Repositories Paper to the Information Futures Commission

AUTHORS: EVE YOUNG (CONVENOR), SHIRLEY SULLIVAN, ROZANA KEKOVSKA, JOE ARTHUR, PETER BODE

A PAPER FROM INFORMATION SERVICES

TABLE OF CONTENTS

What is a Repository? ................................................................. 2
Ten Year Vision for Repositories ................................................ 2
Institutional Repository: How do we get there? ....................... 4
Key Dependencies of this paper .............................................. 6
Bibliography ......................................................................... 7
Appendices .............................................................................. 8
Recommendation

- That there be a coordinated approach to repositories management within the University.
- That all repositories within the University are to be governed by appropriate policies, guidelines and standards.

The existing components of what should constitute a University-wide research information infrastructure should be linked to assist researchers. Such an infrastructure would comprise appropriately configured accessible data, stored in sustainable repositories that provide enhanced data services to researchers, including international standards for curation and preservation.

What is a Repository?

“Digital storage place for the data and research outputs of an educational institution”

“I believe that a mature and fully realized institutional repository will contain the intellectual works of faculty and students – both research and teaching materials – and also documentation of the activities of the institution itself in the form of records of events and performance and of the ongoing intellectual life of the institution. It will also house experimental and observational data captured by members of the institution that support their scholarly activities.” (Lynch 2003, p. 328)

Ten Year Vision for Repositories

1. Clearly delineated policies, frameworks, roles and responsibilities, ensuring appropriate control over storage, access and preservation of institutional data.
2. The internal and external regulatory environment enjoys sufficient breadth and depth for intellectual property and digital rights management to be understood and observed.
3. There is awareness of appropriate licensing, including Creative Commons, and identity management as required, e.g. through participation in the Australian Access Federation.
4. Guaranteed sustainability, reliability and quality of service.
5. An industrial strength repository architecture and infrastructure, with securely maintained networks and servers, to support all research, and other, scholarly data within the University.
6. The University will have compartmentalised repositories which can be holistically searched via a common interface through use of mandated minimum metadata standards.

7. Appropriate hardware and software is in place to support the needs of all repositories, ensuring seamless integration between repositories.

8. There is institution-wide adherence to international standards and protocols for metadata, etc, including those required by the Trustworthy Repositories: Audit and Certification (TRAC): Criteria and Checklist (Feb. 2007) (<http://www.crl.edu/content.asp?l1=13&l2=58&l3=162&l4=91>)

9. Search services and persistent identifiers (permanent URL) enable efficient discovery within collections held in the repository.

10. Defined collection policies clearly outline the scope of repository content.

11. To enable University of Melbourne’s participation in national and international collaborative partnerships, a framework of national, and international, agreements is in place to ensure research data management, integrity and security in an environment where research is conducted and funded on a cross-institutional, jurisdictional, national and international basis.

12. Different data classes for which the University requires repositories:

   - **Institutional repository** – most are open access, containing scholarly, and peer reviewed works. Mandates are now being introduced to encourage researchers to put items of their research into the repository. (See Appendices 1, 2, and 3)

   - **Theses repository** – digital deposit of theses should be made mandatory because theses are a rich source of original scholarship which, in their print form, have been neglected through limited accessibility. It has been shown that requests for deposit do not work, whereas mandates do. The University may need to allow for exceptions, such as work produced in conjunction with a commercial firm, or literature leading to application for a patent. (See Appendix 4 for mandates in Australian Universities and on Harvard College Theses Repository (2008))

   - **Students’ repository** – exemplar output from undergraduates should be accessible to their peers. There has been a need for this expressed by academic staff within the University.

   - **Staff repository** – to enable professional staff to share their work with national and international colleagues.

   - **Learning objects repository** – to share the intellectual output created by the University of Melbourne academic staff to support their teaching.
• **Research data repository** – This is a matter of increasing importance with the growth of e-research. There may need to be multiple, cross-institutional repositories, controlled by a common metadata schema and retrieval interface.

• **Cooperative Research Centre (CRC) Outputs Repository** – There is a need to capture and manage their research output, original research data and administrative records. While there may be cases where data may be shared between multiple owners, the University may be accountable for data control and integrity under various legal instruments.

*NB: It was thought that because repositories cover such a wide range of issues, it was preferable to deal with the broad issues first, then concentrate on the Institutional Repository and issues surrounding the Cooperative Research Centres (CRCs).*

**Institutional Repository: How do we get there?**

- Encourage the academic community to self-deposit. IS needs to provide access to data management resources and the repository as an information resource for researchers. IS recommends that the University develop policy and guidelines on depositing research output (see Appendix 1)

- Economies of scale suggest that, where the academics are depositing work into internationally recognised discipline based repositories, IS harvests the metadata into the institutional repository so the academics only need to deposit in one place.

- Research reporting and compliance plans need to be in place, including integration between the Research Management Program (currently Themis) and the repository.

- Upskill IS staff to create and support sustainable repositories.

