

ETDs in India: Towards a national repository with value added e-theses service

Maitrayee Ghosh

P.K. Kelkar Library, Indian Institute of Technology, Kanpur, India

Abstract

The paper examines the developments in the ETD scene, in particular PhD thesis repositories in India. The purpose is to do a preliminary study and explore the possibilities for creating a national repository for the deposit, discovery, use and long-term care of research theses in an open access environment. The author looks at the current state of deployment of ETD repositories in the academic sector and discusses the subject coverage, number of items, access policy, browse/search option, value added services etc. Despite several limitations, this exercise is valuable and resulted in interesting and thought-provoking information. It certainly raises questions about policies and strategies that national higher education, research funding and policy-making bodies, as well as individual institutional communities within the higher education sector will want to consider. In spite of great interest for e-thesis development and majority of doctoral students in India get their theses produced electronically, no wide scale activity is initiated by government in terms of the storage and dissemination of these materials. Furthermore, adoption of national-level policies on institutional repository development is lacking in India. Till date very few institutes request students to submit electronically their theses and dissertations. This article proposes the construction of a reservoir of extensive doctoral research and an Indian portal, to enable preserving of scientific and technological research materials in the country and a global view of Indian institutional research assets. Few national level institutes like Indian Institute of Science and Indian Institute of Technologies have established ETD repositories and few are at the planning stage. The survey has revealed that digital preservation of Theses and dissertation are already in progress, though some of them are still in a novelty stage. The major problems and concerns reported by respondents are summarized and findings are discussed.

Keyword

Electronic Theses and Dissertations, Indian institutions, ETD consortia, ETD digital libraries, ETD repositories.

1. E-theses and academia

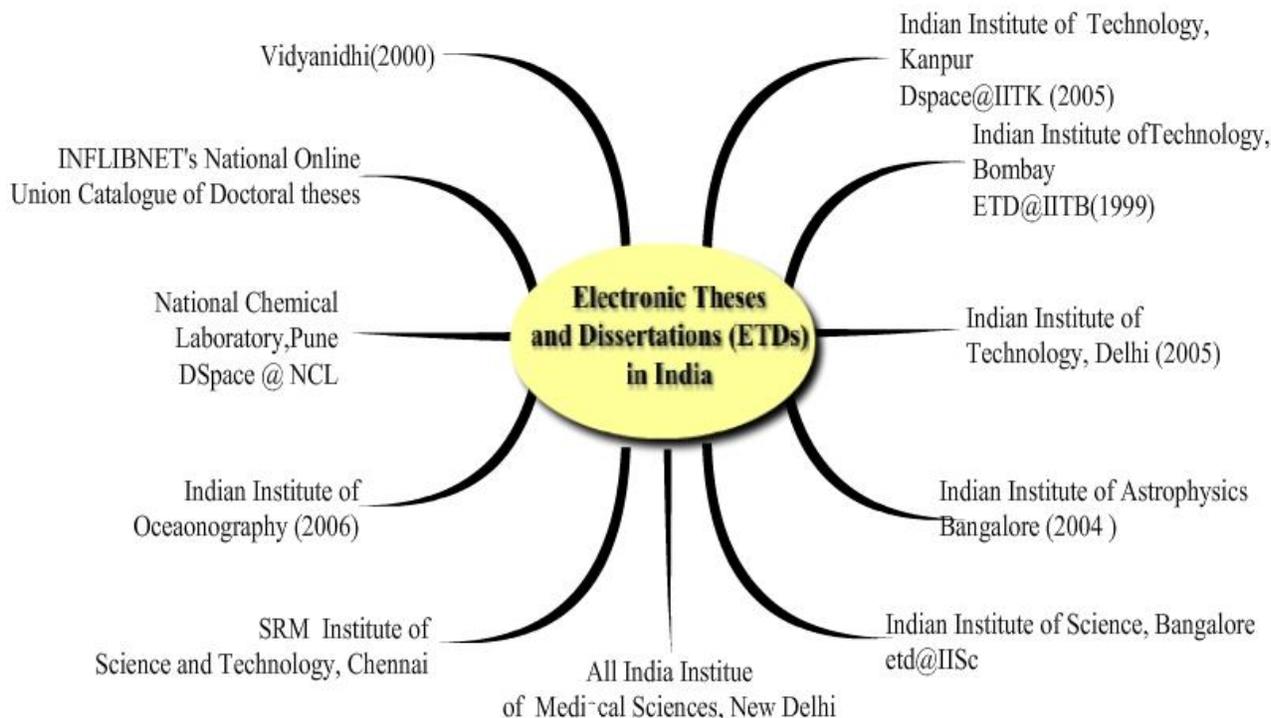
The digital libraries of electronic theses and dissertations (ETDs) are promising to be extremely advantageous to scholars especially in developing countries. The ETD initiatives started in India during late nineties and popularity of this concept is growing rapidly in the higher educational and research institutions to disseminate newly emerged knowledge and expertise. Presently ETDs submitted in Indian universities is mainly in text format, and many libraries have no open access policy. Perhaps the greatest challenge for librarians today, is to develop and maintain sustainable model of open access ETD repositories for users. There has been a realization of the usefulness of the availability of online theses by higher institutions particularly, the elite science and technology institutions. During the last few years more and more research and development institutions and universities are beginning to embrace the idea of creating and maintaining a repository of electronic theses and dissertations. One of the major barriers facing by scholars and researchers in India is lack of access to current literature in their field.

