

# Designing a Digital Databases Page

David Groenewegen  
Digital Resources Librarian, Monash University Library  
David.Groenewegen@lib.monash.edu.au

Simon Huggard  
Systems Manager, Monash University Library  
Simon.Huggard@lib.monash.edu.au

## **Abstract:**

*During 2003 Monash University Library engaged in a major redesign of the library databases web site. The basis of this was a program which used MARC records from Monash's Endeavour Voyager catalogue to generate online electronic journal subject lists. The new databases page was intended to make for easier updating, and to allow us to respond to known user issues. In theory this was straightforward. In practice a number of issues were raised. The paper looks at how Monash resolved these issues, while commenting on the relationship between metadata and technology. Some future directions will also be discussed.*

## ***Introduction***

In March-April, 2001, Monash University Library conducted a library databases usage survey. A clear outcome of this survey was that the current library databases page [1] was inadequate for the needs of our users. It was poorly laid out, too broad in its entry points, had insufficient descriptions of the databases and offered no keyword search facility. Furthermore, it was becoming a burden for library staff to maintain, and was partially duplicated by subject gateway web sites and the library catalogue (Huggard et al, 2002).

The usage patterns revealed by the survey also provided key information about what the new design should look like. Fifty-one per cent of respondents said they used trial and error, and 38% said they used the database name as a means of selecting the database. As an entry point for accessing a database, 67% said they used the databases page itself and 34% said they got there via a search of the Voyager catalogue[2]. Of these, 25% said descriptions of databases were inadequate, with almost as many respondents saying that the arrangement of the databases was confusing.

In partial response to the survey outcomes, the library undertook a portal software trial in 2002. This software provided cross-database searching and allowed for listing and descriptions of databases based on a user's faculty and subject affiliation. This trial ended in late 2002 because we concluded that the software was not sufficiently developed at that stage. The trial did, however, indicate the latent demand for more tightly targeted sorting of our resources (Groenewegen & Huggard, 2003).

After the portal trial finished, it was decided to make a concerted effort to replace the existing library databases page in line with the following objectives:

- Provide a clear entry-point for users wishing to find databases under meaningful subject categories
- Provide a simpler search facility of the subject and descriptive metadata than was offered by the library catalogue
- Provide alphabetical and subject-based listings of all databases
- Provide a technical framework which would provide consistent information about databases which could be used not only on the databases site itself, but also on subject gateway and other web sites of the library or University
- Provide a single point of data entry and description for the databases
- Provide a database driven web site, rather than a manually created and maintained one.

Rather than create a new database of databases, it was decided to make better use of the work already done by our cataloguers. Monash University Library has used Endeavor's Voyager ILMS since 1999. Library policy has dictated that all of the library's resources, both print and electronic, have a record in this system. Since the records of electronic material already contained a good deal of standard metadata, these seemed the most logical source of our new page. Records also contained "e-descriptors" – locally defined metadata that allowed for subject-based sorting of resources.

The approach to the provision of electronic resources via a library catalogue is a well-established one and is acknowledged by many institutions as an attractive method of access.

Herrera and Aldana (2001) have said that “Institutions that have fashioned a database specifically for their electronic databases and another for their electronic journals have created a complex picture for those in user education to convey. From the user’s perspective, one-stop shopping for all library resources is ideal.” However, the approach we have taken gives the users a choice of access point, with electronic resources being available both on the web site and within the library catalogue. Therefore, if users are more comfortable using the web site for access to these resources, they can choose the method they are most comfortable with. The information they are accessing is drawn from the same data source.

Work commenced on the design of the databases page in October/November 2002. Initially we designed some simple web pages to illustrate the preferred layout of the databases pages, with templates for the index page, subject listings, alphabetical listing and detailed individual pages.

The program to pull the necessary information from Voyager and from there to create the HTML pages was written in November/December 2002, with a prototype finalised in January 2003. This prototype was then put on the library intranet and comments were sought from subject and reference librarians.

Originally, we had planned to go live with the new design in February 2003, in time for first semester in March. However, the late release of the prototype, and the growing awareness that further work needed to be done on the Voyager records, meant that the launch was delayed until semester 2. This target was only just met due mostly to the enormous amount of work required to review the existing e-descriptor list, to edit records in Voyager and to incorporate the many comments from Subject Librarians about the design of the new database pages.

