DESIDOC Journal of Library & Information Technology, Vol. 29, No. 2, March 2009, pp. 7-14

Building Subject Gateway in a Shifting Digital World

Usha Mujoo Munshi

Indian Institute of Public Administration
IP Estate, New Delhi-110 012

ABSTRACT

The advancement and widespread adoption of the information and communication technology has expanded the access to information. Alternatively, deluge of digital data, and the rate at which the repositories of information are increasing are not affable because one needs to trace these individually. The paper addresses these apprehensions and focus on interoperability of these digital repositories so that subject gateways could be generated using these repositories to facilitate one-stop shop and easy access to galore of resources contained in these digital repositories, thereby enhancing access to scholarly communications.

Keywords: Digital world, subject gateways, storage and communication networks, digital repositories, digital contents, metadata, open access initiative

1. INTRODUCTION

Storage and communication networks and their widespread adoption by government, business, health care, and others offer the prospect of an age of abundant information about various areas of research in the universe of knowledge. The widespread adoption of the technology has vastly improved the access to a range of varied data sources. Undoubtedly, expanded use will greatly enhance comparative research; and the ability to harmonise datasets over time and geography unarguably will lead to significant improvement in our understanding of societies.

Scholarly community faces a variety of problems in discovery, access and delivery of pertinent information/ data in their respective areas of interest. The users want specialised portals, interdisciplinary links, intelligent integration of all types of information, and data, inclusion of quality information, and single-point access to information/data.

On one hand, a substantial number of organisations have involved themselves generating digital contents of the requisite datasets to address the immediate needs of their user clientele in a more effective and efficient way. On the other hand, it is not only unrealistic but unfeasible as well to think of sufficing the needs of users and overcoming their stresses of yielding good information crop of high relevance and precision with shortest possible mechanism from single or couple of such digital repositories. The answer perhaps could be pooling resources from these digital repositories. The next question that immediately comes to the mind is that what if these repositories are harvested by users, they could be embedded in deep web, thereby unable to retrieve (even) the metadata. Alternatively, with so much of data being digitally captured and put on the web in the repositories and the rate at which these repositories are growing, is adding the information deluge that are not affable because of the fact that one needs to trace them individually. This again is not a sociable endeavour for scholarly community more so as we are shifting more and more towards digital way. This paper is an effort in this direction to address these apprehensions and focus on interoperability of these digital repositories, so that subject gateways could be generated using these repositories to facilitate one-stop shop and easy access to galore of resources contained in these digital repositories, thereby enhancing access to scholarly communications.

2. SUBJECT GATEWAYS—WHY DO WE NEED THEM?

With large amounts of information available on the Internet, users are finding it more and more difficult to access the desired information. Several tools are being designed to help users overcome the problem of 'finding requisite bit of information'. Subject gateways are one such tool designed for a specific user group with interests in a specific subject area. Subject gateways are gaining popularity, as their design arises from the combined efforts

of the subject specialists who give their expert advice on subject details.

As indicated above, that the growing number of digital repositories that are available on the Web, is rendering it difficult for the users to track individual sites in search of necessary information. Many organisational repositories have not been indexed by the search engines due to various factors. As such, requisite mechanism by which the inter-repository resources can be traced and surfed to the top has been the need of the hour. A mechanism needs to be developed by which the repositories can share the resources and work in coordination, to provide a wider domain pursuit to the users. The mechanism which provides the ability to the information systems to work in coordination has been termed as interoperability. Open Archives Initiative (OAI) is one of the landmark efforts to ensure the availability of the metadata of digital resources of many repositories at the users' end.

2.1 Subject Gateways—Definition

Subject gateways, interchangeably also called subject portals, are primarily information gateways, or subject directories. These gateways are portals to web pages and online resources in a particular subject area. They may link to documents, databases, sites, or news items, and often provide some sort of online forum for information exchange. Subject gateways may have a search interface so that you can find what you want, or they may simply be a browseable index of resources. Instead of using a general search engine, using a subject gateway can result in returning of more relevant web pages from the search as they usually link to reliable web pages. While, using a search engine, users need to carefully go for an appraisal process about the information retrieved from the web before relying on it as a basis for further study or research. As such high quality/ recommended resources are being marked in such portals or gateways.

