

Sink or swim

the necessary evolution of the library based on the Australian Digital Theses Program experience

Andrew Wells

University of New South Wales Library

a.wells@unsw.edu.au

Sydney NSW 2052 Australia

www.library.unsw.edu.au

Abstract

In 2002, the Australian Digital Theses Program was formally established by the Council of Australian University Librarians (CAUL). The ADT Program is enabling libraries to demonstrate new roles in scholarly communications in the fast evolving information environment. This has been a major benefit of the ADT Program and a focus for CAUL. Libraries involved in the ADT Program have been able to demonstrate new facilitating and leadership roles with the scholarly community, as well as developing new forms of collaborative activities. This is stimulating other ways of making universities' research available, including the development of institutional research repositories. The ADT Program has provided a timely focus on the hybrid nature of library services. The research community is advocating expansion of the ADT metadata repository to include information about theses in non-digital formats and other types of theses (eg, undergraduate research). CAUL has addressed the value of metadata on ADT by developing agreements for its licensing and commercial re-use. Finally, the ADT Program provides a path for a major contribution to global information infrastructure. This will be achieved through redevelopment of the ADT software to support OAI and metadata harvesting.

This paper briefly covers the development and governance of the program to date, and addresses key issues for academic libraries in rethinking their role based on the ADT Program experience.

Preface

The Australian Digital Theses Program (ADT) began as a Commonwealth funded research project in 1997. Seven Australian university libraries collaborated on the project, using software from the Virginia Polytechnic Institute. The project has now developed into a program of the Council of Australian University Librarians (CAUL) and is becoming recognised as a significant innovation in scholarly communications in Australia. The ADT service comprises a central metadata repository containing descriptions of theses, which are linked to PDF versions of theses hosted on servers at participating institutions. The metadata repository is managed for CAUL by the University of New South Wales Library.

ADT - Progress since ETD 2002

In 2002, a business plan proposing financial, administrative and governance arrangements for the ADT Program was presented to the Council of Australian University Librarians (CAUL). The key features of the plan included:

- All members of CAUL automatically becoming members of the ADT. There are 38 university libraries in CAUL.
- Each member of CAUL paying a levy to cover central costs and maintenance of the metadata repository
- A two-tiered governance arrangement. A policy committee oversees strategic development and contains representatives from CAUL, the Council of Australian Postgraduate Associations, Deans and Directors of Graduate Schools, the National Library of Australia and ProQuest. A technical committee addresses software and standards matters. The technical committee reports to the policy committee. The National Library provides expert advice on preservation and long term access.
- Performance measures addressing growth, participation and relationships with stakeholders
- University of New South Wales to continue managing the ADT program on behalf of CAUL.

CAUL endorsed the business plan in April 2002 and the ADT Program came into formal existence. Key performance indicators have been met or exceeded. Goals for June 2003 included 50% participation by CAUL libraries and over 1000 theses available online. These targets were met in April 2003. Usage of the ADT is generally between 2000 and 4000 hits a day, and over 50% of use comes from outside Australia: the international audience for this service is particularly pleasing.

Developments in the last twelve months include:

- Initial discussions with the Council of New Zealand University Librarians (CONZUL) - maybe this will lead to an Australasian Digital Theses Program. However, discussions are preliminary. CONZUL is demonstrating strong interest in developing a national program for digital theses.
- Acceptance of a metadata policy by CAUL. Each member has extended to CAUL a non-exclusive, royalty-free license to harvest, copy, host and store all or part of the member's metadata in the ADT metadata repository, to allow free access to metadata via the

World Wide Web, and to allow the ADT Program to make the metadata available for international discovery purposes. The policy also addresses commercial re-use of ADT metadata.

- Development of service level agreements between CAUL and the University of New South Wales Library for maintenance of the metadata repository, user support and problem resolution.

These types of developments will maintain the Program as described in the business plan. However, external events are prompting new thoughts about the expansion of the program and its integration with developments in scholarly communication and management of digital assets in universities. These suggest the possibilities of new roles for university libraries. The success of ADT will play a key enabling role in university libraries gaining recognition and legitimacy in the management of institutional knowledge - whether they be learning objects, e-print repositories, commercial electronic publishing or digital theses. Each is a component of a wider approach to digital content management in universities.

The involvement of the Commonwealth Government in developing higher education information infrastructure may enable the further development of ADT and its incorporation in wider programs addressing innovation in digital content management.

Higher Education Information Infrastructure in Australia

Through the collaborative efforts of its libraries, Australia has a well developed information infrastructure, including a comprehensive National Bibliographic Database managed by the National Library and a robust, well-performing interlibrary lending and document delivery scheme. Higher education libraries collaborate in many areas, notably in consortial licensing of electronic information for the sector.

In 2002, the Commonwealth Government took a particular interest in development of the national information infrastructure to support research. The Department of Science, Education and Training commissioned a report on research information infrastructure under its Systemic Infrastructure Initiative.¹ In clause 1.1, it states that the "vision of the Australian Research Information Infrastructure Framework is to facilitate access to information infrastructure resources which will optimise the efforts of researchers in the higher education sector to create, manage, discover, access and disseminate knowledge". The report observes that changes in technology can enable substantial innovation in the processes of creating, disseminating and accessing knowledge, but recommended collaborative approaches to the optimal deployment of new technologies.