Finally, it was agreed that the old databases page would be maintained during 2003 while changes were made to the new databases page, which was launched as an alternative link from the old databases page, in July 2003. The new page also included a survey to find out user responses to the new page, so that it can be further refined over the course of semester 2. The plan was to have the new site fully developed at the end of second semester in preparation for the 2004 academic year. This was achieved, but some work was still needed to conform to the University’s new web style.

### ***The metadata schema***

The usage of resource discovery metadata to enhance the searching and finding of internet-based information has been debated very widely over the last five to six years. Certainly the use of Dublin Core metadata within web pages has allowed for improved and more consistent retrieval of information across whole web sites, using local search engines, and from Internet-based directories and search engines.

The Dublin Core initiative brought the whole area of descriptive cataloguing, classification and subject control into the public domain, after many years of being largely the preserve of librarians working within library catalogues. The massive challenge has been to come up with schema which could cover so many diverse and numerous resources across the whole Internet space; the resulting schema have often only been partially used by web page designers and information specialists.

Controlled vocabulary (such as name authorities for authors) and standardised subject thesauri (such as LCSH or MESH) has been a strong area of expertise and methodology employed by librarians for decades. However, this expertise has not translated into a strong presence in the area of description and resource discovery for Internet-based resources. Often this has been due to gaps in communication, collaboration, information sharing, and the demands of commercial interests between librarians on one side and IT managers on the other side.

The result has been that the focus of IT and web developers has been on the implementation of masses of web-based information, sorted by a whole variety of categories, with a top-level search engine to index the information presented to the user. By using powerful indexing and searching tools on their web site, which can index deep within sites, IT managers have presented the user with a large amount of easily searchable and accessible information. But they have not provided much help on the best ways of navigating around and searching this large amount of information. The result for the user has been to present a similar expectation to that of the modern-day telephone inquirer, who can expect to be put on hold on a telephone system for a long period of time until a customer service representative becomes available to take the call. Similarly, a user of an Internet-based search engine can expect to receive a huge number of results from a search and, due to a lack of time to go through all 2 million results retrieved, chooses a 'hit and miss' approach to the first 10-20 results, hoping that at least one will give them the information they need. Lim and Roberts (2003, p. 433) mention that "they prefer to spend extensive time browsing large result sets than invest time identifying whether their search has been effective, and how it might be improved". This reinforces the need to make navigation and searching on a web site as easy and intuitive as possible for users, so as to not waste their time unnecessarily.

We are not suggesting here that the work of cataloguers and librarians is all that much better presented to the user either. The real domain of the librarian, the Internet-based library catalogue, is often very confusing and difficult to navigate for the user. Roy Tennant (2003, p.28) wrote a controversial, but largely accurate, article in *Library Journal* about the poor design and display of library catalogues. The information in the catalogue is carefully indexed and controlled, but it's often not clear to the novice user what they are searching and why. The mix of resources, ranging from rare collections, realia, bound volumes of journals and recent books, to individual articles, power-point presentations, electronic reserve or course pack articles in PDF format and so on, creates a different kind of Internet chaos within a confined educational space for the user. Similarly, keywords are input, retrieving another meaningless number of records which may or may not have some obvious pointers as to their relevance to the user.

With these problems in mind, it was suggested that we design a replacement databases page, with entry points based on local subject headings used in the Monash Voyager catalogue. Since most of the library's databases were already catalogued and described there, with a lot of work going into summary descriptions, subject categories and format/scope notes, it was felt that the whole databases page could be produced automatically from the library catalogue.

The benefit to the user would be:

- (a) Browsing via controlled subjects (using locally defined electronic resource descriptors – “e-descriptors”)
- (b) Browsing alphabetically
- (c) Simple search page to limit to databases within any keyword (a database of databases).

### ***The Monash University Library E-descriptors List***

The first list of locally defined e-descriptors was produced at Monash in 1996. This list was developed to provide subject-based directory navigation for users of the library’s Electronic Resources Directory (ERD), which was an early “Yahoo” style web-based directory of databases, e-journals, networked CD-ROMs, and web sites such as library catalogues and subject gateways (Smith 1998).

