

AARLIN: An Australian Approach To Managing E-Collection Access

Earle Gow

La Trobe University

Victoria 3086 Australia

Email: e.gow@latrobe.edu.au

Abstract

The Australian Academic Research Library Network (AARLIN) aims to provide seamless access to information resources of Australian University Libraries from the desktops of research staff and students. Over A\$3 million has been granted by government to support the project. A pilot project in 2002 indicated that collaborative use of portal software was a viable component of the system. The Ex Libris portal and SFX software is being installed for use in 21 universities during 2003 and 2004. The pilot project and implementation process suggest that a national approach to information access would be welcomed by University staff and students. The AARLIN project is also providing valuable experience in the development and implementation of collaborative approaches to information resources access and raising issues of technology, management, governance and financing. This paper examines these issues both from the perspective of the overall project and from the perspective of a participating institution and attempts to draw some conclusions about the future viability of the AARLIN approach

Keywords: AARLIN, Australian Academic and Research Library Network, Portal

The Problem

In the traditional print-based library, users usually searched a limited number sources to find the information that they required. These would usually be the library's card catalogue, and/or printed indexes/abstracts. If they failed to find the information that they needed, they would usually turn to the librarian for help. The more enterprising users might also use informal networks to get the required information. Thus the user pathways to information resources tended to be relatively limited and standardised.

With the explosion in the range of digital resources and the growth of the Internet, the information landscape has completely changed in the sense that the library catalogue no longer represents the primary method by which users search for information, and the range of searchable and accessible resources has expanded exponentially. In addition to printed books and journals, users now have access to commercial databases, e-books, e-journals and a proliferation of published e-material available freely on the web. For the user, the information environment has becoming increasingly confusing and non-standard.

Users, confronted with a multiplicity of navigation interfaces and search syntaxes in their pursuit for relevant information, have reacted by relying on a few limited methods of finding information, such as (a) relying entirely on a couple of search engines to search for electronically published material on the web; and/or (b) working within their comfort zone by using only one or two tools which their experience has shown will provide them with a reasonable amount of relevant and credible information.

There are a number of problems with this approach, especially for the researcher. While search engines have generally proven to be quite effective in indexing selected parts of the Web, they are not wholly satisfactory for effective information retrieval for a variety of reasons. Firstly, they index both treasures and trivia, and so it is difficult for the user to be entirely sure that the documents, which they retrieve, are quality ones. Secondly, they index only a fraction of the total number of documents available on the Web. Thirdly, their individual coverage varies widely, and so it might be necessary to use more than one search engine for comprehensive retrieval. Fourthly, the quality of the search results vary markedly with different search engines. Fifthly, there is a lack of vocabulary control and so keyword searches can result in hundreds if not thousands of items being retrieved, much of which may not be relevant. Finally, most search engines are almost useless for searching and retrieving non-textual documents or password controlled sites.

The emergence of e-journals provides a classic example of the challenges that an information seeker now faces, when trying to track down a cited article. In the initial years of e-journal subscriptions, especially when the subscription came as part of an aggregator package (where the journals that were included varied frequently), most libraries avoided creating a catalogue record for each of these e-journals. Over time, libraries eventually opted to create an online list of their e-journals – quite distinct from their catalogue, even though the catalogue had typically been promoted as the mechanism for identifying whether the library offered access to a resource. Thus, if the library subscribed to a paper version and an electronic version of a journal, the catalogue would contain a record to indicate the availability of the paper version, but the user would need to also know that they had to check an additional list: the e-journals list.

More recently, most libraries have chosen to create catalogue records for their e-journals. However, these catalogue records are often minimal records, and lack useful “holdings” information such as the starting and ending year, and details of volume and issues. These records usually contain a URL, which promises to take the user to the full-text for that journal. However, that URL might sometimes take the user to a web page listing issues for that journal; or it might take the user to the publisher’s or vendor’s homepage from which a user would then have to drill down to the journal, the volume, the issue, and the article. Where the user sees a list of issues, they will often find that some of the listed issues only offer Table-of-Contents and possibly abstracts, while some offer full-text. It is not apparent to the user why and when this occurs. A subscribing library typically has no control over this confusing scenario. The supplier’s interfaces (and navigation within them) also vary widely, and again the subscribing library is afforded little or no opportunity to standardise the interface and user pathways that are offered to the user.

