Continuing the Quest for the Quick Search Holy Grail: Oregon State University Libraries'

Federated Search Implementation

Michael Boock, Jane Nichols and Laurel Kristick

ABSTRACT

This article describes the selection, implementation, configuration and assessment of a proprietary federated search product. The decisions made and processes used to facilitate the implementation are outlined. Users' reception of the product based on a usability study and survey responses are presented. Also described are unique and innovative implementations of the software within Oregon State University Libraries and the application of this experience to the creation of a home-grown, open source federated search tool.

KEYWORDS

Federated search, Implementation, MetaFind, Metasearch

Michael Boock, michael.boock@oregonstate.edu, is Head of Technical Services; Jane Nichols, jane.nichols@oregonstate.edu, is Social Sciences/Humanities Librarian; Laurel Kristick, laurel.kristick@oregonstate.edu is Interim Head of Collection Development, all at The Valley Library, Oregon State University, 121 The Valley Library, Corvallis, OR, 97331-3411.

INTRODUCTION

Enabling quick and easy-to-use search and retrieval software is a long time goal of academic libraries. OCLC's 2005 study of patron behavior tells us that our users want simple and speedy access to information with as few clicks as possible (De Rosa 2005). Oregon State University Libraries joined this quest for a simple, easy to use search tool in 2003 when it began exploring options for implementing federated search software. Federated search provides a single search of multiple resources and a unified listing of search results. This article describes the selection, configuration, implementation and assessment of a proprietary federated search product. The decisions made and processes used to facilitate the implementation are briefly outlined. Users' reception of the product based on survey responses and a usability study are presented. Unique implementations of the software are explained. Finally, the Libraries' creation of a home-grown, open source federated search tool is described along with the reasons the Libraries decided to develop their own tool.

Oregon State University Libraries evaluated several federated search products to determine which would be the best fit. Ultimately, the Libraries selected MetaFind from Innovative Interfaces, Inc. The Libraries have used this software since the fall of 2004.

MetaFind began to address several issues surrounding the ability of patrons to quickly locate relevant research materials. Based on interactions at the reference desk and from instruction sessions, librarians knew that students increasingly turned to Internet search engines to meet their research needs because they are fast and easy to use. Searching databases and catalogs and locating the full text of resources is a complex process. Federated searching can simplify and enhance the users' ability to locate relevant resources without librarian mediation (De Rosa 2005). The federated search software also created a way to offer a search box from the

Libraries' home page, allowing users to search without having to choose or locate appropriate resources.

LITERATURE REVIEW

While a number of articles have been written on federated searching over the past five years, relatively few have been published on the selection and implementation of a federated search tool. Almost all of the articles on these aspects of federated searching were published within the past year, after Oregon State University made its own selection and implementation of MetaFind. The only article to predate our experience was written by Gerrity, et al. (2002), a *Reference Services Review* article which focused on the implementation of the Ex Libris LTD products MetaLib, a federated search tool, and SFX, an openURL link resolver, at Boston University.

Colyar (2005) has written a brief article on the process the Louisiana Academic Library Information Network Consortium (LALINC) used to select a federated search tool and link resolver. Curtis and Dorner (2005) include a useful list of resources for evaluating federated search tools in their article.

A few articles on portal implementation include discussion of the federated search piece. Madison and Hyland-Carver (2005) discussed the implementation of a portal at Iowa State University, which included a federated searching component. Feeney and Newby (2005) focused on the selection of databases and the usability testing of a portal at the University of Arizona.

To date, no articles have been published on the selection and implementation process using MetaFind. Several of the articles published in the past year have focused on MetaLib, the

Ex Libris LTD federated search product, and many of these have information useful for the selection and implementation of any federated search tool. Besides the Gerrity, et al. article, Helfer and Wakimoto (2005) discussed the experience of installing MetaLib at California State University, Northridge and Chaffin, et al. (2005) described a similar experience at Colorado State University.

PRODUCT SELECTION PROCESS

In early 2003, a task force was created and charged with evaluating software that would further the Libraries' digitization efforts. Because federated search and OpenURL link resolution software was emerging at this same time, the charge was expanded to include evaluation and recommendation of those software tools. The process from evaluation to selection of these three cornerstones of the Libraries' digital efforts took 1.5 years.

To learn about and evaluate federated search and link resolution software, the task force invited several vendors to demonstrate their products. All of the Libraries' staff were invited to attend the demonstrations, raising awareness of the products and their utility. The Libraries simultaneously held a seminar series to educate staff about tools for accessing digital content.

