

# National Library of Australia IT Architecture Project Report

March 2007

---



# TABLE OF CONTENTS

Table of Contents .....	i
Overview .....	1
Purpose .....	1
Scope .....	1
Benefits.....	1
Credits .....	2
Background .....	3
Context .....	3
Current IT architecture .....	3
Principles.....	4
Achievements .....	4
Future directions.....	5
The problem to be solved .....	6
Challenges .....	6
Inhibitors .....	6
Requirements.....	7
Change 1: Adopt a service-oriented architecture .....	8
Benefits.....	8
Service framework .....	8
Case studies .....	9
Enablers and inhibitors .....	9
Change 2: Single business .....	11
Benefits.....	11
Single data corpus .....	13
Musings .....	13
Enablers and inhibitors .....	13
Change 3: Open source development model .....	14
Benefits.....	14
Enablers and inhibitors .....	15
Conclusion.....	17
Appendix 1: Service-oriented architecture case studies .....	19
Search .....	19
Ingest and Delivery .....	20
Appendix 2 Single business musings .....	23
Wanted resource .....	23
Topic-based searching.....	23
User participation .....	25
Matching and merging.....	26
Branding and marketing .....	26
Partnerships and other issues.....	27

# OVERVIEW

## Purpose

The aim of this report is to define the IT architecture that will be needed to support the management, discovery and delivery of the National Library of Australia's collections over the next three years. The current architecture has enabled the Library to develop a significant digital library capability over the last decade. Now the burden of maintaining and supporting existing systems and services is increasingly hindering us from bringing new services online, improving the user experience, exploring new ideas or responding to technological change. In the meantime, enormous changes are occurring in the broader environment.

## Outcomes

The report identifies a new framework for building digital library services that should address these issues by:

- Implementing a service-oriented architecture
- Adopting a single-business approach
- Considering open-source solutions when these are functional and robust.

## Scope

The changes proposed in this report apply to the Library's core mandate to develop and maintain a national collection of library material and to make this collection available. They deal with the digital library services needing to be in place to collect, to preserve and to provide access to resources in any format. Services needed to support the creation and publication of resources by the Library are dealt with only in terms that would also apply to any creator or publisher needing to contribute resources to the national collection or to reference resources in the national collection in exhibitions, publications and other works. Similarly, corporate services such as human resource management and finance are dealt with only in terms of shared infrastructure such as identity management and authentication.

## Benefits

### *Service-oriented architecture*

A service-oriented architecture is a way of thinking about software as a set of interfaces that can be called to execute a business function. It is becoming widely accepted as best practice in the IT industry where its adoption is being enabled by the emergence of web services based on accepted standards. Implementing a service-oriented approach will result in significant efficiencies through the use of a common shared technical infrastructure that enables innovation supported by an overarching service framework allowing business owners and developers to have a shared understanding of requirements and directions.

### *Single business approach*

Even with a service-oriented approach, the Library's capacity to meet its directions will continue to be eroded as new applications are brought online. As budgets continue to tighten and the Library needs to do more with less, there will come a time when a large proportion of development effort will be spent just maintaining existing applications.

To address this issue, and as part of implementing the service-oriented architecture, it is proposed that the Library regard its digital library services as a single business with a single data corpus that can be deployed in a range of contexts. Rather than developing separate

applications to meet a new requirement, each requirement would be viewed as an enhancement to the business that could be deployed across all relevant business contexts.

This is a significant change to the way the Library currently works. As well as resulting in further significant efficiencies for IT staff, it has the potential to bring library staff together in unprecedented ways to work on problems and ideas and to prototype solutions that enhance the user experience regardless of the point of access.

#### *Open-source solutions*

To achieve further efficiencies, it is also proposed that the Library regularly review the capability of the software products it uses to meet its directions and that, as part of this review, it consider open source solutions where these are robust and functional. For functionality developed in-house, it is proposed that the Library return intellectual property to the public domain.

This is a change from the current policy, which, although it encourages the use of open source software, still reflects a preference for a buy-not-build approach and for licensing models or the transfer of intellectual property to a product vendor.

### **Credits**

IT Architecture Project Team:

- Kent Fitch (Technology & Architecture)
- Paul Hagon (Web Publishing)
- Simon Jacob (Collection Access)
- Alexander Johannesen (Web Publishing)
- Ninh Nguyen (Collection Infrastructure)
- Judith Pearce (Feasibility & Standards)
- Mark Triggs (IT Services)

## BACKGROUND

### Context

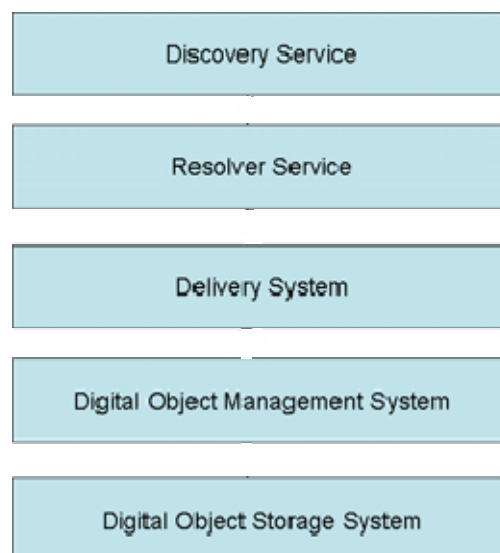
A primary legislative mandate of the Library is to develop and maintain a national collection of library material (including a comprehensive collection of library material relating to Australia and the Australian people) and to make this national collection available<sup>1</sup>. In practice, the national collection is distributed, with the national and state libraries sharing a deposit role for Australian materials and all libraries focusing on the specific needs of their constituencies for overseas materials

For more than thirty years, information technology has been a major enabler for fulfilling this mandate. The establishment of the Australian Bibliographic Network to support the development and maintenance of a national union catalogue in 1981 was a key milestone, as was the implementation ten years later of an Integrated Library Management System to manage and provide access to the Library's own collection.

Growth in use of the Internet as a publication medium and as a mechanism for service delivery presented significant new challenges in the 1990s. The Library recognised that its collecting mandate had to include Australian electronic publications and defined three levels of collecting: electronic publications the Library itself safeguarded for future access; those that were safeguarded by other agencies; and those that were considered of current interest only and linked to in the catalogue for the life of the publication.

### Current IT architecture

In 1996, as part of the Digital Services Project, the Library developed an architecture to support the collection of electronic publications and the digitisation of materials in traditional formats. The architecture has five loosely-coupled layers: a discovery service layer, a resolver service layer, a delivery system layer, a digital object management system layer and a digital object storage system layer.



---

<sup>1</sup> National Library Act 1960 (<http://scaletext.law.gov.au/html/pasteact/1/761/top.htm>).

## Principles

The following principles informed the development of this architecture and still inform all of the Library's digital library development activities:

- the need to unite the functions of the traditional library with those of digital library services in ways that enable discovery of wanted resources regardless of format;
- the need to describe resources once, as part of collection management workflows in ways that enable re-use of the resulting metadata in a range of local and federated contexts;
- the need to be able to cite content and metadata in ways that are unique, persistent and resolvable;
- the need to support discovery in a range of local and federated contexts in ways that enable delivery even when conditions are imposed on access or analogue processes are involved; and
- the need to manage resources in ways that preserve them and facilitate future access.

## Achievements

Over the last decade, the digital library capabilities of the Library have been significantly enhanced under this framework. In Endeavour's Voyager (now part of the Ex Libris product suite), the Library has acquired a third generation Integrated Library Management System that is used as the source of metadata for the digital object management system layer. PANDORA<sup>2</sup> provides a permanent digital archive for Australian websites and the Digital Collections Manager (DCM)<sup>3</sup> integrated collection management and delivery facilities for its digital still image and audio collections. Both of these services have been developed in-house and use persistent identifiers and a resolver service to enable access to content. Digital objects are stored on file systems that are regularly augmented to meet capacity requirements. Delivery services are supported by a document request management system based on Rélais.