- The institutional repository can become a centrepiece for university publishing, one that allows the University to play a new role in the emerging landscape of open access publishing. In this way, faculty members get quick and easy tools to run peer-reviewed journals for their research communities, and get a foothold in journal publishing, without becoming full-fledged professional publishers with complex marketing and sales operations. (See also Open Access Paper, and EPublishing paper from ESRC)
Research/Cooperative Research Centre (CRC) Repository

Problem: Cooperative Research Centres have been proliferating since the 1980s and have reached the point where they have become the most common mechanism for distribution of Commonwealth research funding. While facilitating research flexibility, the CRC model often leaves data ownership and accountabilities unclear.

Please refer to Appendix 5 for an outline of the recordkeeping challenges associated with the CRC administrative and research records.

The way forward:

Any solutions to the repository issues raised by CRC funding and operations will need to be enacted at a multi-institutional or Governmental scale, as a consequence of shared stakeholder-ship in the research projects. As such, the University’s point of approach is through cooperation with other academic institutions, preferably under the umbrella of the Federal Government.

The University could investigate negotiating standard CRC records and data practices with other universities, whereby the Centre Agent for a CRC undertakes to provide appropriate records storage and management services on behalf of the other partners. These practices could be formalised within the Memoranda of Understanding under which the CRCs are established.

Upon the closure of a CRC, University Records Services could have delegated responsibility to provide storage (and later disposal services) for the Commonwealth-defined administrative record for the remainder of its seven year retention period, with costs to be assumed by the DVC (Research).

A records and data custody agreement to be developed for inclusion in all future CRC agreements. Such agreements would cover:

- Ownership and usage rights to the CRC research output.
- Access to an ownership of the CRC administrative record.
- Custody and disposal obligations regarding both the research data, and research and administrative records.
- Access to the Institutional Repository for research outputs of CRCs where the University acts as Centre Agent and where such deposit is agreed with other partners.
Key Dependencies of this paper

There are a number of papers submitted from IS to the Information Futures Commission that relate to this subject, e.g., Open Access, Digitization, ePublishing, Copyright and a Digital Asset Management System.
BIBLIOGRAPHY


APPENDICES

Appendix 1

_Open Access Policies of Repositories worldwide, including ARC and NH&MRC recommendations._

ROARMAP (Registry of Open Access Repository Material Archiving Policies)
_http://www.eprints.org/openaccess/policysignup/_
Mandatory Research Funding in Australia, however it has been discovered that Mandates work, requests do not.

Two kinds of mandate that work:
Deposit mandates (e.g. QUT, 2004, UTas, Nov2006)
Permission mandates (e.g. Harvard 2008)

Harvard’s is the first adopted policy to focus on permissions rather than deposits

Appendix 2

_Australian University Repositories._

There are currently 38 Australian University institutional repositories active, online, and carried current research documents.

Australian Open Access Repositories and Contacts
Appendix 3

No of items in the University of Melbourne EPrints Repository.

[Graph showing the growth of items in the repository]

Appendix 4

Mandatory theses: Australian Universities.

This table contains our current knowledge of Australian progress towards mandatory institutional policies for (a) research articles, and (b) e-theses. [Link to more information]

Appendix 5

Research/CRC records: (submitted by Records Management)

- CRC agreements with the Commonwealth (funding agreements) currently contain a provision for particular administrative records to be retained for seven years. Under the standard CRC funding agreement with the Commonwealth this applies to just the administration records.

- Many CRCs are conducting research in the field of medicine, where experimental drugs might be administered, and where the long term implications of having done so might not be apparent within seven years. There are almost certainly privacy implications where there are human research subjects.
CRC records are not currently recognised as university records, but it is suggested that consultation with Legal Services be undertaken, as the University probably incurs significant exposure to risks from privacy or confidentiality violations, especially where it is the Centre Agent (headquarters). There is currently significant potential for expensive legal actions based on disputed ownership of research data and/or records. How are these privacy considerations met once the CRC closes down? (i.e., when the money has all been spent)

CRC record ownership is also unclear. Once the centre closes, are the records owned by the researchers, the associated academic departments or schools, the private interests who may also be CRC partners?

NB: CRC based funding is proliferating at a rapid rate (This issue will only get bigger)
## Document Control

<table>
<thead>
<tr>
<th>Title</th>
<th>Repositories paper for Information Futures Commission</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short Title</td>
<td>Repositories paper for IFC</td>
</tr>
<tr>
<td>Project Ref</td>
<td></td>
</tr>
<tr>
<td>Author</td>
<td>Eve Young (Convenor), Shirley Sullivan, Rozana Kekovska, Joe Arthur, Peter Bode</td>
</tr>
<tr>
<td>Status</td>
<td>Working</td>
</tr>
<tr>
<td>Version</td>
<td>V2</td>
</tr>
<tr>
<td>Version Date</td>
<td>07/05/08</td>
</tr>
<tr>
<td>Origination Date</td>
<td></td>
</tr>
<tr>
<td>Change History</td>
<td>* BW placed in template for JE</td>
</tr>
</tbody>
</table>