Meta Search in Distributed Electronic Resources : A Study

M Krishnamurthy

Abstract

The explosion of information on the Internet and information technology in general has created challenges for libraries to focus on developing more effective ways to meet the information needs of users. One practical approach is through customized portals performing simultaneous database searching. This paper presents a new library portal service used at the Indian Statistical Institute Library at Bangalore (ISIB). The key feature of a library portal is to allow searching across multiple databases without having repeat search. This feature is generally referred to as meta-search, parallel search, broadcast search or federated search. Discussion also includes strategies of local customization and the impact on library management in an electronic environment.

Keywords : Electronic Resources, Portal.

0.   Introduction

The explosion of information on the Internet and information technology in general has created challenges and more opportunities to information professionals for redefining their roles for the present and future. Web based information services are receiving much attention from the library community. With the advent of Internet and more specifically the World Wide Web, libraries have undergone a revolution in the way that they build information source collection, operate and provide information services to users. The information professionals are uniquely positioned to play a major role in the emerging information age driven by convergent technologies.

1.   Library Portals

While information is readily available, overwhelming at times, on the World Wide Web, libraries are focusing more into developing more effective alternative ways for users to efficiently navigate among the library resources distributed on library's Websites. One of the approaches to accomplish this goal is through customized portals

performing simultaneous database searching. As Strauss has noted that a portal as a special kind of gateway to Web resources" a hub from which users can locate all the web content they commonly need" (Strauss 2003). The design and implementation of local portal will be the responsibility of individual institutions. Local portals will allow tailored access to a selection of data sets of importance to a particular institution, plus integration with other locally licensed data sets and local products, abstracts and indexing and citation databases, which user has access collection of electronic journals including OPAC (Webb, 1998),

The content, structure, and relationship of information in the information space are most commonly represented as alphabetically based or hierarchical categories. Because of information overload on the Web, hypertext-browsing abilities in locating information may be significantly augmented by content based searching (Chen, *et al*, 1998)

The key feature of a library portal is to allow searching across multiple databases without having repeat search. This feature is generally referred to as Meta search, parallel search, broadcast search or federated search.

2.   Model of Access at the Bangalore Center Library

2.1   ISI Bangalore Library

The Indian Statistical Institute (ISI) has three centers at Calcutta, Bangalore and New Delhi. The Library at the Bangalore Centre (ISIB) established in 1978, has a good collection on Statistics, Mathematics. Quality Engineering/ Management and Library and Information Science it) India. Presently it subscribes to about 300 international journals for its 300 library users including faculty, students and research scholars. The overall collection of books and reference materials, reports, reprints, directories, encyclopedias, CDs etc., is approximately 50,000.

The ISIB Library provides from its homepage Web listings to its journal resources, separately linked according to procurement agreement with content providers (Fig. I). They are, namely, ISIB Online Subscriptions, ISIB Periodical Holdings; *Science Direct* consortium, Union Catalogue of Publications in Bangalore Libraries.

2.2   Web Access to ISIB E-Journal Subscriptions

The first listing is the -E-journals available for ISI Bangalore." The ISIB Library has a total of 278 online journal subscriptions since 2000. This listing, written in html format, provides access to ISIB licensed electronic resources either to the full-text articles or to the abstract of the articles. Linking data to these electronic resources are embedded in this hypertext file. They include such data as the details of the URL path and sometimes the session information needed to establish authorization at the vendor/publisher server. Through this dynamic linking, access to the Library services is enhanced with immediacy, accuracy and currency of information, which are the most important factors in fulfilling user information needs.

2.3   Web Access to ISIB Periodicals Holdings

The second listing is the "Periodicals holdings of ISIBC Library-2001." Tile ISIB periodical collection has a total of 412 records of holdings since 2001. This listing, also written in html format, provides access to the ISIB journals alphabetically by the first letter of the title word. It includes publisher information and the beginning date of the periodical holding, and bibliographic notes such as the former and/or later titles. This Web list is continually updated in the source code by library staff to incorporate new titles when they are added to the library.

2.4   Web Access to Science Direct Consortium

The third listing is the "Science Direct Consortium: Online Access to Elsevier Science and Academic Press Journals." In 2003, a contract to use Elsevier Science Direct electronic publications was signed by the Indian Statistical Institute to establish a consortium networking the three IST Centre libraries. The consortium was formed to contribute to the development of research through the acquisition of electronic publication to the participating libraries of Indian Statistical Institute using the LAN, in addition to their individual current journal collection. In this consortia gateway, the ISIB Library maintains a total of 133 online journal subscriptions. They consist of consortium subscription to e-Journals from various Elsevier publishing groups such as Pergamon, Saunders and Academic Press.

2.5   Web Access to Union Catalog of Publications in Bangalore Libraries

The fourth listing is the "Union Catalog of publications in Bangalore Libraries." In 1997, a project to promote access to journals available in Bangalore was launched to create a union catalog of current journals subscribed by the major libraries in the region. The result is a formation of a union catalog of current journals compiled by the National Centre for Science Information (NCSI). The catalog consists of five separate annual listings from 1997 through 2001. It provides bibliographic information as well as the holding library information. It also provides a subject search to the journal titles. However, each annual listing is searched independently and the title word searching is limited by adjacent search only.