The report identified three key areas for development of the research information infrastructure:

- Discovery and management of research information - covering improved access to electronic resources, development of subject gateways and research training
- Access to research information resources - directed at acquiring content. An extension of Web of Science backfiles to 28 Australian universities has already been provided as part of this initiative
- Creation and dissemination of Australian research information - including e-print repositories, digital theses, electronic presses, digitisation of Australian publications and publications, and access management (ie, the authentication and authorisation challenges for access to online resources)

The ADT falls into the last category. Work is underway to bring these projects into a more comprehensive and integrated program - there are clear risks in undertaking these projects in isolation from each other. The report makes clear that these innovations are not designed to disrupt or replace current patterns of scholarly communication, although it observes the economic problems currently experienced by university libraries in acquiring content from commercial providers. In clause 9.6.5 of the report, it makes the point about these innovations complementing existing schemes for scholarly communication and embraces inclusion of Australian research information in the developing global information infrastructure: "the development of an overall strategy is required to improve access to Australian research by supplementing existing publishing mechanisms and taking advantage of emerging technologies ... Underlying these projects is a need to adhere to international standards in areas such as metadata, and to use the Open Archive Initiative protocols."

The most challenging aspect of this is to develop the case to knit together e-print repositories, digital theses, electronic presses and other digital forms of knowledge produced by universities. They can all be regarded as institutional assets: leveraging these assets and collaborating with other universities in exploiting these assets needs further articulation and experimentation. Projects such as D-Space and the University of California's e-scholarship repository are helping in showing us potential paths.

Placing the ADT in relationships with other projects does, however, indicate key features of its further development.

Next steps

The two key service developments are extension of the metadata repository and software redevelopment.

Expanding the ADT metadata repository

The existence of the ADT metadata repository has revived interest in a 'one-stop shop' for information about Australian theses. Bibliographic records for several hundred thousand theses are included in the National Bibliographic Database, which contains over 13 million records. Some data can be found only in local catalogues. The ADT metadata repository seems the obvious place to consolidate these sources of data. While it may seem odd to put metadata for theses which are not available online, researchers and postgraduates have made it clear that they have a need for a comprehensive metadata repository for all Australian theses. Several university libraries are now retrospectively converting theses to digital form instead of copying the thesis when they receive document supply requests. It will be possible then to merely 'hook' the digitised version to the metadata in the ADT metadata repository. At the University of New South Wales Library, we believe on demand retrospective conversion of theses is cost effective, improves access and increases awareness of the ADT Program. Seeking permission from the author can be the most time-consuming aspect of it.

The expansion of the metadata repository will require identification of records in the National Bibliographic Database, their extraction and conversion from MARC21 to the qualified Dublin Core format. Mechanisms to endure updating and synchronisation of the National Bibliographic Database and the ADT metadata repository will need development.

The inclusion of abstracts in the metadata has also been recommended to enhance searching and retrieval.

The scope of theses included in the program are being reviewed. Some have recommended the inclusion of masters coursework and undergraduate theses. This matter will come to the Policy Committee for discussion.

Software redevelopment

The ADT Technical Committee has recommended redevelopment of the software used at local sites and the software deployed for the metadata repository.

Local sites currently use the software based on code developed by the Virginia Polytechnic in 1996. There has been some modification of the software. It enables members to allow self deposit by students, generate metadata, restrict access when required and to store thesis files. The software has served ADT well but it is time to redevelop. Local sites are reporting difficulties in customising the software for local use. In relation to developments in the national information infrastructure, it is critical that the software have the capability to expose metadata to OAI gatherers - compliance with the Open Archives Initiative metadata harvesting protocols is essential. Other university libraries want the ADT software to integrate with their e-print repositories. The Technical Committee

has recommended use of the ePrint software (<http://software.eprints.org>)

The metadata repository currently uses the HotMeta product from the Distributed Systems Technology Centre (DSTC) (<http://www.dstc.edu.au/Products/metaSuite/HotMeta.html>). Again, this has served ADT well but other functionality is needed:

- National Bibliographic Database relationship. MARC21/Dublin Core conversion, de-duplication and batch loading functions are needed
- OAI support
- Searching features which can enable users to separate digital theses from theses only available in print, or with an abstract only.

The Technical Committee has recommended development of new software using open source tools and open source programs that already provide part of the required functionality.

Conclusion

The ADT Program has been successful in terms of its original aims. Research which was difficult to access is now available online: it has found a new audience. Australian universities are beginning to review their thesis submission rules. An increasing number of graduates are becoming confident with electronic publication and appreciating the benefits and implications of this new form of creation and dissemination of scholarly communication. The ADT Program will develop to consolidate these achievements. It has been an important demonstration of innovation led by university libraries working collaboratively.

The next challenge is to position the ADT Program with other innovations in scholarly communication. The digital information environment has produced this challenge. Digital information and digital services are transforming academic libraries around the world. Australian university librarians see that the ADT Program is a catalyst in the transformation of the roles of libraries and librarians in this fast changing environment.

Bibliography

- 1 DEPARTMENT OF EDUCATION, SCIENCE AND TRAINING, SYSTEMIC INFRASTRUCTURE INITIATIVE, INFORMATION INFRASTRUCTURE ADVISORY COMMITTEE : Research Information Infrastructure Framework for Australian Higher Education. Canberra, November 1992 www.dest.gov.au/highered/otherpub/heiiac/report.pdf.