Annals of Library and Information Studies
Vol. 59, September 2012, pp. 155-169

Strategies on e-resources management for smart information systems

M.G. Sreekumar

Center for Development of Digital Libraries (CDDL), Indian Institute of Management Kozhikode (IIMK),
IIMK Campus P.O., Calicut – 673570, Kerala, Email: mgsree@iimk.ac.in

There has been a paradigm shift the world over, the way how information and knowledge is being generated, stored, accessed, disseminated and consumed. Information and communication technology (ICT) is the game changer in this and it has propelled the growth of large numbers of electronic resources across the world. Organizations now spend a major portion of their budget for subscribing to electronic resources such as e-journals, e-books, bibliographic/full-text databases, online services and web portals. It is imperative on the part of the information professional to have a componentized and a multi-system approach to knowledge technologies and information management. Seamless aggregation and meticulous integration of diverse data streams embracing the traditional as well as the electronic information is the most appropriate strategy to be adopted and applied. Switching over to the technology pedestal shall ensure the information systems an exponential growth in the r-resource base, as well as facility to outreach to a wider community at a 24x7 formula. Information systems could also explore numerous resource base enhancement innovations such as the UGC Infonet, INDEST Consortium etc. and several Open Source Digital Library technologies towards making the library vibrant, resourceful, accessible and visible. More over with the advent of an array of ICT applications and with the upsurge and penetration of a vast variety of electronic content, the multitude of electronic information services will indeed attract and potentially influence the users prompting them to step in and stick to the changed scenario. As a measure of the impact on the scientific and research productivity, the increase in the overall scholarly intellectual capital of the institute could be matched against the improvement in research publications and the conferences, symposia, workshops etc. hosted by the organization. Quality improvement on the courses offered by the organization could also be considered.

Keywords: Scholarly Electronic Resources, E-Resources, Content Management, Information management, Digital Libraries, Institutional Repositories

Introduction

The way how information and knowledge is being generated, stored, accessed, disseminated and consumed has radically changed. Technology has been the game changer here and the developments in ICTs, networking and computing that paved the way. The resultant advent of e-publishing opened up several opportunities. There has been a substantial increase in the production of electronic content in the past 15 years. The proliferation of e-publishing across the world has revolutionised the domain of publishing, especially scholarly publishing[1]. The intensity of the impact is mainly directed by the Internet and the WWW. Recent research and developments in E-publishing offer ways to bring in a new equilibrium to the publication ecosystem, scholarly as well as trade. As result of the above, there has been a substantial increase in the digital collections and electronic information in the academic and research libraries. Among these, scholarly e-resources form the major chunk of electronic information sourced by organizations and they are growing at a phenomenal rate. To meet the ever increasing demand from the users, a major portion of the institutional budget is set aside for the subscription of e-resources. Additionally, institutions are getting access to a large number of e-resources through membership in various library consortia.

E-resources predominantly include e-journals, e-books, scholarly databases (abstract, index and full-text), e-reference works, online e-book libraries, value added information portals etc. (Fig. 1). Arranging subscriptions to the multiplicity of online resources that manifest in many different resource types, providing hassle free and uninterrupted access to the e-resources are among the major challenges before the information professionals. As the number of e-resources and e-services have sky rocketed, libraries are facing numerous problems and issues in the provision of their services.

The overwhelming number of web sites and URLs are flooding the library portals and library web pages, making it difficult for information professionals to

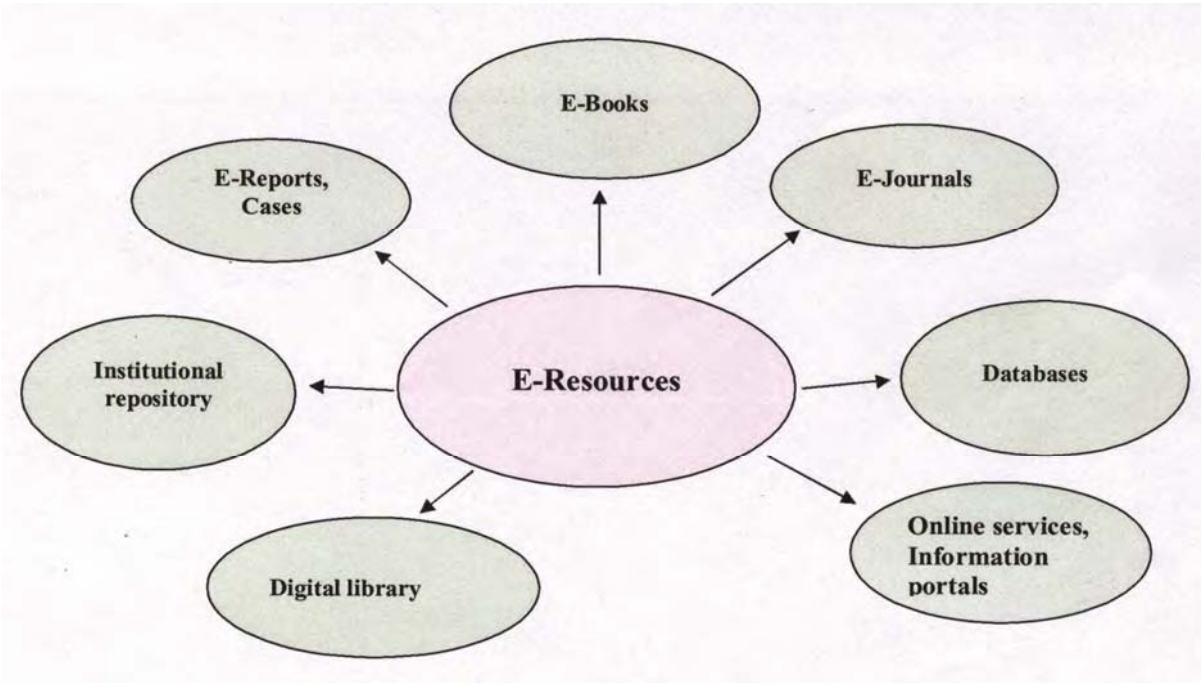

Fig. 1—Sample set of e-resources & services

manage these. There are many different access models being offered by publishers to the ever growing e-resources from all over.

Need for content aggregation and content integration

Looking at the current scenario, one can see a radical change already registered in most of the library's collection building process. There is an amazing penetration of scholarly digital information through a variety of forms such as books (published as such or issued as accompaniment), journals, portals, vortals, reports, CBTs, WBTs, cases, databases etc.[2]. To make matters more complex, the proliferating array of different file formats, standards and platforms in which documents are published, pose a multiplicity of threats to the information professional who is supposed to be the custodian and service provider of these information products once it has found its way into the library. In addition, a new and robust form of information resource coming up across the world in a big way in the libraries, is the Scholarly Institutional Repositories (IR) or E-Print Archives, which are repositories for electronic versions of research papers (self archived by the authors) made available to the scholarly community freely, and hence called Open Access Archives[3]. Considering the above situation, it is practically and technically impossible for a single monolithic digital library approach towards harnessing and embracing the complex and complicated information domain of a library. The all-in-one-container approach is therefore no longer a feasible one. It is therefore imperative on the part of the information professionals to have a componentized and a multi-system approach to knowledge technologies and information management. Seamless dissemination of scholarly information by means of content aggregation and content integration through the following information management techniques is the recommended strategy to be adopted and practiced by the 21st century libraries:

1. Library automation (integration of the traditional library activities/collection);
2. Library portal (seamless integration of the library's entire knowledge base – traditional, paper as well as electronic technologies);
3. Digital library (creation of efficient information systems using digital collections for effective browse, search, retrieval and long term preservation);
4. Open access archive (scholarly Institutional Repositories for worldwide access);
5. Consortia based approach to Collection Development;
6. Licensing agreements and access management strategies.