To confuse the library user further, if a journal is owned both electronically and in paper format, libraries typically have opted to create two separate catalogue records:

- one for the electronic version (with no holdings details indicated, but a tantalising URL to full-text offered), and
- one for the paper version, with holdings, but requiring the user to either visit the library or request it (if they are entitled to that service, and know of the existence of that service).

By the time a user has expended effort to drill down along the e-version supplier's interface, only to discover that there is no full-text available electronically (even though the URL in the library catalogue promised full-text), the user is unlikely to recall that there was a second record in the catalogue that they still have not checked, and which might answer their need.

A further source of confusion to users is in understanding what they are looking at. Previously, the physical characteristics and packaging signalled the genre (a paper journal looked like a journal). Soon after the emergence of electronic versions of indexing tools (such as Medline), some of these tools began to offer the occasional full-text within a subset of records, where the remaining records still only simply offered abstracts. Next, full-text suppliers (such as MCB) began to offer Table-of-Contents and abstracts for additional issues, and variants on these themes have continued.

In addition, libraries often have little control over the content of aggregations, with the result that they pay for and therefore offer their users access to all those versions of the journal. Thus a library might actively offer the user several means of accessing a journal, each with different electronic interfaces (requiring different navigation), as well as in paper format. This cannot but be confusing to the user.

To further confuse the user, e-journal suppliers (such as Elsevier and Blackwells) have begun to enhance their websites by offering users a search option, on their website's front page. It is often unclear to the user what they would actually be searching, and how to interpret the results.

Because of the volatility and other factors in the electronic information products market, libraries have and will continue to change suppliers for a given indexing tool, or e-journal. As well, students and academics may change institutions, and thus may find that, at their new institution, their favourite indexing tool or database is offered via a different interface.

A change in database interface not only means learning a new method of navigating around the service (such as Medline from Ovid or Medline from Ebsco), interpreting field tags of citations, saving citations, combining searches, and so on; but more importantly, it usually involves learning a new search syntax.

While the loss of standardisation of interfaces has largely been beyond a library's control, this mutating library-mediated environment has posed a stay of confidence for some library users. Because the e-environment has added a layer of complexity for users, a majority of them have begun to opt for an Internet search engine as the tool of choice, in place of searching the commercial databases, e-journals, e-books and other information resources acquired by the library at great expense.

A recent study [1], which surveyed incoming tertiary students about their Internet use, found that 73% claimed that they accessed the Internet daily, and an additional 25% access the Internet once a week. Furthermore, 77% claimed that they commence their research using the web, and 67% reported that they find most information for their papers through the Internet.

The Solution

The above discussion shows that the potential user pathways to search for and access information, especially e-information and collections have multiplied exponentially in the past decade. The problems of identifying the most effective

way of organizing collections so that the information contained in them can be readily searched and retrieved are manifold. A random list of these emphasises the problems that even the most seasoned reference librarian faces:

- which sources to use
- how to access these
- how to search them effectively
- how to interpret details offered
- how to navigate around a given paper tool or online interface
- how to track down a cited item
- how to identify the most appropriate information for a particular need, and
- how to identify the most effective way to access that information

AARLIN (Australian Academic and Research Library Network) was established largely to try to solve many of the information seeking problems faced by researchers, as well as the problems faced by individual libraries, in each trying to address the same. The primary purpose of AARLIN is to establish a portal framework to allow users to have seamless access to range of print as well as digital information resources which are “pushed” to them according to their teaching, learning and research needs.

AARLIN Background

The idea of AARLIN arose during a strategic planning meeting of the Council of Australian University Librarians (CAUL) in March 1999. At the time, academic libraries were experiencing difficulties in meeting the needs and demands of the scholarly community, particularly in regard to the information infrastructure to support research and teaching informed by research. These concerns were shared by all Australian universities and were raised in a number of public forums on the ‘crisis’ in scholarly communication. CAUL had been pursuing collaborative resource sharing initiatives and had supported projects aimed at improving access to scholarly literature. Independent reviews of the national information infrastructure confirmed the parlous state of disrepair, the consequences of which had serious long-term difficulties for the national research base.