Based on the vendor demonstrations, the task force created selection criteria for a federated search tool. These criteria included: usability, expected implementation time, installation expertise required and cost. Another significant consideration was using a single vendor for both federated searching and link resolution. The task force believed this would help prevent problems with incompatibility. The task force held follow up conference calls with vendors in order to ask specific questions about the products and in some cases invited vendors for a second visit. Products most carefully considered were ENCompass from Endeavor

Information Systems, Inc., MetaLib from Ex Libris LTD and MetaFind from Innovative Interfaces, Inc. Additionally, early adopters and beta test sites were contacted and asked about their experiences with the vendors' software.

The task force concluded that federated search technology was still in an early development stage. Problems included a limited number of database vendors, high cost of software, lack of vendor and publisher standardization for enabling federated searching of their products and time-intensive implementation. Even though the tools were not fully developed, OSU Libraries decided to move ahead with purchasing a federated search tool and link resolver. User expectations for easier search and retrieval pushed the Libraries to implement a federated search tool that was not yet perfected.

The OSU Libraries purchased and implemented MetaFind from Innovative Interfaces, Inc. for federated search and their WebBridge product for link resolution services. MetaFind was found to be as good as other tools on the market. It did not require a significant investment of staff time to implement – a total of approximately 40 hours - and it was reasonably priced. Moreover, it allowed search boxes to be installed on any web page with very little technical expertise, and it was compatible with other software used by the Libraries. Finally, the Libraries familiarity with the vendors' other systems, staff and customer service was a positive factor.

CONTENT SELECTION

The first implementation step was the selection of content to be searched. It was decided that the subject librarians would be the most appropriate staff for the task because of their intimate familiarity with the Libraries' resources and their role as content selectors. A task force which included three of thirteen subject librarians was charged with facilitating communication

with the remaining subject librarians about which resources to include in the federated search tool and with reviewing existing federated search implementations to determine options for the user interface.

When the task force began selecting databases for inclusion there were no existing guidelines to follow. The task force started with a complete list of electronic resources available to the Libraries and narrowed it from there. This master list comprised approximately 120 databases (citation, full text and reference resources), search engines, and catalogs. In reviewing the list, the need to develop criteria for inclusion or exclusion of resources became immediately apparent. For cost reasons, databases that had a per search charge were excluded. The task force also excluded resources with a small number of simultaneous users in order to avoid turning away users because of simultaneous use restrictions. Since MetaFind only searched web-based resources, a small number of databases hosted on CD-ROM and those that were not web accessible were excluded. Resources that met any of these three criteria were eliminated from the master list.

Next, the task force determined which subscriptions could be included or "profiled" for inclusion in the federated search. Approximately sixty percent of the databases were already profiled by the vendor, including full-text and citation databases from most major vendor platforms and the Libraries' catalog. All of these were eligible for immediate inclusion under the contract terms. However, Innovative had not enabled federated searching of several of the Libraries' databases and it would have required additional cost to include them in MetaFind. Due to the additional cost to include these other databases, librarians reviewed them for potential inclusion. A final selection consideration was to choose a single vendor platform in cases where

the Libraries subscribed to a database from more than one vendor (e.g., the ERIC database on both EBSCO*host* and CSA).

USER INTERFACE

The task force reviewed federated search implementations by other libraries to explore options for the organization and presentation of resources to users. The task force found several different strategies for organizing resources including: organization by subject, by format (e.g. books, articles, reviews), and alphabetical listing. Most institutions organized their resources by subject and this is the structure the task force adopted.

The three subject librarians on the task force met with the remaining subject librarians to discuss MetaFind and get their input on how the tool could be used. As a result of these discussions the librarians decided to distinguish between a basic search page highlighting selected resources and an advanced search page that enabled users to pick and choose from all eligible resources. On both pages the Libraries' catalog was included by default. After the librarians discussed how to present resources on the advanced search page, they decided to map the resources to the subject categories listed in the OSU Libraries' database of databases. The goal was to create a consistent presentation of resources. The eligible resources were grouped broadly by subject discipline which roughly mirrored the University's departmental organization. A partial screen shot is shown below in Figure 1.

