Designing Single-Window Search Service for Electronic Theses and Dissertations through Harvesting

Prosenjit Sarkar, ***Superintendent (Library Services)***
Central Library, The University of Burdwan, Burdwan-713104,
prosenjit.toton@yahoo.co.in, prosenjit.toton@gmail.com

Dr. Parthasarathi Mukhopadhyay, ***Sr. Lecturer***
Department of Library and Information Science
The University of Burdwan, Burdwan – 713 104
parthasarathimukhopadhyay@gmail.com

INTRODUCTION:

Teaching and research are two important aspects of a university education system. Therefore, it is essential to archive, preserve and disseminate the intellectual output of a university. In this context, theses and dissertations may be viewed as important showcases of a university's achievement and contribution in higher education[16]. At present, a large number of universities and higher learning institutes are investing money and efforts in creating, archiving and disseminating the citations as well as theses and dissertations including ETDs (Electronic Theses and Dissertations) in institutional repositories. Obviously without a good search platform, full texts and even citations of these researchers' scholarly works will not be accessible through different digital channel including Internet. The web-based institutional repositories are distributed across the globe and using different software solutions. Fortunately, there is an array of OAI-PMH compatible open source software for archiving digital knowledge objects (e.g. DSpace, E-Print archive, Fedora, Greenstone etc.). These digital media archiving software differ considerably in scope and retrieval functionalities. As a result, users suffer a lot in terms of the load of learning retrieval features of different software. In this regard, this system has designed a prototype union catalogue of electronic theses and dissertations (ETDs), as a single-window search interface to facilitate the retrieval of electronic theses and dissertations on global scale. The

Preprint version of the following paper

Sarkar, P and Mukhopadhyay, P. (2010). Designing single window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies,*57(4):354-364.

aim of this union catalogue is to build an easy search service model using automatic record harvesting to provide access to database of worldwide institutional repositories of ETDs. Of course the repositories need to be compatible with interoperability standard(s) for building a search service on the basis of harvested metadata. Thus we considered interoperability is the process of resource discovery and access. The OAI's Protocol for Metadata Harvesting is a robust solution of interoperability problem, which defines a standard method of metadata exchange.

UNION CATALOGUE OF ETDs:

A union catalogue is a combined library catalogue describing the collection of a number of libraries (see http://en.wikipedia.org/wiki/union_catalog). Union catalogues have been created in a variety of media – networked electronic database is one of them. Union catalogues are useful to librarians as well as researchers, because they help in locating and requesting resources from other libraries, which is a part of library cooperation. Union catalogue of ETDs may assist researchers or library users to search collections of distributed ETDs across the globe. The visibility of union catalogue on ETD may improve considerably with its online availability through the Web. The online union catalogue of ETDs is a system that discovers, organizes and stores metadata records from heterogeneous IDRs[3], and thereby provides a single-window access interface for distributed ETDs.

There are two processes to build a union catalogue, i.e., virtual (distributed) and centralized[5]. In this project, centralized approach has been used to build ETD union catalogue, which is based on Open Achieve Initiative (OAI) framework[12]. The IDRs containing ETDs are different in their subject coverage, software usage, nature of contents and most importantly in retrieval techniques and tools. As a result it is difficult for end users to search these repositories comprehensively that provide free access to ETDs. This situation demands for the development of

a single window search service covering all the IDRs. This single window search service (based on metadata harvesting) is advantageous to researchers and others as it brings them closer to uniform access interface for scholarly information bearing objects[2].

BACKGROUD OF OAI CONCEPTS:

At present time, the Internet has made easy and rapid dissemination of information. Internet has given ample opportunities to research scholar to share their research result among the colleagues in their subject field. But the traditional publication process (in print medium) slow downs the information dissemination or exchange. The Internet has solved this problem. The revolution in Information and Communication Technologies and the need for rapid dissemination of scholarly literature have forced Higher Academic and Research Institutions to build e-print archive or digital library. On the other hand World Wide Web created information management problem, but there was no simple solution to solve this problem. The problem includes persistence, reliability and authenticity[14]. The most important issue was the perpetual access. Another problem is the issue of authenticity of a source producing digital objects. These and some other problems forced Higher Academic and Research Institutions to build e-print archives or repositories of digital objects that called Institutional Digital Repository (IDR) with some enhanced service to the user and as a result thousands of IDRs emerge all over the World. Most of these IDRs developed arbitrarily. The concept of Open Archive Initiative was lying with the emergence of increasing number of e- print archives.

The concept “Open Archive” defines information access at free of cost and without any restriction. The Open Archive Initiative means an Open Archive, which implements OAI Protocol

Preprint version of the following paper

Sarkar, P and Mukhopadhyay, P. (2010). Designing single window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies,*57(4):354-364.

for Metadata Harvesting to allow remote archive to access its metadata using an open standard. A Repository or IDR is synonymous to an Open Archive. In the context of OAI, a Repository or IDR should be network accessible that supports OAI-PMH.