Following a joint meeting of the Executive committees of CAUL and the Council of Australian University Directors of Information Technology (CAUDIT), the development of AARLIN became a joint CAUL/CAUDIT initiative in late 1999.

It is one thing to generate exciting and challenging ideas, another to gain the support and resources required to translate these into an operational system.

A campaign of rallying and seeking support for the project was pursued, culminating in the successful application for funding from the Australian Research Council’s RIEF (Research Infrastructure, Equipment and Facilities) Scheme in 2001, supplemented by contributions from 19 university libraries and the National Library of Australia. This was more of a “proof of concept” project to develop and test the functionality and viability of the AARLIN facility. Expressions of interest were sought from vendors of appropriate software to enable the AARLIN pilot project to proceed. Ex Libris’ Metalib and SFX software was selected and six universities agreed to become test sites to trial the AARLIN prototype.

Academic staff and research students from each university test site were trained to use AARLIN. Valuable feedback, through pre and post AARLIN trial questionnaires, was most encouraging. Equally encouraging was the subsequent receipt of a \$2,823,900 grant from the DEST (Department of Education, Science and Training) Systemic Infrastructure Initiative Scheme, to continue the development of the prototype into an operational system that could be rolled-out to those university libraries that wished to participate.

The post-pilot phase of the project, usually referred to as AARLIN Phase 2, involves 21 participating universities covering all the states and territories of Australia, and is aimed at rolling out of the AARLIN portal to these institutions on a staged basis throughout 2004. At the time of this writing, about 8 universities are planning a soft launch of the AARLIN system to their community by June 2004.

It has taken nearly two years to reach the stage where the universities concerned are now ready to roll out the system. Tasks and activities have included preparation of the RFP (Request for Proposal), selection of the successful vendor, staff recruitment, hardware and software acquisition, software development, meetings and conferences, training of participants and detailed configuration work of the e-resources owned by participants (which total over 2,000 unique resources). This is the largest library research infrastructure ever managed in Australia, and involves rolling out the system to some 340,000 users in 21 institutions spread across the Australian continent.

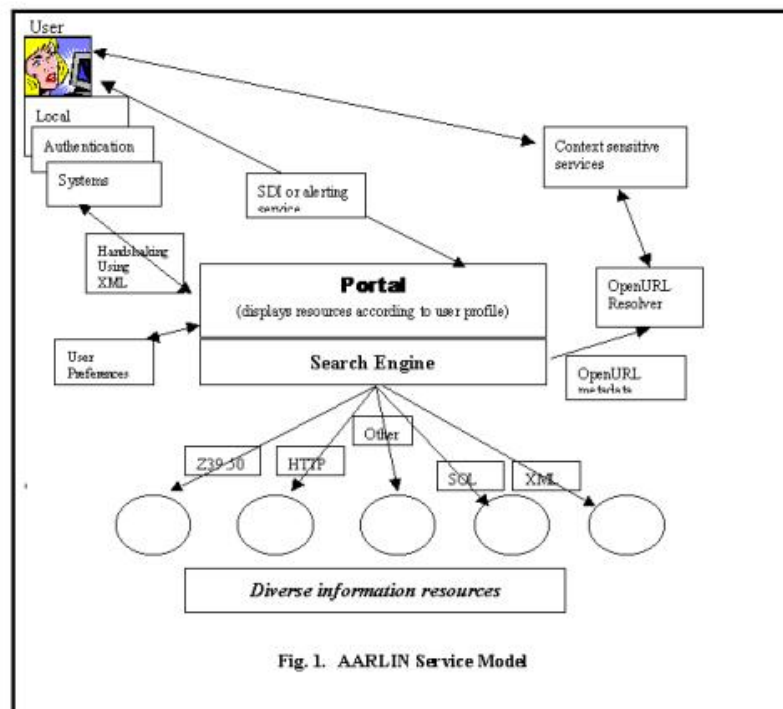
Model Of The AARLIN Portal

The expected outcome of the AARLIN project will be the development of a national portal framework, which will incorporate the following components:

