Journal of Knowledge & Communication Management
Volume 3 Number 1 April 2013 pp. 70-81
DOI: .....



# Aggregation of Library Resources in the Digital World: A One-Stop Shop

**Ashutosh**

*National Institute of Health and Family Welfare (NIHFW), Munirka, New Delhi-110067*
*Email: ashutoshkatyayan@yahoo.com*

**ABSTRACT**

*With the arrival of the network and digital age, the library's holdings of resources have undergone a drastic change in which the proportion of digital resources occupies an instant increase. This article discusses the need for integration of distributed islands of knowledge, i.e., online journals, scattered OA literature, digital repositories, websites, blogs, OA journals, etc., to provide a one-stop shop access to users with the use of existing technologies and techniques. Several integrated models of digital resources that are suitable for libraries have been proposed.*

**Keywords:** Integrated Library Resources, OA Journals, Open-Access Literature, Federated Search, Integrated Search



## 1. INTRODUCTION

Library and information services as well as applications have increasingly become web compliant with libraries, with publishers and information service providers offering browser-based services. However, although information services offer innovative features, it is often found that end users have to put in considerable efforts to find their way around different interfaces and facilities.

Library users face several key impediments to completing tasks involving library resources. Users often need to know which kinds of resources exist or are relevant to their task before they can access them. Once users have some information, they may have difficulty finding related information located outside the current system being searched. These impediments arise because library resources are not adequately integrated. Users need to go to specific resources individually to find information.

This concept of aggregation is not new to librarians, who have been aggregating content into library collections for centuries. Anecdotal evidence suggests that researchers are finding it difficult to find the appropriate resource to support their work from an increasing number of digital and analogue resources (Gow and Roberts, 2003) where on the one hand, there is an increasing range

of information resources in different formats, and with different permissions to access it, and on the other hand there is a loss of standardisation in search interfaces, which clearly causes confusion among users.

A US study has revealed a phenomenon that Brian Hawkins, formerly President of Educause, called "the tyranny of the two searches", according to which users are either too lazy or too ignorant to search for information from more than a single source. However, they tend to use the Internet as the information resource of first resort, and when they fail to find the information they require, they usually do not undertake further searches using other resources like library OPAC (Online Public Access Catalogue) or other relevant databases. Therefore, there is a need to provide users with resource-discovery tools to find accredited and quality information resources, which concurrently match the ease of use of different search engines. This implies the use of a single search interface across multiple and different resources such as OPACs, subject gateways, journals databases as well as single standardised environments within which to view and manage search results.



## 2. PATRONS' NEEDS

Library patrons typically need access to both library holdings and other subject-specific information resources, e.g., open-access publications. They are usually comfortable using different search tools, but have difficulties in identifying and locating the appropriate resources (Herm and Volz, 2008). The interfaces are getting more complex with heterogeneous services offered through web pages, journal sites and repositories among several others. Users usually find it difficult to invest time and efforts in learning all the different options available on these heterogeneous services. Furthermore, it is a well-known issue that patrons are neither satisfied with the interface and retrieval features of the search component of traditional online catalogues nor with the presentation of search results. In addition, scientists and academics want fast and easy access to as many relevant full-text publications as possible.

The scholarly community faces a variety of problems in the discovery, access and delivery of pertinent information/data in their respective areas of interest. Users want specialised portals, interdisciplinary links, intelligent integration of all types of information resources, inclusion of quality information and single-point access to information (Munsi, 2009).

## 3. CHALLENGES/PROBLEMS FOR LIBRARIANS

Librarians have been among the first to recognise that the Internet is not a digital library giving simple and immediate access to information world-wide as some of the apologists of the net believe even today. Moreover, librarians rather quickly recognised the shortcomings of the major search engines for scholarly use (Enderle, 2000).

