Seamless Aggregation and Integration of Diverse Datastreams: Essential Strategies for building Practical Digital Libraries and Electronic Information Systems

M.G. Sreekumar and T. Sunitha
Center for Development of Digital Libraries (CDDL), Indian Institute of Management Kozhikode (IIMK), IIMK Campus P.O., Calicut – 673570, Kerala, INDIA
{mgsree,sunitha}@iimk.ac.in

Key Words (content integration, content aggregation, information systems, practical digital libraries, open source software, digital library software, Greenstone, metadata)

ABSTRACT

In the current information environment libraries need to leverage on the latest digital technologies as well as the traditional paper technologies towards building practical digital libraries and electronic information systems. Digital Libraries built exclusively out of nascent electronic publications such as E-Journals, E-Books, E-Reference Works (WBTs, CBTs etc.), Digital Scholarly works (monographs etc. in the public domain) and Digitized documents conforming to standard digital formats are proving to be an uphill and unfinished task. Perhaps this could be the major reason why the start-success-finish ratio of most of the digital library initiatives, particularly initiated by isolated / individual libraries, are still left at alarmingly low numbers. We find the motivational and emotional bonding among the stakeholders melting down eventually as the digital library development process gets fired up. There are a host of problems the enthusiastic library fraternity face in their digital library development endeavours starting from copyright issues, technology complexities, infrastructure threats, diverse publication types, multiplicity of digital object formats and above all the publishers' stringent policies and monopoly. It is therefore essential that the libraries adopt necessary strategies towards developing digital libraries from the lessons learnt. The monolithic and all-in-one-container approach is no longer feasible and no more advocated. 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 datastreams embracing the print as well as the electronic information is the most appropriate strategy to be adopted and applied. This paper shares IIMK's experience in creating a state-of-art Digital Library Information System by seamlessly integrating and aggregating the print as well as the diverse and distributed digital content penetrating into its knowledge domain. The paper highlights significant features of IIMK's Digital Information System - the content aggregation and the content integration strategies we adopted for designing a scholarship Web Portal and developing a Digital Library using the 'Greenstone'

Open Source Digital Library software. The paper also highlights the role of libraries in promoting Open Access by setting up scholarly Institutional Repositories (IR). In summary, today's digital library information system is to be seen from a much wider and holistic perspective, and provided with a much broadened meaning to hold and put together all the print, digital and electronic information available and accessible to the library.

The New Information Landscape

Libraries today buy, subscribe, license and accumulate information in an unprecedented array of content categories or publication types, and in a rapidly proliferating mix of formats (digital as well as print). A study conducted by OCLC in 2003 reports that the information landscape for the coming 5 to 7 years can be forecasted in four distinct categories [OCLC, 2003]. The study further reports that today's librarians in the new information environment has to equip himself with a variety of skill sets and knowledge keeping in line with the market dynamics and the user demands. The information professional should be capable of handling a set of new genre of information dissemination technologies such as PDAs, Cell Phones, MP3 Players etc. Today's libraries are faced with the challenges of integrating traditional and emerging formats balancing resource allocation between traditional and upcoming technologies and building new information management processes and procedures.

As intermediaries between the challenging, multi-format, multi-media information domain and the ever demanding user community, librarians need to be proactive and concerned about the new information landscape (Figure 1). The four distinct categories of information discussed are popular information, scholarly information, digitization projects and Web resources. 'Popular information' is more or less trade and commercial in nature which comprise the print books and its electronic counterpart, the emerging Print-on-Demand (POD), journals and newspapers (print and online), and the audio/visual media (analogue and digital). The 'scholarly information' is academic and scientific in nature and they consist of the print and their electronic counterparts of books, journals, scholarly articles, thesis and dissertations, course management materials and e-print archives. The 'digitization projects' mainly concern themselves with the vast number of worldwide initiatives on commercial, national, state and local digital library projects. The 'Web resources' form an important stakeholder in the new genre of information resources, providing an array of challenges as well as opportunities to the information professional. The information available in the Web (WWW) could be categorized as surface Web (visible Web) resources and the deep Web (invisible Web) resources. The distinguishing feature of the Web is that it can hold all the above information, and that it can disseminate them ubiquitously across the world crossing the geographical and the time barriers asynchronously.

