide level (such as cross-discipline discovery services; national collection registry; persistent identifier service; federation registry; access policy registry), as well as improving existing and develop new utility services. The human-facing utility services will use the AAF authentication services and ARCS authorisation services (informed by ANDS requirements) for access control. Note that some of the proposed registries (collection, federation, access policy) are essentially enabling; most users will never see these directly but they will contribute critically to things that the users will see. As well as the provision of utility services, ANDS will develop a Federation Utilities National Framework. This will catalyse and support data federation utilities at the level of research communities and their federations. It will also work to develop and maintain community consensus within the research and public sector on a technical framework for the interoperability of data utility services (registries, metadata catalogues, schemas and harvesting guidelines, search services, etc). Development of this framework will be undertaken in collaboration with related UK and US activity. The primary targets for the Utilities program will be content providers and consumers in the data commons, data facility managers and administrators, and research communities building and operating (new or already existing) data federations



---

## ***Discussion***

One of the challenges in the design of the Utilities program has been the relationship between possible Australian services and international equivalents. Anecdotally, most researchers see their primary loyalty as being to their disciplines, rather than their institutions. This means that whatever ANDS chooses to do cannot be inconsistent with disciplinary practice. For some disciplines, this means that the ANDS role will be quite limited. It also means that there may be little value in running (for instance) a narrowly Australian service when the researchers in Australia will be working with colleagues overseas who are using an established international equivalent.

## **Seeding the Commons**

### ***Overview***

The Commons Program will start to seed the Australian Research Data Commons by seeking to make more content available through it. The amount of funding available for ANDS is insufficient to provide or support a repository solution across the entire research-producing sector. As the marginal cost of working with people is high, and ANDS only has limited funding, this program will be highly targeted, providing focussed assistance to a number of target groups. The aim of the program is to ensure that the data and metadata generated within the program targets is captured, stored and made accessible through the Australian Research Data Commons. ANDS will target automated/ semi-automated data and metadata capture to simultaneously improve the quality of what is captured and increase the quantity (by reducing the barriers to capture). ANDS will also aim to recruit information about data collections to be made available through the ANDS discovery services. Once groups have been selected (according to criteria that will be developed in year one) ANDS will fund staff to work within them and the selected groups will provide the long-term storage infrastructure. It is expected that the researchers will need to be key players in this process for it to be effective. In the first year of ANDS the main activities related to this theme will be finding and training these staff (through the Capabilities program described below), and deciding on the most appropriate selection option, and the detailed criteria for selection.

At the same time, in year one ANDS will also concentrate on the recruitment of existing content into repositories, identifying existing repositories of useful content, and making all that content discoverable through the Australian Research Data Commons. Where institutions with valuable existing content do not have the required systems, ANDS will provide small agile teams to set up data repository infrastructure based on already developed toolkits. If demand for this exceeds available capacity, ANDS will develop a transparent process for allocation of ANDS resources. Where repositories (or federations) already exist, this theme will assist with its integration into the Australian Research Data Commons. This may be work performed by staff in the program target (with ANDS advice if needed) or by ANDS staff (building on the technical consultancy expertise in the Capability program). This theme will primarily operate in year 1 (because the focussed assistance program won't start until year 2), but will also continue in years 2 and 3 if there is demonstrated demand for it and it is delivering strategic value.

### *Discussion*

One of the initial decisions made by the ANDS TWG was that ANDS would not provide storage for data (this is one of the ANDS Principles). ANDS assumes that data will be stored in institutional repositories, consistent with the recommendations contained in NHMRC (2007). This is different to the UK discipline data centre approach, and was a deliberate decision. Institutions have obligations under NHMRC (2007) which they are now starting to grapple with, they employ the researchers who produce the research outputs, and they are likely to have a longer-term presence than a funded data centre. On the other hand, each institutional repository won't necessarily have the expertise to curate for the long term a widely heterogeneous assemblage of data objects. ANDS will need to work closely with institutions to ensure that they are aware of the range of curation challenges that they will need to address. ANDS will also need to deal with a very heterogeneous mix of repositories that will be integrated into the Australian Research Data Commons. The solution here will be rely on open standards wherever possible. ANDS will retain specialist expertise to assist with integration issues. Of course, none of this activity will happen in a vacuum, and ANDS will seek to learn from related activity overseas. A number of JISC projects are of interest: the Data Audit Framework<sup>1</sup>, SCARP<sup>2</sup>, and the DISC Datashare project<sup>3</sup>. ANDS is aiming to pilot the Data Audit Framework at a higher level and in less detail and to feed back its experiences with the online tool.