FIGURE 1: RESOURCES GROUPED BROADLY BY DISCIPLINE IN OSU'S METAFIND

☐ All Social Science & Humanities	☐ All Sciences	
☐ **Arts**	☐ **Agriculture**	
☐ Academic Search Premier	☐ Academic Search Premier	
☐ Art Index	☐ AGRICOLA	
☐ ArticleFirst	☐ Environmental Sciences & Pollution Manage	
☐ Grove Art Online	☐ Web of Science	

As the subject librarians worked to group databases, it became apparent that not all of the resources would fit neatly into the subject schema. The librarians puzzled over how to include resources such as library catalogs, conference proceedings, images and newspapers. Because these resources can be used by any discipline they fit into all subject groupings. The research literature offered no solutions; however, reviewing other libraries' implementations helped the librarians devise a strategy for this problem. The librarians chose to create separate categories for library catalogs, conference proceedings, images and newspapers. This solution is visible in Figure 2.

FIGURE 2: CATEGORIES IN OSU'S METAFIND IMPLEMENTATION

☐ **Library Catalogs**
☑ OSU Library Catalog
☐ Orbis-Cascades Summit Catalog
☐ WorldCat

☐ **Conference Proceedings**
☐ PapersFirst
☐ ProceedingsFirst

☐ **Images**
☐ AltaVista Images
☐ Artcyclopedia
☐ Google Images

☐ **News**
☐ Academic Search Premier

During the process of identifying content to be searched from the advanced search page, it became clear that the advanced search page would not necessarily make users' search

experience easier. Federated searching enabled users to search multiple resources simultaneously but they were still required to first select the databases they wanted to search from long lists. The task force decided to make searching easier by offering an option which did not require users to first select the databases to be searched. A powerful feature of MetaFind is that with basic programming a search box can be created to simultaneously search multiple resources. Being able to offer a simple search box to users changed the task force's thinking about how to implement the software. The task force was confident that many of the Libraries' users, especially undergraduates, would prefer to search without having to choose from a list of resources. This perception was based on librarians' interactions with users at the reference desk and in information literacy sessions.

The task force and the librarians decided to focus on creating a quick (or basic) search box populated with pre-selected resources that would quickly yield results. The task force and the librarians envisioned promoting this basic search as a way to meet lower-level undergraduates' need to quickly find full-text materials and books. It was thought that upper division undergraduates, graduates, faculty and researchers would gravitate to the advanced search page because of their familiarity with the Libraries' databases. Again, this reflected librarians perceptions based on their experiences working with students, faculty and researchers in information literacy sessions and at the reference desk.

Once the task force and the librarians decided to focus on creating a better search experience for undergraduates, the task force wanted to push the envelope in terms of where the search box was placed. At that time, most libraries placed a link to a separate federated search page from their libraries' home page. The task force elected to go a step further, and remove a click, by placing the customized basic search box on the OSU Libraries' home page. See Figure

3. The quick search box searches heavily-used, multidisciplinary, full-text databases and the Libraries' catalog. This combination gives users a quick and easy search experience which results in a high number of full-text records. Next to the search box is a link to a help page – "What Is MetaFind?" - which briefly explains MetaFind, describes what it searches and provides a link to the advanced search page.

FIGURE 3: OSU LIBRARIES' HOMEPAGE

PROMOTION AND BRANDING

We took several steps to promote this new service. An article was published in the University's weekly newsletter *OSUWeek* as well as the University's daily email newsletter *OSUToday*. Both are sent to all employees of the University. To inform students, an article was published in the student daily newspaper, *The Barometer*. The service was framed as a new and experimental way to search the Libraries' databases and catalog. We branded the search pages by incorporating a banner that includes links to the Libraries' home page and the Libraries' 'Ask a Librarian' service. Use increased substantially the month the article was published in *The Barometer* (from 174 searches to 7,339), but in the following months, average use fluctuated a

great deal. The median number of searches per month in the first year of MetaFind implementation was 2,512, with higher use tied to the months MetaFind was being used in English Composition classes.

HOW METAFIND WAS RECEIVED

Even though the Libraries were pleased to offer a customized federated search, several issues arose that made the search experience less than ideal. Usability testing and a survey offered from the "What Is MetaFind" page revealed problems users faced when using MetaFind. In August 2004, the Libraries' first functioning federated search page was released to a distance education course for testing. The Libraries also held an informal focus group where student employees and Valley Library staff who had not previously seen or used the product tested it and shared their thoughts. Primary observations from usability testers were that search results were difficult to interpret and that searches were slow, particularly when many resources were queried. Though these were challenging problems to solve, the task force agreed that the customized basic search page offered useful alternatives to searching databases individually.