ETDs consist of masters or doctoral research work that is submitted or archived electronically by an institution, either on an internal network or on the web. According to UNESCO ETD Guide website "An ETD is a document that explains the research or scholarship of a researcher/student. It is expressed in a form simultaneously suitable for machine archives and worldwide retrieval. The ETD is similar to its paper predecessor. It has figures, tables, footnotes, and references. It has a title page with the author's name, the official name of the university, the degree sought, and the names of the committee members. It documents the author's years of academic commitment. It describes why the work was done, how the

research relates to previous work as recorded in the literature, the research methods used, the results, and the interpretation and discussion of the results, and a summary with conclusions.” (<http://etdguide.org/>)

The experts are of the opinion that ETDs can accelerate the scientific research level of the country, increase visibility and improve communication among researchers. Arunachalam (2004) argues that the best way to make scientific research more available worldwide is to encourage scientists to self-archive their research. He further stated there is a need for champion (or champions) in every institution to promote the creation of institutional archives, and persuade scientists to place their papers in them. The recent survey done by author reveals that Central library, Indian Institute of Technology (IIT) Bombay initiated ETD project in 1999, followed by Vidyanidhi at Mysore University in 2000, Indian Institute of Science and Indian Institute of Astrophysics 2004, IIT Delhi and IIT Kanpur in 2005, Indian Institute of Oceanography, Goa in 2006 and so on. Bibliographic data of ETDs are immensely useful for researchers, students and supervisors but they must be able to get them quickly and swiftly. The author’s own institution, P.K. Kelkar library, IIT Kanpur maintains a bibliographic database (with abstract) named as Theses (old database) in addition to DSpace full text repository. IIT Delhi, Central Library has a bibliographical list of PhD theses since 1966 to current, which includes nearly 2000 records accessible through web.

Figure: 1 ETD initiatives in India



Universities in India oblige researchers to submit the print version of the theses to library, additionally few universities/ research institutions require the same version on a CD-ROM or submission online. In spite of number of sporadic efforts there is no single authentic source available where one can locate accurate statistics about the theses produced in India. Researchers often have difficulties in locating the relevant dissertations done by previous students as there is no uniform user interface to conduct a formal search for theses and dissertations. The following sources/databases include information on accepted theses in India:

- ✧ University News – a weekly publication by Association of Indian Universities, New Delhi gives comprehensive list of theses accepted by universities under various subject headings.
- ✧ Library web page or OPAC of individual institution.
- ✧ DELNET (Developing Library Network) database has 44,304 records of theses and dissertations submitted to various Indian universities on different subjects.
- ✧ INFLIBNET (Information and Library Network) union database covers the bibliographic metadata of doctoral theses submitted to more than 130 universities/institutes in India. It has over 150,000 records from all subject areas.
- ✧ “Vidyanidhi” at University of Mysore, hosts bibliographic as well as full text records of theses deposited by member universities.

