

# Theses record exchange: developments in the Australian National Union Catalogue

Roxanne Missingham, Assistant Director General, Resources Sharing and Margaret Kennedy, Director (a/g), Kinetica Database Services, National Library of Australia

**Abstract.** This paper reports on the Australian union catalogue – its role particularly in providing access to all Australian theses; recent research into user expectations; record exchange, in particular implementation of OAI/PMH; development of the new Libraries Australia search interface; development of standards and future direction for increasing national access to digital and print theses.

**Keywords.** National union catalogues; resource discovery; OAI

## Introduction

Australia's libraries have together produced and utilised national resource discovery services of resources in library collections, in particular through the Australian National Bibliographic Database (ANDB)<sup>1</sup>, for many decades. There have been many factors influencing the high degree of cooperation and participation in national resource discovery services, including the large distances between Australian libraries, the complementary nature of many collections and budgetary situations.

Theses are a very interesting example of collaboration. The tradition of recording collections in printed union lists starts with the *Union list of higher degree theses in Australian University libraries* first published by the University of Tasmania in 1959. Over the following years supplements were published regularly until 1988.

The development of an Australian national online union catalogue, from 1981, enabled theses to be recorded in a system which was widely used by university, state, national, special and public libraries. This infrastructure enabled two key issues to be addressed – finding theses and requesting copies through inter-library loan. Access to theses moved from the printed lists to the national union catalogue in the 1980s. In 2004 a major step forward was taken with a new service, Libraries Australia, offering access through innovative use of technology to enable libraries as well as end users to find and get copies of theses.

Through the late 1990s the Australian Digital Theses<sup>2</sup> program encouraged the production, storage and management for access of Australian theses in digital form. This development has significantly improved access to and knowledge of the output of research. The integration of access to print and electronic theses has been a key feature of the national infrastructure. There are new expectations for information seekers of access to full text scholarly literature through the Internet. The online environment has created many opportunities as well as challenges for information seekers and national resource discovery services.

Over the past five years there have been many innovative technological developments which provide new means of sharing or exchanging records with central metadata repositories. The implementation of OAI/PMH harvester technology has added to the range of data exchange mechanisms with the Australian union catalogue and enables significant improvements in automated record distribution.

This paper outlines the development of access to theses through the ANDB, the nature of the service model underpinning access to this database through Libraries Australia, research

into user needs and data and architectural issues. Future directions of sharing of data, including with Google scholar, and emerging standard are then considered.

### **Theses in the Australian National Bibliographic Database/union catalogue**

The Australian library network is supported by the Australian National Bibliographic Database (ANBD), which includes the National Union Catalogue. The service is an essential part of the Library's strategic aim to:

“provide rapid and easy access to the wealth of information that reside in libraries and other cultural institutions – and to break down the barriers that work against this”<sup>3</sup>

The service, released online in 1981, after piloting in 1980-81, was originally designed primarily to support shared cataloguing, by enabling Australian libraries to increase copy cataloguing, resulting in greater efficiencies in technical services. The service was also used for acquisition purposes, for reference enquiries to identify relevant resources and by interlibrary loan staff to identify material held by other libraries. Increasingly, through the development of a web based search interface, the service has been offered to end users, particularly in academic libraries. An interlibrary loan module was implemented to enable management of copy and document requests and manage payments. This enabled efficiencies to occur in interlending, a finding supported by the Interlibrary Loan and Document Delivery Benchmarking Study<sup>4</sup>. The service operates on a cost recovery basis with libraries subscribing for use of the service. The business model and operation of the service is very similar to that of OCLC's WorldCat. Libraries around the world can join the Kinetica service and pay for searches, in addition individual researchers can purchase blocks of searches.