Users observed that search details appear at the top of the screen and the results appear well beneath the top of the page, forcing them to scroll down to find results, creating an extra unnecessary step. Results are displayed as they are returned from each database or resource. The results from a resource or "target" that responds to the request first - usually the Libraries' catalog - are displayed first. Relevancy ranking is not an option and this was commented on during the usability session. It was not clear to our users whether phrase, proximity or Boolean searching was allowed within the search box. From users' feedback, the task force learned that

users thought that the results returned limited information. Though the results include title, author, description, URL, source and database, this was not seen as enough information.

In the search progress section of the search results screen, the resources that are listed are displayed along with the number of hits from each resource, the number retrieved and the status. Users were not immediately certain about how to interpret this information. It was suggested that it would be helpful to be able to view results by database and to indicate which results included full text. Also, users found the flickering web page during search result processing to be distracting. Finally, users said they wanted to be able to return to their original search to refine it, and that they felt that the advanced search screen was too long.

In addition to the usability session, a five question online survey linked from the "What Is MetaFind?" page gathered users' responses. There were a total of 104 surveys collected from November 2004 through November 2005. The survey asked users to respond to the following multiple choice questions and to offer general comments.

1. Status: Undergraduate, Graduate, Faculty or Staff.

2. Frequency of use of library online resources: This is my first try; Once or twice per term; Once a week or more.

3. I want to search multiple databases: Because I want to know which database is best for my search; Because I only need a few things and this is the fastest; I don't want to search multiple databases.

4. This is an improvement over other searches: Yes; No.

5. An open comments field.

Responses to the first question showed that respondents were fairly evenly distributed between faculty (29), graduate students (30) and undergraduates (33). Twelve respondents selected unknown. Undergraduates were underrepresented in the survey results..

The frequency of using the Libraries' online resources was also evenly distributed. Thirty-six respondents used library online resources for the first time, 22 used them 1-2 times per term; 36 used them once per week and 10 did not respond to question two. In retrospect, it became clear that the "library online resources" phrase in question two could refer to either the Libraries' databases and other electronic resources or to MetaFind or to both. The survey authors' intention was to measure use of online resources generally and then compare this to results from question four to attempt to determine whether users saw MetaFind as an improvement. Given the ambiguous meaning of question two, definitive conclusions about the frequency of online resource use could not be drawn.

In response to question three, most respondents (42), indicated that they did not want to search multiple databases. Other respondents used federated search to identify a database for their research (38) or to quickly find a few resources (24). See Table 1: Reason for Searching. Unfortunately, the wording of this question was also slightly ambiguous. It may not have been clear to respondents that "multiple databases" refers to the federated search tool rather than searching multiple databases one at a time.

TABLE 1: REASON FOR SEARCHING

Don't want to search multiple databases	42
Want to know which database is best for my search	38
Need a few things and this is the fastest	24

Although many survey respondents, 44 of 104, said this method of search was an improvement over other searches; a significant number, 33, said that it was not. Users expressed

their reasons for why this method was not an improvement in the comments, which are explored

below. Twenty-seven respondents gave no response to question four.

Respondents became more satisfied with the product as time passed. During the first half

of the survey period, 45% of respondents indicated that MetaFind was an improvement over

other search mechanisms. During the second half of the survey period, 81% of respondents

indicated that MetaFind was an improvement. See Table 2.

TABLE 2: IMPROVEMENT OVER TIME

Nov'04	Yes	0	No	0	No response	21
Dec'04	Yes	4	No	3	No response	3
Jan'05	Yes	8	No	10		
Feb'05	Yes	6	No	6		
Mar'05	Yes	3	No	7		
Apr'05	Yes	2	No	2		
Total Nov'04-Apr'05	Yes	23	No	28	No response	24
May'05	Yes	10	No	0		
Jun'05	Yes	2	No	0		
Jul'05	Yes	3	No	0		
Aug'05	Yes	1	No	0	No response	1
Oct'05	Yes	1	No	2		
Nov'05	Yes	4	No	3	No response	2
Total May'05-Nov'05	Yes	21	No	5	No response	3
Total '04-'05	Yes	44	No	33	No response	27

The final question of the survey asked respondents for comments. During the time period

surveyed, just over 70 comments were made. Comments ranged from a simple "Thank you" to

more in-depth analysis of the search such as, "Every time I use this resource, I like it better.

Although its journal coverage is seeminly [sic] more narrow than many individual article

databases (i.e., fewer hits), it gives an interesting check on gaps of other databases."