- An enterprise portal facility which will allow each member of the AARLIN consortium to direct sets of resources (or “information landscapes”), such as subject-specific lists of databases and web sites, to specific categories of users within their institution, and allow users to create a personalized workspace collecting together their most frequently used resources.
- A simultaneous search gateway demonstrably capable of correctly searching, displaying and retrieving the appropriate records and information from diverse resources using the Z39.50, HTTP and other protocols via a uniform query interface.
- An alerting, or Selective Dissemination of Information service, which can be configured by end users who wish to be kept informed of new resources (e.g. articles or books) in their fields of interest.
- Seamless access to a document requesting facility among the cooperating universities for authorized users, including transfer of relevant user and item metadata.
- Context sensitive linking software using the OpenURL framework to provide end users with “extended” services and deep linking to full text documents.
- The capability of integrating with local authentication systems (including LDAP directories), which will contain appropriate profiling information enabling the respective institutions to “push” relevant information resources or information landscapes to their end users according to the user's profile(s) and the profiles of the group to which the user belongs.

At the same time, end users will be able to refine their information landscapes by adding additional relevant resources themselves.

- Integration with IP-based authentication via EZproxy web server and/or other mandatory proxy authentication systems.
- Selective as well as comprehensive listings of available databases, catalogues, web sites and other electronic information resources on the AARLIN portal interface from which users may make selections for frequent use.
- In the longer term, integration with university wide portals and e-course management systems such as WebCT and Blackboard, using a single sign-on.



As shown in Fig. 1, the AARLIN Service model is built around a national portal framework, which is linked to the local authentication systems of the participating universities. When a user logs on, the portal ensures that the user is an authorised user by communicating with the authentication system of the user's institution. The authentication service passes to the portal metadata relating to the "user's profile" – which may include information such as the research or subject interests of the user, his/her status (e.g. undergraduate, postgraduate, academic staff), and other relevant information. On the basis of that profile, the portal will "push" to the user a suite of relevant information resources (or information landscape). Thus different users would each view a different information landscape. Individual users can further refine their access to the relevant resources by adding or deleting individual items from their default list of resources. The provision of this "push" facility reduces the amount of effort required by users to access relevant information, since many of

them experience increasing difficulty in determining which of the many databases and other electronic information resources have relevance for them.

The portal is linked to institutional authentication systems (many of which are based on LDAP) via a script written by AARLIN staff. This script will map directory information about users against a list of subject-based resources and push those relevant resources to the users.

Another major component of the portal is the common user interface, which allows parallel searching of a diverse range of databases, information resources and websites using multiple protocols. These protocols include Z39.50, HTTP, and SQL. Thus, it would be possible for a user with a single search query to search across multiple citation and full text databases, online library catalogues, Internet search engines, websites and subject gateways, and get a uniform search outcome from this parallel search. The advantage of such a facility should not be underestimated. In fact, a study by Cornell University's Albert R. Mann Library showed that users appreciate the availability of a common user interface to "conceal the complexities of an information landscape characterised by numerous, disparate information resources" and the elimination of the need to master different search interfaces for different databases.[2]. This finding is consistent with the findings of a Monash University survey of databases conducted in 2001 in which many of those surveyed complained that access was too complicated and that multiple search syntaxes caused considerable confusion. [3]

In addition to being able to conduct parallel searching, the portal can pass relevant reference metadata to openlinking software using the OpenURL framework. The openlinking software will use the metadata to "resolve" what types of extended services the user is entitled to according to the user context. For instance, it might display links to full text resources, which the user's library system has subscribed to, but not to other relevant resources, which the user is not entitled to use. Similarly, if there is no full text access, the openlinking software might display a document request form populated with bibliographic metadata, if the user is an academic staff, but will not display this form if the user is an undergraduate and is not entitled to document delivery services.

Finally, the portal will have an SDI or Alerting service, which can keep users informed of new information resources in their fields of interest.

AARLIN can incorporate a number of additional features, although these are not within the scope of the current project. These include:

- Integration with university wide portals
- Integration with e-learning or course management systems like WebCT and Blackboard
- Distributed searching of learning objects repositories or open archive systems (including e-print)
- Chat, email, interactive and video conferencing facilities (to provide human interaction)
- Web delivery of documents
- Provision of 24x7 collaborative reference and help services on a national scale

Consortium Approach

The consortium approach adopted by AARLIN is based on a centralised server model, where software for all 21 universities is stored and maintained by the AARLIN staff. A key feature of this model is that that localisation of the software is undertaken by each individual university with support from the AARLIN staff. During the pilot, configuration work was centralised. On the other hand, in the operational phase of AARLIN, these tasks have been devolved to participants, with AARLIN providing advanced problem solving support and additional configuration as needed. It is important to realise that each participating university has been allocated work to configure specific databases on behalf of all other members. This sharing of configuration work represents another key feature of the AARLIN model.