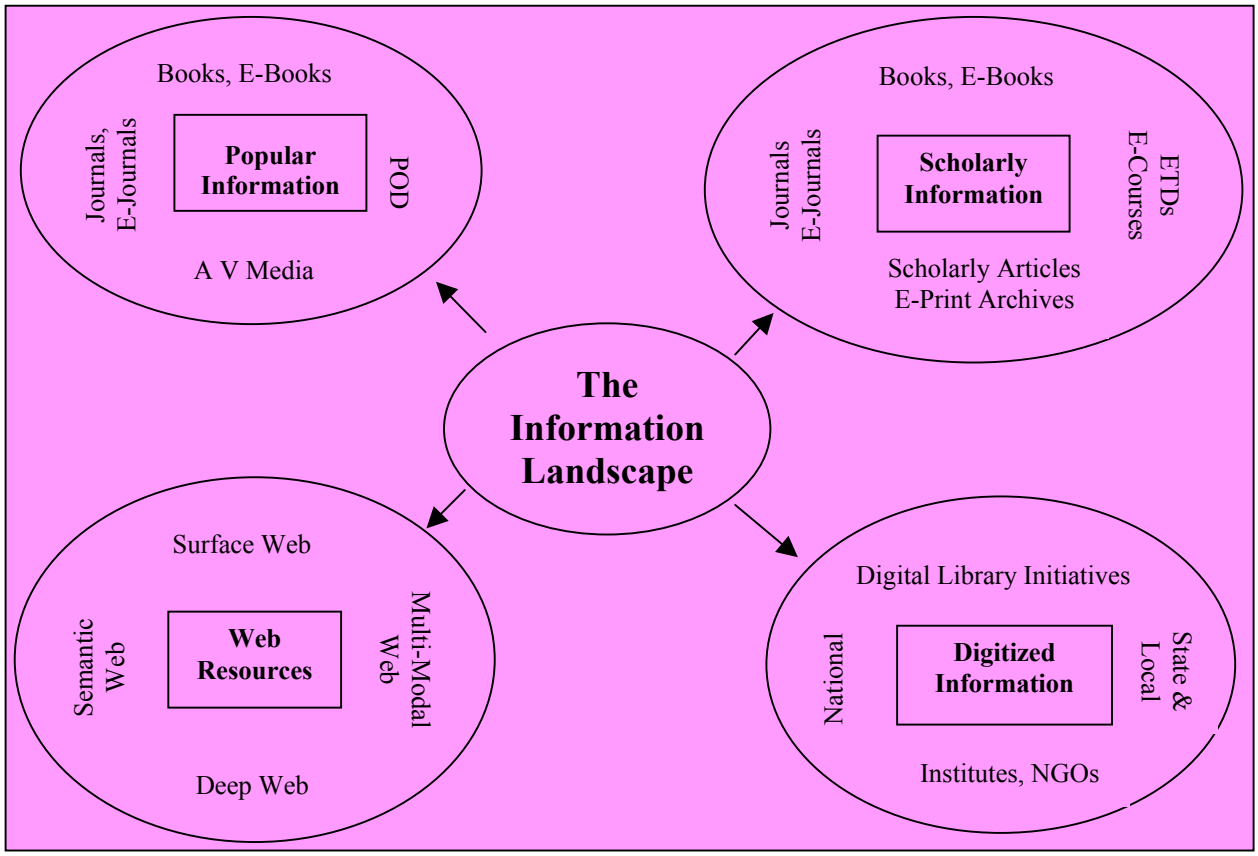

Fig. 1. The New Information Landscape

Features of the New Information Genre

There is a great deal of cultural divide and philosophical deviation between the traditional information resources being handled by libraries for centuries now and the new genre of electronic and digital information being sourced and accessed. In the traditional paradigm, the books and journals bought and subscribed to by the libraries were naturally owned by them, allowing them to make the best use of the resources within the 'fair use' clause or principle. Whereas in the electronic publishing scenario all the traditional belief, approach and understanding about the digital documents that the library purchase / subscribe to, have a world of difference. Libraries get only a license to use the electronic information (books, journals, databases, softwares etc.) while purchasing, and even this license is issued only for a prescribed period of time. Librarians at same time, have the professional responsibility to assure uninterrupted as well as perpetual access to the information subscribed to by the library. Issues of copyright, intellectual property, and fair use are very much important to libraries [Orsdel, 2002]. Considering the prevailing

socio-cultural-economic situations, the libraries need to follow a pragmatic hybrid approach to collection development and presently they collect their resources in both print as well as electronic formats.

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. The content no longer penetrate into our research and academic libraries of higher learning through the traditional print media alone. The recorded knowledge reaches libraries by way of established scholarly publication types. A vast majority (> 70 %) of the collection finds its way into our libraries in a scattered and distributed manner and in a variety of physical, electronic media and digital formats. 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. [Terry, 2000]. 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 [Openarchives]. 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 professional 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 four 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).

1. 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 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.

2. 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 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 Library and Information Centre (LIC) at IIMK (http://intranet.iimk.ac.in/libintra/index.htm) blends digital, audio, video and print media with cutting-edge technology. The LIC aims at providing state-of-art information resources and value-added information services in all areas of management and related disciplines. The LIC is envisaged as the learning resource centre catering to the faculty, students and research scholars. It houses around 20000 books, 360 print journals, access to over 30000 E-Books and over 1070 electronic journals (full-text) sourced directly from publishers (which include EBrary.com, Elsevier Science Direct, Kluwer, John Wiley, Blackwell, Taylor & Francis, ACM, IEEE, and Emerald), and more than 12500 aggregated journal contents sourced through EBSCO and PROQUEST. A host of full-text / bibliographic online databases on scholarly information coupled with a set of national and international databases on companies, industries and the markets are available to users through a local area network (LAN) powered, high-end CD-NET Server having 2 TB storage capacity. The Web enabled Online Public Access Catalogue (OPAC) is available via the LAN.

.