Library Automation

Library automation, taking care of the basic library operations such as Online Public Access Catalogue (OPAC), content acquisition, serials control, information processing etc. is a pre-requisite and it becomes the foundation of a strong digital information system. Library automation should be accomplished based on national and international standards and technologies (MARC, Z 39.50, OAI-PMH etc.), so that it provides scope for forward as well as future compatibility and interoperability. Libraries aspiring for digital libraries and high end digital information systems should ensure that they have accomplished this preliminary requirement.

Library portal as a one-stop-information shop

A portal is a stepping stone or a launch pad for a variety of inter-related Web based service components. Envisaged as a single window and as a single access point to the invaluable knowledge treasure of the library, the library portal is an integrated network of information sources. Besides serving as a web-based interface to the in-house resources, the portal also provides links to strategic sources of information outside. Portal development is more or less a unique combination of art, information technology and library/information science technologies. The portal pages should be highly appealing to the users and the distribution of

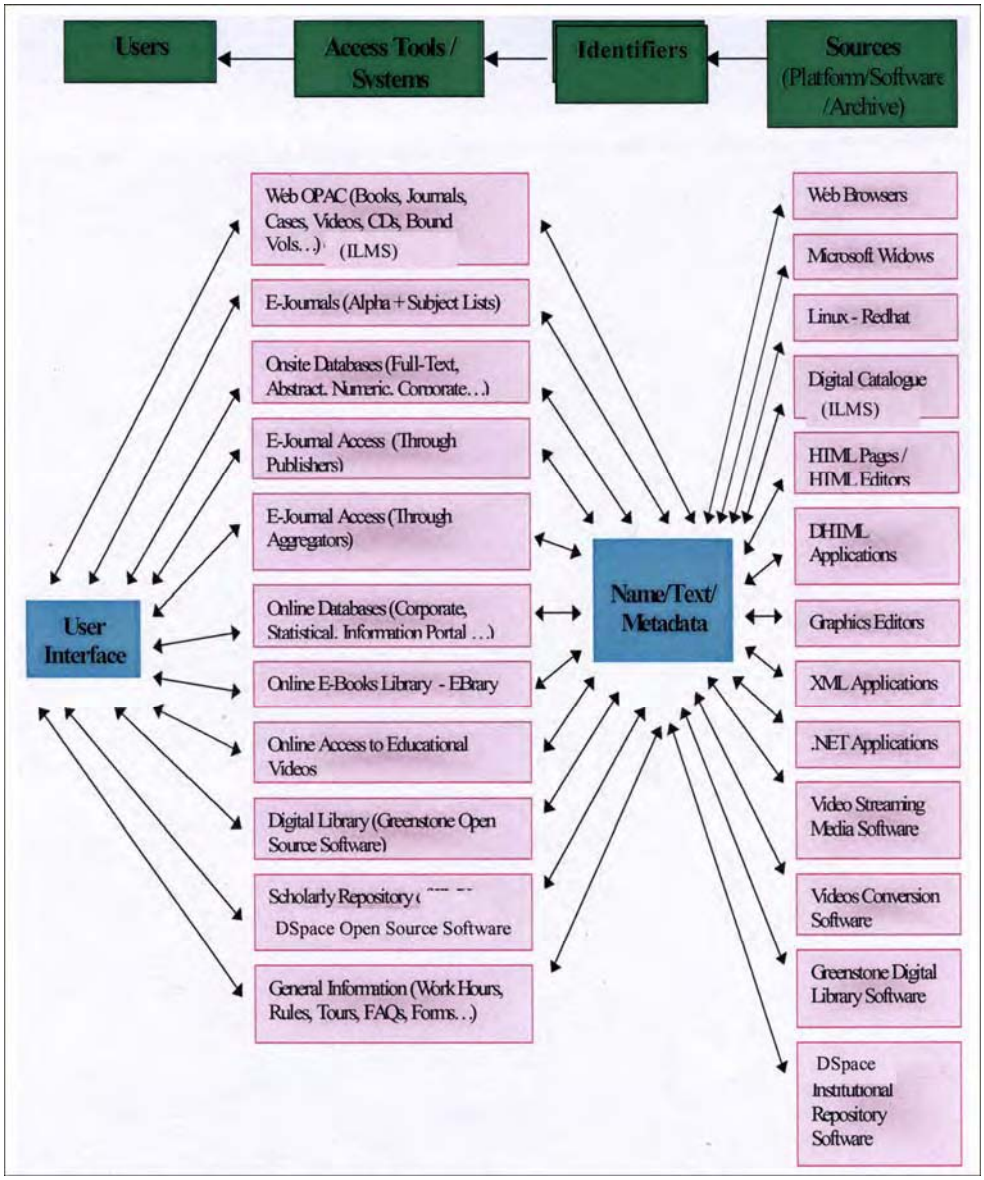

Fig. 2—Content / Information Model of Library Portal

information should be based on subject/service component hierarchies, classification, ontologies/ semantics and with scientific reasoning. Basic principles of Human Computer Interaction (HCI) and fundamental, standard colouring schemes should be complied with. The search services must be user friendly and the hyperlinks provided must be for instantaneous access to a number of strategic locations, contextually relevant.

The portal is accessible throughout the campus, viz., the classrooms, faculty rooms, library, computer center, student hostels and the faculty/staff residence. The library portal has been thus envisaged as a one-stop-shop for all the information resources and information services of the library embracing all the publication types and physical formats. Efforts were made for seamless integration of the content that was possessed by the library which included print, audio, video, online digital resources and onsite digital resources. The content/information integration model adopted for the system is illustrated in Figure 2. Information retrieval in the library portal has been made as user friendly as possible so that even people with less or no computer and Web background can get the best out of the vast treasure of information resources available in the collection. Hyperlink buttons are provided for each and every significant information source or publication type. Simple pop-up and pull-down menus appear as and when these links are clicked. A library portal interface frame is shown in Figure 3.

Users can select the service of their choice based on context relevance such as the library specific information and library rules/regulations, online public access catalogues (OPACs), onsite full-text/abstract/index databases, e-journals, online services such as Web based databases (scholarly as well as corporate) and value added information products, popular and useful hyperlinks, online reference queries, Frequently Asked Questions (FAQ) on library services/activities etc.

Developing digital libraries using open source software

Digital libraries do enable the creation of local content, strengthen the mechanisms and capacity of the library's information systems and services. They increase the portability, efficiency of access,