The e-descriptors list was originally compiled with the faculty areas as well as subject/unit names in mind. Currently Monash has 10 faculties, and about 1500 active units, within its areas of teaching. Many of the subject units are only marginally different, depending on the year-level being taught, or they vary only slightly in their subject coverage. Therefore, when the list of e-descriptors was mapped to Monash subjects or units, it actually amounted to a much smaller list of subjects, about 350 in January 2003. These e-descriptors were used to describe web sites, CD-ROMs, e-journals and many other electronic resources, and are not specifically useful for every database to which the library subscribes. By November 2002, when work began on the programming side of the project, the actual number of databases in the catalogue was about 480 and the number of e-descriptors which had been assigned to one or more databases, was about 225.

Apart from containing a search option, and an alphabetical listing, one of the main features of the new databases page was a listing of all of the subject e-descriptors up front, which would then link to all of the databases relevant to that area of study. Use of the e-descriptors would also allow for more specific lists of resources, since the old databases page was sorted according to the Monash faculties.

We were conscious of going too far the other way. Because we wanted this page to have a simple look and easy navigation, we decided that the use of all 225 e-descriptors was excessive, would make the page far too big and would result in “lists” of a single resource. With this in mind, we worked with cataloguers and subject librarians to cut the list down to a more manageable number. For a front web page to be useful to a user with a 1024x768 pixel computer screen (currently the average screen size used on the Monash web site), you generally have around 40 lines of text available on one single screen. The design of the page being in two columns gave us room for about 90-100 e-descriptors, so it was agreed to cut the list to about 90.

## ***The list of e-descriptors***

Defining the new list proved to be a source of much debate. In the efforts to cut the list down, it was difficult to come up with a definition of what an e-descriptor should be based on. One of the key original definitions, which is still relevant, was that it had to describe an area of teaching at Monash University. However the term descriptor, in the cataloguing sense, meant that cataloguers felt it had to conform to cataloguing standards and remain fairly closely aligned with the Library of Congress Subject Headings, on which the whole list had originally been based. There was also concern that restructures within the university, which result in changes in the names of departments and schools, could render the list obsolete if descriptors that were too “Monash-centric” were used.

Furthermore, they needed to remain well-structured in order to help with keyword, e-descriptor and subject-heading searches within the Voyager catalogue, in which the database records would still be accessible. There was also the complication that terms were not exclusively for use in describing databases, but could be used for any electronic resource (and would continue to be used that way in the Voyager catalogue).

On the other hand, some library staff held the view that the e-descriptors were merely a set of words used to describe Monash faculties, departments and teaching areas, and were a way of navigating on a large web site, and should use natural language and easily understandable concepts/subject words. Therefore it was natural for some subject/reference librarians to want to use a more free-form and uncontrolled vocabulary, as well as faculty/subject-based words, and eliminate some of the more controlled/descriptive e-descriptors.

There were a number of contentious headings. For instance, there was lengthy discussion over whether “Pharmacy” and “Pharmacology” should be separate descriptors, or whether they should be combined as “Pharmacy and Pharmacology”. This was further complicated by the Monash structure – Pharmacy is a faculty, while Pharmacology is a school within the faculty of Medicine, Nursing and Health Sciences. Most of these schools did not get their own descriptors, since it was felt that there would be unnecessary duplication across them (the resources used by Anatomy and Physiology being largely identical). To give Pharmacology a descriptor would be inconsistent, yet as an area of study it does use many resources not essential to other medical research areas. Thus, the process of haggling and negotiation over the culling of the list of e-descriptors took about 3 months to complete.

