

The People Portal project

A paper presented by Judith Pearce at the [Biography and Technology](#) Conference, Australian National University
12 – 14 September 2006.

Judith Pearce

1.1 Keywords

Resource discovery.

1.2 Abstract

Users will increasingly expect to find the information they need through Internet search engines. Seeding these with library holdings has proved problematic. A contextual approach to resource discovery is proving to be a more satisfactory solution. The Library has experimented with this approach through Australia Dancing and Music Australia and is now embarking on a project to build a People Australia service that will act as a hub to resolve identities and link users to related resources.

1.3 Introduction

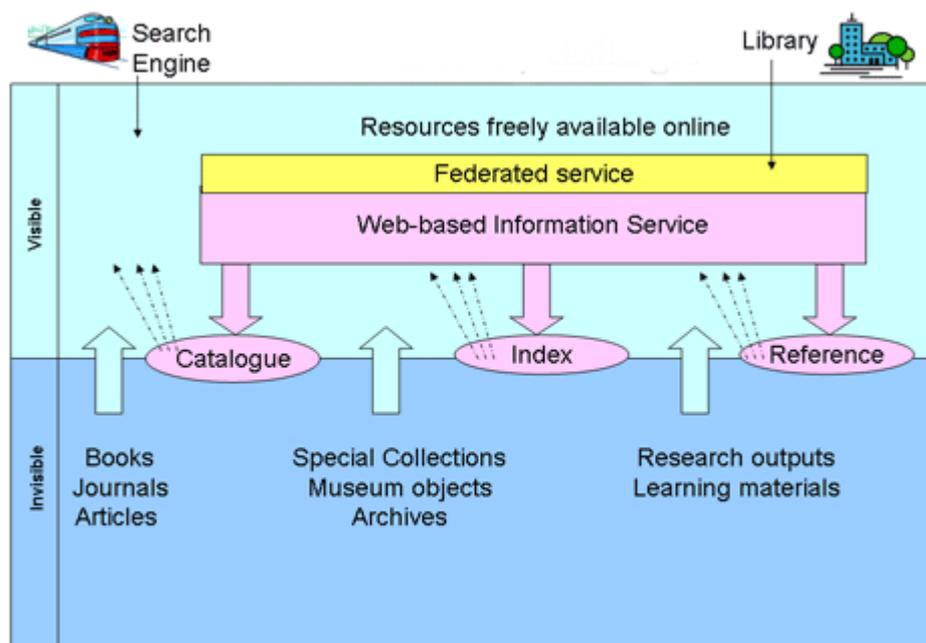
The role of the National Library of Australia is to enable Australians to find the information that they need to understand themselves and their place in the world. To do this the Library wears two hats. It collects, preserves and makes accessible through its catalogue a comprehensive collection of Australian publications and original resource materials; and it also operates as an aggregator, building federated services based on data contributed by libraries and other agencies. [Libraries Australia](#) is one such federated service providing access through the Australian National Bibliographic Database to over 16 million resource descriptions and 40 million items held in Australian library collections.

In this presentation I will talk about some of the resource discovery challenges the National Library of Australia is facing in the Internet age and how biography is being used by the Library to provide a contextual approach to information in services such as Australia Dancing and Music Australia. I'll talk about our plans to generalise this concept in the

form of a People Australia service. Lastly I'll say a little about ways in which users might contribute to the development of such a service.

1.4 Resource discovery in the Internet age

The Internet can be seen as providing access to two categories of resources: those freely accessible online and those that are not. The second category includes books, journals, articles, special collections, museum objects, archives, research outputs and learning materials. Some of these resources are not available online because they only exist in physical format, others because access is restricted.



Internet search engines index the first category of resource but the second category is invisible to them.

Libraries enable Internet access to this material by compiling catalogues and indexes, directories and guides and making these available through web-based information services and through federated services such as Libraries Australia. These services are indexed by search engines but not the databases to which they provide access. Gradually, however, more and more of what we had previously considered to be part of the deep web is becoming available in digital form and thus visible to search engines. Users are expecting that the information they need can be found by 'googling' for quite good reasons.

Information service providers are starting to address the visibility gap by seeding search engines with the content of their catalogues, indexes and guides. It is not really scalable for every library to expose their catalogue records to search engines. Union catalogue providers are better positioned to do this on behalf of libraries because they have already aggregated their holdings. Recognising this, the National Library has been looking at some

time for ways of exposing the 16 million records in the Australian National Bibliographic Database to search engines

This has only been partially successful. Google, for example, has balked at putting large numbers of union catalogue records into its main search engine. Instead it is experimenting with a Union Catalog program based around [Google Scholar](#) and [Google Books](#). The National Library has contributed records to this program and users from the Australian domain may now link to Australian holdings for books discovered in a Google Scholar or Google Books search.