In Libraries Australia<sup>4</sup>, the Library has acquired a means of providing end-user access to the collections of Australian libraries, and support for delivery workflows. Picture Australia<sup>5</sup>, Music Australia<sup>6</sup>, the Register of Australian Archives and Manuscripts (RAAM)<sup>7</sup> and ARROW (Australian Research Repositories Online to the World)<sup>8</sup> exemplify how specialist digital library services might be developed and delivered based on metadata harvested from a range of partner agencies.

All of these services have a metadata repository and search system component based on Inquision's Teratext software. The Australian Bibliographic Database which delivers the Library's union catalogue is developed and maintained through bibliographic utility services provided by OCLC Pica's CBS software and interlending utility services provided by Fretwell Downing's VDX system.

The Library has also had some success enabling the discovery of items in Australian library collections through other pathways, not just its own web-based services. It has done this by making its metadata collections accessible through standard protocols such as Z39.50, OpenSearch and OAI-PMH, by seeding search engines with resource descriptions and images of its digitised collections and by working with Google to make records from the Australian

---

<sup>2</sup> <http://pandora.nla.gov.au/>.

<sup>3</sup> <http://www.nla.gov.au/digicoll/>

<sup>4</sup> <http://librariesaustralia.nla.gov.au/>

<sup>5</sup> <http://pictureaustralia.org/>

<sup>6</sup> <http://musicaustralia.org/>

<sup>7</sup> <http://www.nla.gov.au/raam/>.

<sup>8</sup> <http://search.arrow.edu.au/>

National Bibliographic Database (ANBD) accessible through Google Scholar. It has also looked at the feasibility of providing access to the collection as a logical view of the ANBD and prototyped new models for a national discovery service<sup>9</sup>.

## Future directions

In its Directions for 2006-2008<sup>10</sup>, the Library describes its major undertaking for 2006-2008 as to "enhance learning and knowledge creation by further simplifying and integrating services that allow our users to find and get material, and by establishing new ways of collecting, sharing, recording, disseminating and preserving knowledge".

Five desired outcomes are identified for this period:

- to ensure that a significant record of Australia and Australians is collected and safeguarded;
- to meet the needs of our users for rapid and easy access to our collections and other resources;
- to demonstrate our prominence in Australia's cultural, intellectual and social life and foster an understanding and enjoyment of the National Library and its collections;
- to ensure that Australians have access to vibrant and relevant information services; and
- to remain relevant in a rapidly changing world, participate in new online communities and enhance the visibility of the Library.

Outcome 5 has become a mantra for the Library and informs strategies for achieving all the other outcomes.

---

<sup>9</sup> Library labs (<http://l101.nla.gov.au/>).

<sup>10</sup> <http://www.nla.gov.au/library/directions.html>.

## THE PROBLEM TO BE SOLVED

In spite of the achievements identified above, there is still a huge amount to do over the next few years to position the Library to achieve its directions and to respond to the changes that are occurring in the broader environment.

### Challenges

#### *Collection management and delivery*

The Library's response to the volume of material being created in digital form now needs to be increased by orders of magnitude if the PANDORA Archive is not to become increasingly irrelevant over time. The Library's collection management and delivery infrastructure needs to be extended to support the deposit of electronic publications, to rescue digital content in the collection that is stored on physical carriers, to take regular snapshots of the Australian web domain and to support the mass digitisation of Australian newspapers and journals. There is also a need for an integrated digital repository infrastructure to ensure preservation of and access to content collected through the Library's various management systems.

In the medium term it is unlikely that there will be any significant decrease in the volume of material needing to be taken into the Library in traditional formats. It will be an ongoing priority to make material in traditional formats accessible in digital form, either by digitising it or by acquiring or linking to digital versions. In order to do more with less, staff will need access to workflow systems that minimise the need to re-key data and automate processes as much as possible.

#### *Discovery and access*

To fulfil its mandate to make the national collection available the Library needs to ensure that items in the collection can be discovered and accessed in many different contexts, both inside and outside of the Library's control. This is particularly relevant to achieving Outcome 5. Like many agencies the Library tends to focus on the development of its own web-based services. To remain relevant in an increasingly digital world it needs to take its unique data to other online spaces. To do this effectively, it needs to enhance its record import and export services to support the collaborative development of trusted aggregations of both metadata and full text indexes, to define and market these aggregations and to make them available through standard protocols for re-use by other players.

The Library also needs to continue enhancing its own web-based services to ensure that they deliver a recognisable and competitive product, are easy to use, facilitate learning and knowledge creation and meet user needs. There is a need to consolidate existing services, to improve the capability of searches to deliver results through relevance ranking, clustering and contextualisation, to enable user collaboration in the development and interpretation of content, to ensure a seamless workflow between discovery and delivery and to implement new models for unmediated delivery.

### Inhibitors

Goals to address these needs have been identified in the three-year IT Strategic plan<sup>11</sup> but the burden of maintaining and supporting existing systems and services is increasingly hindering the Library's capability to bring new services online, to innovate and to respond to new technologies. Each new project adds to the number of applications requiring support and hence to the availability of staff to work on new projects.

---

<sup>11</sup> <http://www.nla.gov.au/policy/itplan.html>.



During 2006-2007 alone, it is planned to build three major new federated services - Australian Newspapers Online, Journals Australia and People Australia - and to redevelop ARROW and RAAM. One of the benefits identified for Libraries Australia was that it would provide a generic infrastructure to support innovation and the development of new federated services. In practical terms this has not been achieved.

New services are still being developed as separate applications. Separate solutions are being developed to solve the same problem. Code is not being shared. Enhancements to one service are not immediately able to be applied to others with similar requirements. Services such as RAAM become increasingly more out-of-date as they wait for migration to new technologies. New services such as Music Australia have long enhancement registers. Workflow enhancements that might provide significant efficiencies to the Library have to defer to higher priority projects. At the same time, the cost of recruiting and maintaining staff is rising, so that less can be done with available resources.

## **Requirements**

For the Library to meet its directions for 2006-2008 and beyond, it needs a new approach to the development and deployment of its digital library services. This approach needs to enable the Library to do more with less by making development and support processes more efficient. It needs to support the incorporation of features to improve the user experience that are still lacking in existing services, such as good relevance ranking, clustering, FRBR, annotations and rich relationships. It needs to support a fast response to changes in technology, making it easier to take up and test new ideas and opportunities as they arise. It also needs to support a prototyping environment that enables the Library to look beyond the bounds of current services and ways of doing things, and to tackle some of the things that seem too hard to do now or that it has found too hard to do in the past. These may be what truly differentiate its services from those of other players in an increasingly digital world.

## CHANGE 1: ADOPT A SERVICE-ORIENTED ARCHITECTURE

A service-oriented architecture is a way of thinking about software as a set of self-contained components that can be called to execute a business function. Components can be based on existing software or built from scratch. The service uses mappings to translate messages into the form required by the underlying technology.