The need for librarians to help make sense of interoperable digital information by provisioning resources with care and quality metadata and by connecting users to resources-and resources to

resources-is greater than ever. In order to capitalise on these technologies, librarians must first understand them and be able to relate them to the professional practice of librarianship (Witt, 2010). Challenges can be categorised as follows:

- **Variety of Resources:** Resources are not only growing in quantity but the variety of resources is also changing; the patron gets confused where he/she should start searching for his/her desired information. Library professionals are supposed to subscribe and keep those information resources that can fulfil an organisation's needs and objectives. It becomes a great challenge for librarians to select those resources that are cost effective, authentic, easy to access and can provide better search options to the clientele.

- **Dispersed/Scattered Information Resources:** Libraries are supposed to have different types of resources in their collection to fulfil a patron's needs. The resources may be a catalogue, e-journal database, subject gateways, databases of other libraries, open-access resources, etc., and these resources are scattered on different websites and networks. It becomes a great challenge to pool all information resources at one place and provide seamless access to all of them simultaneously.

- **Changing User Needs:** As technology is changing, the expectations of users from libraries are high. Almost every library user comes with a set of expectations that is defined by their experience of using the web. Users want to access all information resources at one place, so that they can search every source of information within one search. No one wants to search different resources, websites and portals for their required information; they prefer to search only a federated search interface.

- **Leapfrogging Technology:** Information technology continues to advance exponentially. It becomes very difficult for libraries to adopt the latest technology due to both financial and physical resource constraints because employees do not easily change their way of working.

Information managers/librarians have to equip themselves with capabilities needed to link with global trends for the ultimate benefit of information seekers. Professionals can join the revolutionised technological race without sacrificing their conception of traditional libraries. Information providers have to keep themselves abreast with the latest trends if they desire to survive in the present world.

The main challenge for librarians is to provide access to Internet resources and to integrate them into the main function of the library to provide a "one-stop shop" interface.

## 4. SOLUTION: INTEGRATING LIBRARY RESOURCES AND SERVICES/PROVIDING COMMON PLATFORM FOR ALL RESOURCES

### 4.1 Library Website

A library website is a single-user interface for accessing a wide variety of electronic resources both within and outside the library. The importance of any website lies in the currency of information

provided, in its ability to locate information of high relevance and in the provision of a powerful search engine with instant access to full text information. All resources and services can be put on the library website so that users can access all the resources from a single platform.

### 4.2 Consortia Approach

A consortium is nothing but aggregator software, which provides easy access to licensed library electronic full-text resources for library clientele.

Basically, a consortium acts as a mediator between publishers and subscribers for providing online access. They manage electronic resources and provide online access to journals published by various publishers. Aggregators provide access to thousands of online journals from leading scholarly, academic and business publishers, and provide fast and reliable access from a global network of servers to users' desktops around the world (Halijwale, Manjunath and Pujar, (2004).

### 4.3 Federated Search

Information integration over distributed sources is an urgent problem to be solved for providing access to a variety of resources through a common search interface and portal.

The growth of different types of databases, produced by different suppliers, with numerous interfaces and logins means that library users find it confusing when attempting to access information. Library OPACs and web pages have been alienating users with their use of library terminology and by including long lists of databases that users find difficult to select from and search. The needs and expectations of library users, particularly students using academic libraries. The growth of different types of databases, produced by different suppliers, with numerous interfaces and logins means that library users can find it confusing when attempting to access information (Kumar, Sanaman and Rai, 2008).

Federated or cross-database search tools are the correct solution for unifying access to a variety of information resources. It is also known as meta-search, parallel search, broadcast search and a variety of other names; federated search allows users to search multiple different databases from a single interface. These tools can search not only library catalogues but also commercial abstracting and indexing databases, web search engines and a variety of other databases. The query is run through all of the selected resources, and an integrated set of results is returned.

Federated searching resolves this issue and makes these different resources searchable without having to visit each resource individually.

dbWiz is reSearcher's open source federated search tool, with the ability to search all major research databases. dbWiz offers a basic search option, creating a simple "Google-like" search experience, where users enter their keywords and select a search category that has been populated by the library with the most relevant resources to that topic (Stranack, 2007).

