

QUT's Responses to the Australian Code for the Responsible Conduct of Research [Data Management]

QUT Library



Martin Borchert

Associate Director, Library Services

(Information Resources and Research Support Services)



The Australian Code...

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

Section 2. Management of research data and primary materials

... Is a business driver for data management ...

Policies are required

Retention

Secure storage

Record keeping

Identify ownership

Ensure security and confidentiality

Data should be available

The QUT Code...

http://www.mopp.qut.edu.au/D/D_02_06.jsp

Refers to the Australian Code

Provides a policy framework

Elevating section on data management to its own policy

University Policy on Management of Research Data

Soon to be available via
QUT's Manual of
Policies and Procedures
<http://www.mopp.qut.edu.au/>

Endorsed by University
Research and
Innovation Committee

Planning
IP & copyright
Storage
Retention
Record keeping
Privacy and confidentiality
Access & reuse
Disposal

Building Capacity and Capability

Research Data Management Team (High Performance Computing, ITS)

Associate Director with research support portfolio

eResearch Access Coordinator

Research Support Librarian

Data Librarians (coming)

Research Data Management Guidelines

Currently in development

Were considering developing a Data Plan Template, but after consultation with Monash, this will be a Checklist

Keep it flexible and simple

Collaborations

Monash University

Griffith University

Australian National Data Service (ANDS)

Australian Access Federation (AAF)

Queensland Cyber Infrastructure
Foundation(QCIF)

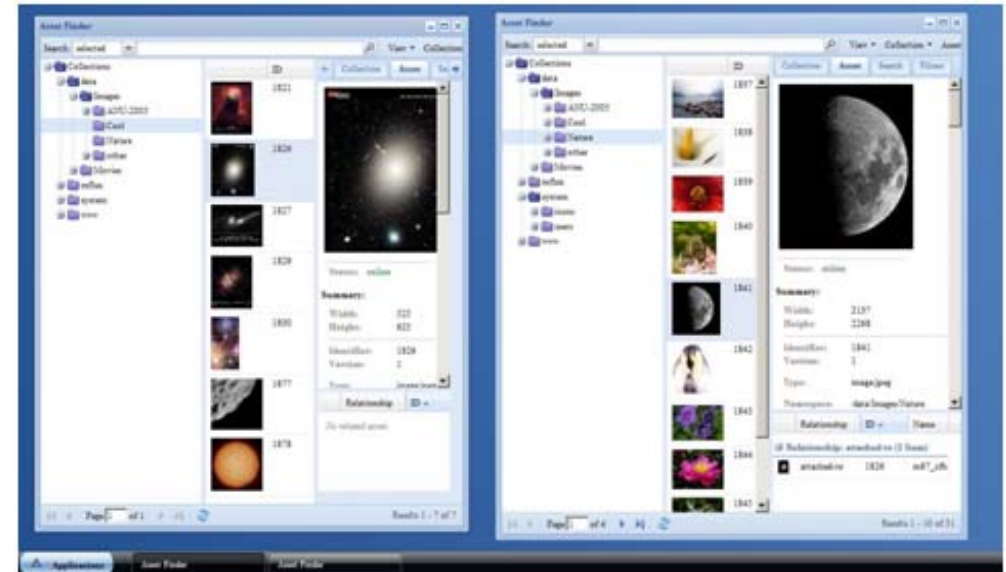
Division of Research and Commercialisation