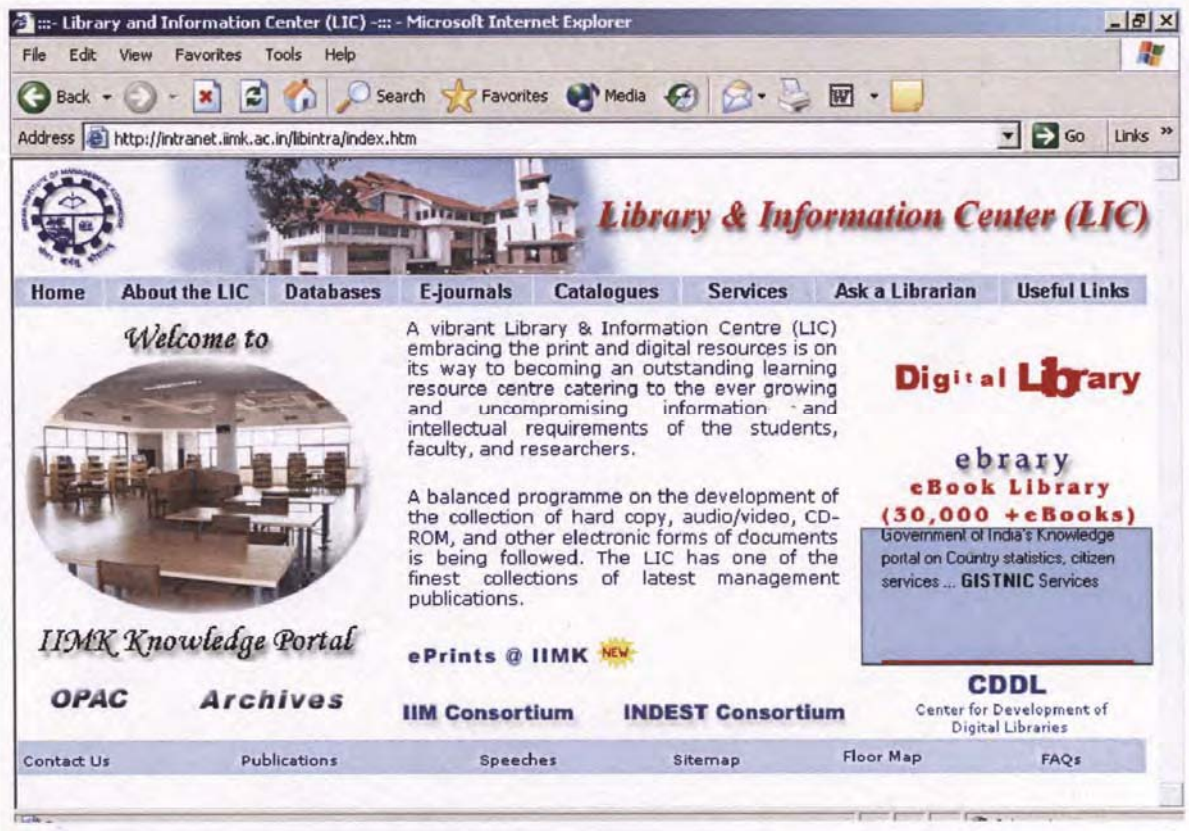

Fig. 3—Interface frame of library portal

flexibility, availability and preservation of content. A well founded and nicely formulated Digital Library project could be a real boost to the library's modernization activities and its endeavours to launch innovative digital information services to the management fraternity. Once the information is made digital, it could be stored, retrieved, shared, copied and transmitted across distances without having to invest any additional expenditure. Value added and pinpointed information at the click of the mouse will thereby become a reality and the library portal could give access to the invaluable collection hosted by the digital library.

Information systems should explore the endless possibilities and opportunities put forth by the open access movements and the open source software philosophies and it should be a chosen decision to either go for a proprietary or an open source digital library software. If the decision is for the latter, the popular Open source software for digital libraries which are in use internationally would be DSpace, Dienst, Eprints, Fedora, CDSWare, Greenstone etc.

From the decade-old field test, Greenstone would outscore the group from the perspective of a practical digital library application[4]. The Greenstone Digital Library Software (GSDL) is a top of the line and internationally renowned Open Source Software system for developing digital libraries, promoted by the New Zealand Digital Library project research group at the University of Waikato and is sponsored by the UNESCO. Greenstone software (versions starting from 2.30) along with Java Run Time Environment (JRE) were deployed for the purpose. The software suite is available at the open source directory 'Sourceforge'[5].

GSDL : Features

The salient features of Greenstone are basically taken from two of the official publications of the software development team appeared in D-Lib Magazine during the year 2001[6] and 2003[7]. Greenstone builds collections using almost popular and standard digital formats such as HTML, XML, Word, Post Script, PDF, RTF, JPG, GIF, JPEG, MPEG etc. and many other formats which include audio as well as video. It is provided with effective full-text searching and metadata-based browsing facilities that are attractive and easy to use. Moreover, they are easily maintained and can be augmented and rebuilt entirely automatically. The system is extensible: software "plug-ins" accommodate different document and metadata types. Greenstone incorporates an interface that makes it easy for people to create their own library collections. Collections may be built and served locally from the user's own web server, or (given appropriate permissions) remotely on a shared digital library host. End users can easily build new collections styled after existing ones from material on the Web or from their local files (or both), and collections can be updated and new ones brought on-line at any time. The Greenstone Librarian Interface (GLI) is a Java based GUI interface for easy collection building. Greenstone software runs on a wide variety of platforms such as Windows, Unix / Linux, Apple Mac etc. and provides full-text mirroring, indexing, searching, browsing and metadata extraction. It incorporates an interface that makes it easy for institutions to create their own library collections. Collections could be built and served locally from the user's own web server, or (given appropriate permissions) remotely on a shared digital library host. The other set of features include OAI plug-in (introduced since the 2.40 version) and DCMI compliance, UNICODE based multi-lingual capabilities and a user-friendly multimedia interfacing[8]. Further more, it has powerful search engine features using 'Managing Gigabyte', 'MG Plus-Plus (MG PP)', Lucene and metadata-based browsing facilities. A very interesting feature of Greenstone is its exhaustive set of well documented and articulated manuals such as 'Installer's Guide', 'User's Guide', 'Developer's Guide', and 'From Paper to Collection' a document describing the entire process of creating a digital library collection from paper documents. This includes the scanning and OCR process and the use of the "Organizer". There is one more interesting documentation 'Inside Greenstone Collections' which clarifies most of the trickier parts of using Greenstone, especially dealing with configuration file for the collection in question. The production digital library version of GSDL as of August 2012 is V. 2.85.

The primary objective of the digital library is to enhance the digital collection in a substantial way, by strategically sourcing digital materials, conforming to copyright permissions, in all possible standards / formats so that scalability and flexibility is guaranteed for the future and advanced information services and are assured to the user community right from beginning. The digital library is to be planned in such

a way that it will integrate and aggregate the existing collections and services with an outstanding user interface. Accordingly, necessary strategies are to be adopted towards working out the digital library system. This implies that the digital library system should have a strong collection interface capable of embracing almost all the popular digital standards, digital formats and software platforms, in line with the underlying digital library technologies in vogue. This is crucial in the case of multimedia integration, which is again important as it should ideally plan to host a digital audio and video library as part of the core library collection. The System Design Architecture of a Digital Library is illustrated in Figure 4.

Emphasis must given to maximize the efficiency and effectiveness of the information access and retrieval capabilities of the system by deploying