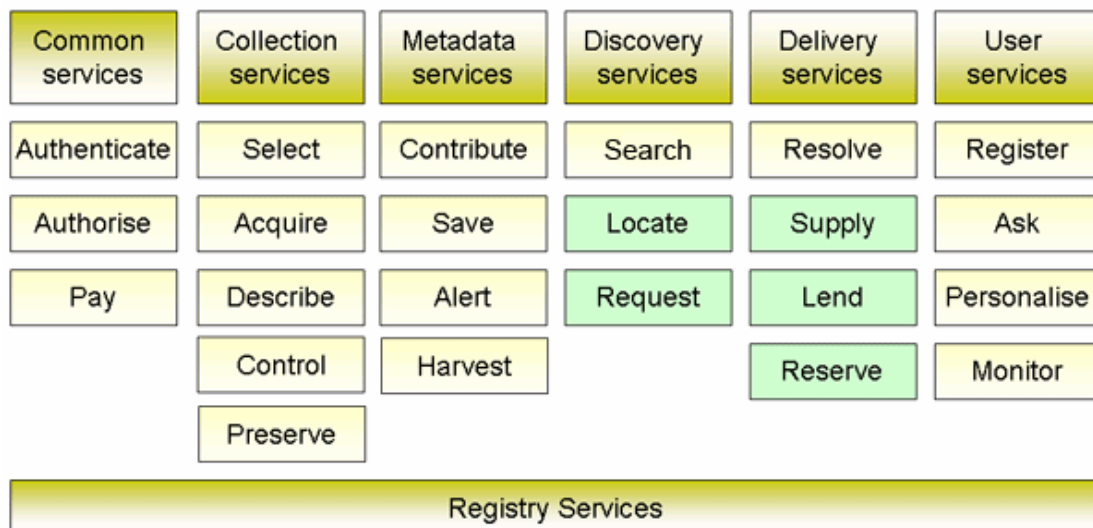
### Benefits

A service-oriented architecture frees business from the constraints of technology by leveraging on existing assets while easily enabling change.

- Services developed once can be re-used in a range of applications.
- Enhancements to a service are immediately available for use by all applications using it.
- Bugs fixed once are fixed for all contexts in which the service is used.
- Interfaces can be easily established with third-party applications.
- Prototypes are easy to develop, supporting innovation and iterative development.
- Functionality can be tested through a web browser.
- Legacy systems can be supported until they are no longer required.
- Underlying technologies can be interchanged without changing the applications.

### Service framework

The efficiencies delivered by a service-oriented architecture can be optimised through an overarching service framework that enables business owners and developers to work together to create maintainable, extensible, compliant systems.



The diagram above identifies a set of high level, abstract services that would need to be supported in a service-oriented approach. These are grouped into six sets.

- Common services - Authenticate, Authorise and Pay - work across applications to identify who the user is, what they are able to do and the conditions that apply and also to manages any e-commerce obligations.

- Collection services - Select, Acquire, Describe, Control and Preserve - support the development and maintenance of the collection.
- Metadata services - Contribute, Save, Alert and Harvest - support the development and maintenance of federated aggregations of content and the sharing of this content with other players. Contribute includes both online and offline methods of contribution, and the contribution of metadata of all kinds, including annotations.
- Discovery services - Search, Locate, Request - support the finding of wanted resources and the transfer of requests for access or use to the resource provider.
- Delivery services - Resolve, Supply, Lend and Reserve - support the delivery of wanted resources, either by resolving directly to the resource once conditions have been met for access, by supplying or lending a copy or by reserving a copy if it is currently in use.
- User services - Register, Ask, Personalise and Monitor - deal with the relationship of the user with the service - enabling the user to register for value-added services, to engage in a dialogue with the service provider in order to get help or provide feedback, to set preferences for their interaction with the service and to monitor their own usage. Monitor also allows the service provider to monitor usage across all users and functions.

The registry service layer provides access to the information about users, contributors, target collections, resource providers, access and use policies and protocol information that needs to be collected and maintained to support these functions.

The Digital Library Federation (DLF) is actively working on the development of a service framework for libraries, based on the example set by the E-learning Framework (ELF)<sup>12</sup>. This work will help the Library to refine its own framework and identify any new protocols and data schemas needing to be supported by its services, the gaps needing to be addressed through a standards-based approach to ensure interoperability with external systems and opportunities for collaborative activities.

## Case studies

Case studies showing how a service-oriented architecture might be implemented for Search and Ingest and Delivery functions may be found in Appendix 1.

## Enablers and inhibitors

Service-oriented architectures are becoming widely accepted as best practice in the IT industry where their adoption is being enabled by the emergence of web services based on accepted standards. For the Library, this is an achievable way of addressing the following issues with staying with the current development approach:

- How to prevent maintenance of applications from absorbing more and more of the available IT resource.
- How to bring new functionality online faster.
- How to improve the efficiency of IT staff so that they can do more with less.
- How to meet user needs in a consistent way.
- How to be responsive to user feedback.
- How to be responsive to technological change.
- How to foster innovation.

---

<sup>12</sup> DLF Service Framework (<http://www.diglib.org/architectures/serviceframe/>). See also Brian Lavoie, Geneva Henry and Lorcan Dempsey, "A Service Framework for Libraries", *DLIB Magazine*, 12 (7/8), July/August 2006 (<http://www.dlib.org/dlib/july06/lavoie/07lavoie.html>).

- How to enable software development as a facilitator of business change.
- How to embrace collaboration in ways that provide a significant return on investment in terms of new capabilities.

One of the highest inherent risks is that business areas and IT do not work together to ensure the re-use of services. The primary control for this is the subject of the next section.

## CHANGE 2: SINGLE BUSINESS

A service-oriented architecture is not a technology that can be implemented out-of-the-box but rather a way of thinking that informs the development process. There are still challenges in agreeing how a service should be implemented across applications; and risks that the new way of thinking will be only partially deployed, with some applications continuing to be developed independently. This risk could be mitigated, and further significant efficiencies achieved, by treating the Library's digital library services as a single business with a single data corpus that can be deployed in many different business contexts.

The single business approach could be implemented at two levels:

- The Library could think in terms of a single business and a single data corpus as part of its strategic and operational planning processes. This would mean that, instead of separate business plans for each new service and separate enhancement registers with competing priorities, there would be one business plan informed by coherent strategies for enhancing the single business. Such strategies might involve the development of new functionality or focus on refining the capability of the business to meet needs in priority areas of interest.
- The IT Division could implement digital library solutions in ways that minimise the number of separate applications needing to be maintained and enable new functionality or refinements developed for one business context to be easily deployed to another.

This document recommends adopting the single business approach at both levels.

### Benefits

#### *Collection management and delivery*

In many ways the Library is already treating collection management and delivery as a single business and reaping the benefits. It has a single system (DCM) that supports the digitisation of both still image and audio materials. Work is underway to build a fully-generalised delivery system for digitised content and a Rights Management System Project is addressing the need to manage access and use across most material types.

Implementing a service-oriented architecture will enable DCM and PANDAS (the PANDORA management system) to share an underlying repository and one could argue that both systems support such separate workflows that they do not need to be regarded as serving a single business. Functionality is converging, however, in areas such as rights management and the requirement to collect electronic publications. There is also a risk that the Newspaper Digitisation Project will deliver a separate but strongly overlapping digital content management solution for newspapers if this is seen as a separate business requirement.

A single business approach to collection management and delivery would enable the Library:

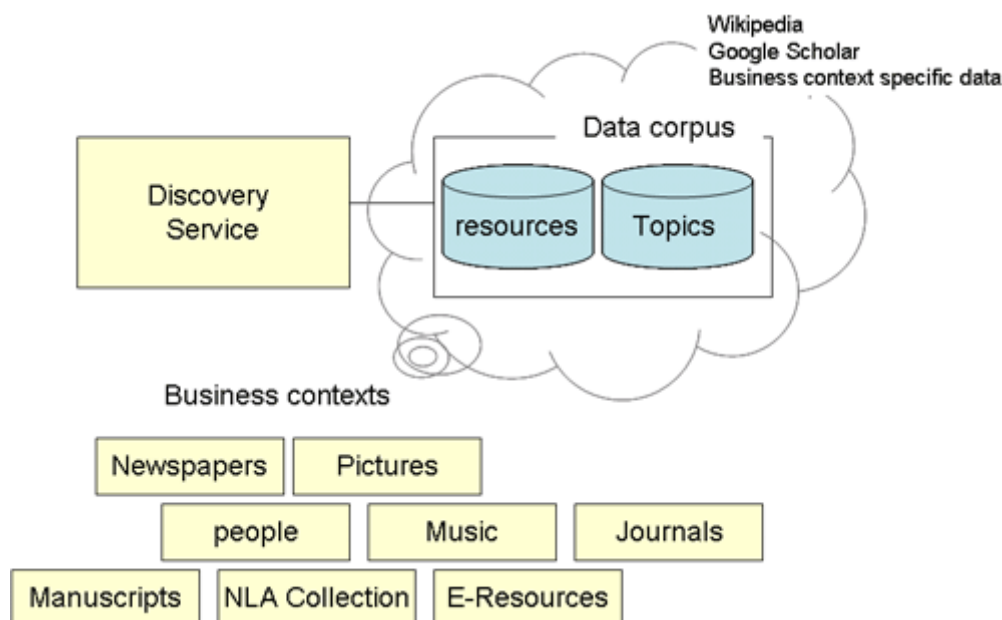
- to replace existing applications over time by a suite of collection management and delivery workflow systems targeted to specific contribution methods and content models; and
- to ensure that metadata and full text indexes are aggregated into appropriate logical views of the single data corpus to support federated resource discovery, regardless of the methods used to collect the content.