ETD initiatives are important because financial barriers prevent researchers in developing countries from accessing the research information they need. University Grant Commission (UGC), an apex body of Government of India is planning to put in place a regulatory policy and an implementation mechanism for maintaining standards, archiving, and accessing of doctoral research from the Indian universities. The setting up of National Knowledge Commission on 13th June 2005 by Planning Commission and its formal launching by Prime Minister of India on 2nd August, 2005 on matters relating to institutions of knowledge production, knowledge use and knowledge dissemination is an inevitable initiative in the current context of knowledge economy. The commission has recommended institutional repositories of research articles, reports, institutional publications and ETDs. One of the easiest and most effective ways to promote open access to research and educational content involves support of ETDs. There are nearly three hundred universities and more than hundred research institutions and deemed universities who can grant doctor or master degrees, so there are increasing number of theses and dissertations every year but most of them are not published and relatively inaccessible. Fox (1998) stated “universities have theses and dissertations, and smaller institutions have bachelor theses and other major papers. These parts of the "grey literature" are largely ignored and disappear from view”.

2. Literature Review

Resources for this review are primarily publications from 1998 to early 2007, gathered in the first instance from citations in online version of Library and Information Science Abstracts (LISA) and Library Literature. To obtain background literature on the subject the search engine like Google scholar (<http://www.google.com>) and BUBL Information Service library and information science journal archives (<http://bubl.ac.uk/archive/lis/>) were used. A search through JESTOR full text database and project muse had also been conducted.

The literature on Electronic Theses and dissertations are fairly recent; documentation about ETD development and initiatives from different countries have appeared through the online proceedings of a series of ETD international symposiums viz Memphis(1998), Blacksburg(1999), Florida (2000) California (2001), Utah (2002), Germany(2003), Kentucky (2004), Australia (2005), Canada(2006).

Most literature discovered in the search focused on case studies of ETD repositories in different parts of the world. Greig (2005) describes the strategies that have been adopted by the staff at Glasgow University Library in implementing an electronic theses database and the challenges faced. Free availability of the electronic theses and dissertations on the web allows anybody to read, download, copy, distribute, print, search or link to the full text, and use them for any lawful purpose. The need of open access to ETDs is discussed repeatedly in the literature. Dobratz (2005) stated “Open Access which means to provide access to publications that are like, electronic theses and dissertations (ETDs) either usually buried in libraries or are not freely accessible by a major part of the scientific community..”. The implementation of ETD programs promotes open access of scholarly communications that in turn leads to improved academic standards. It is clear from the literature that the importance of making available electronic theses and dissertations is clearly understood by all users and institutions.

Electronic Theses and Dissertations: a sourcebook for Educators, Students and Librarians, edited by Edward Fox et.al was published in 2004 includes twenty eight papers authored by ETD experts, project coordinators from countries namely US, Australia, Germany, Portugal and Brazil. It's a useful guide and must read for all stakeholders of ETD projects.

Theses and dissertations are important not only for an individual author, who looks for the maximum visibility and scientific impact of his/her work but also for advisors, academic departments, post graduate programs, and universities (Marcondes & Sayao, 2003).The advantages and attraction of electronic theses and dissertations identified by several authors. According to them ETDs:

1. allow the student to better express ideas and the reader to better understand them
2. provide a unique opportunity to learn more about electronic publishing
3. Promote collaboration between research programs by making research work visible and accessible via a network archive

The importance of adding values to e-theses and various methods that are used in libraries have been discussed by Choudhury and Müller (2006) and the long-term preservation of ETDs by Caplan and Thomas (2006). Copyright ownership is a subject that is vigorously being discussed in national and international conferences.