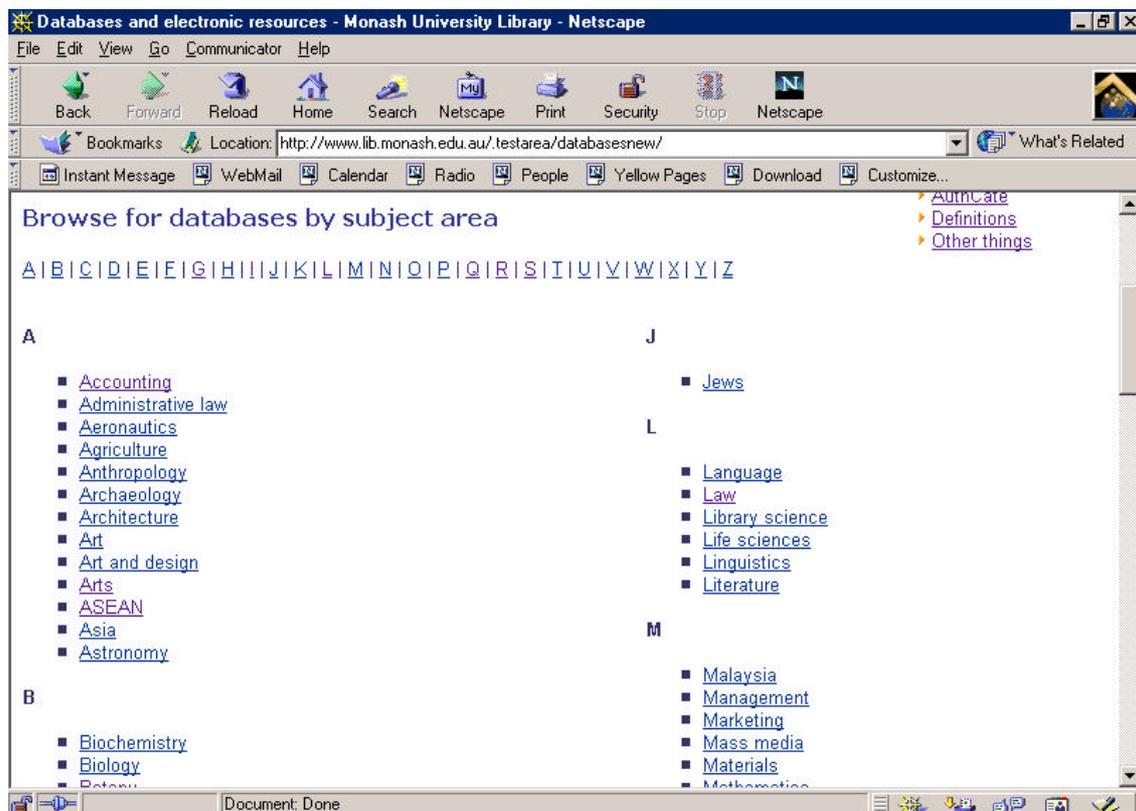
Further complicating the issue was the question of what made something a “database”. The use of the term is somewhat out of date, and most libraries have moved to “Electronic resources”. In the end, the broad definition of “a searchable electronic product” was used, although it could have been applied more rigorously. Future versions of the page will attempt to address this issue further.

## Design of the databases page

The graphic design and layout of the databases web site was something which was very easy to put together, primarily because it was already determined by the corporate design of the Monash University web site, with which all Monash web pages are supposed to conform. Therefore, the layout simply required matching a design with the key elements recommended by our users in the library databases survey of 2001.

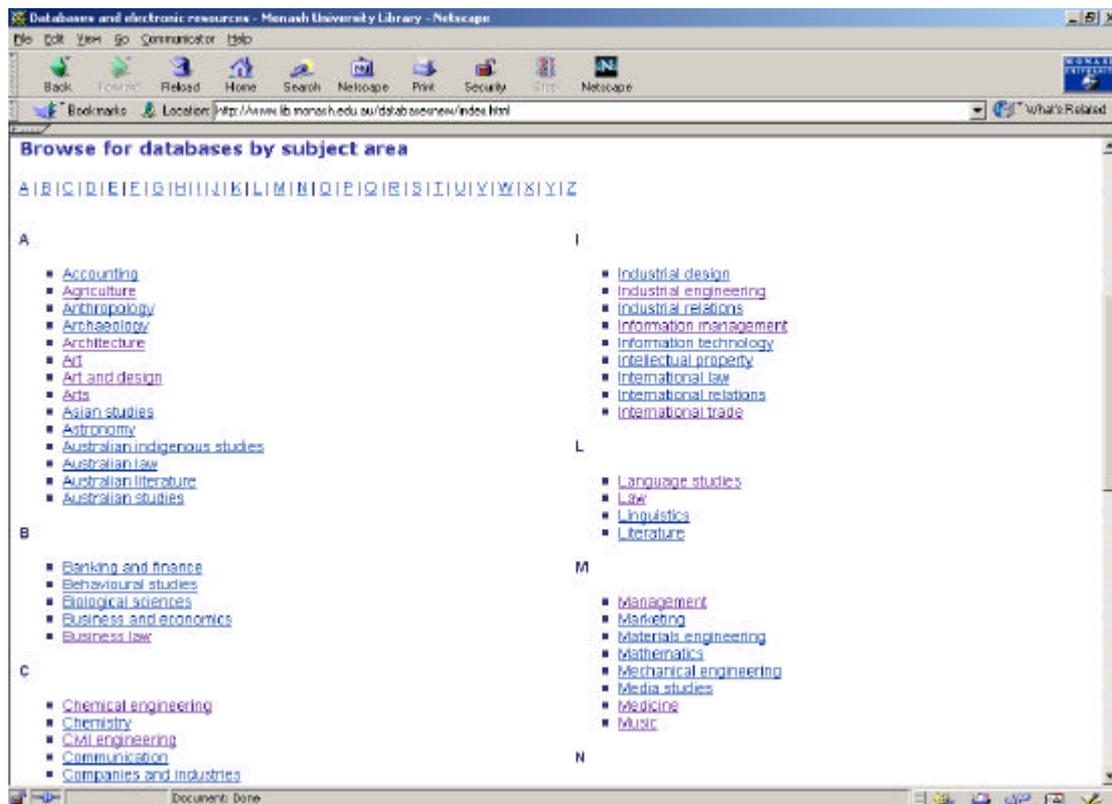
An outline of the technical workings of the site is in figure 1.

