



Reaching new audiences: the People Australia and Picture Australia projects at the National Library of Australia

Pam Gatenby

Assistant Director General, Collections Management,
National Library of Australia

This presentation draws extensively in parts on papers prepared on the topic¹ by Judith Pearce, and on information provided by Basil Dewhurst, Manager People Australia Project (bdewhurst@nla.gov.au).

Meeting: **147 - IFLA-CDNL Alliance for Bibliographic Standards (ICABS)**

Simultaneous Interpretation: **Yes**

WORLD LIBRARY AND INFORMATION CONGRESS: 73RD IFLA GENERAL CONFERENCE AND COUNCIL
19-23 August 2007, Durban, South Africa
<http://www.ifla.org/iv/ifla73/index.htm>

ICABS session on *Libraries and Web 2.0*. IFLA, Durban, 19-23 August 2007.

Outline of paper

In this presentation I will outline the strategic context in which the National Library of Australia plans its collecting and access strategies, and describe ways in which we are incorporating Web 2.0 features into our web resource discovery services as exemplified by a new project called People Australia, and another project that involves collaboration with flickr to bring contemporary images into the PictureAustralia service.

Background

The National Library of Australia has a legislated mandate to build and maintain a national collection of library material, to make that collection available for use, and to provide services that support the Australian library community.

In carrying out these responsibilities, we collaborate with other collecting institutions to develop national research collections in a range of formats that document the history and culture of Australia, we catalogue our collections for access via online discovery services, digitise selected materials, provide onsite and online reference

¹ Judith Pearce, "User collaboration in websites", National Library of Australia Staff Papers, 2006 (<http://www.nla.gov.au/nla/staffpaper/2006/jpearce1.html>)

Judith Pearce, "The people Portal Project", National Library of Australia Staff Papers, 2006 (<http://www.nla.gov.au/nla/staffpaper/2006/pearce2.html>)

services, and have appropriate preservation programs in place to ensure ongoing access to the collections.

We also provide online services that support access to the distributed national collection and underpin resource sharing. Since 1981, the National Library has provided an online national union catalogue, now known as the Australian National Bibliographic Database (ANBD), which is an integral and important component of the national information infrastructure. The ANBD has been built up over many years through dedicated cooperation by Australian libraries and it is regarded as a national asset by the library and research sectors. It is made available for searching free of charge, through the service known as Libraries Australia, and provides access to around 42 million items held by 800 Australian libraries.

The National Library has also developed a number of collaborative national resource discovery services which provide access to subject or format based collections. These are derived from and extend the Australian National Bibliographic Database (ANBD). They include:

- Picture Australia (<http://www.pictureaustralia.org/index.html>)
- Music Australia (<http://www.musicaustralia.org/>)
- PANDORA: Australia's web archive (<http://pandora.nla.gov.au/index.html>)
- the Register of Australian Archives and Manuscripts (<http://www.nla.gov.au/raam/>)
- Australia Dancing (<http://www.australiadancing.org/>)

Strategic context

The current strategic directions and desired outcomes that the Library is pursuing across all areas of activity are outlined in our *Directions for 2006-2008* statement (<http://www.nla.gov.au/library/directions.html>).

Our overarching objective for the period is “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”.

The *Directions* statement identifies five desired outcomes. These relate to our collecting, access, cultural and leadership roles with the fifth desired outcome addressing the need to remain relevant in the rapidly changing, digital information world in which we now exist.

To ensure our ongoing relevance and to provide collections and services that users want, we have in place a number of strategies. For instance, we will:

- adopt new models for collecting materials in digital form;
- provide our users with rapid and easy access to our collections and other resources;
- engage with new audiences, in new online spaces in order to enhance our visibility and encourage new uses of our collections; and
- provide online services that support the creation and sharing of knowledge and ideas by our users.

Web 2.0 services

These strategies are aimed at keeping up with user expectations that have been shaped by the Internet, by the ease of discovery offered by search engines and now, by services that embody the Web 2.0 concept – that is, websites that let people collaborate and share information in previously unavailable ways. Popular examples of websites that offer user participation are eBay, Wikipedia, Amazon and Google.