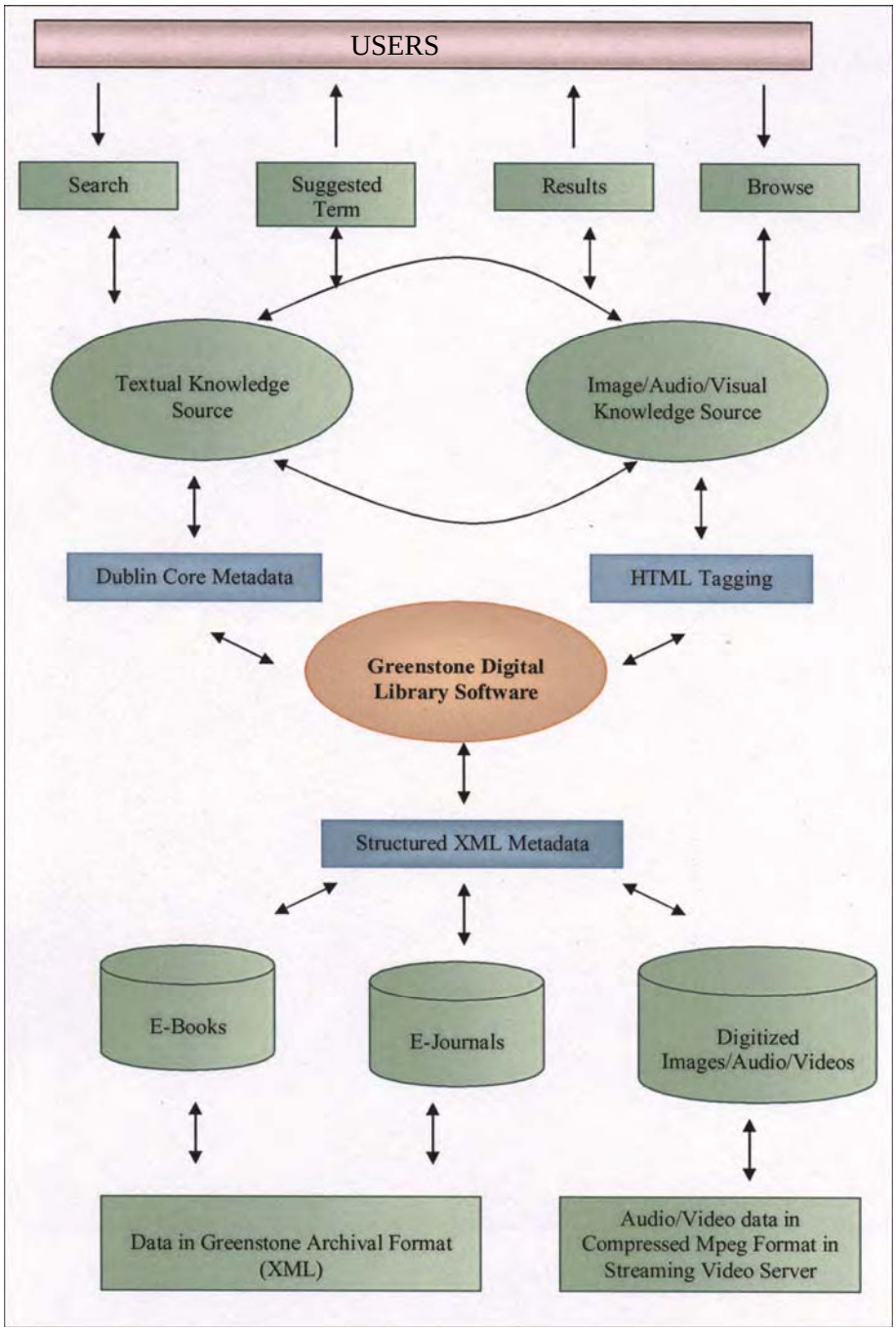

Fig. 4—System architecture of a digital library

cutting edge Resource Description Framework[9] standards for metadata and its XML[10] based encoding. The Internet possesses, in addition to its mammoth proprietary information base, an invaluable wealth and a vast collection of public domain information products such as databases, books, journals, theses, technical reports, cases, standards, newsletters etc., scattered and distributed across the world. This treasure should be explored to its maximum for collection building, based on the source and quality.

Standard workflow patterns are to be identified for the system which include: 'content selection', 'content acquisition', 'content publishing', 'content indexing and storage', and 'content accessing and delivery'. The system may also consider such related issues, viz., preservation, usage monitoring, access management, interoperability, administration and management etc. Dublin Core standards were identified as the desired metadata format and XML as the desired encoding scheme. Snapshot of a sample digital library interface frame is shown in Figure 5.

Trends in digital library service architecture

The latest trends in DL service architecture recommends Service Oriented Architecture (SOA) powered by Web services such as Search/Retrieval via URL (SRU) and Search/Retrieve Web Service (SRW). The open source family of 'Web Services', its open standards and technologies facilitate achieving highly configurable, extensible and dynamic digital libraries.

SRU is a standard XML-focused search protocol for Internet search queries, utilizing CQL (Contextual Query Language), a standard syntax for representing queries. CQL, the *Contextual Query Language*, is a formal language for representing queries to information retrieval systems such as web indexes, bibliographic catalogs and museum collection information. The design objective is that queries be human readable and writable, and that the language be intuitive while maintaining the expressiveness of more complex languages. Traditionally, query languages have fallen into two camps: Powerful, expressive languages, not easily readable nor writable

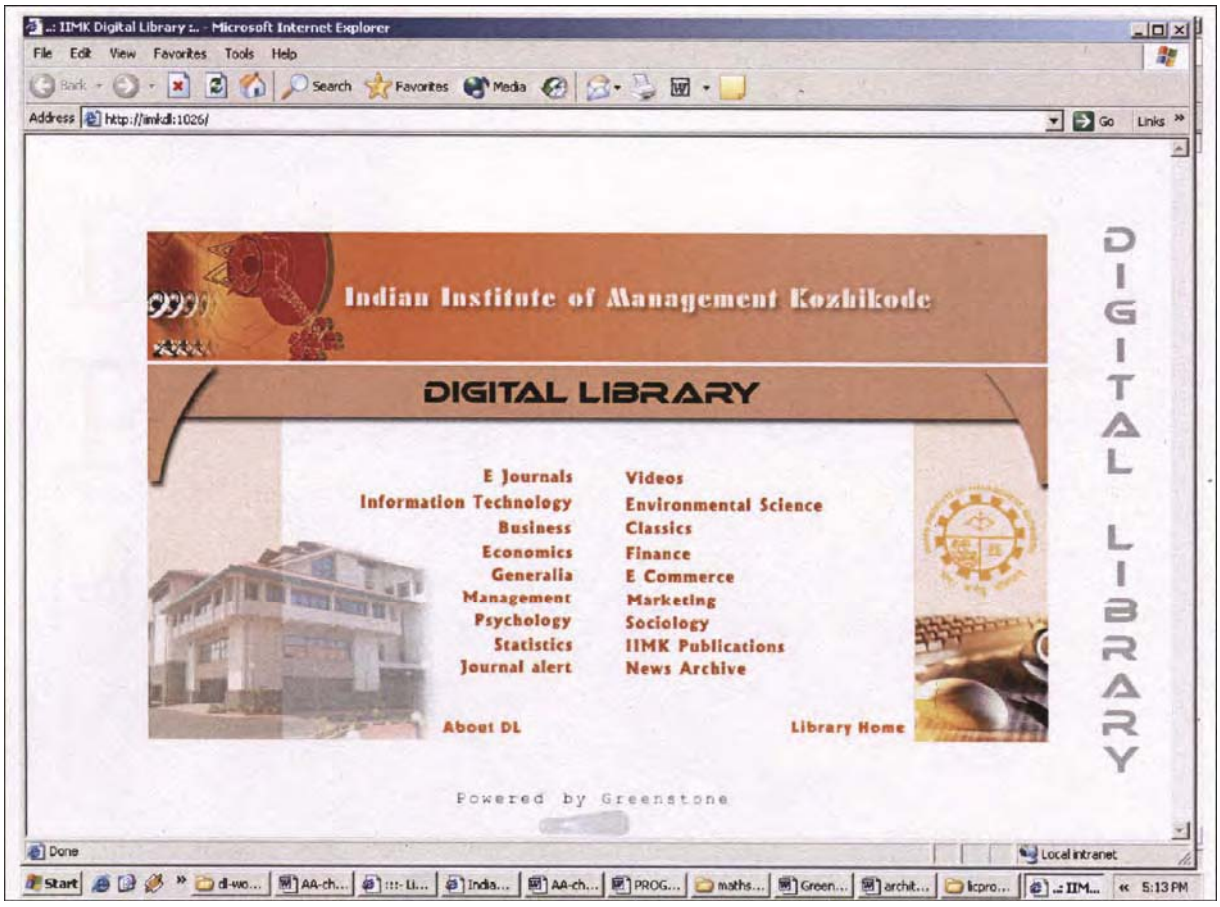

Fig. 5—Sample of Digital Library Home Page

by non-experts (e.g. SQL, PQF, and XQuery); or simple and intuitive languages not powerful enough to express complex concepts (e.g. CCL and google). CQL tries to combine simplicity and intuitiveness of expression for simple, every day queries, with the richness of more expressive languages to accommodate complex concepts when necessary[11].