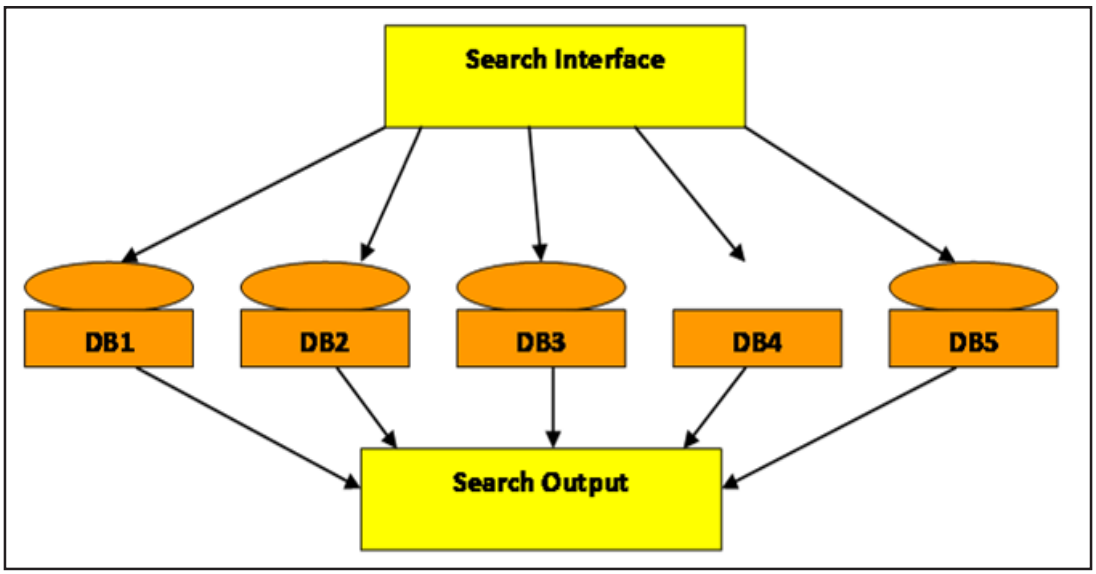


**Fig. 1: Federated Search**

### 4.4 Metadata Harvesting

Since open-access literature is not hidden behind technical access barriers (such as IP restriction), its full text can be indexed and made accessible by conventional search engines; however, although very powerful, these search engines do not offer precise field-based searching of particular bibliographic elements, such as author. The e-prints in a digital archive or institutional repository are described by metadata records (typically in Dublin core format), which provide such information and, using the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), external search systems can retrieve this metadata using a standard protocol, can combine it with metadata from other archives and repositories, as well as can create composite databases that allow users to retrieve information about e-prints from diverse archives and repositories as well as retrieve the full text of e-prints of interest from those systems. Unfortunately, conventional author home pages on the web do not have this capability, and this is a limitation of self-archiving e-prints in this manner.

### 4.5 Web 2.0 Applications

Web 2.0 applications are transforming the web into a space that allows anyone to create and share information online-a space for collaboration, conversation and interaction, a space that is highly dynamic, flexible and adaptable. Web 2.0 services are increasingly becoming embedded in libraries, and librarians all over the world are using Web 2.0 technologies to promote services, share information, as well as engage with users and network with colleagues on a global scale. Most

Web 2.0 services offer options to embed content, either to bring the information into your website or to link up to other Web 2.0 sites, e.g., to integrate your Flickr photos or tweets within your blog. Web 2.0 services can be applied in libraries in many ways, i.e., to communicate, to share, to publish, to collaborate, to integrate, to promote, to create awareness, etc.

### 4.6 Using Web Crawlers

The world-wide web (www) is decentralised, dynamic and diverse; navigation is difficult, and finding information is a challenge today. Web crawler is a tool that solves these problems by indexing and automatically navigating the web. It provides a method to integrate access to website content. Selected websites can be crawled and a searchable index of the web can be created to provide integrated access to users.