A formal definition of the Subject Gateways, as given by the Australian Subject Gateways Forum, is "Web-based mechanism for accessing a collection of high quality, evaluated resources identified to support research in a particular subject discipline". It is a service, which is accessed via a portal, through open standard protocols (such as Whois++, LDAP, Z39.50, Harvest Broker), which allow it to slot in seamlessly behind the scenes. What the end-user sees is web-based, behind the scenes *is* more structured service.[1]

According to DESIRE Project[2], funded by the European Commission, subject gateways find the following description. "Selective subject gateways on the Internet are characterised by their quality control. The core activities of resource selection and description rely on skilled human input (by librarians, academics and experts) and are not activities that lend themselves to automation."[3]

Examples of such gateways include *Scirus* (for scientific information only) at http://www.scirus.com/. It facilitates search across 90 million selected science web pages. If journal resources are also included, one would cross-search 17 million science article abstracts from services such as ScienceDirect, IDEAL, Beilstein on Chem Web, Neuroscion, BioMed Central, US Patent Office, E-Print ArXiv, Chemistry Preprint Server, Mathematics Preprint Server, CogPrints, and NASA.

Other examples include: AGRIGATE (www.agrigate.edu.au) and BUBL (http://bubl.ac.uk/link/sci.html).

3. SUBJECT GATEWAYS IN THE PRESENT STUDY

In the present study, the development of subject gateways or portals is to create such portals by harvesting quality information rich digital repositories created and maintained by various organisations/agencies. These repositories are based on standard practices and using well known standards that facilitate interoperability and thus can allow harvesting of data from across the repositories and provide a one-stop shop for information access to resources that are relevant for information requirements. These digital repositories, generally institutional in character, are not merely the bibliographical type repositories, but substantial portion of such repositories contain full-text documents with passable search features. The resources contained in these repositories are open access resources, and by and large the access to full text is open and free.

However, creating subject portals in this case would need setting up of a Harvester service that is OAI-PMH compliant, so that the metadata can be harvested which provides a conduit to reach the identified relevant documents. But before we actually touch upon the aspect of generating subject portals, it becomes imperative at this point of time to throw some light on critical factors for success of sustenance and execution of gateways.

4. CRITICAL FACTORS

The critical success factors for a subject gateway or a portal broadly encompass the following:

- Collaboration
- The application of standards
- Cross-gateway searching
- Interoperability

4.1 Collaboration

While collaboration is heart and soul of the subject gateways, without which creating a subject portal will not be a portal in true sense of the term, anyway. Unless there are resources that can form integral part of the gateway, otherwise how can a gateway be set-up? Hence North-South and South-South collaboration is a must to do practice required for populating the gateway with high relevance, quality and substantial pool of resources. Collaboration ensures the need to pool efforts early in the gestation of the proposed service. One aspect that is very core to gateways/portals is they are subjected to on-going development process, that may amount to periodical updation/modification/deletion/addition of resources only then such gateways ensure an authenticated repository for scholarly communication.

4.2 Standards

The application of standards is also equally important. The fact that it is only the standard practices that will lead to successful resource sharing in the digital environment necessitates the use of standards, e.g., the standards to be considered for metadata. There are several available metadata schemas such as the Dublin Core being used as a suitable baseline in many cases.

Cross-gateway searching is a vital feature to be facilitated. For the support of cross-gateway searching, there are standards such as Z39.50, for data export or import, and proxies which query repositories simultaneously across multiple servers. Extensions to services such as cross-browsing require the development of other standards, for example in the area of collection description metadata. However, cross-searching and cross-browsing could be facilitated by simple links between existing pieces of information, rather than building whole new structures which individual institutions need to commit to, there may be some hitches in committing this. It may pointed out here that the ubiquitous standard which underpins all of the gateways is the use of the hypertext transmission protocol in web browsers for which the gateways interfaces are designed, particularly the gateway projects in their early stages would not have been successful, without this uniform technology.

4.3 Interoperability

The interoperability aspect, is again a most indispensable issue as that forms part of harvesting data from across repositories, across million of records contained in these repositories. What is interoperability in real sense of the term and how it does what it can? Interoperability is an inexorable in generating the subject portals using harvesting services. Broadly speaking, interoperability is the ability of systems, organisations and individuals to operate together to achieve a common goal. In digital libraries search interoperability is the main concern. Therefore, more emphasis has been given to this aspect in this study.