An early rendition of the databases site, from January 2003, is shown below:

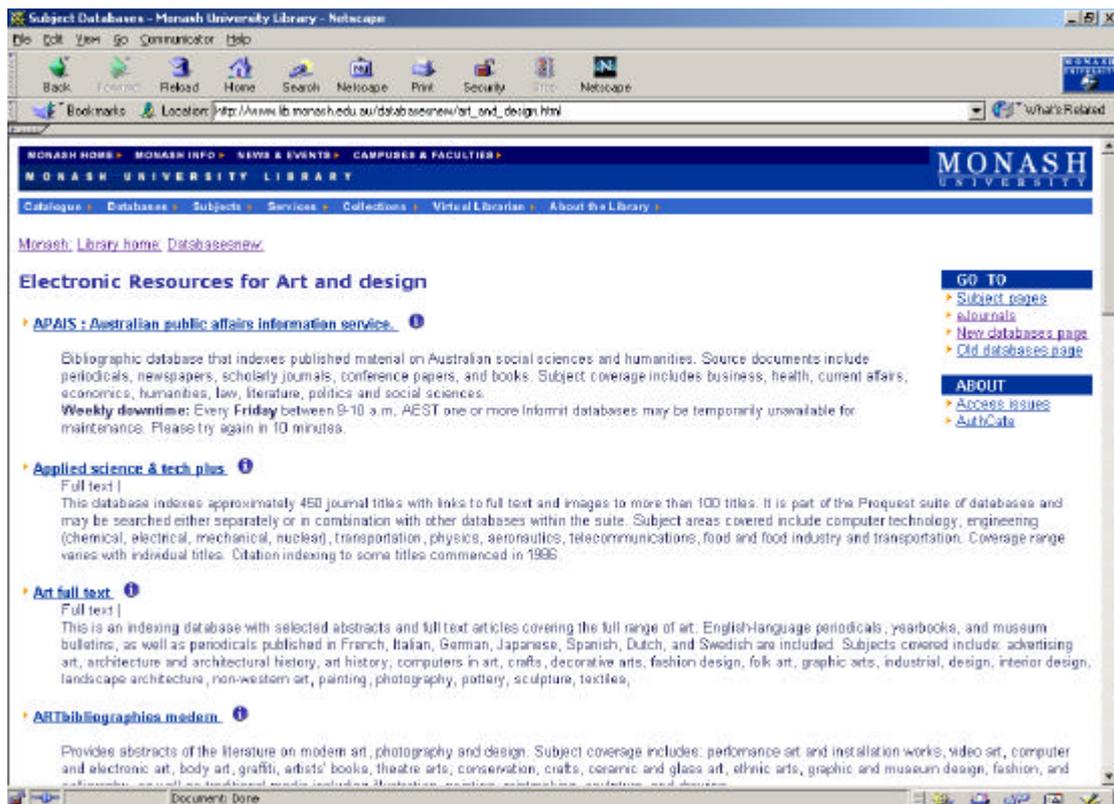


As can be seen from the list of e-descriptors shown, there were a number which reflected older LC headings (such as “Jews”), which were thought to be inappropriate as an entry point. Various country names also appeared, particularly Asian countries such as Malaysia, Indonesia and Singapore, which were consolidated into the subject area “Asian studies”. The only other country designation that was retained was “Australia”, because students have indicated that they struggle to find local material. Various areas of law were also removed, such as Administrative law, Human rights, Corporation law and so on, and these were consolidated under the heading “Law”, on the assumption that most of these lists would be essentially identical, given that most Law resources cover a broad range of the law. Similarly, very specific medical terms were removed and consolidated, as discussed above.