When AARLIN adopted this centralized model staff were not aware whether the skills level required matched the skills level at the participating institutions. Each university's capacity for commitment as defined by the staffing level was also unknown. In the AARLIN model, a memorandum of understanding was required to be signed among the participants. In this agreement or MOU, the responsibilities of the AARLIN Office and AARLIN participating institutions were clearly defined. The extent to which each fulfils its obligations would ultimately determine the failure or success of this consortium model.

Because of the model was based on the sharing of configured resources, it was crucial that the quality of the work of individual institutions met the high standards set by the consortium.

One of the concerns relating to the AARLIN consortium model relates to the scalability of the system. The solution is the acquisition of very large and expensive servers, or the establishment of decentralised servers to service the needs of groups of universities.

Advantages

The major advantages identified in the AARLIN consortium model may be summarised as being:

- Political
- Financial
- Staffing
- Technical

Political. While Australian libraries have participated in many cooperative projects such as interlending, cooperative storage and consortium purchases, the capacity for libraries to collaborate is not acknowledged by Government funding agencies and university CEO's. AARLIN provides tangible and visible evidence of this collaboration, and this will serve the Australian library community generally.

Financial. Our estimate is that it will cost somewhere between \$400,000 and \$600,000 to set up a library portal within a university (including staffing costs). The cost of a national system works out to be less than a third per institution. These cost savings are hardware and software, and most importantly in salary savings.

Staffing. The savings in staff should not be underestimated. The major tasks (some small and one-off tasks, and others large and skills-intensive tasks) are ones which would typically be repeated at each University, and which, within the consortium environment, can be shared.. These include:

- development, by the consortium for the consortium, of a range documentation to assist library staff in the various decision-making processes, and the tasks of implementation;
- development of user-documentation such as online and paper-based HELP-files, promotional material, explanatory and background documentation to assist users, digital self-guided tutorials for use of various parts of the portal software;
- testing of various components of the purchased software on behalf of the group;
- shared server and software maintenance and upgrades;
- increased capacity for effective negotiation with vendors for bug-fixing, software changes, and software enhancements;
- shared costs, skills contribution, and influence on development of additional non-proprietary software features;
- development, by consortium member Universities, of software enhancements, or innovative use and exploration of proprietary-based application software, on behalf of the consortium; and, perhaps most importantly,
- testing, and fixing, or development of parsing scripts, and configuration setups that effectively allow: searching to be conducted against resources which use a range of communication protocols including http, SQL and telnet, and
 - Z39.50;
 - search results records returned from a given target [such as Cinahl on OVID via Z39.50] to be displayed correctly and consistently within AARLIN, using a standardised record format;
 - search results records to be harvested for appropriate openurl metadata, and for the openurl to be created and populated correctly regardless of the range of content that might be in a record;
 - correctly linking from the citation record to the Openurl Resolver and then to the article-level of online full text, wherever online full text is available to that user group;

As in any collaborative project, the potential to create strong and positive relationship cross-institutional relationships is great. There are also staff development opportunities arising both from the ability to learn from each other as well as from the development of additional individual skills. In much the same way as the 24/7 international reference desk concept is appealing to library managers, AARLIN will eventually be able to offer a skilled force of library staff from across 21 Universities, offering a range of expertise related to portal implementation which would be difficult for a single University to cultivate, especially with current University-funding where library staffing levels are so thin that a healthy expanse of additional expertise cannot be built.

Technical. AARLIN has initially installed the software for all 21 Universities onto 1 server, which means that server and software maintenance is reduced. This also provides the added benefit, in the consoritial software setup used, of sharing software and configuration developments with ease.