## **Building Capabilities**

### *Overview*

The Capabilities program will build capability across the research and scholarly communications lifecycle in organisations, systems, services and people. This program will work with the sector to produce a capability maturity model for e-research and information infrastructure. It will provide audit, rating and certification systems and services. ANDS will also work with institutions to coordinate the development of national curricula for capability development. It will coordinate, enhance, and add national focus to institutionally based training initiatives. ANDS will base the activities of this program in the community by establishing a forum drawn from data stewards who are providing retention and access services to research data. These will include research-intensive organisations as well as government instrumentalities holding data of interest to researchers. Staff from this program will collaborate with and support the more targeted activity within the Seeding the Commons program. As cohesive networks of research data are increasingly regarded as an important and enduring part of the collaborative research infrastructure, this program will focus in particular on building the capability of researchers and support staff to contribute to and better exploit national data infrastructure.

### *Discussion*

Because of the size of the sector, ANDS will need to rely on leveraging activity already underway within institutions. Many Australian universities are starting to tackle the need for improved data management practices among researchers and

---

<sup>1</sup> JISC Data Audit Framework:

<http://www.jisc.ac.uk/whatwedo/programmes/digitalrepositories2007/dataauditframework.aspx>

<sup>2</sup> SCARP <http://www.dcc.ac.uk/scarp/>

<sup>3</sup> DISC DataShare <http://www.disc-uk.org/datashare.html>



research support staff. ANDS hopes to provide ways of sharing good practice and learning from relevant overseas activities. A recent JISC report (Swan and Brown 2008) will be particularly relevant here.

## **Benefits**

### ***For Researchers***

ANDS will improve data management by researchers and research communities in the groupings that are the targets of the Seeding the Commons focussed assistance activity. ANDS will seek to give priority to data federating communities where significant improvements to practice and the consequential quantity and quality of accessible and re-usable research data can be achieved. Researchers will also benefit through researchers and research communities being assisted to reuse data within and across disciplines through practice improvements, support for search and access tools, meta-data capture tools and work on implementations that support community agreed frameworks and standards. The result will be enhanced practice by researchers in managing data, the lodgement of increased volumes of high quality data into the commons, and the wider deployment of tools and practices that support effective data reuse.

### ***For Institutions***

Institutions will be assisted to improve their data capture and retention for research data outputs, for a range of purposes, including:

- as an evidentiary record and/or archive holding representing the research activity funded and pursued under the organisation's brand and with its imprimatur;
- to ensure evidence exists in support of published research claims, including the provenance of that data, so that important research can be repeated and verified when required;
- to contribute to the creation and expansion of national data assets;
- to provide certainty of access and ease of use in subsequent or further research by the originators or such assignees as may be determined or agreed (eg. in support of commercialisation);
- to ensure ease of reuse by research collaborations in which the researchers or the institution wishes to participate, particularly where cross-disciplinary;
- as part of the development of profile enhancing collections; and
- as part of an efficient process for contributing research data outputs into general access collections, for discovery and reuse in serendipitous modes. (Australian National Data Service Technical Working Group (2007), p. 12)

They will also be assisted to meet their obligations under the Australian Code for the Responsible Conduct of Research (NHMRC/ARC/UA 2007).

### ***For Government***

The benefits of ANDS for Government are at least twofold.

In the first place, the Australian Commonwealth in general (and the Department of Innovation, Industry, Science and Research in particular) are concerned to ensure that Australian research is as effective as possible. ANDS, through its provision of

improved access to better managed data, will encourage data sharing and re-use, thus facilitating synergistic and serendipitous research outcomes.