The OAI is an international effort to enhance access to e-print archives and also a current important topic in digital library research and implementation. OAI can be regarded as a lightweight transport mechanism to expose and disseminate metadata, as a standard-based technology for document transmission and discovery, and changed component in the changing scholarly communication framework[6].

The network connected IDRs have been built arbitrarily in isolation using different software platform, technologies, protocol, and data structures. These differences prevent in the development of digital library services that enable users to use single search interface to discover resources from multiple Institutional Digital Repositories. The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) is an international effort to address this interoperability problem among distributed IDRs.

INTEROPERABILITY AND CROSSWALK:

Interoperability refers to many diverse aspects of archive initiatives. It includes – metadata formats, architectural framework, their openness to create third-party digital library services, their integration with existing process of scholarly communication, usability in cross-disciplinary context etc.[15]. In the technical arena it is supported by open standards for communication between systems and for description of resources and collections, among others.

There are many definitions of Interoperability have been given. Some examples are given below: -

Preprint version of the following paper

Sarkar, P and Mukhopadhyay, P. (2010). Designing single window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies,*57(4):354-364.

" Interoperability is the ability of multiple systems with different hardware and software platforms, data structures, and interfaces to exchange data with minimal loss of content and functionality"[8].

"Interoperability is the ability of two or more systems or components to exchange information and use the exchanged information without special effort on either system"[1].

In general 'Interoperability' can be defined as the compatibility of two or more different computer systems so that they can exchange data and information and can use the exchanged data and information without any kind of manipulation or loss. Crosswalks facilitate interoperability. A crosswalk is "a mapping of the elements, semantics, and syntax from one metadata scheme to those of another"[8]. At present, crosswalks are common method to achieve interoperability between or among metadata schema. It allows metadata created by one community to be used by another group that employs a different metadata standard. Interoperability and crosswalk ensures exchange of bibliographic data among heterogeneous IDRs across the globe. An example of crosswalks between ETD-MS and Marc-21 (see http://www.ndltd.org/standards/metadata/etd-ms-v1.00-rev2.html) is shown in table 1.

Examples of Crosswalk Specification:

ETD-MS Element	MARC-21 field & subfield	Description
dc.title	245a	A name given in the title page of the resource. e.g. Title of the theses
dc.creator	100a	Author/Researcher who is responsible for writing the theses.
dc.subject	653a	Topic of the content of the theses or keywords.
thesis.degree.name	502a	Name of the degree for which the theses is submitted. e.g. 'Masters' or 'Doctoral'.

Table: 1

Preprint version of the following paper

Sarkar, P and Mukhopadhyay, P. (2010). Designing single window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies,*57(4):354-364.

OPEN ARCHIVES INITIATIVE- PROTOCOL FOR METADATA HARVESTING (OAI-PMH):

An Open Archive is a computer interface to access a collection of data, where the interface must be OAI-PMH (Open Archives Initiative - Protocol for Metadata Harvesting) complaint. Current standard OAI is the Protocol for Metadata Harvesting, which describes how two IDR systems can communicate a stream of structured records from one to the other on a periodic basis. An IDR system that contains the source data and that conforms to this protocol is called an Open Archive.

The Open Archives Initiative was launched in October 1999 in an attempt to address interoperability problem among the many existing and independent IDRs. The aim of this Initiative was on high-level communications among IDRs and the simplicity of protocols[15]. The OAI-PMH is different in comparison to earlier approaches to interoperability. The primary purpose of this protocol is incremental bulk transfer of metadata (harvesting). This protocol does not provide remote searching facility – instead of remote searching, a service provider acquires data from data providers, stores and processes it locally and then provides services to its users[11]. This protocol helps to make possible metadata available as resource discovery. The OAI-PMH is a request-response protocol that allows one computer to collect metadata incrementally over time from another computer, which is called harvesting[13]. The protocol is simple to design and suitable for use in various communities in different contexts. To harvest metadata into the union catalogue of ETDs, this project has adopted OAI-PMH interoperability framework. OAI-PMH facilitates an application independent interoperability framework to share metadata among different heterogeneous IDR systems[4]. The framework is now mainly used for the worldwide consolidation of open access archives that deals with ETDs. Here, the system attempts to design a single window search service for ETDs through harvesting metadata from participating IDRs. It is

Preprint version of the following paper

Sarkar, P and Mukhopadhyay, P. (2010). Designing single window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies,*57(4):354-364.

important to mention here that this single window search service only provides access to metadata – it does not provide the resources themselves. The union catalogue archive of ETDs is designed to function as both service provider and data provider. It harvests metadata from remote IDRs, stores harvested metadata in an internal database and reproduce this metadata through data provider interface. On the other hand this system may also act as data provider for other harvesting tools and systems. There are two participating components in OAI-PMH framework i.e. I) Data Provider; and II) Service Provider.