In the case of collection management and delivery, workflow systems may be delivered by separate applications where the contribution methods and content models sufficiently diverge and where the identified solution has been developed by a third party, for example, the Web Curator Tool as a replacement for PANDAS to support website harvesting workflows.

### *Discovery and access*

The benefits of treating discovery and access as a single business cannot be overstated. It is here that most of the Library's development effort is spent and here that there is most duplication of functionality and most need to improve the user experience if the Library is to remain relevant in a digital age. The Library simply cannot go on the way it has, creating stand-alone applications with strongly overlapping functionality, and achieve its directions. A clear way forward is to build a new single national discovery service that can be accessed through a range of different business contexts.

With a single national discovery service, developers would only need to support one application. Staff would work closely together to identify priorities for the service. Users would have the same opportunities to find relevant information whether they had started the search from a generic search box or from a manuscript or pictures context or from an Internet search engine. The data corpus searched would be the same in each case. The only difference is that the results might give precedence to manuscripts or music or pictures depending on the context.



There would still be a need for projects like People Australia or Journals Australia to address gaps in the information infrastructure but the primary outputs of these projects would be new partners, an enriched data corpus and enhanced functionality that could immediately be deployed to other business contexts.

Instead of redeveloping the same Contribute / Search / Alert / Harvest paradigms for each new application, the Library would be able to invest resources in improving the finding and getting process across all business contexts and in developing support for personalisation and user participation. It would be able to do this in a coherent and cohesive way that crosses project boundaries, through an iterative prototyping process, using laboratory versions to test proposed solutions with real users, and building their feedback into the development and release loop.

## **Single data corpus**

For some time the Library has been thinking about treating the content it makes available through its discovery services as a single data corpus that may be accessed through different logical views. The data corpus could consist of one physical repository or of a number of separate aggregations. Pictures, newspapers and journals may be better managed as separate aggregations to the ANBD, for example. There will also be a need to distinguish aggregations of resources from aggregations of topics (people, organisations, places or subjects).

Treating this data as a single corpus with a range of trusted logical views means that users do not have to search across multiple targets with overlapping content for full recall. The scope of each target can be simply stated and promoted - Australian library collections, our collection, pictures, newspapers, music.

Whether users elect to search the whole corpus or a subset, there is no dumbing down of search results. Tools such as relevance ranking, clustering and assistance with spelling and terminology can be applied to the whole corpus, enhancing search outcomes. In addition, the contextual approach to discovery implemented for Music Australia and being developed for People Australia can be applied across all business contexts and all types of topics.

The corpus itself would also be extendable to aggregations maintained by other stakeholders, including Google Scholar for international journal resources and Wikipedia for topics not included in the Library's own authority files. Each business context would also have target aggregations that would extend the data corpus for that context: for example, the manuscripts context would also report on hits in the National Archives of Australia collection or, for authorised users, RLG's Archival Resources.

## **Musings**

Musings about how a single business approach might be taken to discovery and access may be found in Appendix 2. The second section discusses topic-based searching. It shows how the benefits of the People Australia Project can easily be extended to other business contexts and other topics through this approach. Other sections look at the wanted resource, user participation, matching and merging, branding and marketing and the need for partnerships with Google Scholar and Wikipedia as ways of extending the data corpus.

## **Enablers and inhibitors**

The main enabler for taking a single business approach is that the Library itself has been looking at ways in which it might re-organise itself better to meet its directions and to do more with less. A physical restructure is probably needed less than a new way of sharing ideas, communicating what is happening across the Library and building up the IT literacy of all staff. The single business approach provides a way of doing this by bringing people together to work on solutions to shared problems and by enabling all staff to be involved in testing and evaluating prototypes.

The main risks have to do with acceptance of the single business approach and migration of existing services.

## CHANGE 3: OPEN SOURCE DEVELOPMENT MODEL

The Library's current policy (last articulated in the 2005-2008 Strategic Plan) is:

- to base the development of services on solutions that are available in the marketplace, unless these solutions fall well short of the Library's functional requirements, do not fit the Library's IT environment, are too costly, or involve unacceptable levels of risk;
- to minimise software costs by utilising open source software whenever this provides a functional and robust solution; and
- to minimise maintenance and support costs of in-house developed software by exploring models for collaborative software development, the licensing of software for use by other agencies or the transfer of intellectual property to a product vendor.

The following changes are proposed to this policy:

- to evaluate open source solutions on equal terms with solutions available in the marketplace through a rational costing process; and
- to return in-house developed software to the public domain.

By evaluation on equal terms is meant the use of a rational costing method that takes into account the work that would need to be done to enhance an open source solution to meet the Library's needs, the benefits of that work to the wider community and the lost opportunity costs to the Library itself and to the wider community with a commercial solution if the vendor's development priorities are not aligned with those of the Library.

### Benefits

#### *Collection management and delivery*

The benefits of this approach are already being demonstrated for collection management and delivery through the Library's involvement in the APSR Project (Australian Partnership for Sustainable Repositories)<sup>13</sup> and the International Internet Preservation Consortium (IIPC)<sup>14</sup>. The Web Curator Tool developed by the National Library of New Zealand and the British Library<sup>15</sup> may provide the migration path for PANDAS to a service-oriented architecture. The Global Digital Format Registry<sup>16</sup> will provide the Library with preservation management capabilities it could not have afforded to develop by itself.

#### *Discovery and access*

Similarly, for discovery and access, a recent decision by the Library to adopt the open source product Lucene as the Library's metadata repository and search system will enable the Library to take advantage of further enhancement of this product by an international community with a strong interest in ensuring it remains a robust and functional product.

The Library itself will contribute to this process by enabling its metadata collections to be searched through a Z39.50 -SRU gateway and returning this code to the public domain. This will mean that other agencies wishing to implement new web-based search protocols while still supporting access through legacy protocols will be able to do so, using a best practice service-oriented approach.

---

<sup>13</sup> <http://www.apsr.edu.au/>.

<sup>14</sup> <http://netpreserve.org/>.

<sup>15</sup> <http://webcurator.sourceforge.net/>

<sup>16</sup> <http://hul.harvard.edu/gdfr/>.



Choosing Lucene as its metadata repository and search system rather than a commercial product has also positioned the Library to look at open source solutions for document analysis and the clustering of search results.

#### *Library management system software*

When it comes to a mission-critical system like the ILMS with hundreds of person years of intellectual property invested in its development, it may seem axiomatic to take a buy-not-build approach. Even here, questions are being raised by industry leaders such as Lorcan Dempsey about the future of the ILMS industry<sup>17</sup>. He suggests the following not necessarily exclusive future scenarios:

- Single monolithic system (one of the ILS vendors left standing)
- Vertically integrated system (e.g., with financial, HR, course management, and campus or community portal systems)
- Open Source system (e.g., Evergreen, Koha, the University of Rochester's eXtensible catalog)
- Suite of dis-integrated or interoperable systems
- OCLC custom service suite

The Library has already identified a requirement to replace the OPAC module with a logical view of the NBD but there are other limitations of Voyager that are hindering the Library from achieving its collection management objectives. It tends to be Voyager's support for workflows that dictates the Library's business processes rather than the other way around. Voyager's recent merger with Ex Libris also places the Library at risk of having to replace its current ILMS sooner than it may have planned. It would certainly be of value as part of the planning process to review the capabilities of current open source ILMS solutions and to assess the cost of enhancing the most promising product to meet the Library's needs.