### 4.7 Integrated Access to OA Literature

The open-access research literature is composed of free, online copies of peer-reviewed journal articles and conference papers as well as technical reports, theses and working papers. In most cases, there are no licensing restrictions on their use by readers. Therefore, they can be used freely for research, teaching and other purposes. Swan (2005) in "JISC Briefing Paper on Open Access" has described the importance of OA as "there is accumulating evidence that shows that research articles that have been self-archived are cited more often than those that have not. Across most subject areas there is at least a twofold increase in citation rate. In some subject areas it is even higher. This form of Open Access means that research has much more impact than before. Moreover, the research cycle – where work is published, read, cited and then built upon by other researchers – is enhanced and accelerated when results are available on an Open Access basis".

Providing access to open-access materials has inherent challenges similar to those of other freely available digital works on the Internet. The tasks associated with managing and exposing library patrons to free and open-access electronic resources are every bit as complicated as managing more traditional subscription-based electronic resources.

The problem is "how to expose open-access material to the users." Libraries have to face a number of challenges, which includes cataloguing (or otherwise creating metadata) of OA materials from a pool resources; difficulties in tracking changes in dynamic OA materials; monitoring their availability; lack of adequate coverage of OA materials in indexes, aggregator databases and other conventional finding tools (Bailey Jr, 2006).

Possibly the most prevalent solution is to simply load MARC records into the library's catalogue. However, libraries are leery of loading free resources into their catalogues because of the dynamic nature of OA materials. Hence, it should be made sure that free and open-access resources are stable before their inclusion to the library catalogue.

### 4.8 OpenUrl Link

Another method of managing access to free and OA resources is the OpenUrl link resolver. This is arguably an easier and more efficient way to manage such resources. OpenURL is a mechanism for transporting/encapsulating citation metadata and identifiers describing a publication for context-sensitive linking. OpenURL is a protocol for interoperability between an information resource and a service component, often referred to as a link server, enabling a link to lead to the desired resource.

The link resolver matches metadata from a citation against a database of the library's holdings, often called a knowledge base, and offers the user a menu of links to full text and/or other services (e.g., a document-delivery request form). Then, the link resolver can link the user directly to the desired article or journal from a citation, without requiring the user to visit the library catalogue. If all of the library's journal holdings, including those from aggregated databases and some free titles, are loaded into the resolver, it can provide another form of one-stop shopping (Yu and Breivold, 2008) .

Most commercially available link resolvers are SFX, 360 link from Serials Solution, EBSCO's OpenURL link resolver and WorldCat Link Manager from OCLC. SFX software allows users to link between various web resources such as citation databases, journal articles, web search engines or document-delivery services.

### 4.9 Customised Search Engine (CSE)

In the Internet age, users want information packaged like fast food: served instantly, hassle-free and in bite-sized morsels. Owing to technological advances and the wide dissemination of information, users suffer from information overload and expect their libraries to select the best and organise it effectively for their personal consumption. The world of library technology is changing quickly, driven by fast alterations in the information technology sector. This trend has intensified since the triumph of search engine technology. CSE is a tool used to search a webpage, website or a collection of websites. It can also be used to create a subject guide or pathfinder for resources on a given topic, which are available on the free web. CSE narrows your focus and eliminates quite a bit of the extra information usually contained in search results.

### 4.10 Listing out OA Journals in Different Knowledge Domains/OAJ Portal

The tremendous growth of OA journals has forced libraries to develop means of providing access to these popular resources. Over the past 20 years, methods to connect users to OAJs have taken different shapes fluctuating among a plethora of theories, ideologies and technologies. There are two methods for accessing OAJs. These are open-access journal portals and open-access platforms. These methods act as highly useful access points for retrieving relevant information by the scholarly community (Meera and Ummer, 2010).

### 4.11 Subject Gateways

Since a large amount of information is published on the Internet, users are finding it more and more difficult to access 'requisite bit of information'.

The recent years have seen a number of efforts attempting to leverage information overload. Although the most obvious aspect of information overload is quantity, the problem of information quality has started to receive special attention, particularly when treating the Internet as a source of educational or professional information. Information quality attempts to answer the question: how can one select from "everything" only those information items that meet one's information needs and at the same time carry a certain validity or authority (Zygogiannis, Papatheodorou, Chandrinos, and Makropoulos, 2009).