It may be worthwhile here to throw some light on the aspect of what interoperability means. Priscilla Caplan[4] has defined Search Interoperability as 'the ability to perform a search over diverse set of metadata records and obtain meaningful results'. To bring such interoperability among digital libraries, remarkable effort has been made by OAI. OAI develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content. It is supported by the Digital Library Foundation, the Coalition for Networked Information and National Science Foundation. OAI has its roots in an effort to enhance access to e-print repositories (archives) as a means of increasing the availability of scholarly communication.

4.3.1 Metadata Harvesting Protocol

As the term denotes, the metadata harvesting protocol signifies the set of rules or guidelines harvesting metadata. In order to facilitate metadata harvesting, there ought to be some prerequisites to develop metadata harvesting protocol. As such agreement on: certain aspects like Transport protocol—HTTP or FTP or other such protocol; Metadata format—Dublin Core or MARC or other such format; Metadata Quality Assurance—mandatory element set, naming and subject conventions, etc., and Intellectual Property and Usage Rights is essential.

4.3.2 OAI-PMH

The OAI protocol for metadata harvesting provides an application-independent interoperability framework based on metadata harvesting[5] that can be used by a variety of communities who are engaged in publishing content on the Web. It provides a set of rules that defines the communication between systems like FTP or HTTP in Internet. That is why even though the protocol actually uses HTTP as a transport mechanism between digital libraries, it is popularly known as the "HTTP of Digital Libraries".[6]

There are two classes of players in the OAI-PMH framework: data providers, which administer systems that support the OAI-PMH as a means of exposing metadata, and service providers, which use metadata harvested via the OAI-PMH as a basis for building value-added services.

The protocol based on HTTP and XML was developed with an objective to ensure interoperability between e-print repositories only. Later, in version 1.0/1.1 all document like digital objects were also brought into its purview, and finally the latest version 2.0 supports all kinds of digital resources.

It may be noted that OAI-PMH is not a search engine, or a search tool, or a database. It only provides set of rules to move the metadata not the content of the digital resource from one repository to another. The content remains in the source repository only, and a repository can act both as service provider or harvester and data provider or only service provider or data provider. Besides, the protocol is not restricted only to support simple (unqualified Dublin Core), but can support any metadata schema, which can be provided in XML format.

The use of OAI-PMH is generally recommended due to its very simplistic and convenient features. For instance, implementation is quite simple as compared to distributed search protocol (Z39.50). Besides, search process is faster in centralised search mechanism because the metadata collected from different repositories are kept at local server and retrieved from there at the time of search. On the contrary, in distributed search mechanism (Z39.50) metadata is retrieved from different repositories at the time end-user submits a search query. Moreover, the service providers can modify the display of collected metadata as per their metadata policy. The interaction between service provider and data provider and the process of metadata harvesting is broadly schematically depicted in Fig. 1.

5. GENERATING SUBJECT PORTALS

Generating subject portals or gateways by harvesting OAI complaint information resources in the areas of interest, involve certain key technologies to make the system ready for take off. Some such aspects are briefly touched upon here. As indicated earlier, that the two classes of participants in the OAI-PMH framework are Data Providers and Service Providers. In generic terms, the role of both these participants can be performed.

Firstly, by creating the repositories that comprise of intellectual output of the institutions and are generally termed as institutional repository (IR). Such IRs can be created using open source software. There are several software now available under open source license that comply with OAI metadata harvesting protocol and are released and made publicly available.

Some of these IR software are: ARNO, Tilburg University, The Netherlands (http://www.uba.uva.nl/arno); CDSware CERN, Geneva, Switzerland (http://cdsware.cer.ch/); DSpace, MIT Libraries and the HP Labs, USA (http://wv.w.dspace.org/); E-Prints, University of Southampton, U.K. (http://www.software.eprints.org/); FEDORA University of Virginia, USA (http://www.fedora.info/); i-Tor, the Netherlands Institute for Scientific Information Services (http://www.itor.org/en/toon); and MyCoRe, University of Duisburg-Essen, Germany (http://www.mycore.de/engl/index.html). A list of such software is also available at http://www.soros.org/openaccess/software. It may be mentioned here that some of these software also support search and retrieval of repositories in vernacular languages such as Indic languages using UNICODE. For creation of digital