As part of assessing this cost, the Library needs to look at opportunities lost by continuing to depend on the capabilities of the industry to meet the needs of Australian libraries; and opportunities that might be gained by actively working with an international community to develop a robust and functional ILMS in the public domain. These include opportunities to improve the efficiency of collection management and delivery processes in the National Library itself and in other libraries as well as opportunities to develop and enhance the NBD as a national union catalogue. There could be benefits in prototyping solutions to specific problems such as the time taken to do subject cataloguing and the need for workflows to support a federated approach to authority control.

### **Enablers and inhibitors**

Open source software development models have accelerated with the advent of the Internet. The commercial world is now also starting to recognise it as a way of reducing costs and achieving directions.

The Library is already using open source software for some applications and contributing to the development of open source solutions through a range of collaborative projects. Implementing a service-oriented architecture will make it easier to test the capability of open source software components. The reputation of the Library will be increased as stakeholders see it taking a leadership role in areas where commercial products are not meeting their needs or the needs of their users.

---

<sup>17</sup> "Library systems world", *Lorcan Dempsey's Weblog*, November 13, 2006 (<http://orweblog.oclc.org/archives/001202.html>).

Risks of this change in policy are minimal. The financial risk is low as the Library does not have a history of significant return on investment through the licensing of code. Risks associated with operations and services will be addressed through controls already provided by the Library's project management methodology.

## CONCLUSION

Information technology is a crucial enabler for delivering the Library's digital library services. Over the last ten years the Library has developed a digital library architecture that supports discovery and access to wanted resources regardless of format. This architecture has some strongly separated components. The metadata repository and search system, for instance, uses the Z39.50 search and retrieval protocol, enabling one product that is Z39.50 compliant to replace another. In other areas the architecture is less modularised. In both DCM and PANDAS workflow systems are tightly coupled to the underlying repository. In addition. Within each component, there are also many more functions that could be seen as self-contained and built in ways that would enable them to be shared between applications and delivered through different technologies as technologies change.

In the IT industry these issues are being addressed by adopting a service-oriented approach. The Library is well positioned to take such an approach because of the number of standards that are already in place or under development to support interoperability in a global information environment. The Library's IT Division has already started to implement a service-oriented architecture, beginning with the services required to support currently scheduled projects. This will change the way the Library specifies requirements and builds and maintains applications.

There will need to be more planning and peer review at the start of projects to determine what services are needed and how they might need to be enhanced to support the new requirement. Services will be built iteratively, with early versions only delivering the functionality that is immediately required. Short development timeframes for work packages will allow for prototyping, frequent business and user testing and experiential learning. Once a suite of services are in place, development times will be shortened as services start to be shared between applications. Business analysts will draw on a set of core use cases. It will be easier to prototype and iteratively test solutions. As a result development teams will be able to work more efficiently and to size, cost and schedule projects with more confidence.

Even so, the capability of the Library to meet its directions will continue to be eroded as new applications are brought online. As budgets continue to tighten and the Library needs to do more with less, there will come a time when a large proportion of development effort will be spent just maintaining existing applications. To address this issue the Library needs radically to rethink how it might continue to fulfil its core mandate to develop and maintain a national collection of library material and make it available.

This report recommends that the Library regard its digital library services as a single business with a set of clearly defined products. This is a significant change to the way the Library currently works. Rather than developing separate applications to meet a new requirement, each new requirement would be viewed as an enhancement to the business. Projects like Music Australia and People Australia would result in new partnerships, an enhanced data corpus and new functionality that would become immediately available to other business contexts.

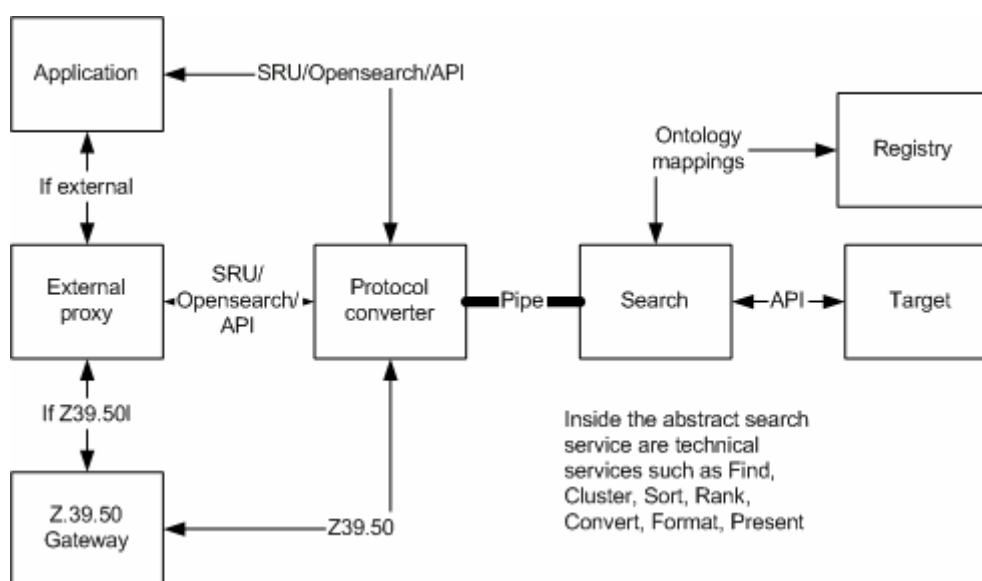
Coupled with this recommendation is a quite significant change to the Library's software development policies. Although the Library does use open source software, its current policy is to prefer solutions available in the marketplace unless they significantly fail to meet the Library's needs. The IT Division now proposes to compare open source solutions on an equal basis with solutions available in the marketplace, using a rational costing method and to return intellectual property developed by the Library to the public domain.

These three strategies will position the Library to bring new functionality online faster and to meet user needs in a more consistent way. The Library will be able to respond more easily to user feedback and technological change. There will be more opportunities for innovation through prototyping and beta releases, raising the profile of the Library in the community. Library staff will work together to develop ideas and identify priorities and to be informed about what the Library is doing. Time will be needed for experimenting, learning, training, there will be some slippages with scheduled projects in 2007 to start implementing the new architecture, but opportunities for earlier deliverables as new services come online. Over time, the full benefits will start to be realised, with services only needing minor configuration changes to be adapted to meet a new requirement.

## APPENDIX 1: SERVICE-ORIENTED ARCHITECTURE CASE STUDIES

### Search

All of the Library's federated search services are currently accessible to third party systems through the Z39.50 search protocol. However, it is not possible to use the metadata in these collections without implementing a Z39.50 client and this is not a trivial exercise. This is inhibiting the Library from looking at new ways in which data could be used and combined, both in its own on-line spaces and at other points of user need. There is no coherent strategy for promoting the Library's metadata collections to new players and staff do not have a shared understanding of the importance of this need in terms of meeting Outcome 5.



A service-oriented approach to search would address this problem by making the capability to support multiple search protocols as both a requester and a responder part of the digital library infrastructure. In the diagram above, a single search service supports requests and responses in a range of protocols by means of a converter. When a new target is registered, mappings are made to a single internal protocol. This target then becomes available through all the supported protocols. When a new protocol needs to be supported, changes are needed in just one place for all registered targets to be enabled.

The SRU standard is the most likely candidate for the internal protocol. It has an extension capability that can be used to carry requests for services like clustering or ranking based on user preferences that are not natively supported. These services will distinguish the Library's web-based applications from those of third parties searching the same target. However, using SRU as the internal protocol also positions the Library to standardise the extensions in order to offer these services to third parties.