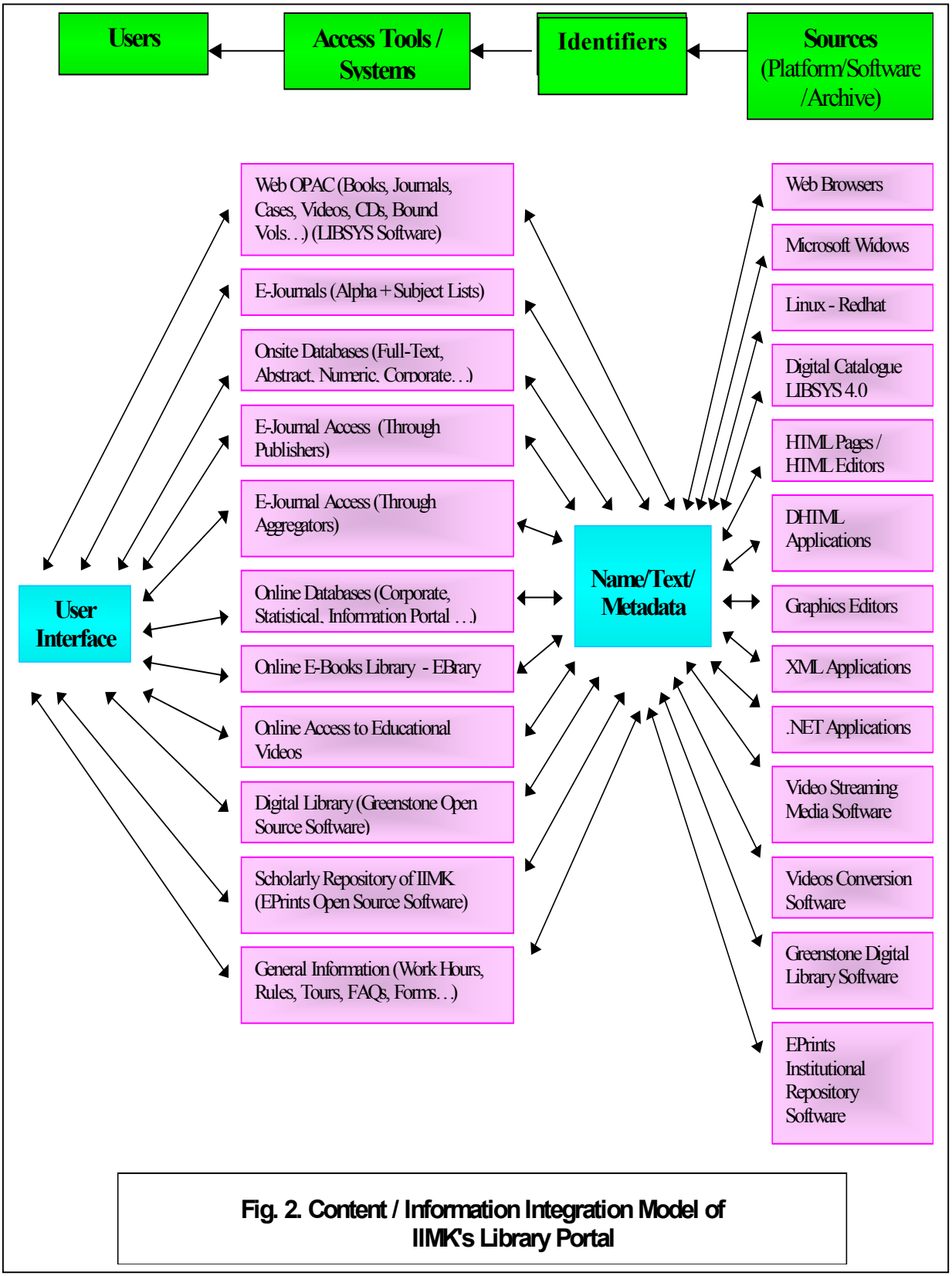

Fig. 2. Content / Information Integration Model of IIMK's Library Portal

Further more, the LIC hold thousands of CD-ROM publications on contemporary issues. The audio / visual unit in the LIC hosts over 225 educational videos covering a wide range of disciplines in management. The Library follows barcode / smart-tag based transactions and it is fully computerised. Smart-card based access control systems and Smart-gate based E-Security systems are the other distinctions of the library.

Launched as early as 1999, the IIMK Library Web Portal renders its entire set of information resources and services through a 'single window' [IIMK Library Portal]. A good Internet connectivity (2Mbps) and a state-of-art structured Local Area Network (LAN) with a giga-bit speed and fibre-optic network backbone supports the information access / retrieval requirements of the users on a 24 X 7 formula. 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 at IIMK. 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.

Fig. 3. Interface Frame of IIMK Library Portal

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. The Library Portal interface frame is shown in Figure 3.

3. 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, flexibility, availability and preservation of content. The Digital Library project sponsored by the Ministry of Human Resources Development, Government of India, has given IIMK 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 has become a reality at IIM Kozhikode and the Library Portal is now giving access to the invaluable collection hosted by the Digital Library.

IIMK is an ardent proponent of the Open Access and the Open Source Software philosophies and it was a chosen decision not to go for a proprietary digital library software. Accordingly we evaluated some of the popular Open Source Software for digital libraries, which were in use internationally. 'Dienst', 'Eprints', 'Fedora', 'Greenstone' etc. were the candidates for the preferred software. Obviously Greenstone outscored the group and we formally adopted the software for creating the IIMK digital library [Greenstone]. 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, headed by Dr. Ian H. Witten, 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' [Sourceforge].