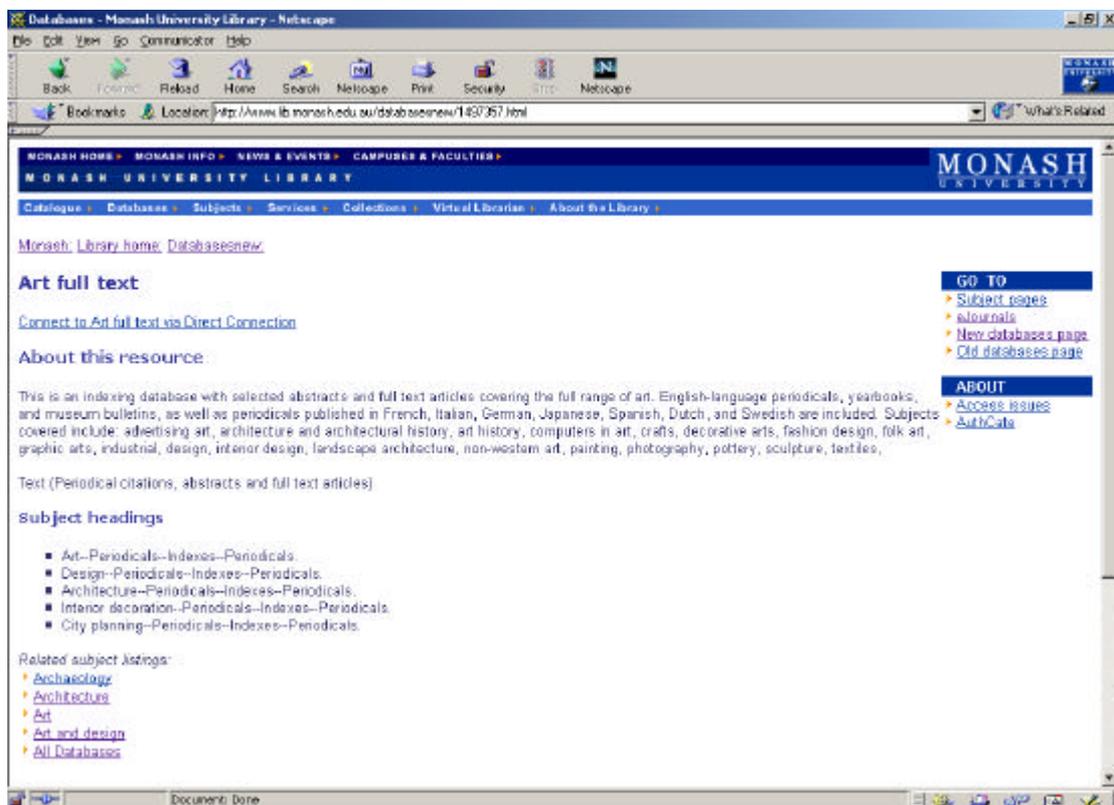
A more up to date version of the site, from August 2003, is shown below.