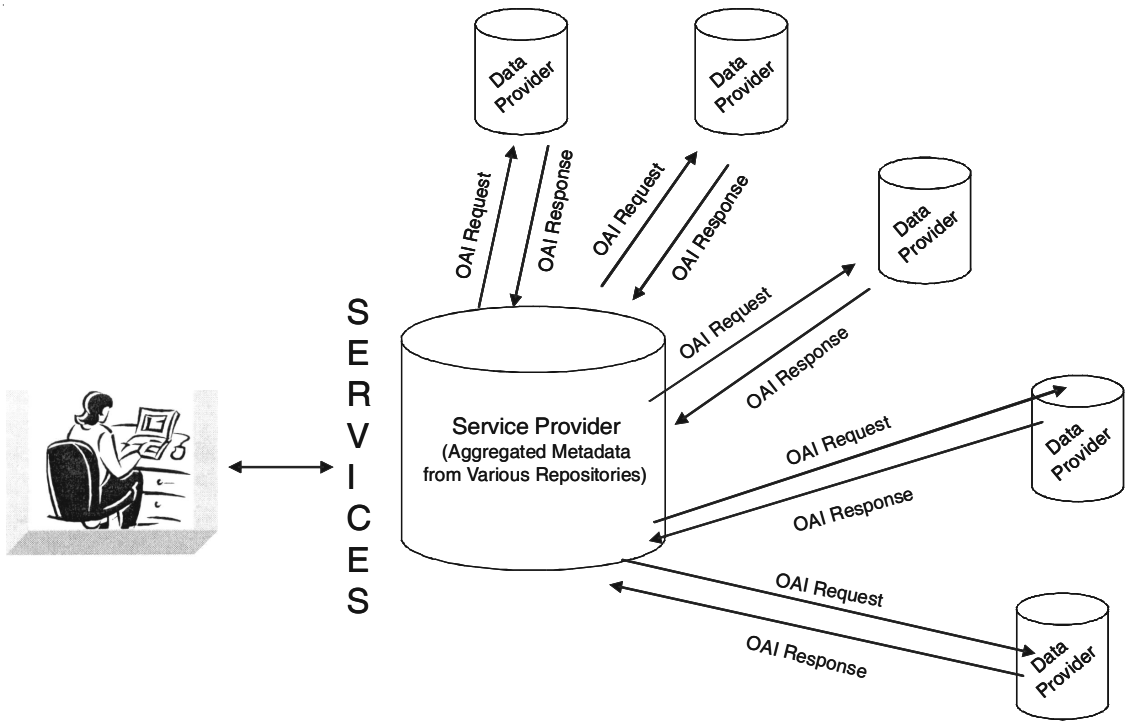

Figure 1. A schematic representation of transactions between service provider and data provider.

libraries in vernacular languages using UNICODE, the Unicode needs support from various quarters such as operating systems, programming languages, application software, word processors. For instance in the case of compatibility of UNICODE and Dspace, the requisite parameters are supported. Linux supports Unicode; PostgresSQL supports Unicode; Java supports Unicode. You only require fonts to display in the browser and you need to modify Tomcat files.

Once these repositories have been created, they ought to be OAI compliant, and interoperable, so that others can also harvest them and use. Secondly, by harvesting the quality IRs created by others by setting up harvester using harvesting services. The second point is more important for the present study as our main focus of the study is based on this. This can be achieved as follows.

It becomes imperative to talk about the key concepts associated with the harvesting services. What is a harvester and what it does is a question that we need to answer.

5.1 Harvester

A harvester is a client application that issues OAI-PMH requests. A harvester is operated by a service provider as a means of collecting metadata from repositories; where a repository is a network accessible server that can process the OAI-PMH requests in the scheduled manner. The data provider manages the repository in order to expose metadata to harvesters.

To allow various repository configurations, the OAI-PMH distinguishes between three distinct entities related to the metadata made accessible by the OAI-PMH.

(i) Resource: A resource is the object or "stuff' that metadata is "about".

(ii) Item: An item is a constituent of a repository from which metadata about a resource can be disseminated. That metadata may be disseminated on the fly from the associated resource, cross-walked from some canonical form, actually stored in the repository, etc.

(iii) Record: A record is metadata in a specific metadata format. A record is returned as an XML-encoded byte stream in response to a protocol request to disseminate a specific metadata format from a constituent item.

It may be worthwhile to throw some light on two other important aspects, i.e., unique identifier and selective harvesting.

Unique Identifier: A unique identifier is used in OAI-PMH request for extracting metadata from the item. It explicitly identifies an item within a repository. Though, items may contain metadata in multiple formats, however, the unique identifier maps to the item and all possible records available from a single item share the same unique identifier.