Pavani and Southwick (2004) named four essential components for establishing ETD consortium as ETDs, Coordination, Cooperation and integration. They identified five common objectives of ETD consortia are:

1. To help institutions start projects by raising the awareness on the importance of ETDs and by tutoring the first steps on implementing local ETD digital libraries
2. To discuss and create standards to allow interoperability
3. To share Software solutions
4. To offer training activities
5. To build union catalogs

Brazil Digital library for Theses and Dissertations architecture is discussed by Southwick (2006); in this model the data providers are universities and research centers and Instituto Brasileiro de Informa ao em Ciencia e Tecnologia (Ibict), a Brazilian government institution playing the central role and acting as a service provider; metadata is harvested from the data providers to create a central metadata repository in Ibict. Similarly CALIS, China has developed a central portal for ETDs and follows centralized retrieval and distributed full text model. All the joining universities collect their own theses and dissertations database and submit it to the Tsinghua university central server. Readers can get full text through Tsinghua university commutative borrowing system or getting contact with the original university library. The copyright is controlled by the local system of the joining members. The portal provides users with a simple and intuitive interface for searching and browsing through a merged metadata collection of theses and dissertations (Jin, 2004)

Korea Institute of Science and Technology Information (KISTI) Electronic Theses and Dissertations (ETD) system is a national digital library of ETDs in South Korea (Zhang et.al, 2001). Shin (2006) discusses about dCollection server which is a digital repository system implemented in 2003 as a result of collaboration among leading college and university libraries in South Korea. In KERIS, each library inputs the metadata using dCollection and produces the ETDs. The original digital files are deposited in the local university library's server with the metadata, and the URL of metadata and digital files are stored on the dCollection server at KERIS.

Vijaykumar (2005) reported 86% university librarians, and 60% PhD supervisors in Indian universities have ETDs awareness about whereas only 56% researchers know about ETDs. He opined "The less awareness among academics shows that enough awareness programs are required before initiating ETD programs in Indian campuses. He further mentioned that 28% Universities in India have started developing ETD repositories.

A National Conference on Digital Library and e-Theses (NCDLET) was organized by Jadavpur University, Kolkata, India in January 2005. The objectives of NCDLET mentioned in the promotional brochures are:

- ✧ To showcase Digital Library Initiatives and ETD examples in India
- ✧ To evolve appropriate organizational models for Consortia of Digital Libraries
- ✧ To initiate the ETD movement in India by raising the awareness levels about the potential of ETDs in contributing to scholarship
- ✧ To prepare a vision document for the digital library and ETD movement in India

The INDEST consortium has been at the forefront in the provision of electronic resources to India's premier Science and Technological institutions made recommendation to set-up OAI compliant digital archives in the member institutions. Under INDEST EXTENDED proposition online submission of theses and dissertations proposal was discussed in INDEST Workshop in December, 2005. It was further proposed that INDEST consortium should set-up a harvesting service at its headquarter to harvest metadata from its member institutions.

Several conference presentations have provided updates on the progress of ETD repositories in India. There are reports on existing institutional repositories such as the All India Institute of Medical Sciences, New Delhi, SRM institute of Science and technology, Chennai, etd@IISc, Bangalore etc.

No studies are found that investigate possibilities for a central reservoir of e-theses in India, and propose a model which includes all the existing library consortia/ networks in operation to serve as data provider and make a complete collection. Therefore, it can be stated that, present study is the first attempt which looks into the problems and possibilities for further developments in ETD repositories in India.

3. Scope and objectives

An understanding of how ETD digital library systems are functioning is important to assess the strengths and weaknesses of systems and thus point to directions for improvements. Such an understanding is also useful to identify the best practices and develop guidelines future development. Although there is sizeable number of ETD digital libraries in India and maintaining a steady growth but there is no single study undertaken which looks into how digital libraries of ETDs is functioning, what are the constraints in the process, and how it can be improved. This study seeks to fill the gap. It is an attempt towards a comparative study of prominent digital libraries of ETDs with high level description of key aspects and proposed a model for national repository of ETDs.