There are two professional ways for making subject-specific Internet resources accessible: by implementing limited-area search engines like customised search engines or by creating a subject-specific gateway. Both approaches can be combined and, this may, perhaps, be the solution for the future. The most promising way is the concept of subject gateways. Subject gateways are one such tool designed for a specific user group with interests in a specific subject area.

It provides the only opportunity for a user to get an overview of the existing Internet resources of relevance for his specific research field rather quickly.

Hence, subject gateways can be defined as Internet services, which support systematic resource discovery, providing links to resources (documents, objects, sites or services) predominantly accessible via the Internet. The service is based on resource description. They are typically databases of detailed metadata (or catalogue) records, which describe Internet resources and offer a hyperlink to other resource (Das, 2001).

### 4.12 Reinvent the OPAC Having Integrated/Federated Search

Traditionally speaking, the OPAC is a tool allowing people to search for and identify things in or available from libraries. Today, people's expectations regarding search and access to information have dramatically changed.

To cater to users' dramatically changing expectations, libraries have to redesign their OPACs by adding new features and the latest technology to provide user-interactive search interfaces. A number of approaches are being used by different library software. Browsing, faceted navigation, relevance ranking, did you mean, recommendation feature, Web 2.0 applications (tagging, RSS), book jacket images, search history, search within results, manage references and create bibliography are some of these features.

Endeca, a product from commercial market, provides few outstanding features like "most popular" sort option, "relevance ranking", "spell correction" and the "more titles like this" features. Polaris

provides an ILS with features such as faceted navigation, relevancy ranking, "did you mean" feature, enriched content and book jacket images.

The MIT Library Catalogue provides Save/Email search results (Fig. 2) and the Search Within Search (Fig. 3) features through which users can manage their references and refine their search results.