CQL Query

Formerly known as Common Query Language, and later changed to Contextual Query Language, a CQL query consists of either a single search clause, or multiple search clauses connected by Boolean operators. It may have a sort specification at the end, following the 'sortBy' keyword. In addition it may include prefix assignments which assign short names to context set identifiers[11].

SRW is a variation of SRU. Messages are conveyed from client to server, not by a URL, but instead using XML over HTTP via the W3C recommendation Simple Object Access Protocol (SOAP), which specifies how to wrap an XML message within an XML envelope. The SRW specification tries to adhere to the Web Services Interoperability recommendations (WS-I). WS-I is an open industry organization chartered to promote Web services interoperability across platforms, operating systems and programming languages. The organization help customers to develop interoperable Web services by providing guidance, recommended practices and supporting resources. Specifically, WS-I creates, promotes and supports generic protocols for the interoperable exchange of messages between Web services[12].

SRW is a web service for search and retrieval. SRW provides a SOAP interface to queries, to augment the URL interface provided by its companion protocol SRU. Queries in SRU and SRW are expressed using the CQL. Standards for SRW, SRU, and CQL are promulgated by the United States Library of Congress.

Library discovery services

Presently a typical library of a higher learning organization subscribes to thousands of e-journals and e-books, tens of hundreds of databases, in addition to having books as well as other documents in print form. Patrons, especially the demanding ones, may consult the library's services but certainly go to Google also, for help. In such a situation, life will be pretty easy if users can start their discovery with a search engine provided by the library itself, which will direct them to the most appropriate content for their search that the library has as well as subscribed to by the library, rather than leaving them to try their luck with Google or anything else. Discovery services do have an edge over the federated searches and other such previously available tools. They aim to index all the library's licensed electronic content, as well as integrate the indexes pertaining to the non-digital/ traditional content. The major difference with discovery searches is that they don't search the library's local indexes and instead, they search a centrally held database of publisher content, thereby allowing more complete indexing, much faster searching, and more complex integration of result sets. In addition, rather than seeing multiple results for the same article if it's held in several different databases or abstracting and indexing services, these search tools display a composite result for each item. Linking from the list of results to the full text of an article is also controlled by the library through their link resolver system, and hence a user is directed to the most appropriate full text copy a library holds, generally the version of record on the publisher's own platform. Some of the prominent library discovery services include *'Summon'*, from ProQuest's Serials Solutions business, EBSCO's *'Discovery Service'*, and Ex Libris' *'Primo Central'*, all of which aim to put the library back at the center of search and information discovery[13].

Institutional Repositories (IR) as Open Access Archives (OAA)

An archive is a generally accepted synonym for a repository. A repository is a network accessible server that holds scholarly digital content or eprints. Scholarly Archives or Institutional Repositories are established medium to communicate peer reviewed (post-prints) and non-peer reviewed scholarly literature (pre-prints). There are basically three types of scholarly archives in vogue, viz., author archives, institutional archives and subject archives. Subject archives are also called as central archives. According to Stevan Harnad, open archiving is just self-archiving the articles the author has published in (peer-reviewed) non-OA journals. Hence it neither bypasses nor replaces peer-review. It has nothing to do with changing peer review. Self-archiving is a way of supplementing non-OA journal access with an OA

version for those would-be users whose institutions cannot afford the non-OA journal.

Institutional Repositories are scholarly archives which draw their strength from the Open Access (OA) movement[14]. An institutional repository (IR), also known as an 'e-prints archive', is a digital archive of the research output created by the faculty, research staff, and students of an institution and accessible over the Internet to end-users both within and outside of the institution[15]. As a facility it consists of hardware, software and procedures to capture, organize, archive, disseminate and manage digital research resources of the institution. IRs provide a simple, Web-based mechanism to researchers to deposit ('self-archive') and access their research publications. IRs have the potential to bring significant benefits to educational and research institutions such as:

- Improved visibility and impact of the institutional research - since the archive is freely accessible on the Internet, more researchers can consult and cite the research publications;
- Help in establishing priority to the research findings;
- Rapid communication of research;
- Long-term preservation of research papers;
- Integrated view of the institutional research which is otherwise distributed over a large number of external sources;
- Foster increased collaborative research projects and attract project grants;
- Improved research knowledge management;
- Create value-added services: Individual, department-wise online publication lists, support citation linking, etc.

Thanks to the Open Access movement catching up the world over. It is all about democratizing the scientific intellectual capital, which often draws its energy from publicly or privately funded research. It is not against proprietary or peer reviewed scholarly journals.

There are numerous advantages that OA boasts while they campaign worldwide. Authors as well as Institutions can derive a number of benefits out of Archives. For authors, instant dissemination of the fruits of their long years of rigorous research to a global audience is the first and foremost. OA papers get increased visibility through novel models of harvesting done by search engines such as the Google, Citeseer etc. and the interoperability among similar archives achieved through the Open Archives Initiative (OAI) Protocol for Metadata Harvesting (PMH) are unparallel value additions to OA Archives. While more visibility leads to more citations, one's research impact naturally gets scaled up. Authors are therefore attracted to come to OA Archives. Additional benefits to self archiving include the assurance of the long term preservation of their articles and the facility to have a proper control as well as meticulous monitoring of one's own Publications.

For institutions too, a long list of advantages and benefits invite them to OA. Firstly, the institute's archive, popularly known as Institutional Repository helps in pooling the organization's Intellectual Capital into one central place which is otherwise scattered, distributed and unnoticed. The archive therefore serves as a one-stop-source or a single access point for the research output of an institution. It provides ample scope for introspection as to whether the institute is going in the right direction on its research activities. Necessary strategies and meticulously designed action plans could be charted out based on the feedback. Institutional repositories facilitates instant generation of research reports and thereby saves a valuable amount of time otherwise spent unwanted. Most importantly the archives ensures long term preservation of its scholarly materials with the help of Open Source softwares and Open Standards of data models and data structures.

Institutional Repository (IR) Softwares

There are many world renowned free and open source Institutional Repository (IR) softwares available such as EPrints, DSpace, FEDORA, VT-ETD, ARNO, i-TOR, CDSware etc. They are issued either under GNU public license or the BSD license and can be downloaded from their own sites or open source software directories such as SourceForge [Sourceforge]. Each of the software has a host of features, unique facilities and excellent capabilities, which the users could explore and experiment. Further, by using the Open Archives Initiative (OAI) interoperability protocol for metadata harvesting (OAI-PMH), content in these repositories can be easily shared at metadata level to establish a single-point cross-indexing and search service at national/ international level. The OAI-PMH protocol enables integration of content in these repositories with

worldwide cross-archive search services. As discussed earlier, this paper focuses on the DSpace software.

DSpace

DSpace is a digital asset management software jointly developed by Hewlett-Packard and MIT Libraries, and it is arguably one of the appreciated open source software deployed worldwide for building digital institutional repositories that captures, stores, indexes, preserves, and redistributes content in digital formats. DSpace provides the institutions and universities operate an open access and interoperable institutional repository at the local level. It is also intended to serve as a repository back up for future development to address long term preservation and remote/online access issues. The system was launched during late 2002 as a live service hosted by MIT Libraries, and the source code made publicly available according to the terms of the BSD open source license, with the intention of encouraging the formation of an open source community around DSpace[16].