To understand the current status and identify the issues to be considered for developing a national ETD reservoir for India, nine digital libraries of ETDs are evaluated. The examples are by no means exhaustive but efforts have been made to give wide subject coverage. The digital libraries are chosen from premier institutes of India viz. Indian Institute of Technologies, Indian Institute of science, Indian Institute of Astrophysics, National Chemical laboratory Pune, etc. who have a good digital library infrastructure and sizeable no of e-theses. Majority of them have registered either with openDOAR, ROAR or other harvesting registry like OAIster etc.

The purpose is to examine the current status of Digital library of ETDs in India on the need of developing an e- theses portal and interoperable national ETD repository with unifying access. The ETD policies that are currently in place at different Institutional repositories have been discussed to give a complete outlook of this paper.

The author has proposed the following:

- ✧ A seamless ETD system for all universities/research institutions, which are already member of national networks or consortia viz. INDEST, INFLIBNET as well as libraries attached with colleges or small organizations with underdeveloped ICT infrastructure, can join either as a data provider or simply access the e-theses by paying membership fee.
- ✧ A model for national e-theses repository

4. Methodology and selection of sample

This preliminary study concentrates on the state of the art ETD digital libraries in India and seeks to identify a suitable model for national repository of ETDs in India. The research method used in this study is the triangulation of three different methods (1) e-mail questionnaire (2) face to face interviews/ telephonic conversation and (3) content analysis of ETD websites. Triangulation increases the depth of understanding and investigating the unknown phenomenon by using multiple data gathering techniques. The questionnaire is chosen in addition to web page evaluation since the location of the targeted respondents is geographically scattered across the country. E-mail was the main tool in distributing the questionnaire. The project coordinators were interviewed to find out if they would like to share their experiences/opinions on ETD projects. Permission was asked to schedule the meeting and conduct face to face interviews with them. The content analysis of websites was done in addition to questionnaire and interview to collect up to date data. Some of the points used as a basis for the selection of digital libraries for the evaluation are:

1. The importance, relevance, credibility of the organization hosting the digital library
2. Type of collection and users
3. The access facilities, submission policies and user interface

To identify the parameters that need to be looked at to evaluate the services provided by Digital libraries of ETDs, a thorough literature review was completed. The Registry of Open Access Repositories (ROAR) and Directory of Open Access Repositories (OpenDOAR) were consulted but they do not represent a complete picture of Indian ETD repositories as the registration is not mandatory but voluntary. There are twenty four institutional repositories from India listed in Registry of Open Access Repository (ROAR, <http://roar.eprints.org/>); but only five have maintained ETDs; further, three of them have registered under e-theses viz. IIT Delhi, NCL, Pune and IISc, Bangalore. Although published literatures indicate that there are more than a dozen ETD digital libraries in India but time constraints and geographical location didn't permit author to visit all the libraries. To serve the purpose of this study, nine digital libraries of ETDs who have either responded to the questionnaire or their web pages are accessible on WWW is selected for this study. The time limitation to circulate the questionnaire to libraries is only from middle of March-May 2007. Additionally, limitations such as few ETD web pages were not accessible during the period of study, limited response to questionnaire etc. lead to not having large sample size for meaningful analysis.

4. A snapshot of major ETD projects/initiatives in India

The following table summarizes the special features of ETD digital libraries collected through literature search, and information on web sites:

Table: 1 ETD digital libraries in India

Institution/ ETD Digital Library	Criteria for selection
Indian Institute of Technology, Kanpur -Dspace@IITk	-Highest number of full text theses (PhD, M. Tech and M. Des) in the area of Engineering sciences, Pure sciences, Bio sciences, Humanities and Social sciences. -a bibliographic database (with abstract) of 9219 PhD, MSc, M.Tech theses accessible through internet. - Campus access to full text theses and Mandatory CD submission.