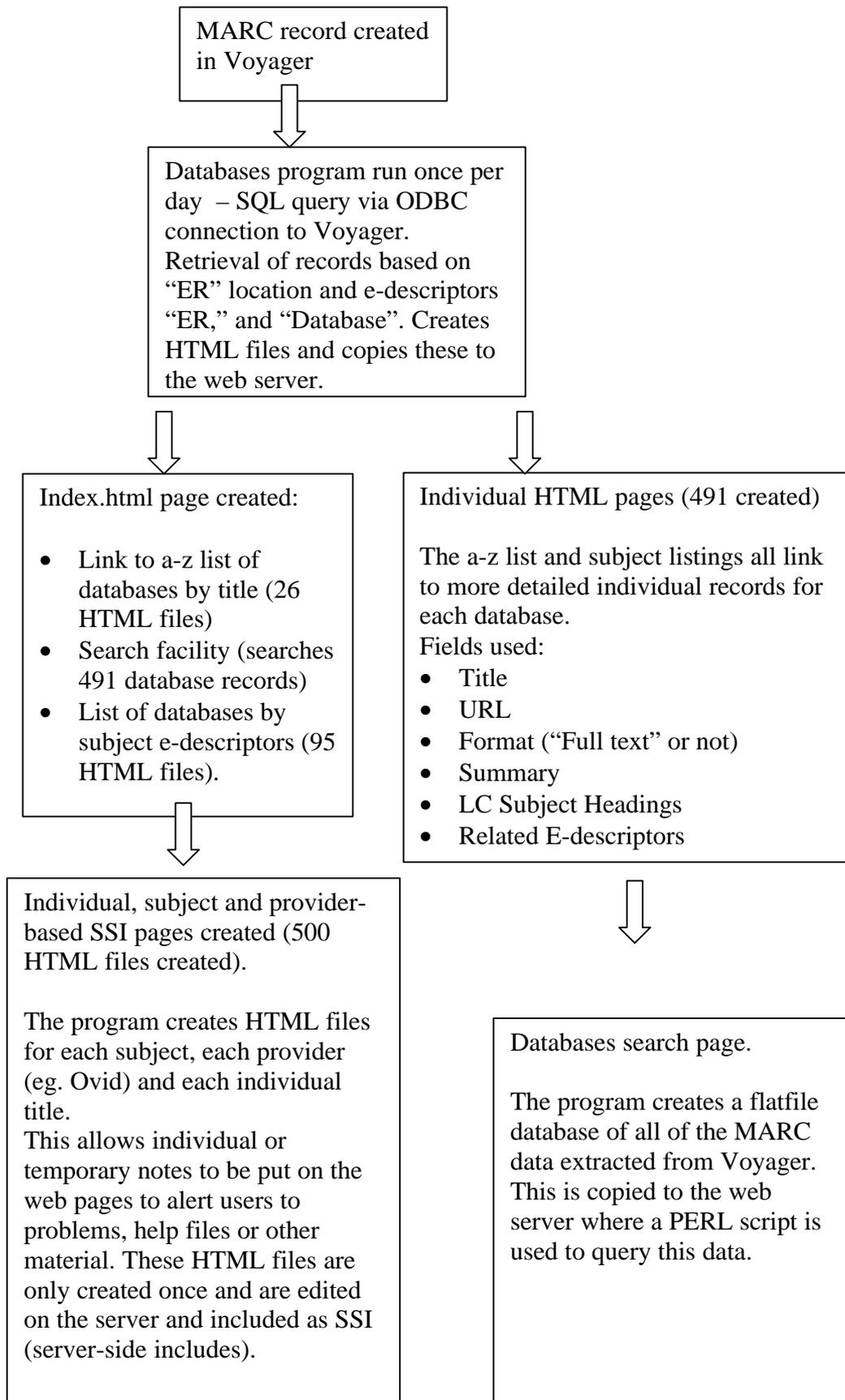
If a user clicks on one of the entry points, they see a listing of all of the relevant databases by subject area with a brief description of each, as shown below:



The details on this page come from the Title, URL, summary and format fields in the MARC record. A link to further information is contained behind the **i** ("i") information button. This takes the user to the full description for a database as below:



**Figure 1 - Technical design**



The creation of the databases page programmatically, and from a relational database consisting of detailed descriptive records, meant that we were able to produce a sizeable web site quite quickly, record maintenance done in a single location. The programs used to create the pages were very similar to those already written for the library e-journals site [3] and were able to be rewritten and maintained without too much effort. Apart from the initial cataloguing work done to ensure integrity and that standards were followed for every database record, the ongoing maintenance in Voyager required almost no changes to staff procedures or staff levels, as the design fitted into the current work practices within acquisitions, cataloguing and systems.

The inclusion of individual pieces of text, loaded on to web pages via Server Side Includes (SSIs) (see figure 1), allows for rapid updating outside the regeneration cycle of the page. It also allows for the inclusion of information that might not be suitable for MARC records – for instance downtime and access messages, links to alternative resources and notices of upgrades. These can be added and removed with minimal effort, and without the intervention of cataloguers or programmers.