Features & functions

DSpace is a 100% open source software and is freely available for download from the open source software directory SourceForge. The software has been built on a strong architecture supported by state-of-art digital library technologies and embracing almost all latest trends in information sciences. It provides the users, especially the librarians and system administrators, every freedom for building, managing, customizing, administering and Internet publishing world class institutional repositories and digital libraries. Its major features include the ability to accept all forms of digital materials including text, images, video, and audio files. Possible content includes scholarly articles and preprints, technical reports, working papers, conference papers, books, e-theses, multimedia publications, Datasets: statistical, geospatial, matlab, etc. Images: visual, scientific, etc.; audio files, video files, learning objects, bibliographic datasets, reformatted digital library, collections, Web pages etc.

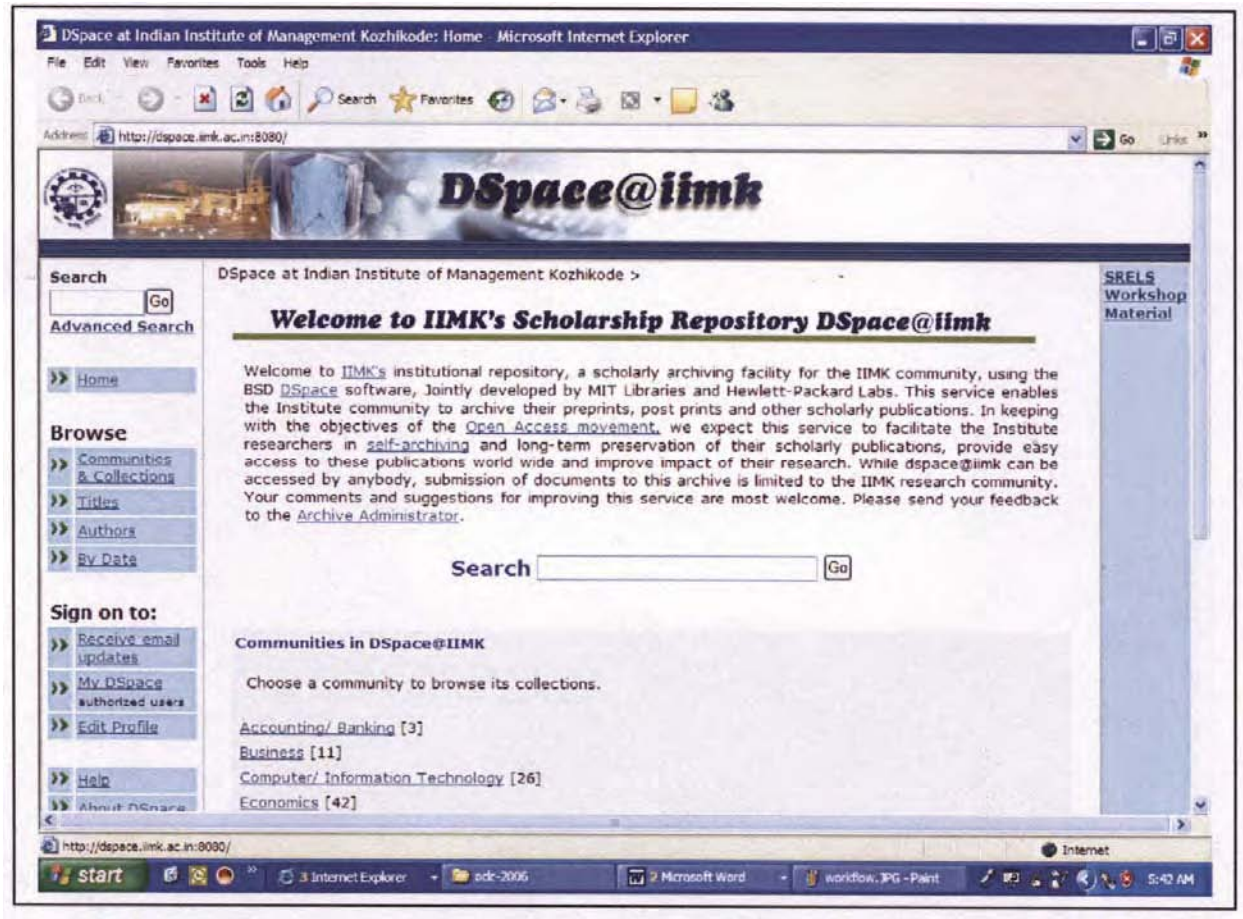

Fig. 6—Sample of DSpace repository

For enhancing the resource discovery features, DSpace supports Dublin Core metadata unqualified element set as well as provisions for the qualified Dublin Core metadata registry. The software allows the communities and users to publish their articles remotely on the archives. It has CNRI 'Handles' support for Persistent URLs (PURL) which assigns and resolves persistent identifiers for each digital item. Interoperability is another salient feature of DSpace, and it supports the Open Archives Initiative's Protocol for metadata harvesting (OAI-PMH) V2.0 as a data provider. OAI support was implemented using OCLC's OAICat open-source software to make DSpace item records available for harvesting. DSpace uses the versatile Lucene search engine for full text searches. Lucene search engine is a part of Apache Jakarta project, and brings along laudable search features like 'fielded', 'Boolean', 'exact term', 'proximity', 'wild cards', 'fuzzy', 'range', 'boosting terms' etc. DSpace supports unlimited exporting/importing of digital content, along with its metadata in a simple XML-encoded file format. The database management system used is PostgreSQL which supports transactions between Oracle as well as MySQL. DSpace enjoys international acceptance across the world and it provides a customizable Web interface. The workflow process for content submission, the decentralized submission process, the remote publishing facility are regarded the unique features of DSpace. Most importantly it is Open URL compliant also[16]. Snapshot of a sample institutional repository interface frame is shown in Figure 6.

Institutional policy on open access

Setting up an institutional repository is not a big deal now a days. But arriving at a suitable and feasible open access policy at the institutional level is a Herculean task and this need the active participation of the information professionals and the scholarly community of the institution. We need to do a bit of scouting, and if necessary, little lobbying also towards this. A reasonable amount of guidance on this and also on submission guidelines, author benefits, copyright issues etc. could be well seen at the IR at the Indian Institute of Science (IISc) set up by NCSI[17].

E-Resource licensing agreements and access management strategies

Arranging subscriptions to the multiplicity of online resources that manifest in many different resource types, providing hassle free and uninterrupted access to the e-resources are among the major challenges before the information professionals. As the number of e-resources and e-services have sky rocketed, libraries are facing numerous problems and issues in the provision of their services. The overwhelming number of Web sites and URLs are flooding the library portals and library web pages, making the lives of librarians, a misery. There are many different access models being offered by publishers to the ever growing e-resources from all over.

In the case of e-resources, users' access goes to multiple pointers and thereby it becomes complex. Users in general prefer not to come to the library and expect access to library materials from remote sites. They expect to access all the resources seamlessly. Some of the recent genre of electronic materials do not have their physical counterparts at all. Further more, access to the e-resources is no longer controlled by the library. The terms of access and delivery are dictated by publishers, aggregators or vendors. Providers resort to different methods and technologies towards authenticating access to their products. Lack of uniform access delivery standards that could be provided to the users, customized and configured locally, is a critical issue faced by the library fraternity currently.