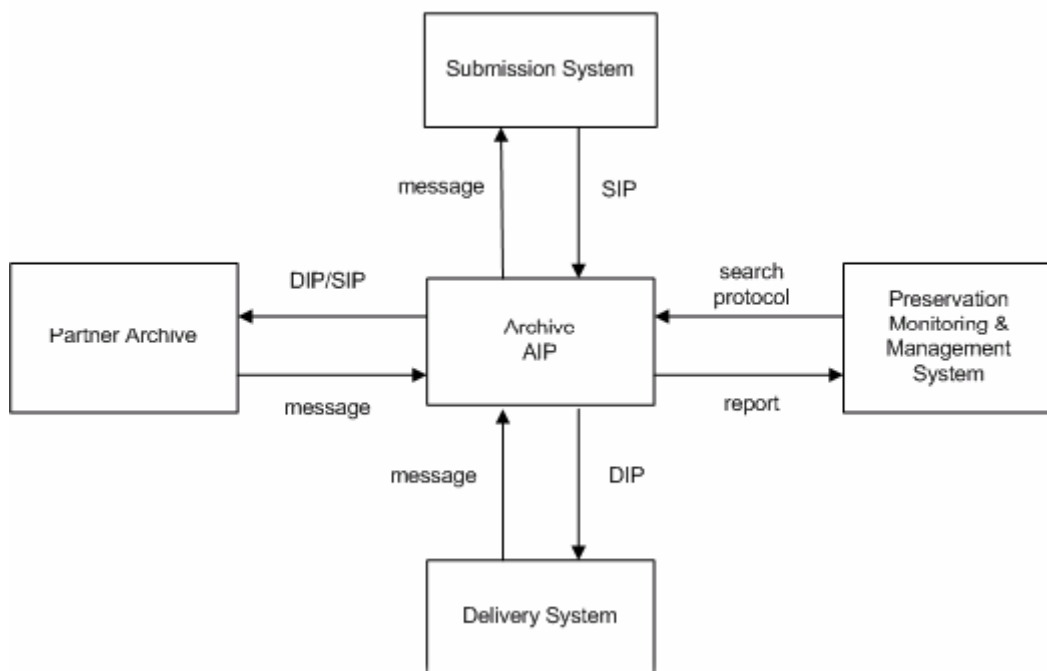
The approach illustrated above for Search can also be extended to other services in the Library's service framework in ways that will support current and legacy protocols. The benefits of having a native level of support for standard protocols in the architecture cannot be overestimated. A standards-based service-oriented approach for core services such as Contribute, Alert, Harvest and Request will allow protocols such as SRU Update, RSS, OAI-PMH and OpenURL to be supported across all applications. It will also ensure that these

protocols are part of the Library's way of thinking when training new staff or prototyping new requirements; and that gaps in standards are identified and addressed through a testbed approach, as part of the development process.

## Ingest and Delivery

For historical reasons, PANDORA and DCM have been developed as two separate applications, with separate data models and separate Ingest and Deliver services. This has made it hard to decide which system to use for new kinds of materials and new contribution methods. PANDAS has on its enhancement register a requirement to support the deposit of electronic publications in a range of formats. DCM is being enhanced to support the management of subscription service datasets and electronic publications on physical carriers. Meanwhile, the Library has begun using the Open Journal System (OJS) software to assist groups to publish Australian journals on its website. It has a licence for the VTLs Vital software as part of its participation in the ARROW Project, with plans to use this software for an independent scholar's repository. There are also requirements to support mass digitisation for Australian newspapers and journals, with a concomitant requirement for workflow support. For all of these collections, there is a need to implement a preservation management regime that is file format-based and independent of the system used for the lodgement or capture of material.

A service-oriented approach to ingest and delivery would enable the systems used to support collection management and delivery workflows to be separated from the underlying repository. The diagram below shows the various systems that would need to interoperate in the proposed architecture and the interfaces between them.



Submission systems would use the Ingest service to pass a Submission Information Package (SIP) to the archive. Delivery systems would request a Dissemination Information Package (DIP) from the Archive. Partner archives would submit or request packages as part of a

transfer of content between one system and another. Preservation monitoring and management systems would use a search protocol to request reports from the archive about the status of content in particular file formats.

The Library now has a good understanding of the protocols and data schemas needed to implement this framework - METS for structural metadata, ALTO for OCRd text, PREMIS as a framework for preservation metadata, XACML for access and rights management and ISO 2146 as a framework for registry services. Through its involvement in the APSR Project it has developed a an Australian PREMIS profile based on METS and through the IIPC it has been involved in the development of the WARC format for archived websites and open source tools for the large-scale archiving of websites.

The Library is currently considering FEDORA for its repository software solution but what software is used is less important from the architectural point of view than a service-oriented approach with standard interfaces that will allow this software or some of its components to be replaced if a new technology better meets the Library's directions.



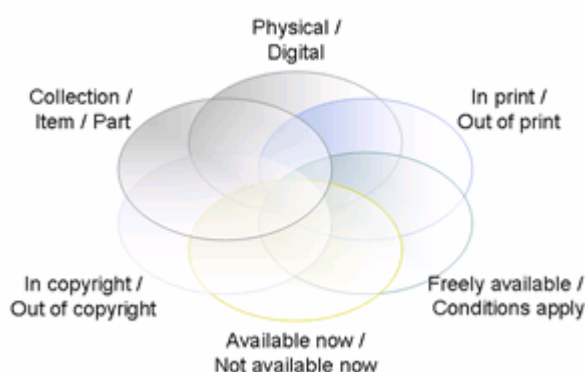


## APPENDIX 2

### SINGLE BUSINESS MUSINGS

#### Wanted resource

A useful simplification when talking about the digital library business is to think in terms of "the wanted resource". The diagram below shows that physical and digital objects share characteristics such as bibliographic level, whether something is still available from the original publisher or not, whether it is in or out of copyright, whether it is freely available or conditions apply and whether it is available now or in use for some purpose.



A collection manager's capacity to manage a wanted resource should not be limited by these characteristics. They will determine how the item will be described, stored and preserved. However the required specialisation only needs to occur at the point where separate workflows take over; for example, when an item is needs to be digitised or ingested into a digital repository or requires preservation action specific to its carrier and format. Similarly, a user's capacity to find and get wanted resources should not be limited by these characteristics. They will determine which type of delivery service will be used, whether the delivery service needs to support functions such as authentication, authorisation and payment and how the item will be delivered. However the required specialisation only needs to occur at the point where separate workflows take over; for example, when an item is lent, not copied, and therefore has to be returned.

#### Topic-based searching

In Libraries Australia, users can search or browse authority files to find the preferred forms of headings and navigate to linked resources. However, the workflow to do this is not user friendly. It is unlikely that more than a handful of users would even know that this functionality is there. Browsing authority files is more integral to the discovery workflow in the Library's catalogue but users still need to know what search options to select and the navigation to linked resources is cumbersome. The default keyword search does not exploit the reference structures in authority files.