DATA PROVIDER:

Data Providers administer systems that support the OAI-PMH as a means of exposing metadata[4]. A data provider maintains a repository that allows external online access to metadata through OAI-PMH. Any data provider archive must satisfy some basic requirements. The archive must have an online interface and a web server that can be used for the purpose of protocol.

SERVICE PROVIDER:

Service Providers use metadata harvested via OAI-PMH as a basis for building value-added services[4]. A service provider is an entity that harvests metadata from data providers in an attempt to provide higher quality services to users. Once the OPACs (Online Public Access Catalogues) of institutions are made OAI-complaint, the base URLs of such OPACs are to be registered with an OAI-based service provider. The service provider will then harvests metadata from the registered OPACs (Data Provider) on a periodic basis.

There are some advantages of implementing OAI-PMH framework are given below[17]: -

I) It provides a new model for scholarly communication in the Internet era. OAI solves the

problem of Interoperability;

II) To implement OAI-PMH is easy and simple;

III) It is open; and

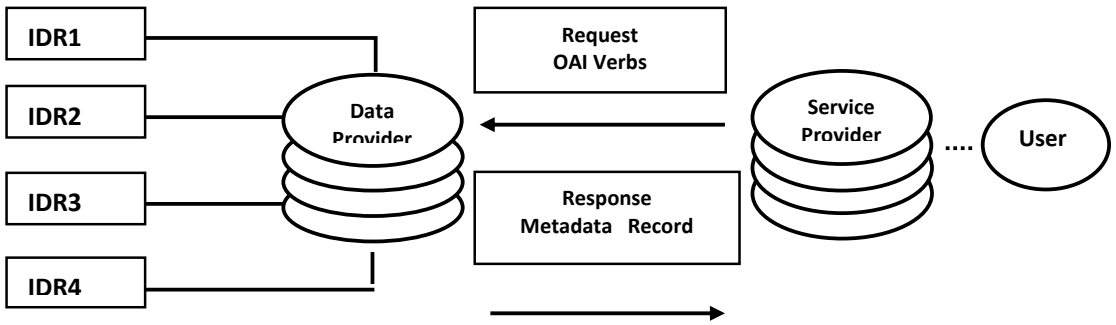
IV) It uses current web standards wherever applicable.

The OAI-PMH supports six verbs that are used to send requests to the Data Providers. Table 2 describes this.

Response to these OAI-PMH verbs request are given below

Request verb	Response function
Identify	To retrieve information about description of IDR, e.g. name of the repository, the protocol version, administrator's email etc.
ListMetadata Formats	To retrieve the metadata format from the IDR.
ListSets	To retrieve the sets and subsets from an IDR.
GetRecord	To retrieve an individual metadata record from an IDR.
ListIdentifiers	To retrieve unique identifiers from an item.
ListRecords	To harvest records from an IDR.

Table: 2



Functions of Data Provider and Service Provider in OAI-PMH Framework

Preprint version of the following paper

Sarkar, P and Mukhopadhyay, P. (2010). Designing single window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies,*57(4):354-364.

Fig.1

DESIGNING FRAMEWORK:

Design and development of harvesting framework for ETDs requires following steps, strategies and planning. There are three major components -

i) Development of software framework;

ii) Selection, installation and configuration of harvesting tool; and

iii) Selection of IDRs and collection of essential attributes for harvesting (name of the IDR, resource URL, base URL, etc.).

The prototype harvesting framework named as UniTHES, is based on open source software and open standards. This framework, like many other systems is based on LAMP architecture. The versions of different software are as follows –

- PHP >= 4.2.x (including PHP 5.x); Microsoft IIS requires PHP 5.x
- MySQL >= 3.23.23 (including MySQL 4.x/5.x)
- Apache >= 1.3.2x or >= 2.0.4x or 2.0.5x /Microsoft IIS 5.x or 6.x
- **Operating system (OS):** Any OS that supports the above software, including Linux, BSD, Solaris, Mac OS X, Windows (preferably NT based Windows flavors).

As a whole, the use of the open source software in developing the framework depends on a structured methodology. The steps related to the process of the designing harvesting framework may be divided into three major groups[7].These are as follows –

Group A: LAMP related activities

Preprint version of the following paper

Sarkar, P and Mukhopadhyay, P. (2010). Designing single window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies,*57(4):354-364.