Access management is becoming more and more an important concern for libraries. It provides controlled and mediated access to the library's e-resources and online services. It defines which users exist and what roles they have. It comprises creating and defining the list of users with access to the various e-resources and assigning access right in the form of roles. The most important elements of any E-Resource license agreement are the determination and definition of the authorized users and how the e-resources will be accessed.

The predominant access management methods practiced are predominantly IP authentication and User ID/ Password based access for a small number of resources. User ID/ Password based access method is among the earlier models to emerge which is the popular access authentication system. Being popular and simple, most of the service providers use this model for access authentication. But when the number of resources and users have increased, this method became a problem for libraries. Another problem with User ID/Password is that users have to deal with

many passwords for a wide variety of services. User ID/Password method provides less security as it is not easy to keep the password secret for long period. As a result, IP authenticated access method become more popular. IP Authentication is one of the most commonly used authentication mechanism. IP filtering should be performed by both library and service providers. To make IP filtering work properly, library should provide the vendors the range of IP addresses of the network of the library or the parent organization. For a wider audience, IP based access is mostly preferred, as the users need not have to bother about the User ID and Password every time.

IP based access authentication is a good choice if the users are closely located and covered by a single network. This method allows (i) seamless access (ii) usage statistics for the institution (iii) greater security as there is no misuse of usernames and passwords (iv) access to all computers thereby releasing other terminals and staff time (v) direct recognition of institutional networks by publishers and vice versa[18].

Though the IP based access authentication has been found to be an appropriate and a suitable access model for the on-campus community, though the main limitation of this method is that it can not be used for remote users. For the off-campus user community the IP based access to e-resources will not be possible, as they will not be able to resolve the Institute IP while sending the URL requests from outside the campus. Thus there became a need to deploy alternative mechanisms that meet the needs of both local and remote users.

Single Sign – on (SSO) and Remote Login

Presently most of the information systems provide a vast array of electronic resources through the library portal and its internal web pages. In addition, A to Z pages such as alpha lists or subject-based pages of databases and e-journals with their persistent URLs are placed in the portal. As the number of e-resources increases, it becomes difficult for the libraries to maintain these web pages. Each service provider follows their own access management system. Similarly when the access to certain resources are provided through third party access gateways such as Ingenta, Athens etc., users have to be authenticated multiple times. The situation becomes worse when each one of these systems uses its own authentication mechanism and authenticator.

Single sign-on systems (SSO), also known as Reduced Sign on (RSO), addresses this much efficiently and successfully. SSO refers to a type of authentication system where a user only has to be authenticated once, and is able to access multiple services where he has access permission, without having to enter his password again. SSO system helps the users to access the resources very fast and secured. It also provides quality response time as well as user satisfaction[19].

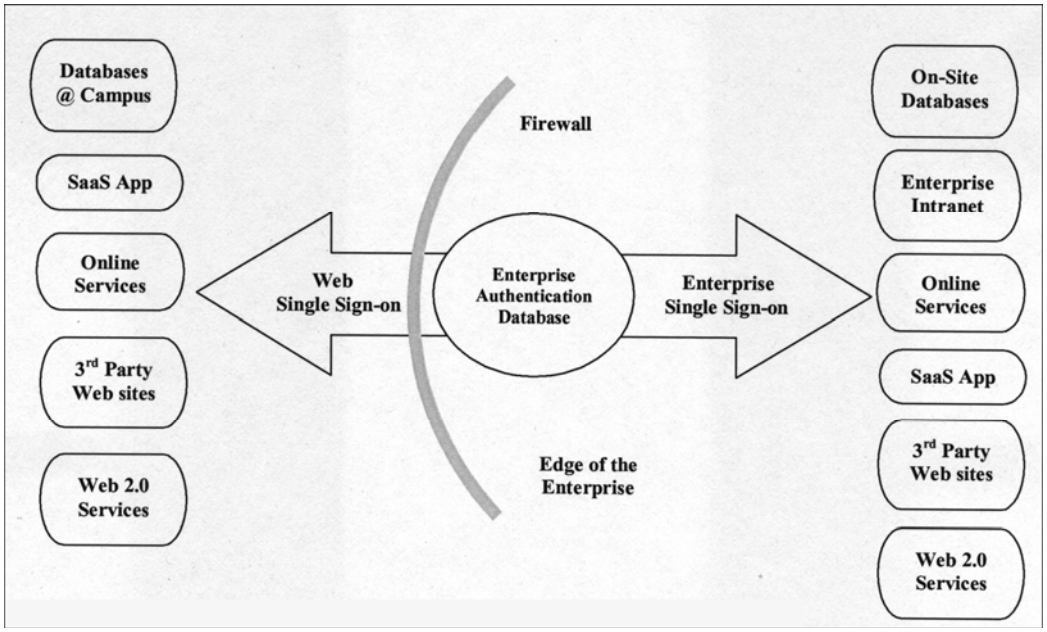

Fig. 7—Architecture of the single sign-on remote access system

Single sign-on systems provide multiple benefits to libraries by relieving the users from remembering different user names and passwords; time being spent for declaring identities; supporting conventional authentication such as Windows credentials (i.e., username/password); saving on IT costs on authentication related enquiries; hassle free secured access provisions and centralized reporting for compliance adherence. The advantage of SSO is that it uses the existing centralized authentication servers and it combines this with techniques such that the users will have to enter their credentials just once only[20]. Fig. 7, shows the single sign-on architecture in an institutional setting.

This model envisages ensuring of a seamless as well as hassle free access to the vast reservoir of digital on-site and remote online resources, to the user community, regardless of whether they are on-campus or off-campus. Users will have to login only once for getting access to the resources.

Working principle

The remote login based SSO setting work basically in two scenarios. The first option would be a subscription/cloud service which operates through a third party site, in which registered users of an organization will be remotely accessing their licensed content using a two-factor identity secured authentication, such as User ID and Password. The parent organization of the users will work with the service provider to introduce them for the first time so that users could take it forward themselves. The second option, which is seen as the preferred one, is by installing the SSO server on site and allowing the users to login from remote locations. Users will first register with the system and the system will enroll them as authorized members. Each time users will login to the proxy server whenever they want to access the institutional licensed content. The working model of an SSO Remote Access System is illustrated in Fig. 8.

Products and solutions

There are a number of remote access authentication models that prevail in the market and are being used by organizations with varying degrees of satisfaction. There are a few open source solutions and prominent among them include Shibboleth. Prominent commercial solutions include Ezproxy and Athens. EZproxy is a web proxy server program extensively used by libraries to give access from outside the library's computer network to restricted-access websites that authenticate users by IP address. EZproxy helps provide users with remote access to web-based licensed content offered by libraries. EZproxy provides a number of benefits to the libraries. It is an industry leading, robust middleware solution for remote user authentication; it connects to