<p>Indian Institute of Technology, Bombay etd@IITB</p>	<ul style="list-style-type: none"> -One of the core members of INDEST consortium -Above three thousand full text theses and dissertations available through intranet. The ETD digital library is developed by using Greenstone software. -The bibliographical records are accessible through web; One of the core members of INDEST consortium. -Mandatory online submission and a signatory of NDLTD.
<p>Indian Institute of Technology, Delhi</p>	<ul style="list-style-type: none"> -Around two thousand PhD theses bibliographic records are available (year wise) since 1964. -Registered in ROAR as e-theses repository. -Host institution of INDEST consortium and a model site for both e-print and e-theses. -Thirty e-theses are available for campus access up to content page.
<p>Indian Institute of Science, Bangalore 2004 etd@IISC</p>	<ul style="list-style-type: none"> -etd@IISc is a joint service of NCSI and IISc which allows Open access to full text theses -registered in ROAR as e-theses repository, also included in OAIster harvesting register. -Core member of INDEST consortium
<p>National Chemical Laboratory, Pune</p>	<ul style="list-style-type: none"> - A research, development and consulting organisation with a focus on chemistry and chemical engineering. - Registered in ROAR as e-theses repository. - Full text theses and accessible through web
<p>Indian Institute of Astrophysics, Bangalore</p>	<ul style="list-style-type: none"> - A premier institute devoted to research in astronomy, Astrophysics and related physics. -Full text open access digital repository on Astrophysics and allied subjects - registered in ROAR
<p>National Institute of Oceanography, (NIO) Goa</p>	<ul style="list-style-type: none"> -The NIO library has been recognized as the National Information Centre for Marine Sciences (NICMAS) -The Digital repository Service (DRS) is Compliant to OAI-PMH V 2.0 and has more than 500 theses and dissertations. - Contents of DRS is harvested by OAIster (Global harvester for archives), Google Scholar, Avano (Marine and aquatic science harvester from France) and CASSIR (Indian archives harvester at IISc, India).
<p>Vidyanidhi, University of Mysore</p>	<ul style="list-style-type: none"> -Multilingual database having 15,000 records in Hindi and 600 records in Kannada, more than 50,000 metadata of PhD,

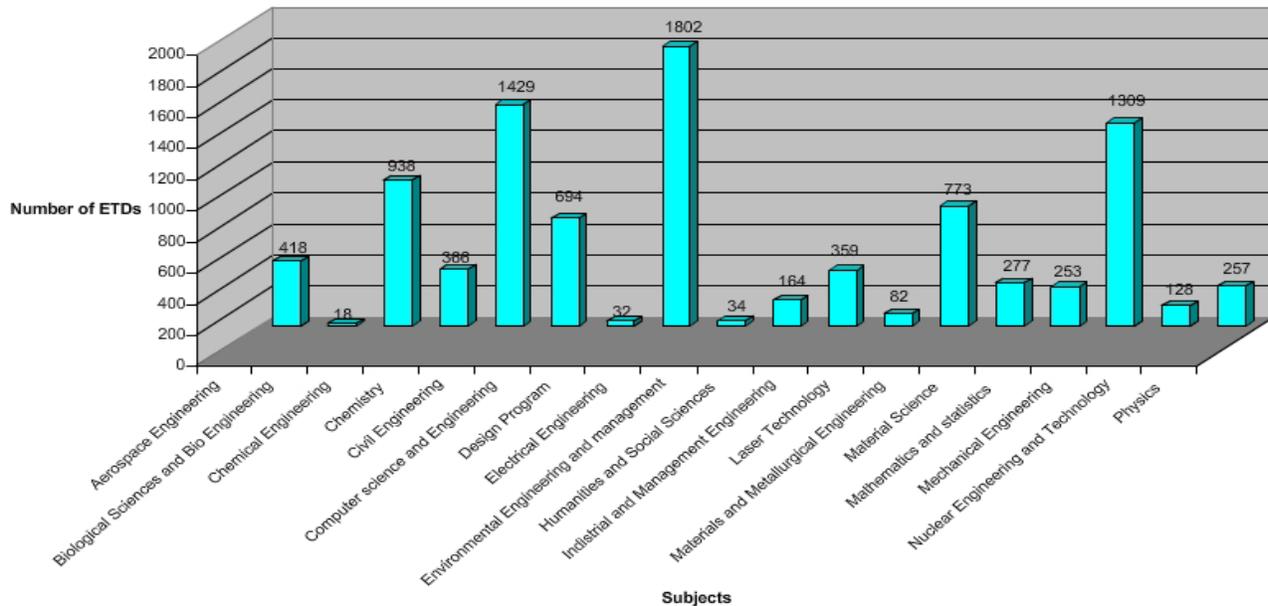
<p>INFLIBNET, Ahmedabad Theses online database</p>	<p>M.Phil, MA, M.Sc, M Tech, ME, MD, M Com., MBA, MCA, MLIS, M.Ed., theses and dissertations. -It is known as ETD consortium of Indian universities</p> <p>Information and Library Network (INFLIBNET) Centre is an autonomous inter-university centre of University Grants Commission (UGC) involved in creating infrastructure for sharing of library and information resources and services among academic and research Institutions. It works collaboratively with Indian university libraries and has the theses database with over 1,50,000 records of the doctoral studies submitted to the Indian universities in all discipline. Efforts are in progress to include abstracts in this database.</p>
--	--