PKP harvester 2.X is based on AMP architecture. So, we have installed Apache, MySQL and PHP before installation of PKP harvester. Although there is no hard and fast rule, the order of installation sequence of this manual is as follows:

Apache (The Apache httpd server is a powerful, flexible, HTTP/1.1 compliant open source Web server):

- Installation of Apache;
- Testing of Apache; and
- Apache Configuration and Control.

PHP (PHP is an open source server side scripting language):

- Installation of PHP;
- Configuration of PHP.

MySQL (MySQL, the most popular Open Source SQL database, is developed, distributed and supported by MySQL AB):

- Installation of MySQL;
- Initialization of MySQL Server;
- Creation of database, user and manage permission.

Testing of AMP Links through Scripts:

- Testing PHP-Apache Link;
- Testing PHP-MySQL Link.

Group B: Harvester related functions:

Preprint version of the following paper

Sarkar, P and Mukhopadhyay, P. (2010). Designing single window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies,*57(4):354-364.

This group includes two major activities:

i) Installation of PKP harvester;

ii) Configuration of PKP harvester.

iii) The installation process of PKP harvester is quite easy. It requires two sets of information – (a) login name and password for the administrator and (b) database details (name of the mysql database, user of database and password of the database user).

The configuration processes are divided into three groups:

a) Site management (configuration of site-specific details, language, crosswalk, plugins and reading tools);

b) Archives (creation of archives, managing created archives); and

c) Other administrative functions (layout, customization etc.).

Group C: Repository related activities

The most important task of the administrator is to set up archive(s) for metadata harvesting. The UniTHES started with thirty OAI-PMH compatible open access IDRs containing electronic theses and dissertations. For selecting IDRs containing ETDs, some pre-determined criteria were under consideration, i. e. IDRs should have ETDs in English language; each IDR have more than 10 records and should have OAI-PMH base URL. In this research project DOAR (Directory of Open Access Repositories)[9] and ROAR (Registry of Open Access Repositories)[10] were consulted for selecting IDRs. The intrinsic attributes of these repositories are as follows –

List of Harvested IDRs in the Union Catalogue of ETDs

Sl. No.	Name of the Sponsoring Institution, Country & URL	Name of the IDR	OAI-PMH base URL	No. of Items
1	University of Lethbridge, Canada, http://www.uleth.ca/	DSpace at the University of Lethbridge	http://www.uleth.ca/dspace-oai/request?verb=Identify	15

Preprint version of the following paper

Sarkar, P and Mukhopadhyay, P. (2010). Designing single window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies,*57(4):354-364.

2	Cambridge University Library and Computing Service, University of Cambridge, United Kingdom, http://www.cam.ac.uk/	Dspace at Cambridge	http://www.dspace.cam.ac.uk/dspace-oai/request?verb=Identify	29
3	Learning & Informaiton Services, University of Wolverhampton, United Kingdom, http://www.wlv.ac.uk/	Wolverhampton Intellectual Repository and E-theses	http://wlv.openrepository.com/wlv-oai/request?verb=Identify	63
4	University of the Witwatersrand (Wits), South Africa, http://www.wits.ac.za/	DSpace at My University	http://witsetd.wits.ac.za:8080/oai/request?verb=Identify	898
5	University of Johannesburg (UJ), South Africa, http://www.uj.ac.za/	DSpace at the University of Johannesburg	http://ujdigispace.uj.ac.za:8080/dspace-oai/request?verb=Identify	53
6	Caltech (California Institute of Technology), United States, http://www.caltech.edu/	Caltech Theses and Dissertations	http://theses.caltech.edu/ETD-db/OAI2/oai?verb=Identify	2601
7	COM•DTU (Department of Communications, Optics & Materials), DTU (Technical University of Denmark), Denmark, http://www.dtu.dk/	System Competence Area Document Server	http://sysdoc.com.dtu.dk/oai2d.py/?verb=Identify	22
8	University of Washington Structural Informatics Group (UW-SIG), University of Washington (UW), United States, http://www.washington.edu/	UW-SIG Publications	http://sigpubs.biostr.washington.edu/perl/oai2?verb=Identify	211
9	University of Auckland, New Zealand, http://www.auckland.ac.nz/	ResearchSpace at The University of Auckland	http://researchspace.auckland.ac.nz/dspace-oai/request?verb=Identify	1437
10	Napier University, United Kingdom, http://www.napier.ac.uk/	Repository@Napier	http://researchrepository.napier.ac.uk/cgi/oai2?verb=Identify	48
11	Waikato Institute of Technology (Wintec), New Zealand, http://www.wintec.ac.nz/	Wintec Research Archive	http://researcharchive.wintec.ac.nz/cgi/oai2?verb=Identify	172
12	Victoria University of Wellington, New Zealand, http://www.vuw.ac.nz/	ResearchArchive at Victoria University of Wellington	http://researcharchive.vuw.ac.nz/oai/request?verb=Identify	444
13	University of St Andrews, United Kingdom, http://www.st-andrews.ac.uk/	DSpace at St Andrews University	http://research-repository.st-andrews.ac.uk/dspace-oai/request?verb=Identify	429

Preprint version of the following paper

Sarkar, P and Mukhopadhyay, P. (2010). Designing single window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies,*57(4):354-364.