These types of services have certain characteristics in common.

- They are based on large data repositories, often with some of the data existing previously in a form that was not useful or easily accessible.
- The data can be combined with data from other services for new and unexpected purposes; for example, Google Maps can be embedded in your own web pages.
- User participation is encouraged through the contribution of data and through tagging, comments, ratings and reviews, for instance. A community often forms around user participation services that operates on a trust model.

The pace of change seen with services that encapsulate Web 2.0 principles, such as flickr and Wikipedia, has significant implications for libraries as users now expect a degree of inclusiveness and interactivity from web sites. In shaping our strategic directions to respond to these expectations and to ensure our ongoing relevance, we have looked to popular Web 2.0 services to provide models for developing our own resource discovery services.

In this paper I'll outline the approach we are taking with two services, People Australia and PictureAustralia.

Using biographical information to support discovery of library resources

As with all libraries, one of the greatest challenges we are now facing is finding new, effective ways of making sure people can discover our library resources and then use them. One approach we have experimented with as an alternative way of making Australian collections visible through the main Google search engine, involves seeding it with biographical entries that allow users to navigate to linked resources. This contextual approach is demonstrated by our Australia Dancing and Music Australia services.

Australia Dancing is a federated directory of resources held in Australian collections related to dance in Australia. Its main purpose is to guide users to resources but the primary objects for discovery are people, companies and performances. It is a stand-alone service that can be searched and browsed, but users can discover entries through Internet search engines. The entries usually appear high up in search engine result sets so the service attracts a high number of users.

Similarly, Music Australia is a federated resource discovery service specialising in Australian music. It consists of two linked databases – one of resource descriptions, and the other of information about people and organisations (referred to as “party” information) sourced from a number of contributors. This information is encoded

using a schema developed especially for the service – the Metadata for Australian Parties Schema (MAPS). These entries establish a web presence for a “party” and enable users to discover related resources.

The need for a people portal service

A number of services now exist or are being developed which provide online access to Australian biographical information with notable examples being the Australian Dictionary of Biography Online, the Dictionary of Australian Artists Online, and the two National Library services mentioned above. There are also valuable sources of information hidden in name authority files, finding aids and resource descriptions created by a range of institutions and projects around the country.

The National Library considered that a portal service providing one-stop shop access to all this biographical information, with links to related resources, would be a valuable national service to provide. It would make it easier for researchers and others to identify data dealing with the same person or organisation and to access and share the information, and it would address the issue of duplicative creation of information about people that is currently occurring through different projects. In addition, as demonstrated by the Australia Dancing and Music Australia services, it would provide an effective resource discovery approach to the resources associated with the biographical entries, many of them held by Australian libraries.

In deciding to develop a people portal service, we were aware of the potential of the Australian Name Authority File to be repurposed and extended to support it. The Australian Name Authority File is a database of 120,000 personal names and 90,000 corporate names used in the Libraries Australia service to collocate resource descriptions under a single form of name, and to provide alternative forms of name and cross references for searching purposes. As such, it is an authoritative means of identifying Australian people and organisations that has been collaboratively developed and maintained by Australian libraries in online form since 1981.

People Australia

The People Australia project commenced in mid-2006 and is being managed by Basil Dewhurst (bdewhurst@nla.gov.au). A detailed project proposal is available at <http://www.nla.gov.au/initiatives/peopleaustralia/PeopleAustraliaFeasibilityStudyWebsite.doc>.

The aim of the People Australia project is to develop a service that enables users to find information about Australian people and organisations and get associated resources, such as works by and about the person. Information about a person could be provided for instance, by linking to photographs, newspaper or journal articles, oral history interviews and reference sources.

The service will be based on entries from the Australian Name Authority File (ANAF), which will be clustered with other biographical entries created by partner agencies in their domains and it will provide links to related resources. Interfaces to partners’ content will be established and the People Australia “clusters” will be made accessible to search engines.