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 [Witten, 2001] and 2003 [Witten, 2003]. 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 [Unicode]. Further more, it has a powerful search engine ‘Managing Gigabyte’ Plus-Plus (‘MG’ PP) and metadata-based browsing facilities. A very interesting feature of Greenstone is its exhaustive set of well documented and articulated manuals (http://www.greenstone.org/cgi-bin/library?e=p-en-docs-utfZz-8&a=p&p=docs) 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 primary objective of the IIMK digital library was 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 was planned in such a way that it will integrate and aggregate the existing collections and services with an outstanding user interface. Accordingly, necessary strategies were adopted towards working out the digital library system. This implied 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 was crucial in the case of multimedia integration, which was again important as it was planned to host a digital audio and video library as part of the core library collection. The System Design Architecture of the IIMK Digital Library is illustrated in Figure 4.

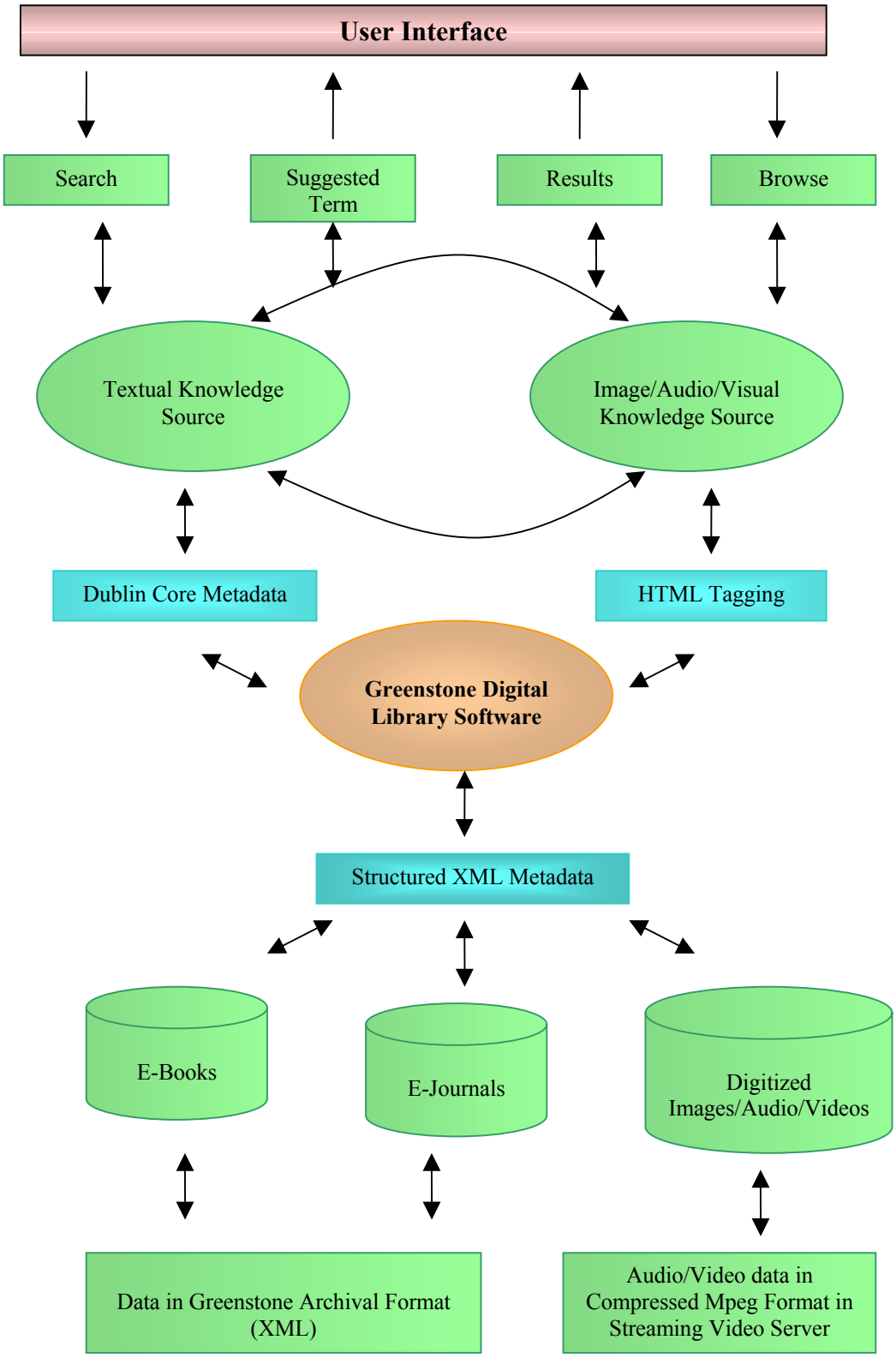

Fig. 4. System Architecture of IIMK Digital Library

Emphasis was given to maximize the efficiency and effectiveness of the information access and retrieval capabilities of the system by deploying cutting edge Resource Description Framework [RDF] standards for metadata and its XML [XML] 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 was explored to its maximum for collection building, based on the source and quality.

Standard workflow patterns were identified for the system which included 'content selection', 'content acquisition', 'content publishing', 'content indexing and storage', and 'content accessing and delivery'. The system also considered 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. A snapshot of the IIMK Digital Library interface frame is shown in Figure 5.

IIMK DL Collection

The Digital Library Project at IIMK took off during September 2001. The first four months were mainly devoted to infrastructure build-up such as hardware and software, digitization strategies, strategies regarding collection building, and finalization of the variety of content categories. Since the campus intranet and the Internet connectivity were robust, there were no worries on those