3.   Analysis of Need

The cooperative environment of library resources has created an environment that the ISIB Library can serve with a host of journal resources available in the region. The development of the Library's homepage reflects the active participation of each component of the Web listings. However, users often need to navigate among and between the various Web pages to locate the journal holding or full-text articles in question. Since the search is mainly conducted as a browsing operation, knowledge of title words is required as key word or truncated searches are not possible. There are also issues of high maintenance for the upkeep of these hypertext-markup lists,

Perhaps the biggest difficulty is the lack of normalized data across the various Web listings. For example, forms of abbreviation were not standardized for name of the journal title, publisher, place of publication, or library code of ownership. As a result, the search, regardless of how thoroughly conducted, would be futile and meaningless. As electronic and print resources are increasing to overlap in various journal collections, the best practice for providing an efficient and accurate access is inevitably through a library portal that is customized to navigate simultaneously across tile various lists.

4.   Technologies and Strategies for Portal Customization

4.1   Normalization of Metadata in a SQL database

The Web pages at the ISIB homepage were created to display individual listings of journal resources. In order to achieve normalization of data, a MS Access database was created as the main resource registry. Four tables were then constructed sharing many identical metadata field names and information, such as Title, ISSN, Place of Publication, URL, Subject, etc.

Data from the Web lists at the original ISIB site were downloaded as sources to populate the corresponding registry. For the union catalogs of current serials from 1997 through 2001, the five years data were merged into one registry. A Visual Basic application was composed to eliminate duplicates and consolidate the holdings for multiple years.

Spelling and use of abbreviations were standardized through the registry. This is a crucial feature for producing accurate and meaningful results for the end users.

4.2   Programming for Simultaneous Searching

As searches are performed in multi-user situations, the best practice is to install the metadata registry over a MS SQL server, which allows better transaction throughput over the Internet (Mischo and Schlembach, 1999).

Unlike the Library's original Web access to the library resources, the new Web applications are typically composed of several Active Servers Pages (ASP). The embedded VB scripting commands process and parse the user search strings into appropriate SQL commands to retrieve data across the four tables defined in the metadata registry.

The new information portal provides a gateway to the entire spectrum of library journal resources. However, it also allows end-users to choose among them as desired for more efficient search strategy. Fig. 2 shows the newly designed ISIB Library search aid as a gateway to its journal resources.

The entered search term, computes and retrieves a number of matches simultaneously from a pre-selected list of information sources, i.e., in this case all are selected: Journals at ISIB Library, *Science Direct* Journals at ISIB Library, Periodicals Holdings at ISIB Library, and Union Catalog of Bangalore Libraries. Users may proceed to choose from here with one single click to locate more details about the journals of interest.

5.   Evaluation of the New Library Portal Service

With the advent of digital libraries and of wide area network, enormous amount of textual information is made available all over the world. Searching and browsing are the two resource discovery paradigms mostly used to access this information. To improve access to the information stored in our library portal we incorporated a library search aid.

5.1   Efficiency of Searches

The advantage of simultaneous searching across multiple sources of library information is evident. For users, the new library portal design presents a gateway to navigate fruitfully from a simple step rather than making a series of search attempts indiscriminately across the previous Web lists. The biggest improvement of this approach is exemplified particularly in searching the 5-year union catalogs into one database, as it eliminates the need to make multiple searches one at a time as required previously.

With the new approach, collection scope is well defined for users to make intelligent decision for making their selection. Resource sharing is better served as interlibrary loan staff can easily determine from where their request is best to be filled.

5.2   Maintaining and Updating Issues

One of the major concerns in providing library resources is the maintenance issue. Information is only as good as it is up-to-date. Consequently, the foul Web listings on the ISIB Library homepage present several challenges in maintaining the information current. For example, new title information must be updated in hypertext language. It is also quite cumbersome to make changes, such as library holdings, to the existing files.

In contrast, the new library portal performs a uniform SOL searching across a database with uniformly defined metadata tables. Each table has identical field structure so that information can be kept in standardized format. Normalization of title words, library codes and abbreviations improves the accuracy of search matches as well.

6.   Conclusion

The basic concept of the library portal is to provide databases, and localization. It is therefore necessary for the library to organize the information in order to reflect institutional or consortia information resources licensed by and made available to the users. The efforts to enhance access involve both public as well technical services in order to maintain a successful performance. We feel that good planning should be low maintenance but highly efficient one. Our solution to present a library portal service achieves this goal with ASP technologies, a registry accessible over a SOL server, and a uniform metadata structure of tables to produce effective simultaneous searching of matched results.

7.   References

1.   Schatz, B.R., et al. Internet browsing and searching. User evaluations of category map and concept space techniques, Journal of the American Society for Information Science. Vol.49; 1998; p582-603,

2.   Mischo, W.H. Library portals, simultaneous search, and full-text linking technologies. Science and Technology Libraries. Vol.20; 2001; p133-147.

3.   Mischo, W.H.; Schlembach, M.C. Web-based access to locally developed databases. Library Computing. Vol. 18: 1999; p51-58.

4.   Strauss, H. Web portals: the Future of Information Access and Distribution, The Serials Librarian. Vol.44: 2003; p27-35.

5.   Webb, J.P. Managing licensed networked Electronic resources in a university library. Information Technology Libraries. Vol.17; 1998; p198-206.

About Author

Dr. M Krishnamurthy   is working as Librarian in Indian Statistical Institute, Bangalore.
E-mail : krish@isibang.ac.in, mk13murthy@hotmail.com