Finally, the effort involved to setup and implement a suite of off-the shelf portal software, while typically represented by vendors as being plug-and-play, is in

reality a significant time and labour commitment for any University embarking on this journey. Some library staff involved in the AARLIN Pilot even equated this “implementation” to the implementation of a new ILMS system. As this genre of product matures, some aspects of the implementation have become smoother; however, the processes and skills required, coupled with the tools available within the various suites offered by vendors, means that it is still a significant and challenging commitment for a single University library: one that some libraries simply could not afford, if solely reliant on their own efforts.

Disadvantages.

There are also some disadvantages to the consortium approach. One, frequently mentioned, is the loss of local autonomy. This consortial setup, in a small number of cases, obliges universities to agree on some elements of the shared software over which they might have preferred to retain autonomy. The extent of this is relatively low, and participants have not indicated that they feel disadvantaged by this. The issue is more one of the culture of autonomy, and a library’s traditional workflows and decision-making processes..

The central server model (which we expect to modify as usage loads increase) has raised some interesting challenges regarding costs of Internet traffic, at a time when these costs are being pushed back to users’ faculties and schools, in some Universities. We not yet know whether this represents a significant disadvantage or whether it is an issue only in extreme circumstances. AARLIN is working to resolve charging for Internet traffic, as well as seeking a cost-effective solution to address the increased Internet traffic that arises in a central-server model.

The central server model also has the potential to magnify a single point of failure. In that connection, we have tried to ameliorate this by creating a fail-safe system, whereby back up servers will take over if the main server fails.

A possible disadvantage relates to delays that might occur in dealing with problems by the AARLIN Office, which has to carry a larger share of the support burden than would occur in a single institution installation. In a non-consortial approach, training would usually involve up to 12 staff from a single university, with a range of backgrounds, skills and expertise. AARLIN has adopted a policy of training the trainer aiming to devolve training responsibilities to each university. This has not always proved to be successful.

For expertise to be fostered within the consortium, however, AARLIN must first provide an effective and supportive consortial infrastructure. The strength of success in fostering this expertise will also affect the ease with which each participant University is able to implement the initial AARLIN system and subsequent enhancements.

Legal And Administrative Framework

Funding for AARLIN Phase 2 will cease at the end of 2004, after which new sources of funding will have to be found. To ensure the sustainability of the system, the AARLIN planners are taking steps to develop a business plan.

Apart from developing a business case, the Plan will also need to consider various organisational and administrative models to run and manage the AARLIN system.

Preliminary thinking indicates that three alternatives will need to be considered, viz.

- getting one of the participating universities to run the AARLIN system, either as an independent unit or an entity within a larger departmental organisation
- outsourcing to an Applications Service Provider
- establishing a jointly owned company

The general consensus appears to be that at least for the next three years, the cheapest way to continue with the AARLIN system is to continue to run it as a unit within La Trobe University, and for participants to meet the running costs on a “subscription basis”. In the longer term, other alternatives will probably be pursued. In any case, the legal and administrative structures can be quite complex, and will need to be examined thoroughly by the participating universities.

Concluding Remarks

While the AARLIN system aims to assist users by providing the equivalent of a “google-like” portal to search resources which are “pushed” to them on the basis of their teaching, learning or research profiles, the fact that the implementation of the Z39.50 protocol by vendors is still haphazard and imperfect means that the goal of seamless access has yet to be achieved. Over time, there is no doubt that the AARLIN system will be improved as vendors improve their implementation of their offerings and relevant search protocols. The significance of AARLIN is that it takes the commitment of libraries to the provision of quality services to another level of sophistication. Furthermore, the aggregation of effort in the consortial effort provides palpable evidence that this is a way forward for academic libraries.

References

- [1] Waldman, M. (2003). “Freshmen’s use of library electronic resources and self-efficacy”, *Information Research*, **8(2)**.
<http://InformationR.net/ir/8-2/paper150.html>
- [2] Payette, Sandra D. & Rieger, Oya Y. April 1997. "Z39.50: the User's Perspective". [online] *D-Lib magazine*. <http://www.dlib.org/april97/cornell/04payette.html>. [Accessed 14 October 2002]
- [3] Huggard, Simon, et al. 2002. “Monash Library Database Usage Survey”. In: *VALA 11th Biennial Conference and Exhibition: e-volving Information Futures, 6-8 Feb 2002*. Melbourne: VALA, p.313-342.