fronts. The IIMK DL was slated to be populated with E-Books, ('born digital' as well as digitized), E-Journals, management cases, technical reports, working papers, faculty publications, teaching notes, CBTs, WBTs, educational videos / audios, value adding the existing OPAC, speeches, presentations etc. It took around two / three months for us to install and get proper command over the software. The collection comprise multiple digital formats such as HTML, XML, PDF, PS, RTF, JPG, GIF, MPEG etc. The DL has now got a sizable collection of over 500 E-Books covering almost all areas of business & management, literature, classics, and Information Technology. Cross collection pointers were given to project Gutenberg (classics and literature) and to other GSDL example collections (social science, environmental science) and hence the virtual collection strength is over 15000 E-Books. The existing collection also include over 1000 E-Journals (gateway access provided through Greenstone, but users can search on the title – free text, subject keywords, publisher, or other metadata), sourced from Elsevier, Kluwer, John Wiley, Blackwell, Emerald, Taylor & Francis, ACM, IEEE etc. The Videos collection streams out over 225 world classic educational videos, which is again a distinctive specialty.

Fig. 5. IIMK Digital Library Home Page

IIMK has plans to host the IIM Consortium (led by the six IIMs in India) digital archive and studies are underway on the feasibility of putting the IIM Consortium digital content under Greenstone [IIM Consortium]. A rigorous collection development drive is being worked out with most of the popular electronic publishers specializing in business and management. One of the major bottlenecks to the speedy collection development is the non-availability of proprietary and trade literature in Greenstone acceptable formats such as HTML, Word, or PDF. Most of these publishers put their materials in their own proprietary E-Book reader formats, from which the text extraction becomes almost impossible.

Considering the special features of the IIMK digital library, during September 2004, 'Greenstone' had accepted it into their Examples Collection (http://intranet.iimk.ac.in/cgi-bin/library) among around 25 such Digital Libraries across the world who could get into the Greenstone examples collection (http://www.greenstone.org/cgi-bin/library?e=p-en-home-utfZz-8&a=p&p=examples) [IIMK Digital Library].