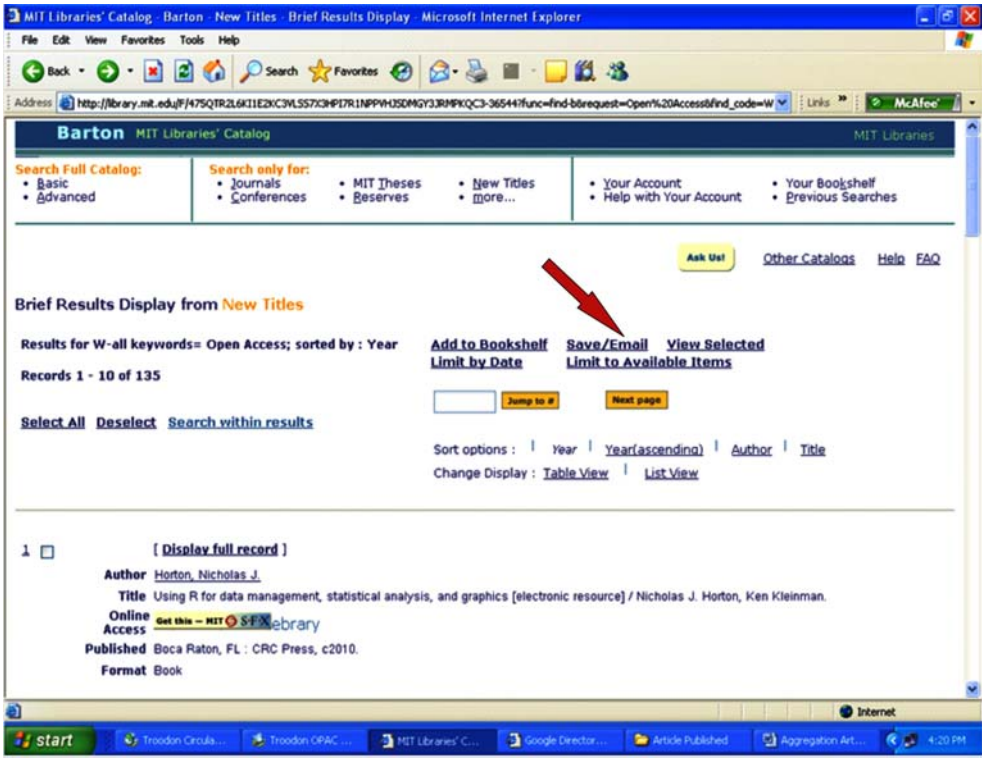


**Fig. 2: MIT Library Catalogue Providing the "Save/Email Search Results" Feature**

These catalogues with such interactive interfaces allow users to search for all materials through one search interface. These have been referred to as a Google type of interface and are viewed as a disservice to the user. These catalogues create a transparent system, enabling library users to get their work done more quickly and efficiently.

## 6. CONCLUSIONS

There is a diverse range of information resources in a variety of formats from a large number of differently organised collections and databases with different and sometimes idiosyncratic access

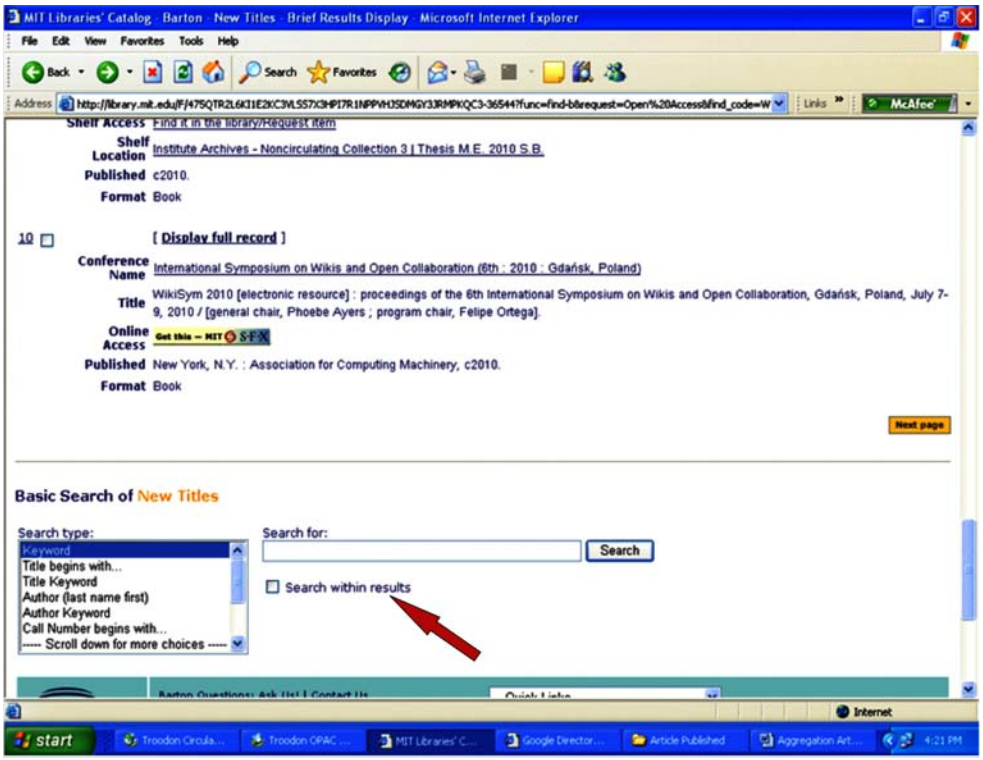


**Fig. 3: MIT Library Catalogue Providing the "Search Within Result" Feature**

protocols. The need to develop improved discovery and access 'tools' or mechanisms has become a high priority.

The main task for librarians in integrating Internet resources is to reconcile the demands for new technical and organisational solutions, which subject-specific websites require with the traditional function of a research library.

Users are often attracted to information systems and services that offer innovative features in browsing and searching collections. However, experience shows that usually users do not have the time and maybe the patience to learn complicated features on different information services, to extent that they only get used to some services and continue to use only those just to avoid having to learn and memorise how to navigate in different interfaces. The unified approach to different information services presented in this paper simplifies the procedure and provides access to different information in one interface. There are, of course, efficient tools and services offered in the commercial domain. Libraries are in need of deploying some open-source tools and toolkits

to offer a low-end and yet robust approach to build a unified searching system most suitable to academic, research and such other information systems that involve searching across heterogeneous information services.