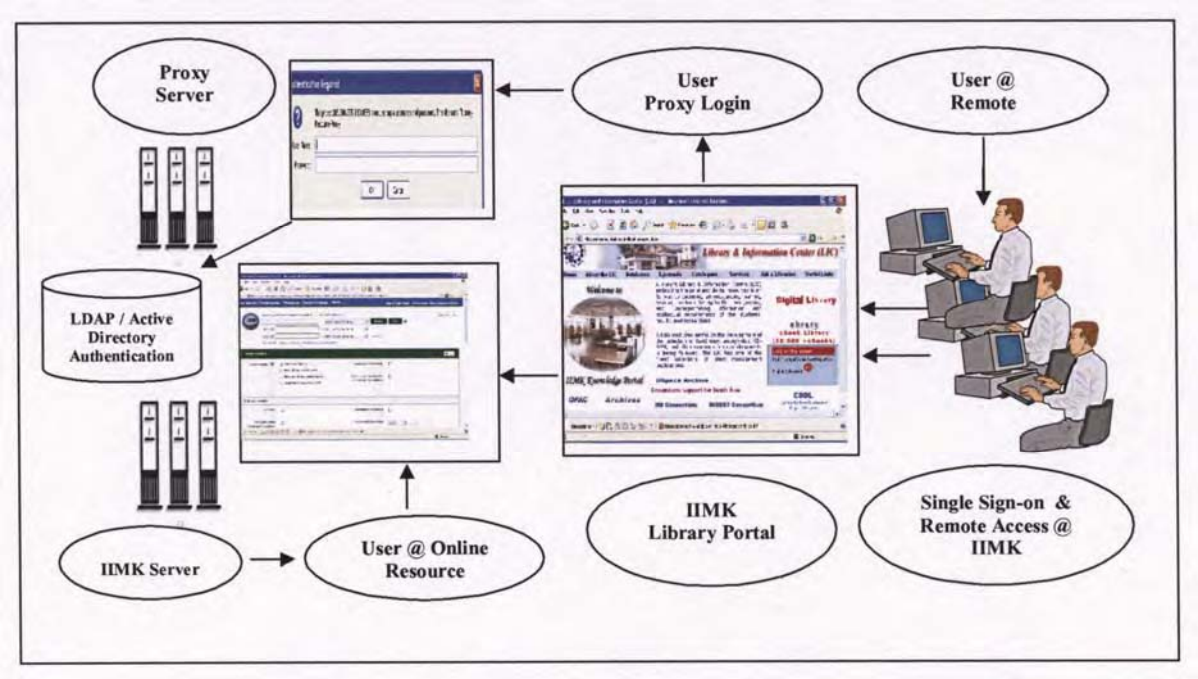

Fig. 8—Working model of single sign-on remote access system

a large number of content providers; it connects to a wide variety of authentication services (including LDAP, SIP, and Shibboleth) which reduces the number of authorizations/passwords and provides a better end-user experience; it is an easy to setup and easy to maintain program[21].

Athens is a popular access and identity management service, offering single sign-on, to protected resources combined with full user management capability. Athens allows individuals to use single sign-on to access web resources and reduces the administrative burden for librarians and information managers. Classic Athens provides organisations with the tools necessary to create and manage usernames for single sign-on access to protected web resources. Major benefits of Athens are: secure single sign-on access to multiple web-based services; devolved administration facilities at organisation level; remote access user accounts; encrypted bulk user account upload services; highly scalable services - supports millions of user accounts; replicated, load balanced and fully resilient architecture. A recent extension of Athens, the OpenAthens framework, enables organisations to manage user identity and enable secure access to online resources through a single sign-on identity[22].

Shibboleth is standards based, open source software package for web single sign-on across or within organizational boundaries. It is software that implements Security Assertion Markup Language (SAML) protocols, separating the functions of authentication and authorization. Major goal of shibboleth is to allow users to access internal and external resources seamlessly using a single, institutionally controlled identity. Major benefits of shibboleth include: (1) reducing the time needed to manage and access to protected resources such as sharing resources among several institutions and managing a large number of accounts; (2) increased security (Single Sign On – SSO – do not need to remember multiple passwords, acquire information about the users from reliable providers, etc.); and (3) interoperability with similar standard-based solutions[23].

Conclusion

In the new information environment there has been a vast outgrowth of e-resources, and recent developments in technology have registered an unprecedented influence on the user community favouring the deployment of IT enabled services and in furthering the e-resources in libraries. In order to achieve a substantial increase and a face-lift in the information resource base, libraries need a tremendous and a drastic reengineering of their processes and procedures and should aggressively participate in library consortia. The situation prevailing in the current library scenario reveals that it has reached a critical mass both in terms of the digital collections as well as the number of users. Libraries today face the unprecedented challenge of managing an array of content spread across a host of publication types and in a rapidly proliferating mix of formats. With the availability of the state-of-the art information technology solutions and the Web revolution, libraries are now better off in terms of easy access to more information and in the provision of user friendly value added services. It is therefore imperative that libraries resort to access solutions which are simple and easy for the users to cope with. Time has now ripen to a stage that we need to think beyond the framework of digital libraries, and be concerned more about the information strategies for future direction, at least to take care of the immediate requirements. What is important at this juncture is a complete reengineering of our present practices and a strategic positioning of our future direction towards information management policies leading to E-Scholarship. In the current scenario it is almost impossible for a single monolithic digital library approach towards harnessing and embracing the turbulent information domain of a library. The all-in-one-container approach to information management is therefore no longer feasible and advocated, and hence a multi-system approach to knowledge technologies and information management is the recommended alternative. Seamless dissemination of scholarly information by means of content aggregation and content integration through robust automated library systems, info-rich scholarship library portals, componentized open digital libraries and through the fast catching up scholarly electronic publishing mode of institutional repositories, are the modern information management strategies and techniques to be adopted and practiced.

References

1 Sreekumar M G and Sunitha T, Community acceptance of E – journals as a robust print surrogate and scholarship supplement: The IIMK experience. In Proceedings of the

International Conference on Digital Libraries (ICDL 2006). Ed. By TERI, New Delhi, India. 06-08 December 2006. pp. 531-541.

2 Terry A A, How today's technology affects libraries' collection choices, *Computers in Libraries*, 20 (6) (2000) 51-55.

3 Open Archives Initiative (OAI). <http://www.openarchives. org>

4 Greenstone <htttp://www.greenstone.org>

5 SourceForge.net (world's largest Open Source software development website) <http://www.sourceforge.net/>

6 Witten I H, Bainbridge D and Boddie S J, Greenstone: Open-Source Digital Library Software, *D-Lib Magazine*, 7 (10) (2001) 1-16.

7 Witten I H, Examples of Practical Digital Libraries: Collections Built Internationally Using Greenstone, *D-Lib Magazine*, 9 (3) (2003) 1-15.

8 Unicode. <http://unicode.org>

9 RDF. <http://www.w3.org/RDF/>

10 XML. <http://www.w3.org/XML/>

11 SRU. <http://www.loc.gov/standards/sru/>

12 SRW. <http://www.loc.gov/standards/sru/sru1-1archive/ srw.html>

13 Wiley / Blackwell Library discovery services: A better way to find what you are looking for. http://blogs.wiley.com/publishingnews/2010/12/22/library-discovery-services-a-better-way-to-find-what-youre-looking-for/

14 Openacess. <http://www.soros.org/openaccess/>

15 NCSI Workshop on Institutional Repositories. <http://www.ncsi.iisc.ernet.in/indest-ncsi-ir/>

16 Dspace Wiki. < http://wiki.dspace.org/index. php//EndUserFaq >

17 ePrints @ IISc < http://eprints.iisc.ernet.in >

18 Armstrong E A, Croft V F, Kok V T and Lener E F, Electronic Resources Access: issues and resolutions in two academic libraries. <https://research.wsulibs. wsu.edu: 8443/jspui/bitstream/2376/2025/1/Armstrong Electronic resources.pdf>.

19 Authenticationworld.com , <http://www.authenticationworld.com/Single-Sign-On-Authentication/>

20 Wikipedia <http://en.wikipedia.org/wiki/Single_sign-on>

21 Wikipedia <http://en.wikipedia.org/wiki/EZproxy>

22 Athens <http://www.athens.ac.uk/>

23 Paschoud J and Garibyan M, Shibboleth for new generation access management (UK Perspective), (2005) <http://gatetest.library.lse.ac.uk:8080/dspace/bitstream/1988/ 2783/4/jpmg_IADIS2005 paper.doc>