Researchers

Research Data Management Repositories

Using Arcitecta Mediaflux
digital asset
management system

<http://www.arcitecta.com/mediافلux/>



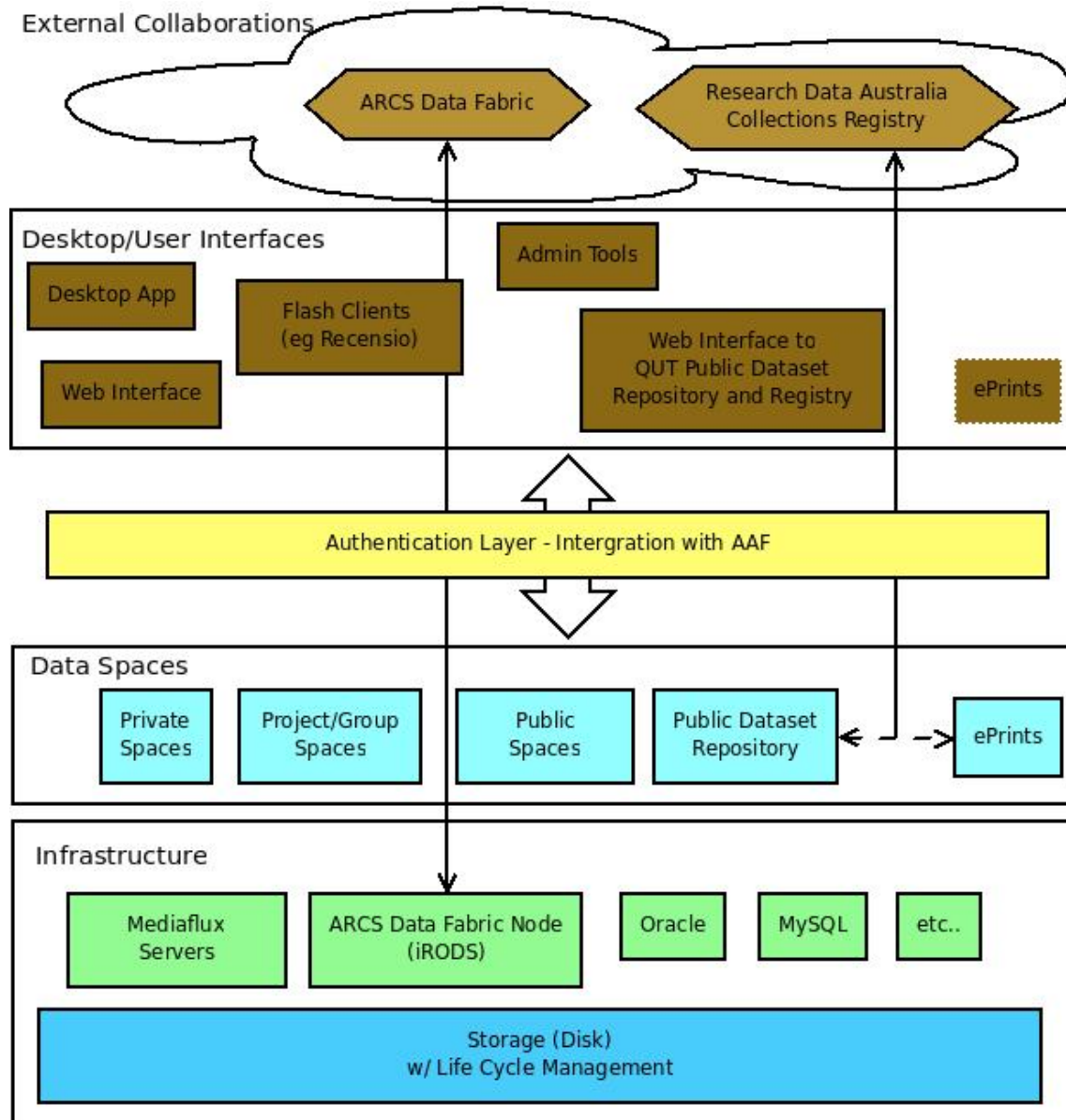
Mediaflux screen shot

Research Data Management Repositories

Using iRODS – *Integrated Rule-Oriented Data System* developed by the Data Intensive Cyber Environments research group