4. Creating a Scholarly Institutional Repository (IR) using Open Source Software

Institutional Repositories are scholarly archives which draw their strength from the Open Access (OA) movement [Openaccess]. An institutional repository (IR), also called as '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 [NCSI Workshop]. 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. OA literature is digital, online, free of charge, and free of most copyright and licensing restrictions [Suber]. 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. According to Stevan Harnad, the proponent of Open Access, it 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 [Harnad].

Establishment and management of IRs could be achieved by libraries with the availability of several free, open source repository software packages like DSpace, Fedora, VT-ETD, EPrints etc. 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.

The immense potential of the interoperable institutional repositories has prompted IIM Kozhikode library to capture its intellectual capital using the open source software 'EPrints'. The GNU EPrints software (version 2.1) was downloaded from the EPrints software site 'http://software.eprints.org' [EPrints]. EPrints was further installed in a Redhat Linux machine with the help of a shell script provided by the National Center for Science Information (NCSI), Bangalore (http://www.ncsi.iisc.ernet.in). The EPrints software along with more than a dozen other supporting software were successfully installed, configured and customized according to the local needs of IIM Kozhikode. The EPrints Database was named 'eprints@iimk' and was given the Uniform Resource Locator (URL) "http://eprints.iimk.ac.in'. A static Internet Protocol (IP) number was allocated specifically for the purpose. The repository is in the making with the research publications of the Institute which include pre-prints of articles, book chapters, management cases, technical reports, theses etc. The repository is now available in the Internet for users to browse, search, or retrieve. Author, Subject, Title, Source, Year etc. could be browsed, searched and retrieved easily. Faculty, researchers and staff of the Institute have been invited to submit their research papers into the archive. The library also helps authors to submit their paper as mediated submission. The interface frame of the institutional repository of IIMK (http://eprints.iimk.ac.in) is shown in Figure 6.

Fig. 6. Institutional Repository Home Page of IIMK

Conclusion

An aggregated and integrated approach to the complex, yet challenging information management strategies to be adopted and practiced by the libraries for the 21st century is illustrated in this paper. These are shared based on our real life experience at IIMK over the past seven years. The changing landscape of the information paradigm indeed poses an array of problems and challenges to the information profession fraternity, the scholars / patrons and the scholarly publishing community. Yet, it brings along with it an unprecedented host of features and opportunities as well. The electronic information environment has many salient features which the paper paradigm cannot boast of. Among the major challenges include the information professionals' emergent need to acquire the necessary skill sets and working knowledge on the cutting-edge information science and information technology areas and in leveraging them in a contextually relevant manner. 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 next decade. 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.

Acknowledgement

The authors acknowledge their thanks to the Director, Indian Institute of Management Kozhikode (IIMK) for the encouragement and support.

References

1. OCLC Report on Five-Year information format trends. 2003
<http://www.oclc.org/reports/2003format.htm>

2. Orsdel, Lee Van; Born, Kathleen. 2002
Doing the Digital Flip.
Library Journal, 127 (7): 51-55.

3. Terry, Ana Arias. 2000
How today's technology affects libraries' collection choices.
Computers in Libraries, 20 (6): 51-55.

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

5. IIMK Library Portal <http://intranet.iimk.ac.in/libintra/index.htm>

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

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

7. Witten, Ian H. et al. 2001
Greenstone : Open-Source Digital Library Software
D-Lib Magazine, 7 (10): 1-16.

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

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

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

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

12. IIM Consortium. <http://intranet.iimk.ac.in/libintra/iimconsortium.htm>

13. IIMK Digital Library. < http://intranet.iimk.ac.in/cgi-bin/library>

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

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

16. Suber, Peter. <http://www.earlham.edu/~peters/fos/overview.htm>

17. Harnad, Stevan. <http://www.ecs.soton.ac.uk/~harnad/intpub.html>

18. EPrints. <http://software.eprints.org>

19. IIMK Institutional Repository. <http://eprints.iimk.ac.in>