The service supports:

- cooperative cataloguing, enabling Australian libraries to reduce the costs of their cataloguing by using records created by others;
- support for interlending, allowing Libraries to share resources by borrowing (or receiving copies) of library materials;
- immediate access to online resources (with the exception of those restricted by licensing);
- support for cooperative collection development, enabling libraries to reduce the amount of duplication in their collections; and
- access to the collections of Australian libraries for individual researchers, enabling identification of relevant material in Australian libraries and online.

Libraries Australia<sup>5</sup> enables Australian libraries to access and use international union catalogues for resource sharing and cataloguing. Libraries Australia provides access to databases such as the RLG Union Catalog, OCLC's WorldCat and the national union catalogues of New Zealand and Singapore. It also enables libraries to order materials from suppliers such as CISTI and Infotrieve, in addition to ordering from other Australian libraries.

The service is utilised by approximately 1100 libraries. With more than 40 million holdings, approximately 14 million bibliographic records including over 574,000 electronic resources, it forms an essential tool for Australian libraries in all sectors – public, special, academic, technical and further education, health, corporate, law, state and national. Over 6.5 million searches were undertaken on the service in 2003/04.

Coverage of Australian theses is strong – regular surveys have been undertaken to investigate reporting of theses and any issues. Surveys have been carried out in 1996, 1998 and 2002<sup>6</sup>. All university libraries were again contacted in 2004 about thesis reporting issues. The 2002 survey found that 30 of the 39 libraries recorded all their higher degree

print theses on the NBD. Many libraries also recorded minor or undergraduate theses on the NBD. Those 8 libraries who recorded some of their print theses most commonly had partial coverage because of retrospective backlogs that could not be cleared due to insufficient resources or current backlogs caused by delays in implementing new ILMS software. There are currently 158,000 theses on the ANBD, with a significant rise in digital theses. A very strong relationship exists both with university libraries and ADT.

### Australian National Bibliographic database coverage of Theses

	Total print theses	New print theses added	Total digital theses	New digital theses added	Digital theses % of total new theses
2000	120,147	7737	1442	319	3.96%
2001	129,837	9690	1746	304	3.04%
2002	138,020	8183	2775	1029	11.17%
2003	145,436	7416	3500	725	8.91%
2004	151,854	6418	4127	627	8.90%
2005 (to end June)	154,424	2744	4247	122	4.26%