Comments also reflected the respondents' level of experience with the Libraries'

resources such as, "Very nice interface...a few glitches, but overall a neat resource. I can't access

all of the estimated number found - is that just a limitation of the technology? It would be great

to be able to search for a given phrase in all of the library's resources at once (e-journals, etc...)

even if it took some time for the operation to complete... but maybe there are too many variables

there.”

FUTURE DIRECTIONS

When MetaFind was selected in 2003, the Libraries recognized that federated search

technology was not fully developed and that standards for federated search tools had not yet been

created. Using this product was a temporary solution that quickly enabled the Libraries to offer

federated searching of databases and the Libraries’ catalog. Experience with the product also

familiarized staff with the technology and enabled the Libraries to consider some of the options

for organizing and selecting resources to be offered in a federated search.

A prominent feature of the OSU Libraries Strategic Plan (2004) is for the Libraries to be

“as easy to use as Google and other search engines” and to “provide faculty and students with the

information they require - whenever and wherever they require it” (Oregon State University

Libraries 2004). A key action plan in the first year included the creation of a cutting edge digital

library with a focus on enabling quick, easy and transparent access to the Libraries’ resources.

This focus of the Libraries’ strategic plan and the deficiencies of commercial federated search

tools such as MetaFind prompted the Libraries to experiment with the implementation of a

home-grown federated search tool and link resolver. “LibraryFind”, the Libraries’ homegrown

federated search tool, currently in beta, provides an easy to use search interface that combines

interface and design flexibility, the ability to interface with other systems, access to selected

local and licensed information, and continuous improvement based on use metrics, ongoing

usability studies, and user surveys (Reese 2006) (Frumkin 2006).

Google and Yahoo! have quick response times. While libraries are not yet able to match

the speed that patrons are used to from commercial search engines, libraries have identified and

implemented several methods that cut search response times considerably. For example, when

licenses allow, OSU Libraries are harvesting content from data providers and storing it locally to

speed retrieval. With a single click LibraryFind makes full-text available from the search results

through the use of a locally-developed link resolver with which it interfaces. See Figure 4.

Improvements to search result displays are being made based on the usability data

collected during the implementation of MetaFind. Many of the changes are easily accomplished

with LibraryFind but were not feasible with the vendor software. For example, search results are

prominently listed on the screen – directly beneath the search box. Also, information about the

databases that are being searched and the number of hits from each database is easily viewed on

the right-hand side of the screen. This frees up the most valuable screen space in the center.

FIGURE 4: BETA VERSION OF LIBRARYFIND
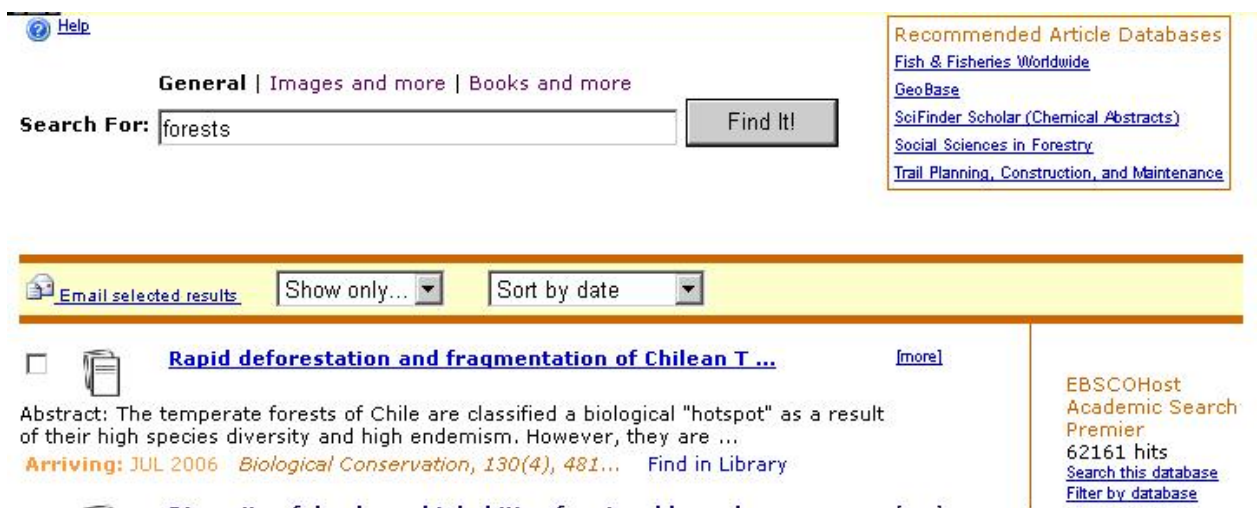
LibraryFind displays search results using a locally developed relevancy algorithm that continuously self-improves based on retrieval analysis. The algorithm will be revised based on usability study results, database, journal and article-level quality indicators, individual preference or based on patron type. Searches can also take advantage of pre- and post-coordinated material type (articles, books, newspaper articles, images) filters.