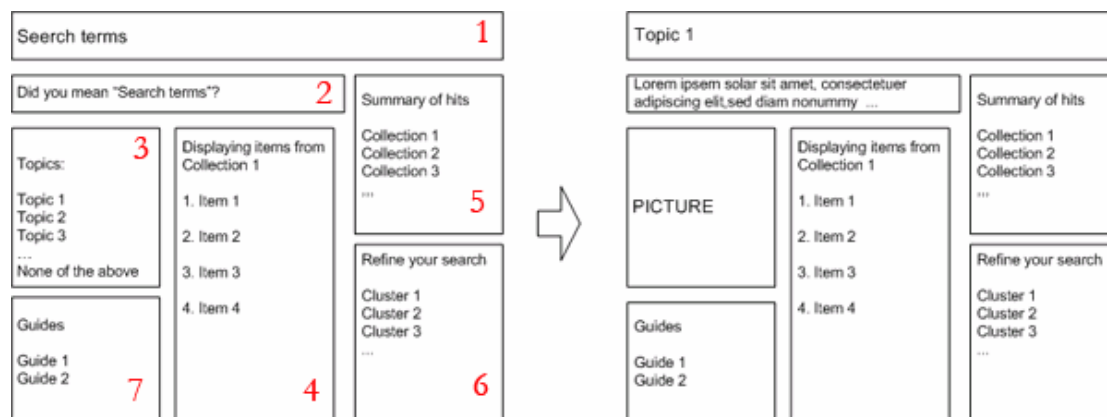
Earlier this year, the Library demonstrated the importance of relevance ranking to successful search outcomes and implemented a relevance ranked search result in Libraries Australia. A group is now investigating how to cluster search results in ways that will help users to refine their search. This work has raised a number of questions about the Library's trusted aggregations, how they would be presented to users in different business contexts and how the system could help users to find the right terms to search from a simple keyword search.

In the meantime, Music Australia has been implemented with a function that exploits the Australian Name Authority File to enable the discovery of people and organisations as well as resources. A separate parties database has been developed that contains records from the Australian Name Authority File and biographical entries contributed by partner agencies. Searches are conducted across both the resource and parties databases and presented in the form of a single results page, from which users can elect to go to detailed information about a party or a resource. This contextual approach to information is now being explored in a more comprehensive way through the People Australia Project.

A real danger now exists that the Library will provide four different ways of exploiting authority files in its services, each implementing a different paradigm, some less successful than others and none really solving the problem of enabling successful search outcomes regardless of where the user has started their search. The Library is currently in a very good position to change this outcome, with opportunities to leverage off scheduled projects such as People Australia and RAAM and to extend the thinking already done through Music Australia and the Catalogue Access Review Working Group. By getting library staff across all programs to think about discovery as a single business extraordinary benefits can be realised. As an early priority the Library needs to look at how the new search paradigms piloted in Music Australia and planned for Libraries Australia and People Australia can be extended to all search contexts, including the new NLA Catalogue.

A promising approach seems to be looking at search as a two-step process. The first step retrieves results by searching for both topics and resources. If a wanted resource is ranked high on this page the user can step through to full details about the resource and how to get access to it. Else the user can refine or extend their search by selecting a wanted topic.

The diagram below gives a very early idea of what the results page might look like and how it would differ from the topics page:



In this model, the intermediate results page is divided into seven zones:

- Zone 1 displays the user's search terms.
- Zone 2 provides a spell checking facility.
- Zone 3 ranks topics that meet the search criteria.
- Zone 4 ranks resources from the default collection that meet the search criteria.
- Zone 5 reports hits against all available collections.
- Zone 6 provides options for refining the search based on clusters of information in the retrieved resource descriptions.

- Zone 7 lists guides that might be relevant to the user's need if they had conducted a broad search such as family history.

The topics page has a similar structure but in the model zones 2 and 3 are no longer needed so this space can now be used for a scope statement with links to known entries on the topic and a picture illustrating the topic if one has been assigned. There are opportunities for the topics page to be further refined by what the system knows about the topic. For instance, if it is a person, users could have options to limit the results to resources by or about the person.

In this model topics are not limited to Australian parties as in Music Australia or People Australia. The full set of authority files are searched to provide the ranked list of topics. This search is also not limited just to the Library's authority files. It searches the full text of biographical entries harvested for Music Australia and People Australia and could also search other sources such as gazetteers and the Wikipedia. The existence of a heading in the authority file may be sufficient for a search to be successful. The page may meet the user's need solely through the resources retrieved by the search. The model also incorporates the idea of allowing users to say that none of the listed topics meet their needs and to create a new topic and add a scope note and picture as part of the process.

With careful design, this model could render irrelevant the concept of local authority files, or authority files limited to a specific logical view with implications for how authority control is performed, both within the Library and nationally.

## **User participation**

The idea of allowing users to say that none of the listed topics in a results page meet their needs and to create a new topic is just one of the ways in which users might collaborate with the Library in the development of content and its interpretation. A user might also want to resolve duplicate topics on result pages, merge resource descriptions, move a topic or resource higher in the result set ranking, take a resource out of an FRBR cluster, correct metadata or rate, annotate or tag a topic or resource. These are the kinds of things that users are expecting to do on the Internet. A page that does not allow some kind of interaction is now starting to look inert. New technologies such as Ajax are making it easy to change data displayed in a browser without having to reload the whole page.

The interesting thing about these scenarios is that they are not all instances of annotation. Some involve changes to topics and resource descriptions. This is an area where libraries have traditionally maintained strong control but there is an army of users who could become passionate advocates for improving the quality of content if this was facilitated through the interface. Cataloguers, reference librarians, project partners and subject experts all have skills that could be brought to bear.

There is a number of things that could be done in the architecture to position the Library to experiment with different layers of user collaboration. These include the implementation of authentication and authorisation services that support a range of different trust models, the capability to view previous versions of a record and to roll back to a previous state if a mistake has been made and the capability to treat some kinds of changes as suggestions.

Because of the federated nature of the Library's services, enhancements to metadata made through the Library's interfaces may need to be shared with contributors and vice versa. There are also other ways in which users might collaborate to enhance content. For instance, cataloguers or subject experts might work in their own online spaces, using standard protocols to update the federated data corpus. A general user creating a new topic might do so through the Wikipedia.

## **Matching and merging**

The Library's ability to take risks with user participation and to roll back changes might be enhanced by an alternative "soft" approach to matching and merging records. Instead of doing matching and merging on ingest, contributed records could be retained and combined for searching and display purposes. This would have to be done not only for searching and display but also for any type of product the Library exchanges with other partners or services.

The need to exercise caution with the current approach both to ensure a quality product and because it is not trivial to roll back changes means that high thresholds are set for matching and merging and considerable staff time is needed for review of near matches. The extent to which "crowd sourcing" might assist in ongoing quality improvement of the data corpus is something that the Library needs to explore and the soft approach might enable this. Another benefit of the soft approach is that the provenance of the data would not be lost and this could make the NBD more amenable to functioning as an OPAC.

Regardless of where matching and merging is undertaken it incurs a relatively similar load on resources and for some events such as global corrections on the database there could be some adverse implications for the capacity and rate of these changes. The successful outcomes of any matching or merging process is contingent on the robustness of the programs used for determining duplication, the consistency and quality of the data provided by contributors, and the interfaces that are provided to support human intervention in matching and merging.

The Library plans to move to an FRBR view of bibliographic resources to provide searchers with a more useful search result set and to display bibliographic resources in context. The same matching rules used to combine records at the FRBR manifestation level can be applied to create a synthetic MARC record for close copy cataloguing purposes if required. Under current conditions Australian cataloguers would still want accurate and discrete manifestation level records for copy cataloguing.

Any move to implement a soft matching process would have to take into consideration the storage and performance implications, how the quality of cataloguing products could be maintained, how global change would work and how legacy data would be handled.

## **Branding and marketing**

The single business approach supports the development of separate business contexts that can have their own home page and behaviours tailored to the kind of material being searched. Stylesheets can also be used to provide a different look if required. There will be constraints on the extent to which each of these "services" can use different paradigms for discovery and access but this will contribute to a consistent user experience.

As a general principle, each business context should enable users to extend their search in a seamless way to other kinds of material if required. Topic-based searching and reporting of search results in other collections will be a primary strategy for achieving this.

Below is an example of how the newspapers context might look using a single business approach:

Did you mean: William Charles Wentworth?