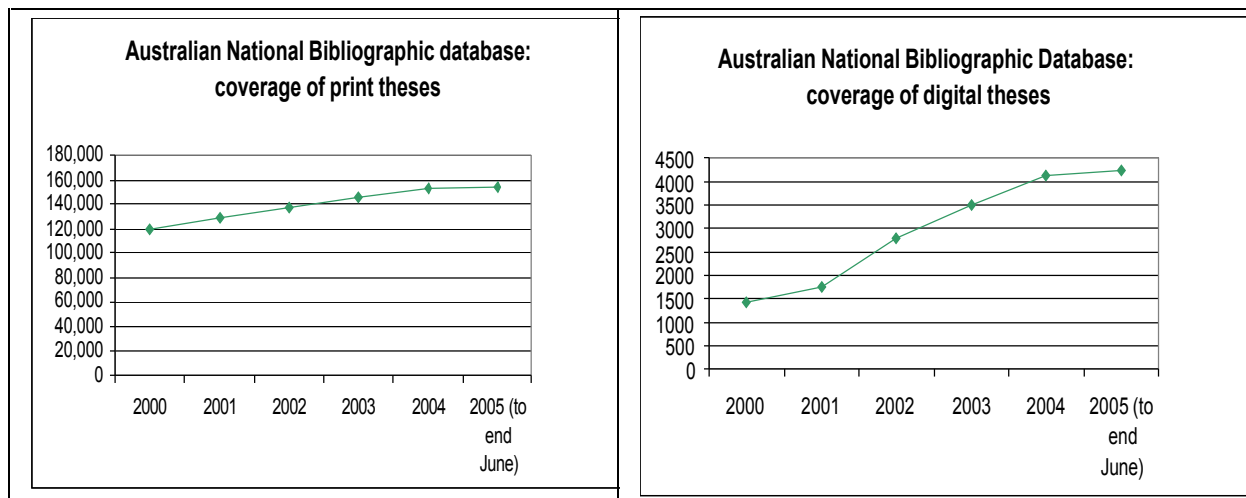


Figure 1: Increasing coverage of theses: print vs. digital

Overall the growth in digital theses has been very significant in recent years. One of the critical steps in considering the national catalogue as a mean of accessing Australian theses is to understand user needs. In conjunction with redeveloping the service the Library undertook a pilot project to better understand user needs in accessing information in Australia.

### But what do users want?

Over the last two years the Library has undertaken several projects to identify the key needs of end users - researchers, students and the general public. The first such project was the pilot Information Australia service, which offered easy Google-like access to the ANBD, PictureAustralia, full text current Australian journals and a document supply service provided gratis for the six libraries in the pilot. The pilot has enabled Australians to freely access the national union catalogue without the impost of charging for access or document delivery. Information Australia provided a more user friendly interface than the existing Kinetica search interface, and provided federated searching:

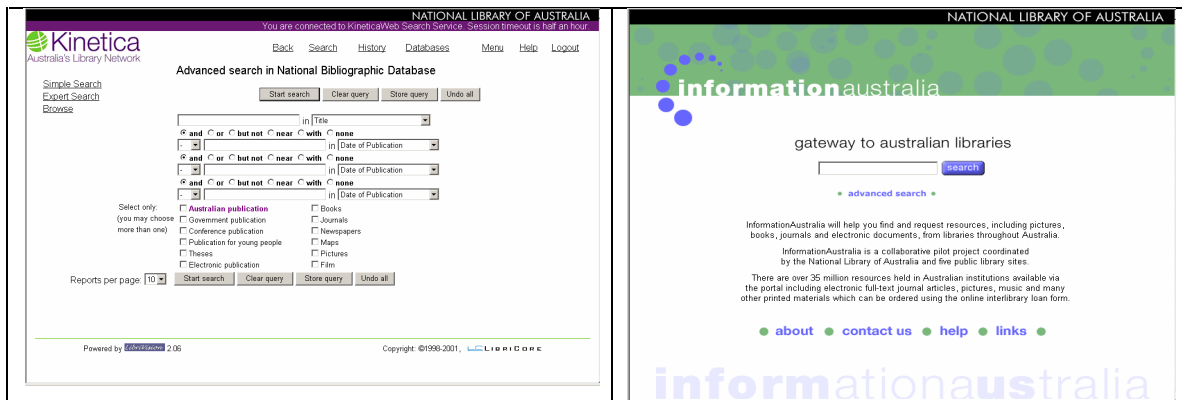


Figure 2: Comparison of Kinetica and Information Australia basic search screens

The pilot included extensive evaluation based on:

- an online survey of users;
- feedback from participating public libraries; and
- an external report based on focus groups and interviews with searchers.

Last year the Library conducted research<sup>7</sup> into the information seeking behaviour of Australians using focus groups and interviews in public, state and university libraries. Participants were able to discuss their general information behaviour and then focused on their expectations of access to the Australian library network. Public library users were selected who had used Information Australia. University based participants and independent scholars had used the Kinetica search service to access the ANDB, some through their local catalogue OPAC via Z39.50, others via the Kinetica Search interface. While the groups were relatively small, their composition was based on a carefully selection to establish a representative cross section of user communities.

The sample consisted of:

- Three focus groups with end users of the Kinetica Service at the National Library, the Australian National University and the State Library of Victoria;
- Three focus groups with clients of the InformationAustralia pilot service at Brisbane City Council Library (QLD), Yarra Plenty Regional Library (Vic) and Southern Tablelands Regional Library (NSW);
- Seven structured telephone interviews with users of the InformationAustralia pilot in Thuringowa (QLD) and Northern Rivers (NSW).