In parallel with these activities, the Library has been looking at alternative ways of making Australian collections visible through the main Google search engine. In particular, it has been experimenting with a contextual approach to resource discovery through two of its services – Australia Dancing and Music Australia - by seeding search engines with biographical entries that allow users to navigate to linked resources.

1.5 Experimenting with the contextual approach

[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 also discover entries through Internet search engines. Australia Dancing entries usually appear high up in search engine result sets. Because of this, for such a specialist service, it attracts a high number of users.

Australia Dancing was not really conceived to be a dictionary of biography. It was designed on the assumption that:

- The biographical information in the entry may itself meet the user's need.
- An entry can be published without linked resources
- The biographical information does not have to be complete to enable the discovery of resources.

It was thought that there would be benefits in publishing an entry as soon as it contained sufficient information to be useful rather than waiting until resources were available to create a full biography or until the subject was sufficiently established to have a body of associated research material. In practice, the business owners of Australia Dancing prefer not to publish entries until they have a complete set of biographical details. This works for a hand-crafted service like Australia Dancing but not for the much bigger Music Australia service where the Library has now scaled up its experiments with the contextual approach.

[Music Australia](#) is a federated resource discovery service specialising in Australian music. It consists of two linked databases: a database of resource descriptions derived from the Australian National Bibliographic Database and a database consisting of information about people and organisations sourced from a number of contributors including the Australian Name Authority File. This information is encoded using a schema developed especially for the service – the Metadata for Australian Parties Schema ([MAPS](#)). The term 'parties' is

used to cover both people and organisations. Full entries for parties co-exist in Music Australia with entries derived solely from the Australian Name Authority File. These entries do not contain much information, but they are still a means of establishing a web presence for the party and enabling users to discover related resources.

The Australian Name Authority File is a database of 120,000 personal names and 90,000 corporate names used by Libraries Australia to collocate resource descriptions under a single form of name and to provide alternative forms of names and cross references for searching purposes. As such it is an authoritative means of identifying Australian people and organisations that has been collaboratively maintained by Australian libraries in an online form since 1981.

1.6 The need for a broader portal service

A number of services now exist or are under development which provide access to biographical information, some of which also link to related resources. Prominent examples include the [Australian Dictionary of National Biography](#) and the [Dictionary of Australian Artists Online](#). There are also valuable sources of information about people and organisations hidden in name authority files, finding aids and resource descriptions. For some time staff at the National Library have been toying with the idea of repurposing and extending the Australian Name Authority File to provide a portal to scholarly biographical information from a range of sources like these and to support authoritative links to related resources.

There many enablers for doing this. Music Australia has provided a testbed for exploring many of the architectural and business issues. A range of national and international projects are exploring similar issues. There is a growing body of biographical entries in machine-readable form. The Music Australia testbed has shown that content is not needed to give a party a web presence and enable users to find related resources. A significant investment has already been made in the Australian Name Authority File, which is well-positioned to be extended to provide an identity resolution service for Australian parties based on persistent identifiers. The idea of applying persistent identifiers to parties has been taken up by the Publishing Industry, which is promoting the development of an International Standard Party Identifier ([ISPI](#)).

A project has now been initiated by the Library with the following objectives:

- To implement a new People Australia service based on the Australian Name Authority File that clusters biographical entries created by partner agencies, with links to related resources.
- To identify partner agencies and set up appropriate interfaces to their content.
- To make People Australia “clusters” accessible through search engines.

The service will be based on a database of people and organisations with records contributed from a variety of sources including the Australian Name Authority File. Incoming 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, with local identifiers for each contributed record and a national identifier for the collocated cluster.

As part of the project the use of the MAPS Schema will be reviewed to make sure that the Library is taking a standards-based approach. Existing schemas and models are still struggling conceptually with the functional requirements for authority records, seeing their primary use in relation to discovery of resources. More work is needed to develop an international standard that supports the discovery of parties based on shared characteristics such as gender or field of activity, independent of roles that party might play in relation to the creation or subject of resources.

It is anticipated that a range of people will use People Australia, including general information seekers, researchers, teachers and students, reference librarians, government departments, family historians, hobbyists and special interest groups, community groups and journalists. Some will use the service directly, others will find entries through search engines, others will access the information it contains through third-party services that use standard protocols to interact directly with the repository or to harvest data for re-use.