**Possibly related searches:**  
 For *Wentworth, William Charles, 1907-*  
 See [W. C. Wentworth \(William Charles\), 1907-2003](#)  
 For *Wentworth, William Charles, 1790-1872*  
 See [W. C. Wentworth \(William Charles\), 1790-1872](#)  
 For *Milton, Viscount (Charles William Wentworth Fitzwilliam), 1786-1857*  
 See [Charles William Wentworth Fitzwilliam, Fitzwilliam Earl, 1786-1857](#)  
 For *Norris, William Vernon Wentworth*  
 See [W. V. W. Norris \(William Vernon Wentworth\)](#)

**Guides:**

- Early explorers of NSW
- Freemen and convicts
- Liberal politicians

16 results in Newspapers

1. Sydney Election. [News]  
 THE MAITLAND MERCURY AND HUNTER RIVER GENERAL ADVERTISER, August 2, 1843  
 By (From our Correspondent.) - (From our Correspondent.) - Sydney, Monday Evening, To-day, at twelve o'clock the Mayor seared on the hustings, and declared that William Charles Wentworth and Robert Lowe, Esquires, were duly elected to represent the City of Syd- ney in the... 366 words Page: 2
2. Sydney News. [News]  
 THE MAITLAND MERCURY AND HUNTER RIVER GENERAL ADVERTISER, July 24, 1856  
 The National Banquet. (Abridg'd'es it's'ns'p'vs, .T'aj,45'1'. The publi' banquet in commemoration of the anagorision of i'Responsible Government took place on Thursday evening, at the 'Prince of Wales Theatre, Sydney, which was tastefully decorated for th'e co'sp'io'n. .The whole of the arrangements were... 1464 words Page: 4
3. Government Gazette. [OFFICIAL APPOINTMENTS AND NOTICES]  
 THE MAITLAND MERCURY AND HUNTER RIVER GENERAL ADVERTISER, August 16, 1848  
 FRIDAY, AUGUST 11, 1848. His Excellency the Governor has been pleased to appoint the undermentio'ne'd gentlemen to be magistrates of the territory and its dependencies, viz.:George Everett, Esq; of Wandsworth; Ch'as'es'el Mac'ronald Esq., of Falkner Plains, Charles William Marsh, Esq.,... 991 words Page: 1
4. Dinner To G R Nichols, Esq., M L C. [News]  
 THE MAITLAND MERCURY AND HUNTER RIVER GENERAL ADVERTISER, April 19, 1854  
 o'b ;ITio'v'p'v'n'ns'a publicdinner was given i' o'statements to Mr. (J. i. Nichols, me the L-g-islative Council for the Nor- \* unner took place at the Northumberland ti \* el, %%'s'P'ot M'i land, and wis'a very elegantly ; .vefully'v and not... 5197 words Page: 2
5. Crown Lands Beyond The Settled Districts. [News]  
 THE MAITLAND MERCURY AND HUNTER RIVER GENERAL ADVERTISER, August 13, 1851  
 Runs Obtained By Tender. (From a Supplement to the Government Gazette) I Crown Lands Office, Sydney, 24th July, 1851. It is hereby notified, for general information, that the teniders of the undermentioned parties having been accepted for the runs of crown lands, specified in... 1975 words Page: 2
6. Government Gazette. [News]  
 THE MAITLAND MERCURY AND HUNTER RIVER GENERAL ADVERTISER, October 26, 1844  
 Friday, October 18. At eleven o'clock on Thursday, the 21st day of November next, the Colonial Treasurer will put up to auction, at the Colonial Treasury, Sydney, the licenses to occupy, under the regulations of 21st August, 1841, the following portions of land,... 1444 words Page: 1

Results in other collections:

- People 18
- Bibliographic 1566  
 non fiction(250), biography (42), theses (4), juvenile (5), ...
- Pictures 211
- Music 32  
 recordings (5), scores (3)
- Maps 15
- Film 8

Refine your search:

- Newspapers more...**
- The Argus 2511
  - Maitland Mercury 1823
  - Melbourne Truth 1511
  - Sydney Gazette 1223

- Article Types more...**
- General News 4318
  - Editorials 1250
  - Letters 123
  - Sport 23

- Dates more...**
- 1810-1830 1265
  - 1831-1850 2180
  - 1851-1899 871
  - 1971-1990 3423

The Library's current applications are branded and marketed separately. Although it will be possible to do this for each business context, the Library will need to think hard about whether this is really meeting user needs. Branding and marketing the Library's discovery business as a single service will simplify business planning for the Library and choices for users, while still allowing the different business contexts to be promoted separately as products of the service.

The infrastructure will still support all three approaches - the development of separate applications, the deployment of a business context with separate branding and the deployment of a singly-branded business in different business contexts. The Library will need to look closely at the configurations needed to meet its directions, taking care that the applications and contexts it promotes are meaningful to users. In doing so, it will need to think of logical views of the data corpus and access to them through machine interfaces as being significant products in their own right and promote them accordingly.

If it is decided to retain an existing service as a separate application, it will still be important to migrate it to the service-oriented architecture to minimise support costs and reap the benefits of sharing technical services.

## Partnerships and other issues

Traditionally, the Library's business partners have been other libraries, booksellers and publishers and metadata aggregators and document providers. In recent years the Library has increasingly engaged with other collecting agencies and also with search engines, as a means of making items in the national collection discoverable, and with online bookshops as a means of making them accessible. In an increasingly digital world the Library may find new partners in unexpected places. Certainly, its data may be used in ways it had not anticipated. The Wikipedia and Google Scholar may emerge as two important new partners with whom the Library may need to establish a formal business relationship: the Wikipedia because of its

importance as a source of information in topic-based searching, and Google Scholar because of the need to find an alternative to metasearching for discovering journal articles.

The Wikipedia is constantly growing and developing its business and content model. In a recent development it is enabling Wikipedians to input biographical details in a structured form in entries about people. The Deutsche Bibliothek<sup>18</sup> has exploited this feature to provide links to records in its catalogue from German Wikipedia entries. This is just one of the ways in which the Library might take responsibility for an Australian subset of the Wikipedia in order to deploy it effectively in its own topic-based searching and to enable access to items in Australian collections from Google and Wikipedia searches.

The Journals Australia project will help the Library to enhance its data corpus with metadata and full text indexes for Australian journals but it will not address the issue of enabling access to the content of overseas journals held in Australian collections either physically or by being made accessible through subscription services. The higher education sector has been trying to address this issue through metasearch portals that allow users to search across large numbers of databases but it is becoming increasingly clear that this model is not really working. The Library's current strategy for providing access to journal articles has been to describe aggregations at collection level through its *E-Resources* service<sup>19</sup>. Even here there are issues with enabling access to collections the Library's users are authorised to access from other points of need such as the catalogue, Libraries Australia or Google Scholar.

At the moment Google Scholar's coverage is not exhaustive and a search of a general subscription database like Academic Search Premier and Google scholar for the same topic will find very different results. On the other hand, Google Scholar provides free access to the content it does index and it has a service enabling users to link through to library holdings. It may be that collection-level access through an *E-Resources* business context together with the use of Google Scholar as an extension to the Library's data corpus for journal articles may do enough to meet this need, particularly if the Library works actively to help make Google Scholar more comprehensive in its coverage.

In the Library's own discovery service, it would probably be sufficient to know the user's affiliated library through cookies to pass the URL of a licensed resource to the affiliated library's OpenURL Resolver Service. The user would then either get direct access to the resource or be prompted for local authentication if a higher level of authorisation was required. This would require the Library to encourage libraries to set up a resolver service. Another service the Library might offer libraries is to package up their subscription holdings in the format required by Google Scholar on their behalf. This would be a good way of ensuring that there was current volume and issue data in NBD holdings statements. A task to derive E-Resources from the catalogue has been on the Library's books for a number of years. One could imagine providing library-specific views of E-Resources as one of the products enabled by encouraging libraries to record their subscription holdings on the NBD.

---

<sup>18</sup> Jakob Voss, "Metadata with Personendaten and beyond", *Transwiki:Wikimania05/Paper-JV2* (<http://meta.wikimedia.org/wiki/Transwiki:Wikimania05/Paper-JV2>).

<sup>19</sup> <http://www.nla.gov.au/apps/eresources/action/front>.