14	AUT Library (Auckland University of Technology Library), AUT (Auckland University of Technology), New Zealand, http://www.aut.ac.nz/	repositoryaut	http://repositoryaut.lconz.ac.nz/dspace-oai/request?verb=Identify	56
15	Deutsches Elektronen-Synchrotron (DESY), Germany, http://www.desy.de/html/home/index.html	DESY Deutsches Elekronen-Synchrotron	http://pubdb.desy.de/oai2?verb=Identify	298
16	Universiti Putra Malaysia, Malaysia, http://www.upm.edu.my/	PSAS IR	http://psasir.upm.edu.my/cgi/oai2?verb=Identify	432
17	Indian Institute of Astrophysics, India, http://www.iiap.res.in/	Indian Institute of Astrophysics Repository	http://prints.iiap.res.in/oai/request?verb=Identify	88
18	Oxford University Library Services (OULS), University of Oxford, United Kingdom, http://www.ox.ac.uk/	Oxford University Research Archive	http://ora.ouls.ox.ac.uk:8080/fedora/oai?verb=Identify	826
19	Open Access Institutional Repository at the Robert Gordon University, United Kingdom http://www.rgu.ac.uk/	The Robert Gordon University	http://openair.rgu.ac.uk/dspace-oai/request?verb=Identify	307
20	IGIDR Library, Indira Gandhi Institute of Development Research (IGIDR), India, http://www.igidr.ac.in/	Kautilya@igidr	http://oii.igidr.ac.in:8080/dspace-oai/request?verb=Identify	58
21	Johns Hopkins University (JHU), United States, http://www.jhu.edu/	JScholarship	http://oai.library.jhu.edu/oai-dspace/request?verb=Identify	61
22	Monash University Library, Monash University, Australia, http://www.monash.edu.au/	Monash University ARROW Repository	http://oai.lib.monash.edu.au/oaiprovider?verb=Identify	735
23	Information Sciences Library, University of Pittsburgh (UP), United States, http://www.pitt.edu/	Minority Health Archive	http://minority-health.pitt.edu/perl/oai2?verb=Identify	1019
24	University of Wisconsin Systems (UW), United States, http://www.wisc.edu/	Minds @ UW	http://minds.wisconsin.edu/oai/request?verb=Identify	3696

Preprint version of the following paper

Sarkar, P and Mukhopadhyay, P. (2010). Designing single window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies,*57(4):354-364.

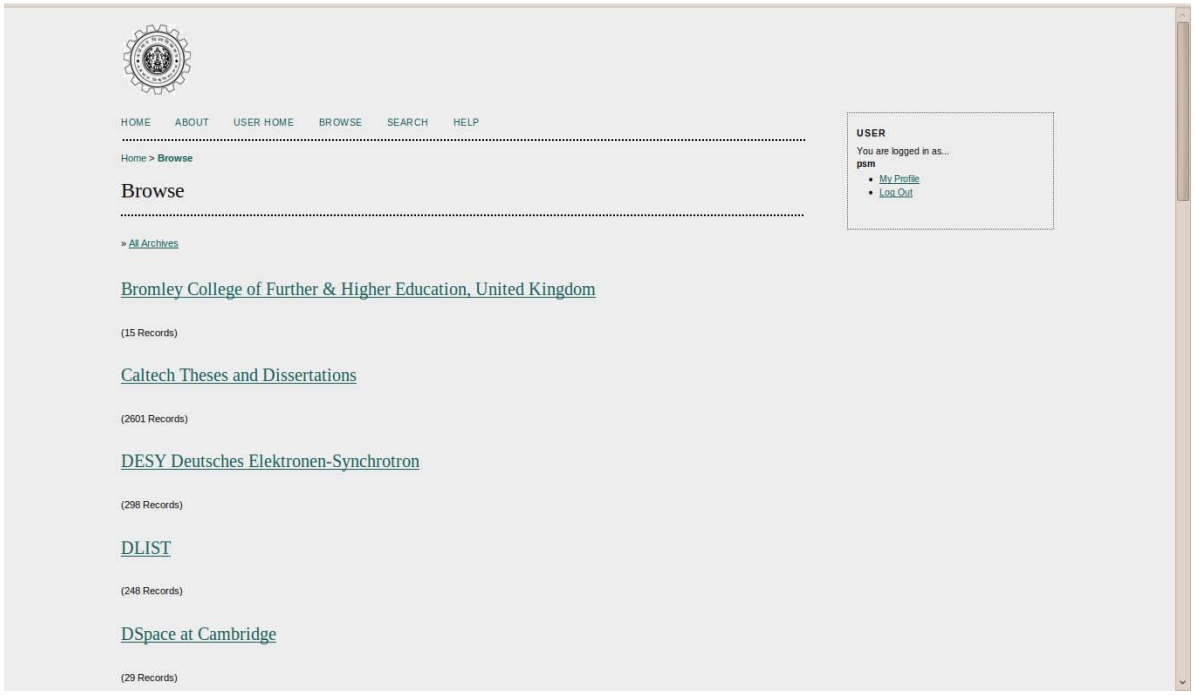
25	University of Leicester Library, University of Leicester, United Kingdom, http://www.le.ac.uk/	Leicester Research Archive	http://lra.le.ac.uk/dspace-oai/request?verb=Identify	206
26	Kansas State University, United States, http://www.k-state.edu/	K-State Research Exchange	http://krex.k-state.edu/dspace-oai/request?verb=Identify	1181
27	Northeastern University, United States, http://www.northeastern.edu/	IRis	http://iris.lib.neu.edu/iii/oairep/OAIRepository?verb=Identify	1254
28	Oregon State University (OSU), United States, http://oregonstate.edu/	DSpace-DEV at Oregon State University	http://ir.library.oregonstate.edu/dspace-oai/request?verb=Identify	97
29	SFU Library (Simon Fraser University Library), SFU (Simon Fraser University), Canada, http://www.sfu.ca/	SFU Institutional Repository	http://ir.lib.sfu.ca/dspace-oai/request?verb=Identify	4076
30	Drexel University, United States, http://www.drexel.edu/	iDEA:Drexel E-repository and Archives	http://idea.library.drexel.edu/dspace-oai/request?verb=Identify	466