It is clear that libraries need to understand the dynamics of service delivery and make sincere attempts to provide speedy access to scattered online resource from a single interface.

Libraries need to adopt and use existing open-source aggregating solutions (OpenUrl Link Resolver, OJS, CUFTS, dbWiz, Metata Harvester, etc.) to provide better, one-stop shop access to library resources.

## REFERENCES

Bailey Jr CW (2006). Open Access and Libraries. Collection Management 32(3/4) (2007): 351-383. Available at: http://digital-scholarship.org/cwb/OALibraries2.pdf (accessed 08 October 2010).

Das PK (2001). Subject Gateways: The Clever Way to Information. Available at: http://ir.inflibnet.ac.in/dxml/bitstream/handle/1944/107/cali_26.pdf?sequence=1.

Enderle W (2000). The Integration of Internet Resources into a Library's Special Subject Services: the Example of the History Guide of the State and University Library of Goettingen. LIBER Quarterly, 10, 342-366. Available at: http://webdoc.gwdg.de/edoc/aw/liber/lq-3-00/lq-3-00-s342b366.pdf (accessed 09 October 2010).

Gow E and Roberts K (2003). The AARLIN Project: Integrating the Research Information Infrastructure. Available at: http://www.aarlin.edu.au/publications/documents/IFLA.pdf (accessed 09 October 2010).

Halijwale SS, Manjunath GK and Pujar SM (2004). Electronic Journals: Modalities for Providing Free Access. Annals of Library and Information Studies, 51(1), 82-85.

Herm K and Volz S (2008). The Library Search Engine: A Smart Solution for Integrating Resources Beyond Library Holdings. ZIB-Report 08-37. Available at: http://opus.kobv.de/zib/volltexte/2008/1127/pdf/ZR_08_37.pdf (accessed 08 October 2010).

Kumar S, Sanaman G and Rai N (2008). Federated Search: New Option for Libraries in the Digital Era. International CALIBER-2008.

Meera BM and Ummer R (2010). Open Access Journals: Development of a Web Portal at the Indian Statistical Institute. The Electronic Library, 28(4), 540-554.

Munsi UM (2009). Building Subject Gateway in a Shifting Digital World. DESIDOC Journal of Library and Information Technology, 29(2), 7-14.

Stranack K (2007). The reSearcher Software Suite: A Case Study of Library Collaboration and Open Source Software Development. Serials Librarian, 55(1/2), 117-139.

Swan A (2005). Open Access: Briefing Paper. Available at: http://www.jisc.ac.uk/uploaded_documents/JISC-BP-OpenAccess-v1-final.pdf (accessed 07 October 2010).

Witt M (2010). Aggregating Web Resources. Available at: http://americanlibrariesmagazine.org/ore (accessed 20 August 2010).

Yu H and Breivold S (2008). Electronic Resource Management in Libraries: Research and Practice. Information Science Reference. New York, pp. 224.

Zygogiannis K, Papatheodorou C, Chandrinos K and Makropoulos K (2009). Automatic Web Resource Discovery for Subject Gateways. Available at: http://www.ionio.gr/~papatheodor/papers/NSCR-hcii.pdf (accessed 09 October 2010).

## About the Author

**Ashutosh** is currently working as Assistant Librarian at National Institute of Health and Family Welfare (NIHFW), New Delhi. He previously worked at Institute for Defence Studies and Analysis (IDSA), New Delhi. He holds his Master in Library and Information Science (MLISc) from IGNOU in 2009. He has qualified UGC-NET in the year 2012.

His area of interest includes information organisation, information retrieval, and knowledge management. His current article focuses on aggregated access to distributed information resources. Ashutosh can be reached at: ashutoshkatyayan@yahoo.com or ashutosh@nihfw.org.