VALUE-ADDED SERVICES AND CONTINUOUS IMPROVEMENT

The Libraries will continue to conduct usability studies and will constantly refine the LibraryFind software based on the collected data and use analysis. Services will be added to take advantage of future technological innovations. Examples of features that are currently being developed include bookmarking of search results, citation exporting, and search query spell checking.

LOCALLY-PRODUCED DIGITAL COLLECTIONS

Even though the Libraries make their collections of locally-produced digitized content and born-digital content easily harvestable via Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), they are still required to pay to incorporate these local collections in MetaFind. LibraryFind freely includes this locally produced digital content, presenting users with information from the institutional repository, ScholarsArchive@OSU, and various digital collections alongside citations from databases and the online catalog.

CONCLUSION

The Libraries' experience with MetaFind has been an important step in offering federated search to its users. Despite limitations of the software, the implementation and use gave staff and users important experience which has fostered the creation of new tools. The Libraries anticipate that the implementation of the locally developed federated search tool and link resolver will provide the ease of use that our users have come to expect from search engines and will promote enhanced access and increased usage of library materials.

REFERENCES

Boock, Michael, Laurel Kristick, and Mary Caughey. 2006. Implementation of MetaFind at Oregon State University. Presented at the Northwest Innovative User Group Annual Conference, October 19, 2004, in Portland, Oregon. http://hdl.handle.net/1957/223 (accessed May 12, 2006).

Chaffin, Nancy, Kevin Cullen, and George R. Jaramillo. 2005. Cross-Database Searching: The Implementation of MetaLib. *Technical Services Quarterly* 22 (4): 39-53.

Colyar, Nancy. 2005. The Process of Selecting Federated Search and Link Resolver Products. *Louisiana Libraries* 68 (2): 13-14.

Curtis, AnneMarie, and Daniel G. Dorner. 2005. Why Federated Search? *Knowledge Quest* 33 (3): 35-37.

De Rosa, Cathy. 2005. *Perceptions of Libraries and Information Resources: a Report to the OCLC Membership*. Dublin, Ohio: OCLC.

Feeney, Mary, and Jill Newby. 2005. Model for Presenting Resources in Scholar's Portal. *portal: Libraries & the Academy* 5 (2): 195-211.

Frumkin, Jeremy. 2006. MetaSearch. Presented at Oregon State University Libraries Management Retreat, February 7, 2006, in Newport, Oregon.

Gerrity, Bob, Theresa Lyman, and Ed Tallent. 2002. Blurring services and resources: Boston College's implementation of MetaLib and SFX. *Reference Services Review* 30 (3): 229-241.

Helfer, Doris Small, and Jina Choi Wakimoto. 2005. Metasearching: The Good, the Bad, and the Ugly of Making It Work in Your Library. *Searcher* 13 (2): 40-41.

International Federation of Library Associations and Institutions. 1998. Functional Requirements for Bibliographic Records - Final Report. http://www.ifla.org/VII/s13/frbr/frbr.htm (accessed May 12, 2006).

Madison, Olivia M. A., and Maureen Hyland-Carver. 2005. Issues in Planning for Portal Implementation: Perfection Not Required. *Journal of Library Administration* 43 (1-2): 113-134.

Open Archives Initiative. 2004. The Open Archives Initiative Protocol for Metadata Harvesting. http://www.openarchives.org/OAI/openarchivesprotocol.html (accessed May 12, 2006).

Oregon State University Libraries. 2004. *Oregon State University Libraries Strategic Plan 2004.* http://osulibrary.oregonstate.edu/strategic_pw.pdf (accessed May 12, 2006).

Reese, Terry. 2006. Metasearch: Building a Shared, Metadata-driven Knowledge Base

System. *Ariadne* 47. http://www.ariadne.ac.uk/issue47/reese/ (accessed May 12, 2006).