Table 2

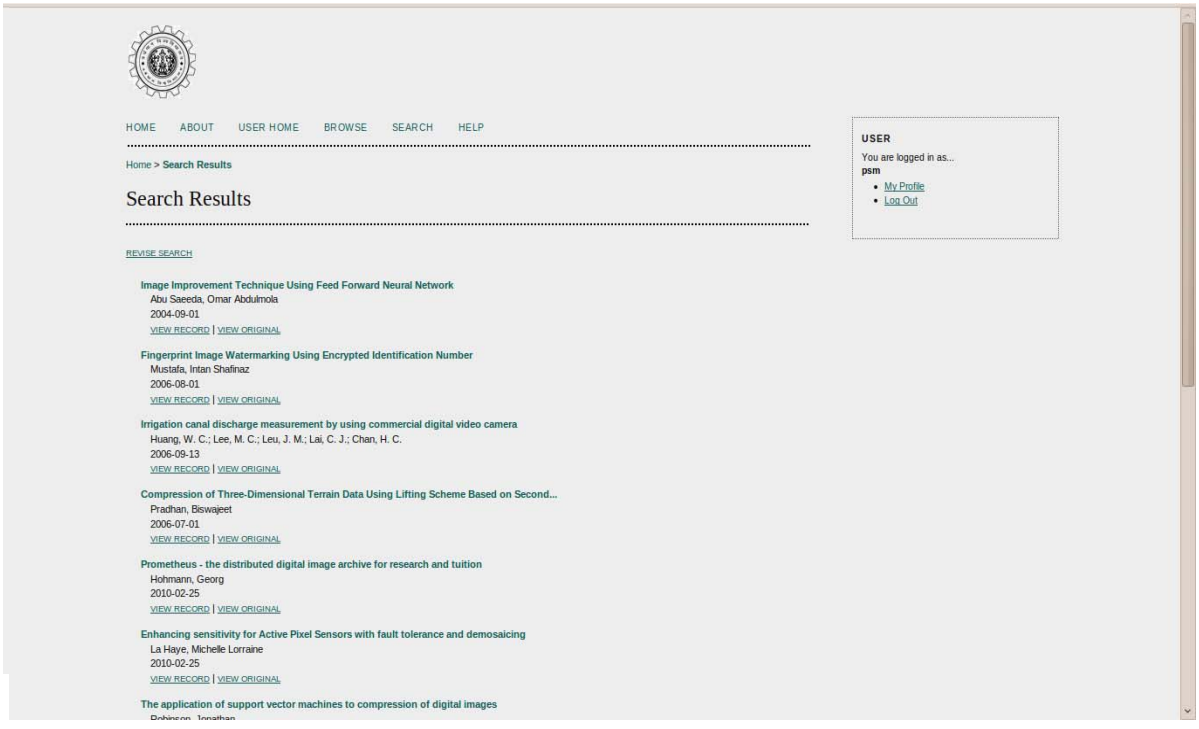
UniTHES so far harvested metadata from thirty repositories and agglomerated a total of more than twenty one thousand open access ETDs. It supports localized searching of metadata elements in two modes – simple and advance. Users can limit search in a single repository or a group of repositories (by default a given search session includes all the available repositories). Search can be filtered by DC metadata elements like title, author, date range, language etc. UniTHES can harvest metadata from DCMES (Simple Dublin Core metadata), MARC 21 bibliographic format and ETD-MS (metadata for electronic theses and dissertations).