Selective Harvesting: Selective harvesting allows harvesters to limit harvest requests to portions of the metadata available from a repository. The OAI-PMH supports selective harvesting with two types of harvesting criteria that may be combined in an OAI-PMH request: datestamps and set membership. Harvesters may use datestamps to harvest only those records that were created, deleted, or modified within a specified date range. To specify datestamp-based selective harvesting, datestamps are included as values of the optional arguments.

5.2 OAI Tools

There are several OAI-implementation tools, which can be broadly grouped according to their role in the OAI-implementation process: [7-9]

- OAI data provider tools and scripts,
- Digital library systems, and
- OAI validation and harvesting systems.

5.2.1 OAI Data Provider Tools and Scripts

A wide array of technological resources serves the needs of OAI data and service providers, ranging from simple scripts to entire systems. OAI data provider tools and scripts expose metadata via the OAI-PMH.

When selecting tools for managing metadata, consider not only whether they meet minimum OAI requirements (simple or unqualified Dublin Core format) but also whether they include other features for increasing the quality of shareable metadata. Some tools and scripts in this group are:

- Metadata Migrator Tool (http://metacluster.library.emory.edu/mosc/upload.php)
- OAI Static Repositories and Gateways (http://www.openarchives.org/OAI/2.0/ guidelines-static-repository.htm)
- OAI FileMakerPro Repository Gateway (http://cic harvest.grainger.ui uc.edulfmpgateway/
- OCLC OAIC (http://alcme.oclc.org/oaicat/)
- Open Journal Systems (http://www.pkp.ubc.ca/ojs)

- Virginia Tech DLRL Projects (http://www.dlib.vt.edu/projects/OAI)
- eXtensivle Stylesheet Language Transformation (http://www.w3.org/TR/xslt)
- ZMARCO (http://zmarco.sourceforge.net)

5.2.2 Digital Library Systems

Digital library systems employ tools for creating, editing, managing, and sharing metadata. The following are a selection of these systems with strong OAI-data provider capabilities. More vendors are now providing open source versions of their tools. Some of these are:

- ContentDM (http://www.contentdm.com)
- Collection Workflow Integration System (http://scout.wisc.edu/Projects/CWIS).
- Digital Library eXtension Service (http://ww.dlxs.org)
- DSpace (http://www.dspace.org)
- Ex Libris DigiTool (http://www.exlibrisgroup.com/digitool.htm)
- Fedora (http://www.fedora.info)
- Greenstone (http://www.greenstone.org/cgi-bin/library)

5.2.3 OAI Validation and Harvesting Systems

OAI validation and harvesting systems collect and display shareable metadata. Validation systems may be used to explore metadata records before they are harvested. Minimally, harvesters gather metadata records so they can be indexed, searched, and browsed. The easier the harvester is to use, the more compatible it is with current systems. Ideally, harvesters should search sets rather than entire repositories and use resumption tokens. Some of these systems are:

- OAITransform (http://www.dlxs.org/products/archive-by-CDROM/11a/Lib/src/bin/o/oaister/oaitransform/OAI Transform)
- Perl O-O Harvester (http://www.dlib.vt.edulprojects/O*AI*/software/harvester/harvester.html)
- PKP Open Archives Harvester (http://www.pkp.ubc.cal pkp-harvester/)
- Reap and OAI Harvester Object Library (http://sourceforge.net/project/showfiles.php?group_ id=47963&package _id=46165)
- Repository Explorer (http://www.dlib.vt.edulprojects/OAI/software/explorer/explorer.html)

5.3 Open Source Metadata Harvesting Tools[10]

There are several OAI Harvester software. A list of the major OAI harvester software is given below for reference.

- Arc (http://arc.cs.odu.edu/)
- Citebase (http://citebase.eprints. org/cgi-bin/search)
- CYCLADES (http://www.ercim.org/cyelades/)
- DP9 (http://arc.cs.odu.edu:8080/dp9/index.jsp)
- MelND (http://www.meind.del)
- METALIS (http://metalis.cilea.it/)
- my.OAI (http://www.myoai.com)
- NCSTRL (http://www.ncstrl.org/)
- Purseus (http://www.perseus.tufts.edu/cgi-bin/vor)
- Public Knowledge Project–Open Archives Harvester (http://pkp. ubc. calharvester/)
- OAICA T (http://www.oelc.org/research/sojtware/oai/cat.htm)
- OAI Repository Explorer (http://re.cs.uct.ac.zal)
- OAIster (http://oaister.umdl.umich. edu/o/oaister/)
- OASIC (Open Archvies en SIC) (http://oasic.ccsd.cnrs.jrl)
- OAIHarvester (http://www.oelc.org/research/software/oai/harvester.htm)
- DLESE OAI Software (http://dlese.orgloailindex.jsp)