It is important to note that partner agencies will be significant users of the service as well as contributors of data. It will provide a hub for interlinking between services. It will enable developers of biographical services to share the effort of creating entries, it will act as a means of resolving identities and it will be a source of data for other aggregators. Like many agencies, the Library tends to think in terms of its own digital spaces when describing a proposed new service and its users, but the system interfaces to the People Australia repository and to the Library's other aggregations will be of equal importance in ensuring the success of the project.

Work has not begun yet on the detailed analysis but here is a very early mockup of how an entry might look in the People Australia user interface:

National Library of Australia

PeopleAustralia
The ones we have records for

Login | Logout | About | Contact | Help

Search: Go

Home | Advanced Search | Browse | History | Saved Records | Saved Queries | Alerts

[Gilmore, Marshall](#) [Gilmore, Meredith](#)

Gilmore, Dame Mary Jean (1865-1962)

Resources:
Found [519 resources](#)

Limit to:
[Books](#)
[Journals](#)
[Newspapers](#)
[Manuscripts](#)
[Music](#)
[Oral history](#)
[Pictures](#)
[Theses](#)

Results by year:

2000	<input type="checkbox"/>
1990	<input type="checkbox"/>
1980	<input type="checkbox"/>
1970	<input type="checkbox"/>
1960	<input type="checkbox"/>
1950	<input type="checkbox"/>
1940	<input type="checkbox"/>
1930	<input type="checkbox"/>
1920	<input type="checkbox"/>
1910	<input type="checkbox"/>
1900	<input type="checkbox"/>

Other names:
Gilmore, Dame Mary
Gilmore, Dame Mary Cameron
Gilmore, Mary Cameron
Gilmore, Mary Jean

Born:
16 August 1865,
Woodhouselee, NSW,
Australia

Died:
3 December 1962, NSW,
Australia

Fields of Activity:
[autobiographer/memoirist](#)
[columnist](#)
[contemporary-affairs](#)
[commentator](#)
[essayist](#)
[journalist](#)
[pacifist](#)
[poet](#)
[schoolteacher](#)
[socialist](#)
[utopian](#)

GILMORE, Dame MARY JEAN (1865-1962), writer, was born on 16 August 1865 at Mary Vale, Woodhouselee, near Goulburn, New South Wales, eldest child of Donald Cameron, a farmer, born in Inverness-shire, Scotland, and his native-born wife Mary Ann, née Beattie.

[More from the Australian Dictionary of Biography Online...](#)

Other biographical entries:
[Australian Trade Union Archives](#)
[Austlit, Australian Literature Gateway](#)
[Encyclopedia of Aboriginal History](#)

Resources

By About Online only Prefer my libraries

1. Papers of Mary Gilmore [manuscript]. Gilmore, Mary Dame, 1865-1962. 1883-1962. 16 boxes. held by 1 library [ONLINE]
2. Papers of Dame Mary Gilmore [manuscript]. Gilmore, Mary Dame, 1865- 1962. 1948. 1 cm (1 folder). held by 1 library
3. Mary Gilmore / selection and introduction by Robert D. Fitzgerald. Gilmore, Mary Dame, 1865-1962. [Sydney]: Angus and Robertson, [1963] 165 p. held by 39 libraries, including ANU
4. Mary Gilmore : a tribute / by Dymphna Cusack, T. Inglis Moore and Barrie Owendon. Cusack, Dymphna, 1902-1981. Sydney : Australasian Book Society, 1965. 223p. : ill. held by 55 libraries, including ANU

This mockup was developed for the feasibility study conducted in 2004-2005. This version reflects new thinking about the need for 'trusted aggregations' to support linking to related resources – the Australian National Bibliographic Database for books, MusicAustralia for music resources, PictureAustralia for pictures and future aggregations for links to journal and newspaper articles. Work that the Library is currently doing on relevance ranking and clustering result sets will further affect the design of this page.

A primary function of the People Australia service will be to identify and collocate information about the same party and to disambiguate parties sharing the same name. This has always been a challenge for national authority files. In another paper on [user collaboration in websites](#) I talk about how Web 2.0 services are now providing new models for harnessing the collective endeavour of online communities to achieve this purpose. Cataloguers, reference librarians, curators, subject experts, the general user and parties themselves may all have knowledge that can contribute to the development of the service and be highly motivated to do so, given appropriate easy-to-use interfaces and trust models.

1.7 Conclusion

Much work and thinking still needs to be done to develop the People Portal concept. Music Australia has shown us that the technology to do this will be relatively simple. The challenge will be to develop a sustainable business model for the service based on extensive consultation with potential partners.