The Commonwealth Government is also keen to ensure that taxpayer funded research is accessible to the taxpayers that provided the funding. This is consistent with other initiatives in the UK and EU around improved access to Public Sector Information (PSI). This accessibility for publicly funded research outcomes is the subject of a Commonwealth Government Accessibility Framework currently being developed. While not everything that is discoverable through the Australian Research Data Commons will be publicly accessible, much will. There is also the potential for ANDS to facilitate greater engagement between the public and the researchers, both as consumers and in some cases co-creators of the research.

### **Beyond ANDS**

The review of the NCRIS Roadmap has already indicated a need to move beyond the ANDS approach of relying on institutional repositories. The final report (National Collaborative Research Infrastructure Strategy, 2008) has foreshadowed the need for a national data grid to provide for long-term preservation of data. A dedicated high performance network would link the nodes in this grid, allowing researchers to move data rapidly from instrumentation to computing resources and to institutional storage. Interestingly, this proposal is foreshadowed to extend to partnering with research organisations for the development of institutional nodes of the storage grid. This would be on the condition that the storage is used exclusively for research data; that the institutes co-invest in the infrastructure; that each institute publishes and adopts a data management plan; and that each institute ensures its researchers use and abide by the data management plan. Of course, there is no way of anticipating when this will become reality, or whether it will be funded.

At the same time as ANDS commences its work, two related initiatives are underway in the UK and the US. The United Kingdom Research Data Service<sup>4</sup> is undertaking work to develop an understanding of UK's current and future research data service needs, work with other UK stakeholders to identify priorities for action, develop a number of scenarios for a possible service, and develop a detailed business plan for the preferred option(s). The Interim Report from this activity (UKRDS 2008) recommends a model with a number of similarities to the approach taken by ANDS. In the US, the National Science Foundation is about to announce the successful projects to be funded under the first round of the Sustainable Digital Data Preservation and Access Network<sup>5</sup> (DataNet) call. The DataNet call aims to build sustainable infrastructure by creating a new type of organization that the NSF does not believe exists today. They are looking for librarians, archivists, and computer/computational/information scientists who will together to build excellent infrastructure for science and/or engineering, while engaging deeply with intended users; domain scientists will be full partners in the process.

### **Conclusions**

The Australian National Data Service is taking an approach that tries to pragmatically blend learning from existing and preceding initiatives in Australia and

---

<sup>4</sup> UKRDS - <http://www.ukrds.ac.uk>

<sup>5</sup> DataNet: [http://www.nsf.gov/funding/pgm\\_summ.jsp?pims\\_id=503141](http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503141)

overseas, while working within the constraints of the available budget. The task of achieving the hoped-for cultural change in data management practice is significant, but there appears to be a readiness in the various sectors to engage with the ANDS program of activities. As we learn from our early experiences and finetune our activities over the next three years, we in Australia hope to be able to continue to learn from our colleagues in the data management community internationally and share our learnings and outputs in a spirit of collaboration.

## Acknowledgements

The author wishes to acknowledge the members of the original ANDS Technical Working Group (who contributed so significantly to developing the vision articulated in *Towards the Australian Data Commons*), the members of the ANDS Organisational Network (for their enthusiasm and engagement with the task), and the members of the ANDS Project Management Committee (for their commitment to work collaboratively to bring ANDS into being).