5.4 Open Archives Harvester: Installing and Harvesting

In this study we would attempt to focus on one of above indicated OAI Harvetser—Public Knowledge Project (PKP)—Open Archives Harvester (http://pkp.ubc.ca/harvester/) for discussion purposes. It may be pointed out that many organisations have set up this harvester to harvest OAI compliant repositories for facilitating subject gateways in their areas of interest. A list of such implementations is available at http://pkp.sfu.ca/harvester-list.

It may be pointed out here that there are some difficulties or challenges encountered in harvesting metadata. For instance a metadata harvester needs to start and stop retrieval at arbitrary points in the set to allow parts of large collections to be downloaded. Much data is corrupted and does not parse well using a tool and at

times most of the work requires transforming broken data into usable sets.

5.4.1 PKP Harvester

The PKP Open Archives Harvester is a software used to accumulate and index freely available metadata, providing a searchable, web-based interface. It is open source, released under the GNU General Public License. It was created and is maintained by the Public Knowledge Project, in Vancouver, Canada. Originally developed to harvest the metadata from Open Journal Systems articles and Open Conference Systems proceedings, the Harvester can by used with any OAI-PMH-compliant resource. It can harvest metadata in a variety of schemas (including unqualified Dublin Core, the PKP Dublin Core extension, the Metadata Object Description Schema (MODS), and MARCXML). Additional schemas are supported via plugins.

PKP search engine, the Harvester version 2.x includes many search goodies. Searching is highly scalable, it creates an inverted index for searching. Flexible search interface allows simple searching and advanced searching using crosswalk fields from all harvested archives. Advanced searching of archives that share the same schema is possible using fields as defined in the schema. When creating crosswalks for searching, administrator can define elements—text, date, or HTML multiple select interface widgets.

PKP graphic user interface uses impressive graphic interfaces, having an imposing design. It is easy to navigate the site and find the administrative tools to harvest archives and search. PKP user interface with CSS and template-based HTML facilitates easy customisation.

The entire site, including harvester and search engine works out of the box. The site installs easily into a LAMP-based server (Linux-Apache-MySQL-PHP4) without writing configuration files or installing non-LAMP dependencies. The code is readable and well documented. This software is flexible, so it can be installed on Windows using open source tools alone.

Harvesting in PKP can be managed through a web page from a server running PHP and MySQL. In the browser adding an archive and running it is simple. The PKP OA harvester allows any institution to create their own metadata harvester, which can be focused specifically on gathering information from/or for their research community.

By setting the harvester, a union catalogue of potential digital resources is basically created, which provide access to these digital resources by harvesting their descriptive metadata or records using OAI-PMH (the Open Archives Initiative Protocol for Metadata Harvesting). These digital resources can be picked on selective basis depending upon the subject areas of interest of the users or simply as per the institutional mandate. It may be indicated here that generally these resources do not surf to the top during navigation using general search engines; these resources are often hidden from search engine users behind web scripts. These resources also known as the deep web or embedded web. The owners of these resources share them with the world using OAI-PMH. The harvester facilitates harvesting metadata of documents from other repositories and mounting it on a local server, while the actual document rests in distributed mode on the respective servers. The best example of such a facility can be accessed at http://www.oaister.org/which currently provides access to 11,682,009 records from 777 contributors[11]. Other examples include http://ir.isical.ac.in., which is a pack of ISI repository and information services (IRIS) in which HORUS (Harvesters for Open Repositories with Unlimited Search) is a suit of information services based on OAI-PMH (Open Access Initiative Protocol for Metadata Harvesting) HORUS collects metadata from various digital repositories dealing with subjects like computer sciences, biological sciences, etc. and provides a single-stop search engine for full-text resources in the respective subjects.

The philosophy of identifying the repositories lies in the fact that once you have identified the repositories on selective basis meeting your desired information requirements, and indexed them on to your site, the automatic refurbishing is taken care of as and when additional/new records are added by the data provider in the respective repositories.