The research identified some very interesting trends in information seeking behaviour:

- most participants from the public and university sectors used the website and their library catalogues as first points of departure to find information;
- those in the university sector most often used their local library catalogues as the first point for searching, a smaller number used national catalogues and databases, such as Scifinder; and
- independent scholars also used websites and catalogues, however this group appeared to use a wider range of resources to search, perhaps due to the depth of their research.

Expectations of online access to information on the collections of libraries are now very high:

- Public library users want to be able to find and request material online from where ever they are. They also increasingly are using the catalogue to identify material which they will purchase if required. One participant commented: *“go straight to X Library, put an item on hold or place an order – but waiting list can sometimes be very long. Buy book if important.”*;

- University participants expected access to Australian collections to be available in a simple and intuitive manner and found that the current process failed to enable easy access to the ANBD; and
- Independent scholars expressed considerable interest in federated searching and preferred to use a federated search service than single databases.

Interestingly the barriers perceived in terms of obtaining access to collection materials once they were found through the ANBD were seen as very significant. Charging for interlibrary loans and opaque interlibrary loan processes were cited by users in all three groups as significant barriers to accessing collection materials.

This is a very powerful indication that the traditional national union catalogue model of providing a resource discovery service with holdings information but not easy ordering paths to suppliers including libraries and commercial suppliers is flawed. The two key themes emerging were that users of a national resource discovery service believe that access means “find” *and* “get”, not just the ability to find where a publication is held, and that for many “get” includes borrowing from a library and the ability to order or purchase online if a user wants to keep the publication.

When asked about the interfaces for searching the ANBD and access to Kinetica or Information Australia, comments clustered around four main issues:

- Content issues
- Ease of use
- Visibility and knowledge of the service
- “Getting” issues.

Individuals highly valued the service and indeed a number found the access to theses, integrated with access to published research, enabled them to identify theses that they did not know existed.

### **Theses and the Australian National Bibliographic Database: record contribution and exchange**

Australian university libraries are able to contribute records of theses to the ANBD:

- online through a cataloguing client or via a web form;
- off line through batch files which can be ftpd or harvested (using OAI/PMH).

Many theses records are catalogued online through the cataloguing client. The software used is shortly to be replaced as a part of the Kinetica Redevelopment Project. The OCLC PICA CBS software will be implemented later this year, offering improved functionality and performance.

Increasingly libraries are using automated processes to contribute both MARC records and DC metadata records. A graphical representation of record supply options follows:

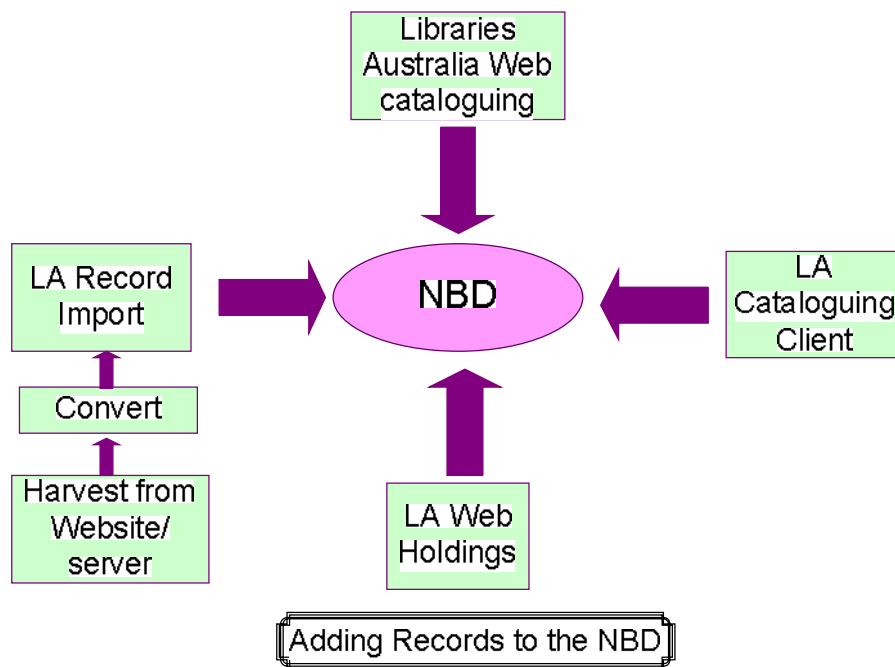


Figure 3: Record contribution methods

While cataloguing online is most usual for union catalogues, the National Library has been using OAI/PMH for harvesting records to collect descriptions for a variety of services, including PictureAustralia and the ANBD. The Library has developed a record collection system based on compiling a profile of the participant agency to manage the record ingestion process. Files of records are "harvested" using OAI-PMH or ftp from contributors. The participant agency profile records details of the contributing library, authentication information, place and method of file acquisition and collection events<sup>8</sup>. The collection of records is then generated by the profile.

When files of records are acquired from contributors (in XML), they are converted to MARC using the intermediary format of Metadata Object Description Schema (MODS), which enables the XML records to be enriched. This conversion is written for each contributor based upon test records sets. Often quite extensive record testing is required, as records for different formats of material (such as serials and monographs) may have different issues. Separate profiles were required for each record supplier because of the wide variance in content and quality of records. The OAI-PMH harvester workflow is also being used to obtain records for special materials (for example, records of music resources from the Australian Music Centre). These records are converted for use in the ANBD and also for supply to MusicAustralia.

### Libraries Australia

On 8<sup>th</sup> December 2004 the Library launched the new search interface for the ANBD – Libraries Australia, the first deliverable of the Kinetica Redevelopment Project. The redevelopment project aims to:

- introduce improvements for customers in areas such as search functionality and online performance (response time);
- give the Library an improved capacity to innovate, and thus introduce further improvements in the future;
- integrate the Kinetica Chinese Japanese and Korean (CJK) service into the core system;
- develop new search interfaces to meet the needs of individuals; and

- obtain the best value for money solution, taking into account functionality, risks and costs.

There are three search interfaces – a simple search based on a Google-style approach with Boolean functionality, an “advanced” search offering the ability to select indexes for searching from drop down boxes and a command search which supports the Z39.50 Common Command Language (CCL) query syntax. The simple and advanced search interfaces also include limits, enabling the user to easily restrict searches to resources of various formats, such as newspapers, theses, pictures, music; or types, such as Australian, online, government or conference.



Figure 4: Libraries Australia basic search screen

Some of the features of the service include:

- *Improved finding* :through easy to use search interfaces;
- *Better results*: The new search interface includes more sophisticated help, for example misspelling English language words, countries, places or names will result in a suggestion from a dictionary of over 270,000 terms (see Jspell 2004). These suggestions are hyperlinked to enable users to then immediately search with a single click. The search service also supports searching via common synonyms, for example a search for “~car” will also search for the terms automobile, auto, motorcar and machine (see WordNet 2004).
- *Improved accessibility*: Compatibility with adaptive technologies and compliance with the W3C level A guidelines is essential for the service. Testing has been undertaken to ensure that the service is fully accessible.
- *“Improved “getting”*: For users, as identified in the research, there is an expectation that they are able to “get” as well as “find” resources. Building a “get” service into Kinetica is essential to the delivery of a relevant service for end users as well as libraries. When a user finds a record they have four options available to “get” the resource:
  - online – providing direct access to the online resource, where this is available (currently approximately 500,000 resources are available online), or to online pages relating to the resource;

- through their library – including checking for availability from their local catalogue where this is possible;
- through other libraries – again with the ability for checking of availability and, where the subscribing library agrees, the ability to place an ILL request electronically; and
- through other suppliers including bookshops, document supply services such as Infotrieve, CISTI and SCIPRO or the National Library's Copies Direct service.