Preprint version of the following paper

Sarkar, P and Mukhopadhyay, P. (2010). Designing single window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies,*57(4):354-364.



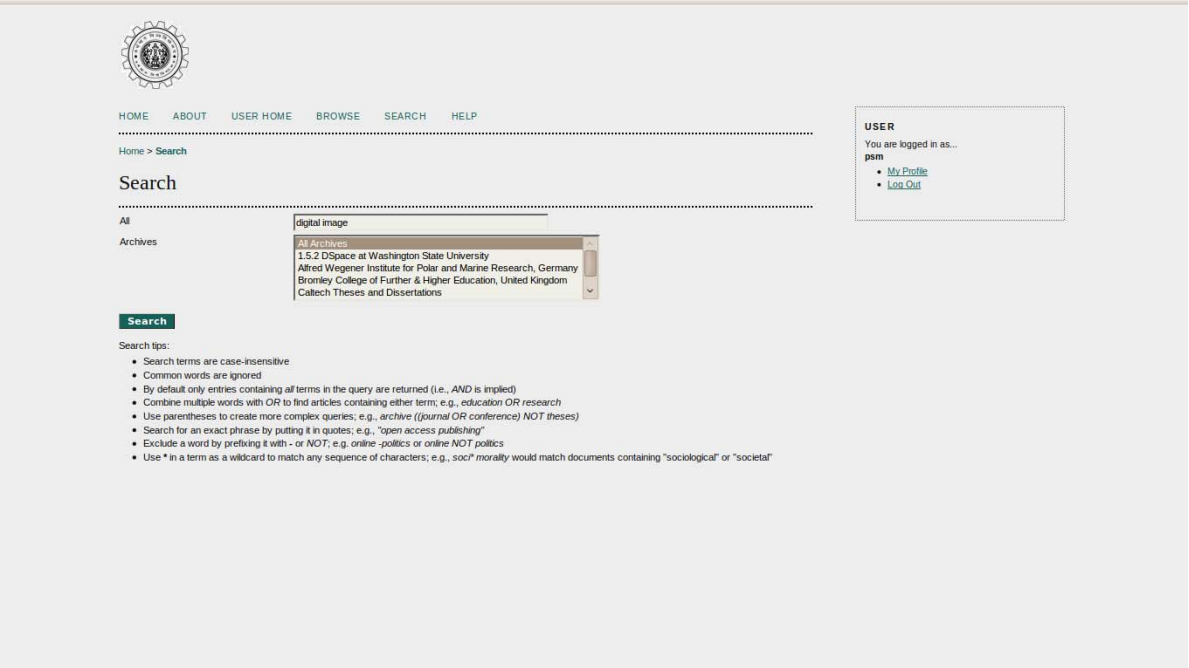
A snapshot of browsing ETD archive



Preprint version of the following paper

Sarkar, P and Mukhopadhyay, P. (2010). Designing single window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies,*57(4):354-364.

A snapshot of search results from ETD repository



A snapshot of search queries

CONCLUSION:

UniTHES is beginning of a new era of localized resource discovery services that can harvest and update metadata from different OAI-PMH compatible open access ETD repositories. At present it includes thirty large-scale open access repositories (based on carefully crafted selection criteria) dedicated to the domain of ETD (in different disciplines). This system has the capability to include new OAI-PMH compatible ETDs and to update records for existing ETDs

Preprint version of the following paper

Sarkar, P and Mukhopadhyay, P. (2010). Designing single window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies,*57(4):354-364.

for the purpose of developing a comprehensive single window local search service for open access ETD repositories.

References:

1. CC:DA (2000). Task Force on Metadata : Final report, *June 16,2000.* RetrievedOctober14, 2009, from http://www.libraries.psu.edu/tas/jca/ccda/tf-meta6.html

2. Cole Timothy W, (2003). Using OAI: innovations in the sharing of information. *Library Hi Tech*, Vol. 21, No. 2 (2003): pp.115-117.