Incoming name records will be matched against existing records using information from the full text of entries and from linked resources. Matching entries will be collocated rather than merged and each contributed record will be assigned a local identifier and a national identifier will be assigned to the collocated cluster.

Results will be presented in the form of a directory entry, with links to additional information about the person that might be available from partner agencies, and to resources by and about the person listed in the Australian National Bibliographic Database and in other target search services.

For instance, a search on Mary Gilmore, a famous Australian author, poet and journalist who died in 1962, will result on in entry that provides:

- a brief biographical note;
- a list of other forms of her name that have been used;
- links to sources of biographical information such as the Australian Dictionary of Biography Online;
- a list of associated resources found in linked services such as Libraries Australia, Picture Australia and Newspapers Australia with options for viewing the list by format and type of resource; and
- a link to getting options for those resources that are not available online.

Partners

The People Australia service will develop relationships with partner repositories from a number of sectors so that the service is able to provide users with access to and disseminate information about the widest range of Australian people. A large number of potential partners have been identified but the first phase of the project will involve working with the University of Melbourne which manages archives relating to women's history, trade unions, and scientists; the Dictionary of Australian Artists Online managed by the University of New South Wales; the Australian National University which manages the Australian Dictionary of Biography Online and the State Library of Victoria, who will contribute their Artist Authority File.

Partners will be significant users of the service as well as contributors of data. People Australia will provide a hub for interlinking between different services and enable developers of biographical services to share the effort of creating and resolving entries. It will also be a source of data for other aggregators.

Target Audiences and uses of People Australia

The People Australia service is being developed to appeal to a wide audience and to meet the needs of anyone wanting information about Australian people or organisations. However, some groups should find the service especially useful, for example;

- researchers,
- teachers and students,
- reference librarians,
- government departments,
- family historians,
- hobbyists, special interest and community groups, and

- journalists.

Users will range from those interested in detailed research to those interested in specific information. They might use the website to:

- find information about a specific person or organisation;
- discover relationships between people and organisations;
- discover or locate people or organisations who share common characteristics, e.g., gender, role, cultural heritage, field or time of activity (for example, find Australian writers born in the 19th century as opposed to Australian composers or scientists);
- navigate to resource descriptions provided by a range of discovery services;
- navigate to more specific portals such as MusicAustralia for more detailed information (for example, works composed by, or performed by, or conducted by, a person); and
- obtain new information and resources relating to any person or organisation through an alerting service.

For some users (for example, primary or secondary school students) a biographical or historical note may be sufficient to satisfy their needs and they may not navigate to associated resources. Others may expect to find a list of resources by or about the person.

Some, such as genealogists and journalists, for whom the person or organisation is the major focus, may make People Australia their first choice for discovering resources. Other users may discover People Australia entries through an Internet search. Others may start by accessing a subject-specific service (for example, Music Australia) and then exploit links from that service to People Australia to find more information about a person and their activities in other domains.

Typical usage scenarios include the following.

 <p>Scenario 1</p> <p>A journalist is writing a retrospective piece on Bob Hawke for the weekend magazine of a national newspaper.</p>	<p>The journalist wants to know:</p> <ul style="list-style-type: none"> • When he joined the Labor party • When he became president of the ACTU • The dates of his term of office as Prime Minister • His major achievements. <p>She also wants to find:</p> <ul style="list-style-type: none"> • Online images of Hawke, and whether the image can be re-published • A list of recent biographies of Hawke • Where his Prime Ministerial archive is held • Whether he has a personal website • Whether there are other online services which provide more information.
--	--