Initial work in cataloguing involved editing of every database record in the catalogue (some 491 records) to re-assign the e-descriptive metadata, check links, input revised summary information, and amend or input new date coverage information. The process also uncovered a number of resources that had never been catalogued, or which had very old and out of date records.

### ***Benefits of a database driven web page***

#### (a) Flexible design

Because the program could mix and match the data any way that was required, this meant it could produce its output in many different ways. Currently it produces alphabetical lists, subject-based lists, a search page for finding relevant databases based on keywords in all fields of the records (including summary LC subject headings, format and so on) as well as detailed individual web pages for each resource. Due to the hierarchical use of the subject e-descriptors, it was easy to create links from individual database records back to other related subject-based lists.

#### (b) XML output

The program also produces output in XML format to allow compliance with future systems at Monash (currently the main web servers at Monash are not able to display XML-based data directly), as well as providing a potential method of pushing content to relevant areas in the my.monash portal [4]. At the time of writing this paper, a lot of work was still required to provide a useful XML schema and DTD for this data.

#### (c) Simplified updating

New resources can be added to the databases page without any HTML or cataloguing expertise, using a web input form for the cataloguing of electronic resources [5]. This form allows a subject librarian (or more often, the library's electronic resources team) to immediately input a record into Voyager from a web-based form, which then produces a record in the databases page on its next automatic update – all without any knowledge of MARC, cataloguing or HTML editing.

## ***Benefits for library staff***

Catalogue records were suddenly exposed on the library's home page, and were fixed immediately if any errors were obvious. This exposed a lot of very old records which had been forgotten about, but also increased the visibility of the work being done by our cataloguers. The existing records contained a wealth of useful information, most of which was underutilised by Voyager users. We hope that this process will increase awareness of the good work our cataloguers do.

The Voyager catalogue was enhanced through better summary descriptions of databases being provided by reference/subject/faculty librarians, who were dissatisfied with original descriptions in the records. These staff members also gained a higher level of ownership of these descriptions.

## ***Conclusion***

Creation of this new page entailed far more work than originally envisaged. Our original assumptions about the suitability of the existing catalogue records proved to be overly optimistic – we had assumed that little work would be required, but in the end virtually every record was edited, in some cases more than once. However, the end result of this is a more accurate and up-to-date catalogue, with a greater incentive to ensure that the records are continually maintained, as they will be reflected in the main user entry point to electronic resources.

Additionally, we have been able to address the concerns of our users, and still produce a result that requires less work to maintain than the manually coded page used in the past. Further evaluation of the site will be undertaken to ensure that it meets future needs.

## Endnotes

- [1] Actually a collection of around 30 pages, but known within Monash as the “Databases page”
- [2] Users were able to nominate more than one entry point.
- [3] <http://www.lib.monash.edu.au/e-journals/>
- [4] <http://my.monash.edu.au>
- [5] <http://www.lib.monash.edu.au/intranet/techserv/forms/fastcat.html>

## References

- Groenewegen, D. and Huggard, S., (2003) “The answer to all our problems?” In: *Information Online*, January 21-23,  
<http://www.alia.org.au/conferences/online2003/conferencepapers/groenewegen.htm>
- Herrera, Gail and Aldana, Lynda (2001). “Integrating electronic resources into the library catalog: a collaborative approach” *Portal: Libraries and the Academy*, Vol 1, No. 3, 2001. pp. 241-256.
- Huggard, S. et al, (2002) “Monash Library database usage survey” in: *VALA 2002, Proceedings of the 11<sup>th</sup> Biennial Conference & Exhibition, 6-8 February 2002*, pp.313-342,  
<http://www.vala.org.au/vala2002/2002pdf/23Hugard.pdf>
- Lim, Edward & Roberts, Kate (2003). “Over the rainbow : finding information in the virtual learning environment”. Paper presented at *Educause in Australasia 03*, 6-9 May 2003, pp.431-439
- Smith, Lisa (1998). “The Monash University Library Electronic Resources Directory.” *LASIE* Vol 29, No 1, March 1998. p. 55-62.
- Tennant, R. (2003) “Library catalogs : the wrong solution”. *Library Journal* v 128 no 3 (Feb 15 2003) p. 28,  
<http://libraryjournal.reviewsnews.com/index.asp?layout=article&articleId=CA273959>