3. Jayakanth F, B Sharada and Minj F, (2007). An OAI-based approach to build and maintain union catalogue of OPACs. Retrieved September 3, 2009, from http://drtc.isibang.ac.in:8080/xmlui/bitstream/handle/1849/403/Francis_jayakanth.pdf?sequence=1

4. Lagoze C et al., (2002). The Open Archives Initiatives Protocol for Metadata Harvesting : Protocol Version 2.0 of 2002-06-14. Retrieved October 14, 2009, from http://www.openarchives.org/OAI/openarchivesprotocol.html

5. Macgregor G and Nicolaides F, (2005). Towards improved performance and interoperability in distributed and physical union catalogue. Retrieved May 10, 2009, from http://www.emeraldinsight.com/Insight/viewPDF.jsp?contentType=Article&Filename=html/Output/Published/EmeraldFullTextAr

6. Mischo W H, (2004). United States Federal Support for Digital Library Research and its implications for digital library development. In J. Andrews, & D. Law (Eds.), *Digital Libraries : Policy, Planning and Practice* (p. 11). Aldershot, England: Ashgate.

7. Mukhopadhyay P, (2006). Five laws and ten commandments : the open road of library automation in India.In Proceedings of the National Seminar on Open Source Movement - *Asian Perspective, xxii, Roorkie, 2006.*. IASLIC, Kolkata, 2006. (pp. 27-36) Retrieved August 22, 2009, from https://drtc.isibang.ac.in/handle/1849/409

8. NISO (National Information Standards Organization), (2004). Understanding Metadata. Bethesda, MD : NISO Press. Retrieved October 14, 2009, from http://www.niso.org/standrds/resources/UnderstaningMetadata.pdf

9. OpenDOAR. Retrieved August 20, 2009, from The Directory of Open Access Repositories - OpenDOAR: http://www.opendoar.org/find.php?search=&clID=&ctID=6&rtID=&cID=&lID=25&rSoftWareName=&submit=Search&format=summary&step=20&sort=r.rName&rID=&ctrl=new&p=1

10. ROAR. Retrieved August 20, 2009, from Registry of Open Access Repositories: http://roar.eprints.org/?action=home&q=&country=&version=&type=theses&order=name&submit=Filter

Preprint version of the following paper

Sarkar, P and Mukhopadhyay, P. (2010). Designing single window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies,*57(4):354-364.

11. Suleman H and Fox E A, (2003). Liveraging OAI harvesting to disseminate theses. Retrieved August 22, 2009, fromhttp://www.emeraldinsight.com/Insight/viewPDF.jsp?contentType=Article&Filename=html/Output/Published/EmeraldFullTextArticle/Pdf/2380210211.pdf

12. Suleman H, (2004). NDLTD union catalog project. In E A Fox … et al.(Eds.), *Electronic Theses and Dissertations : a sourcebook for educators, students and librarians* (pp. 73-77). New York: Marcel Dekker.

13. Suleman H, (2004). Networked Digital Library of Theses and Dissertations. In E A Fox … et al. (Eds.), *Electronic Theses and Dissertations : a sourcebook for educators, students and librarians* (pp. 49-72). New York: Marcel Dekker.

14. Suleman H, (2002). Open digital libraries. Retrieved July 9, 2009, from http://scholar.lib.vt.edu/theses/available/etd-11222002-155624/unrestricted/odl.pdf

15. Van de Sompel H and Lagoze C, (2000). The Santa Fe Convention of the Open Archieve Initiative. Retrieved October 22, 2009, from http://www.dlib.org/dlib/february00/vandesompel-oai/02vandesompel-oai.html

16. Wong I S and Li Y -O, (2009). *Creating a virtual union catalog for Hong Kong dissertations and theses collection.* Retrieved October 22, 2009, from http://www.emeraldinsight.com/Insight/viewPDF.jsp?contentType=Article&Filename=html/Output/Published/EmeraldFullTextArticle/Pdf/2630270212.pdf

17. Yu S -C, Chen H –h and Chang H –w, (2005). Building an open archive union catalog for digital archives. Retrieved May 10, 2009, from http://www.emeraldinsight.com/Insight/viewPDF.jsp?contentType=Article&Filename=html/Output/Published/EmeraldFullTextArticle/Pdf/2630230

18. Mukhopadhyay, P. (2004a). Measuring Web Impact Factors: A webometric study based on the analysis of hyperlinks. *Proceedings of the IASLIC XXI National Seminar on Information Support for Rural Development* (pp. 411–425).

19. Mukhopadhyay, P. (2004b). Organization and dissemination of digital objects through web and CDROM: a framework for Indian libraries. *Proceedings of the International Conference on Digital Libraries,* (p. 470–8).

20. Mukhopadhyay, P. (2006b). VidyaOnline: Design and Development of a FOSS based Virtual Learning Environment on Library and Information Science at Vidyasagar University, West Bengal. *Proceedings of the Conference on ICT for Facilitating Digital Learning Environment* (p. Paper–A).

Preprint version of the following paper

Sarkar, P and Mukhopadhyay, P. (2010). Designing single window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies,*57(4):354-364.