<p>Scenario 2</p>  <p>A teacher in a remote Indigenous community is leading his students through an assignment on the challenges and achievements of Indigenous sporting heroes.</p>	<p>The teacher wants his students to discover:</p> <ul style="list-style-type: none"> • A list of Indigenous sports stars • What Indigenous group they belong to • Whether they grew up in remote areas • Outlines of their sporting careers • Major awards and honours they have received <p>And to access online:</p> <ul style="list-style-type: none"> • Pictures and news stories about their heroes.
<p>Scenario 3</p>  <p>A nationally syndicated radio program mentions an upcoming Art Gallery of New South Wales exhibition on the artist Margaret Preston, resulting in heightened interest in Preston's work.</p>	<p>A UK tourist who will be in Sydney during the exhibition wants to:</p> <ul style="list-style-type: none"> • Read about Margaret Preston's life • See some of her work online and order prints • Find the most recent publication on her work and order it through Amazon.
<p>Scenario 4</p>  <p>The Coleambally local history group wants information on significant people born in the district before the town was officially opened in 1968 so that they can construct a float for their local historical pageant.</p>	<p>The Coleambally local history group wants to find a list of:</p> <ul style="list-style-type: none"> • People born in the district before the town was officially opened in 1968 • What areas of life they were significant in – politics, farming, music • Any major honours and awards; including military achievements • Brief descriptions of their lives • Online images so that they can design costumes for pageant actors.

User participation in People Australia

In contrast to maintaining traditional library name authority files, there will be no attempt to resolve conflicts between parties with the same name. Instead, Web 2.0 services provide models for harnessing the collective endeavour of online communities that could be adopted to achieve this. A range of users, from cataloguers to subject experts to the general public, may all have knowledge that can contribute to the development of the service and to clarifying ambiguities created by people with the same or similar names. Easy to use interfaces and support for user contribution

will be a feature of the service in order to encourage user participation in its development.

The service will also create web services that support the sharing of data in ways that encourage the development of new services, minimise duplication of effort and preserve the intellectual property and context of entries developed by partners.

Standards based approach

The People Australia service will be based on open standards that support interoperability and sharing of data. An OAI-PMH (Open Archives Initiative for Metadata Harvesting) Harvester that will support the harvesting of information from repositories that support OAI-PMH will be implemented and it will also allow “party” data to be made available to other OAI Harvesters. A z39.50/SRU/SRW interface will allow the querying of partner repositories using these search protocols and will accept compliant queries and provide responses. The repository entries will also be made available to search engines for indexing.

People Australia will link to a number of internal and external services to support resource discovery and delivery of associated resources. For instance, it will link to Libraries Australia to gather updates to names in the Australian Name Authority File, to resource discovery target services available through the Libraries Australia Z39.50 interface, to the InterLibrary Resource Sharing Directory and to document delivery systems to enable users to place requests for copies of resources they want. Interfaces will also be supported with partner systems to harvest data from them for the People Australia repository and to search their repositories as target services.

Timeline for development of People Australia

There are four project stages to development of People Australia, as follows - analysis and design, scheduled for completion in July 2007; development, scheduled for completion in the first quarter of 2008; soft launch; and production service launch. The final stages should be completed by June 2008, when the service will be available to the public. Additional partners will be brought on between the soft launch and the launch of the production service.

Picture Australia and flickr

Another project that the National Library of Australia has in place that demonstrates Web 2.0 characteristics, enables people to contribute images to our PictureAustralia service (<http://www.pictureaustralia.org/index.html>) through flickr, the web based photo sharing service now owned by Yahoo! (<http://www.flickr.com/>). This project advances our strategic aims of exploring new models for collecting resources and of providing services that support the creation and dissemination of knowledge by users.

PictureAustralia is a federated service that lets users search across a number of online pictorial collections at the same time. It began in 1998 as a pilot and now provides access to over one million pictures provided by 44 participants. Until the collaboration with flickr commenced, the service provided access primarily to historical material. Following a user survey in 2004 that found that 43% of users were

looking for contemporary images, the Library decided to work with the flickr service to achieve this. This would also enable us to engage with new user groups, another of our strategic objectives.

flickr was chosen as a suitable partner because it already had a large following of Australians, because it encouraged the use of metadata to aid discovery and because it had a programming interface that could easily be mapped to Dublin Core, the metadata schema used to contribute records to PictureAustralia.

How the collaboration works is quite simple. Two PictureAustralia groups were set up on flickr in the first instance (<http://www.flickr.com/people/92276616%40N00/>), one to collect images relating to Australia Day, the other a more general group based on people, places and events. (The launch of the pilot was timed to coincide with Australia Day (26 January) 2006.)