For bookshops that support a Web Services interface such as Amazon (see Amazon 2004), the new service is able to automatically check availability using an ISBN or author/title search against the bookshop's database. Users can then link directly to that bookseller from Libraries Australia in order to purchase the item online.

Significantly the project has enabled an increasing number of users of libraries to be able to easily access the nation's resources. The majority of Australian universities and an increasing number of state and public libraries are providing unmediated access for their users. The new subscription model has enabled this to occur without financial barriers.

### **Standards for record descriptions**

In 2005 guidelines were developed for record descriptions for theses for contribution of records to the ANBD (<http://www.nla.gov.au/kinetica/download/thesesrev.doc>). The importance of developing network standards to ensure theses could be found successfully, was identified through consultation with university libraries. The process also revealed some deficiencies in the MARC format. The guidelines emphasise the importance of including a URL in the 856 field to supporting access to digital theses. Consultation with university libraries is key to ensuring that the guidelines are implemented and regularly reviewed to ensure they remain relevant.

### **Future directions**

Libraries Australia and the ANBD provide national resource discovery and access services for Australian theses, enabling these resources to be surfaced within a national union catalogue and as a special set of resources. Continuing to increase the coverage of Australian theses through cooperation with university libraries is critical. There may be opportunities to increase the range of digital theses available through digitising theses on demand (from ordered copies). Collecting these resources and records will be important to enable future users to access the full text at the click of a mouse.

The move to electronic theses has provided some challenges in record standards and record sharing, and close collaboration with university libraries and ADT provide a framework for continuing to collaboratively build the national resource discovery infrastructure for theses.

Implementing new technology such as OIA/PMH enables records to be supplied easily and conveniently. Developing the new Libraries Australia interface has enhanced the ability of Australians, through their library network, to find and get theses. Future directions for Libraries Australia include investigation into relevance ranking and implementation of Functional Requirements for Bibliographic Records (FRBR) to further enhance and simplify access. A free public search interface is also planned for 2005. Opportunities to promote access to theses, such as through Google scholar and other resource discovery channels will also be pursued as a part of the Library's activities to enhance access to Australian research.

The critical support provided by the National Library for:

- discovery of theses and related research resources;



- ordering of copies of theses;
- sharing of information and resources

provides a basis for further collaboration and developments to support the national research infrastructure.

---

<sup>1</sup> For more information on the Australian National Bibliographic Database see <http://www.nla.gov.au/kinetica/databases.html#nbd>

<sup>2</sup> <http://adt.caul.edu.au>

<sup>3</sup> National Library of Australia, *Directions 2003-2005*, <http://www.nla.gov.au/library/directions.html>

<sup>4</sup> National Resource Sharing Working Group 2002 Interlibrary Loan and Document Delivery Benchmarking Study, [http://www.nla.gov.au/initiatives/nrswg/illdd\\_rpt\\_sum.html](http://www.nla.gov.au/initiatives/nrswg/illdd_rpt_sum.html)

<sup>5</sup> For more information on Libraries Australia see <http://www.nla.gov.au/kinetica/librariesaustralia.html>

<sup>6</sup> Survey of Australian Theses Contributions to the National Bibliographic Database Final Report November 2002, <http://www.nla.gov.au/kinetica/austthesessurvey.html>

<sup>7</sup> For more details of the project and research see Missingham, R., Wilson, F., Smith, J. (2005) "Evaluating the public library portal", paper presented to the Information Online 2005 conference, Sydney, <http://www.nla.gov.au/nla/staffpaper/2005/missingham1.html>

<sup>8</sup> For more details see Missingham, R. (2004) "Reengineering a National Resource Discovery Service", D-Lib magazine, vol. 10 no. 9, September 2004, <http://dlib.org/dlib/september04/missingham/09missingham.html>