6. PKP SOFTWARE INTEGRATION: AUGMENTING IRS WITH HARVESTER SUPPORT

Using OAI Harvester for metadata harvesting, the repositories can be located and harvested for scholarship purposes. It is necessary to exchange knowledge between repositories; avoiding heterogeneity issues and giving an interpretation of metadata harvested. While setting up of harvester supports harvesting OAI-compliant digital repositories created by other organisations/ institutions/ agencies, it also facilitates integration of institutions intellectual output, so as to make provision for searching across repositories including the ones created by users. One model for implementing the complete set of PKP software is for libraries to create an integrated site that brings together their institutional repositories, along with Open Journal Systems, and Open Conference Systems. In addition, the Open Archives Harvester is used to index and search the content of the site and also harvesting OAI-compliant digital repositories, thereby making it part of a global knowledge exchange.

By doing so, on one hand, this allows the library to offer faculty members a way of depositing their work, wherever it is published or whatever state it is in, as well as supporting their involvement in the running of journals and conferences. It also creates a single point of search for all that the organisation is doing along these lines.

Taking example of an open source software such as Dspace for creating digital libraries comprising the institutional output or what is generally termed as IR, it is possible to integrate Harvester for providing single-point access to galore of digital repositories, the ones you have created or which others have created.

As indicated earlier, a Harvester is basically a client application that issues OAI-PMH requests It is operated by a service provider as a means of collecting metadata from repositories. DSpace has data provider functionalities. Thereby augmenting DSpace with the harvester support and then, importing DSpace in the broader context of IR is feasible and more ingenious. For instance, a good example of a library integrating this set of open source software is the UTSePress (http://epress.lib.uts.edu.au/) at the University of Technology at Sydney, which provides access to the integrated services: the services bundled here are Institutional Repository (using DSpace), Journals (using OJS), Conferences (using OCS), and Search (using PKP OA Harvester). Other examples may include, IRIS (http://ir.isical.ac.in/) set up at Indian Statistical Institute, Kolkata which has several components that include ORION (Open Repository of Isi Online), and an IR comprising intellectual output of ISI faculty and scholars. The IR is created using Dspace open source software. Besides there is HORUS as discussed above which is basically harvesting service of open repositories using PKP-OA harvester. In addition there is also provision for federated search.

7. CONCLUSION

To capitalise on availability of open access digital repositories, and developing new services involving the broadest possible group of information service providers is of utmost priority. As such collaboration to research is the mantra of the day. Such initiatives beyond doubt will help to progress towards a stronger national information channel infrastructure for an integrated research libraries and information network. Where these resources are being developed—across institutions/across regions—and facilitated would provide for enhanced access to scholarly communications that were hitherto not accessible due to various factors. The potential to integrate services is very strong. The subject gateways have established a new trend in collaboration, which can only revitalise library service provision in the future. As change is inevitable, we need to move with this change to support information landscape for research and development or scholarship purposes. The only constant in the changing environment of provision of bytes has been the use of metadata itself. Everything else is subject to change.

REFERENCES

1. Australian Subject Gateways Forum (ASGF) http://www.nla.gov.aulinitiatives/sg/gateways.

2. DESIRE Project. http://www.desire.org/.

3. Selection Criteria for Quality Controlled Information Gateways, Work Package 3 of Telematics for Research project. http://www.lub.lu.se/~traugolt/OIR-SBIG.txt.

4. Caplan, Priscilla. Metadata Fundamentals for All Librarians. American Library Association Editions. 2003.

5. The Open Archives Initiative Protocol for Metadata Harvesting Protocol Version 2.0. http://www.openarchives.org/OAI/openarchivesprotocol.html#Item.

6. Rhyno, Art. Using Open Source Systems for Digital Libraries. Westport: Libraries Unlimited. 2004, pp. 26.

7. Kellogg, D. Open source OAI metadata harvesting tools. *California Digital Library* http://www.cdlib.org/inside/projects/preservation/recall/harvesting.pdf.

8. Tennant, R. Specifications for metadata processing tools. *California Digital Library,* 2004 http://www.cdlib.org/inside/projects/harvesting/metadatatools.htm. Visited on 13 May 2007.

9. Tennant, R. Bitter harvest: Problems and suggested solutions for OAI-PMH data and service providers. *California Digital Library,* 2004. http://www.cdlib.org/inside/projects/harvesting/bitter_harvest.html.

10. Open Archives Forum: Information Resource Database. http://www.oaforum.org/oafdb/listdb/listsoftware.php#top.

11. OAIster. http://www.oaister.org/.