Users upload their images to flickr and add them to one of the PictureAustralia groups. Once a week the Library harvests the metadata and thumbnails and adds them to PictureAustralia. Since the pilot was launched over 18,000 images have been contributed to PictureAustralia in this way. In PictureAustralia the images show up in result sets side by side with images from other collections. A low level of moderation is carried out by staff who check for inappropriate images before they are brought into PictureAustralia.

The flickr images go into a kind of half-way house where they remain accessible through PictureAustralia for the life of the image itself on flickr, without any specific agency needing to take them into its collection. This is enabled by the PictureAustralia architecture, where only the metadata and thumbnails are harvested. Users link to flickr to see the full image. Significant images are, however, being selected for the National Library of Australia's collection. The creators of these images are contacted to obtain a high resolution version of the image, which is then taken into the Library's digital archive, to safeguard for future access. A high resolution version is needed because of the bandwidth limitations of uploading images to flickr.

Unsurprisingly, metadata quality is an issue, not so much because contributors are not using a standard thesaurus (PictureAustralia members already use a mixture of thesauri) as because sometimes they are not giving images a sufficiently descriptive title or assigning enough tags. Guidelines on tagging are provided on the Picture Australia flickr site and Picture Australia administrators monitor the images and email contributors with advice as required. This has improved the quality of the metadata. As a sense of community develops it is expected that there is likely to be a level of self-education and management through the discussion forums associated with the PictureAustralia groups.

The outcomes from the flickr project have so far been very positive. There is no doubt that the pilot project has increased PictureAustralia's coverage of topical issues. It has done this in a very Web 2.0 fashion by harnessing the collective efforts of people already capturing events as they happen for their own purposes. Some of the contributors are now also actively taking pictures with contribution to *PictureAustralia* in mind. This has led to the introduction of many more full-colour

images and also to some interesting juxtapositions of old and new pictures with the same subject. Groups have already formed on flickr independently of this project, to collect Australian images around particular themes and it has been possible to invite photographers contributing to these groups to contribute their pictures to Picture Australia as well. There has been little need to moderate images, although some have been removed as out of scope, for example, pictures of pets.

Besides contemporary images, there has also been an unexpected increase in heritage images and this is a rich area for further mining. The thing that most distinguishes PictureAustralia from Google Image and flickr, is its historical material. The flickr halfway house approach will enable the Library to invite the uploading of more historical images to flickr and their contribution to PictureAustralia than it has been able to take into the print collection in the past. For instance, strict collecting policies have been applied until now to family history materials because of the costs of bringing these into the collection.

The visibility of PictureAustralia has certainly been increased by the pilot, as the new contributors search for their own images and become fascinated with the juxtapositions that they find. In a survey recently conducted of participants in the project, 39% had not heard of PictureAustralia before, 64% heard about the project through flickr and other flickr members and 69% went to Picture Australia to see their images. There has been a significant sustained increase in use of PictureAustralia as a result.

Future directions

To sustain the interest shown by Australians to participate in the pilot project, new *PictureAustralia* groups that focus on targeted subject areas will be identified and the Library will continue to provide guidance on the content that is being sought and how it should be tagged. However, we will need to make PictureAustralia easier to search and find relevant images as the repository is already very large and will increase significantly in size as the flickr collaboration proceeds.

We are looking at ways of enhancing access to the content through ranking and clustering of result sets based on the cataloguing and tagging data provided with images and the geocoding of images. However, in an image-based service, getting this right depends on the quality and consistency of the metadata. Discovery fails when an image is not appropriately or sufficiently described; and it is here that users are well-positioned to add tags and comments that might reveal previously unknown information about the image such as the location or name of a subject. To address the growing user expectation of participation in web services, the Library will later this year start experimenting with collecting annotations of all kinds; for instance, tags comments, notes, ratings and reviews. We will also consider how to manage this data and how to share it to support ranking and clustering of result sets or the browsing of images deemed most interesting or relevant, both at the local and aggregated level.