## References

- [report]American Council of Learned Societies Commission on CyberInfrastructure for the Humanities and Social Sciences (ACLSCCHSS) (2006). *Our Cultural Commonwealth*. Retrieved July 25, 2008 from [http://www.acls.org/uploadedFiles/Publications/Programs/Our\\_Cultural\\_Commonwealth.pdf](http://www.acls.org/uploadedFiles/Publications/Programs/Our_Cultural_Commonwealth.pdf).
- [report]Association of Research Libraries (2006) *To Stand the Test of Time: Long-term Stewardship of Digital Data Sets in Science and Engineering*, Association of Research Libraries, 2006. [www.arl.org/pp/access/nsfworkshop.shtml](http://www.arl.org/pp/access/nsfworkshop.shtml)
- [report]Australian National Data Service (2008). *Interim Business Plan*. <http://ands.org.au/andsinterimbusinessplan-final.pdf>
- [report]Australian National Data Service Technical Working Group (ANDS TWG) (2007). *Towards the Australian Research Data Commons*. Retrieved July 24, 2008, from <http://www.pfc.org.au/bin/view/Main/Data>
- [book]Borgman, C.L. (2007). *Scholarship in the digital age: Information, infrastructure, and the Internet*. Cambridge, MA: MIT Press.
- [report]Digital Curation Centre. (2005). *Digital curation and preservation: Defining the research agenda for the next decade*. In Report of the Warwick Workshop, November 7-8, 2005. Digital Curation Centre: Warwick, UK. Retrieved July 27, 2007 from [http://www.dcc.ac.uk/events/warwick\\_2005/Warwick\\_Workshop\\_report.pdf](http://www.dcc.ac.uk/events/warwick_2005/Warwick_Workshop_report.pdf)
- [book chapter]Hey, T., Trefethen, A. (2003), A. "The Data Deluge: An e-Science Perspective," *Grid Computing: Making the Global Infrastructure a Reality*.

(Chichester: Wiley, 2003), pp. 809-24.

[www.rcuk.ac.uk/cmsweb/downloads/rcuk/research/esci/datadeluge.pdf](http://www.rcuk.ac.uk/cmsweb/downloads/rcuk/research/esci/datadeluge.pdf)

[internet site]Joint Information Systems Committee (2008). *Information Environment*. Retrieved 25 July 2008 from

[http://www.jisc.ac.uk/home/whatwedo/themes/information\\_environment/](http://www.jisc.ac.uk/home/whatwedo/themes/information_environment/)

[internet journal]Lynch, C.A. (2003) *Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age*, ARL Bimonthly Report, Vol 226 (February 2003), pp. 1-7. [www.arl.org/newsltr/226/ir.html](http://www.arl.org/newsltr/226/ir.html)

[report]Lyon, L (2007). *Dealing with Data: Roles, Rights, Responsibilities and Relationships*.

<http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/publications.html#2007-06-19>

[report]National Collaborative Research Infrastructure Strategy (2008). *2008 Strategic Roadmap for Australian Research Infrastructure Exposure Draft*. Retrieved July 25, 2008 from <http://www.ncris.dest.gov.au/NR/rdonlyres/B6230B56-DDF0-42DC-8BA7-C227976343EE/22780/ExposureDraft2008StrategicRoadmapforAustraliaResearch.pdf>

[report]NHMRC/ARC/UA (2007). *The Australian Code for the Responsible Conduct of Research*. Retrieved November 4 from

<http://www.nhmrc.gov.au/publications/synopses/r39syn.htm>

[report]National Collaborative Research Infrastructure Strategy (2008). *Strategic Roadmap for Australian Research Infrastructure*. Retrieved November 4 2008 from

<http://www.innovation.gov.au/ScienceAndResearch/Documents/Strategic%20Roadmap%20Aug%202008.pdf>

[report]National Science Board (2005). *Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century*, National Science Foundation, 2005. [www.nsf.gov/pubs/2005/nsb0540/start.jsp](http://www.nsf.gov/pubs/2005/nsb0540/start.jsp)

[report]Swan, A and Brown, S (2008). *The skills, role and career structure of data scientists and curators: an assessment of current practice and future needs - Report to the JISC*. Retrieved Nov 4 2008 from

<http://www.jisc.ac.uk/publications/publications/dataskillscareersfinalreport.aspx>

[report]UKRDS (2008). *Interim Report on the UK Research Data Service Feasibility Study*. Retrieved November 4 2008 from

[http://www.ukrds.ac.uk/UKRDS%20SC%2010%20July%202008%20Item%205%20\(2\).doc](http://www.ukrds.ac.uk/UKRDS%20SC%2010%20July%202008%20Item%205%20(2).doc)