Storage resource broker

ARCS data fabric node

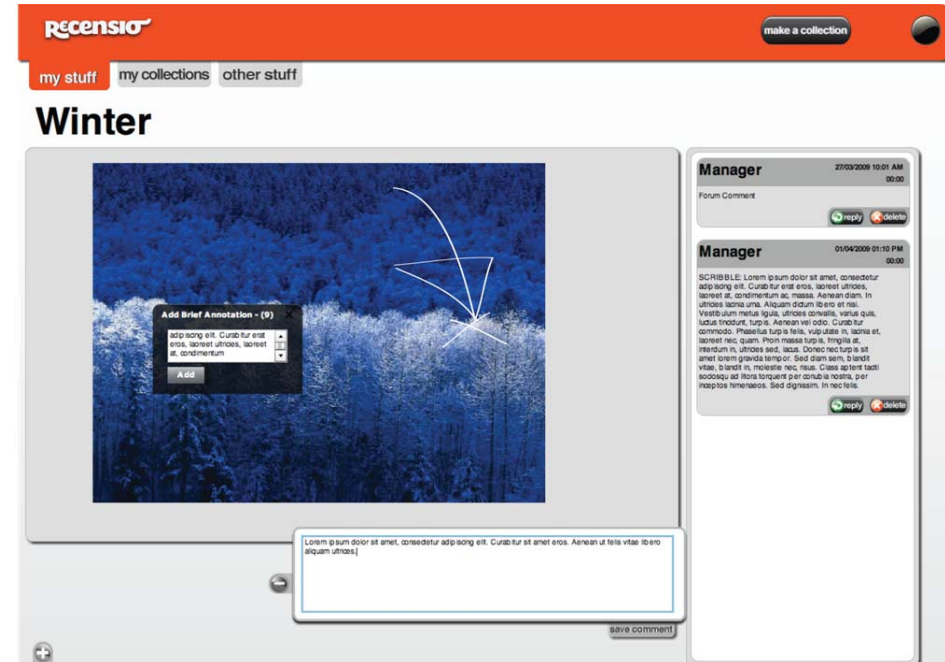


Architecture representation

Research Data Management Repository

Pilot Projects

1. Creative Industries - Resensio - PhD data using Mediaflux
2. Learning spaces
3. Medical engineering data using ActionScript and MySQL
4. Ground water data using Drupal
5. Upcoming electron microscopy images



QUT's Resensio provides for annotations

Enterprise-Wide Storage Solution

We call it eStore

Tiered storage area network (SAN) and network
attached storage

Seamless storage space available

Awareness and Skills #1

2007 Data Management Survey with Monash and UQ (Henty, Weaver, Bradbury, Porter, APSR, 2008) available on QUT ePrints
<http://eprints.qut.edu.au/14549/>

Findings included...

- Nearly all researchers have digital data created in the course of their work
- The few researchers who claim not to have digital data either genuinely do not have any or, more likely, do not recognise what they have as digital data, probably because it is text
- There is an extraordinary range of non-digital data being collected, with implications for the need to digitise at some future time
- The size of a digital collection does not seem to be of importance to most researchers, except for those who have significant data storage requirements
- Researchers use a wide range of software, largely proprietary. The range is made up of a small number of core applications and a much larger number of specialist applications. This has implications for later data curation
- Researchers currently do not recognise the implications of their software choices and later access to their data
- Researchers use many different means of storing and backing up their data, often using storage media which are unreliable and short-lived
- Most researchers do not have research data management plans, although they do recognise the need for them
- Training is sought in areas related to data management planning, either prior to a project or after, digitisation and data rescue (for older materials)
- Most researchers are responsible for their own data management which may vary from haphazard to highly organised
- Most researchers are willing to share their data and in many cases already do so. They would like an easier means of doing so
- Most researchers see their data as having value beyond the immediate project
- Only a small proportion of researchers use the grid or high performance computing. Many more researchers consider themselves as conducting eResearch, suggesting that the two are not necessarily connected
- Most researchers are unclear about the intellectual property regime which governs their research
- There is some negativity among researchers about data management, which is seen as another bureaucratic requirement being imposed on their time. This

Extract from the
Henty et al report, 2008

Awareness and Skills #2

2009 eResearch Practices Survey

Broader, but some focus on data management

Example questions...

4. How would you describe your skill level in the following *data management* practices?

Please answer for each row.

Select your skill level by selecting the most appropriate CIRCLE button, and indicate if you would like training or more online resources by selecting the SQUARE on the right.

	NOT APPLICABLE to my research	UNAWARE	AWARE: No experience	BASIC SKILLS	COMPETENT	EXPERT	Would like TRAINING or ONLINE RESOURCES
Preparing a data management plan	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Managing your digital data storage needs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Managing privacy and confidentiality of your data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Managing the legal issues of your data (copyright, contracts, licences)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Backing up your digital data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Retaining your digital research data according to legislative and funding body requirements	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Disposing of your digital data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Using pre-existing data for your own research	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Assigning descriptors or metadata to your research datasets	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Preparing a data exit plan (e.g. for retiring or departing academics on your team)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>

5. How would you describe your skill level in *managing* the following types of data for your research?

Please answer for each row.

Select your skill level by selecting the most appropriate CIRCLE button, and indicate if you would like training or more online resources by selecting the SQUARE on the right.

	NOT APPLICABLE to my research	UNAWARE	AWARE: No experience	BASIC SKILLS	COMPETENT	EXPERT	Would like TRAINING or ONLINE RESOURCES
Fieldwork data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Experimental data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Survey data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Interview data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Clinical data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Laboratory notes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Data collected from sensors or instruments	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Data automatically generated from or by computer programs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Digital audio files	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>
Digital video files	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>

Sample survey questions and response options using KeySurvey

Awareness and Skills #3

Seminar series for library and IT staff on eResearch skills, including data management

Overview of research at QUT

What is data management?

Why is it important?

Using Mediaflux

Conducting data interviews

Using RIF-CS metadata schema

Awareness and Skills #4

eResearch awareness and skills workshops for researchers

Data management seminars planned

One-on-one consultations with HPC staff

Liaison Librarian referrals

Finally ...

Lots already started...

Lots more to do...

Lots more to learn...