5.1 Theses and dissertations at P K Kelkar Library, IIT Kanpur

As a core member of the INDEST Consortium PK Kelkar Library of IIT Kanpur shares in the vision to create a digital library with local service inter-faces forming part of a global information community for clients in IITK. The mission of PKKL is to fully utilise and develop the information resources on Science and Technology for the purpose of promoting education, research and lifelong learning as a contribution to development. The library has a Bibliographic database of theses and dissertations of more than 9000 records with abstract, developed and maintained by Documentation division accessible through library web page. The abstract files are in PDF, which is the most authoritative online format. In majority of the cases the author abstract/ synopsis were scanned but with few exceptions when abstract were not available in the hardcopy and subsequently prepared by researchers/students from similar subject area or in charge of the section. There are various ways to find information in this database using the search option: Author, Title, keyword, subject etc; alternatively one can conduct keyword search using Boolean operators. It is also possible to search the database under advisor or supervisor. With the emergence of full text scanning of theses and dissertations project in 2005, when IITK joined in the million book project for universal digital library and took up the job of retrospective conversion of theses and dissertations for developing a digital library of ETDs using Dspace software, then the entry of records in the bibliographic database was suspended. The scanning of hardcopies took almost a year at Indian Institute of Information Technology, Allahabad. Subsequently, metadata was extracted from the old theses database supported by iitKLAS (Indian Institute of Technology, Kanpur Library Automation System) software. Now institute is putting efforts to make open access to full text theses/dissertations, and if this happens then Indian Institute of Technology, Kanpur will be the first IIT in India to have full text open access repository.

Figure: 2

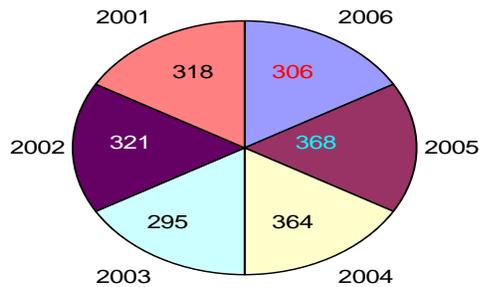
Subjectwise distribution of ETDs at IIT, Kanpur



During the last six years the number of theses and dissertations added to IITK collection including PhD, MSc, M.Tech and M. Des is shown in the following chart:

Figure: 3

Submission of TDs in IIT Kanpur

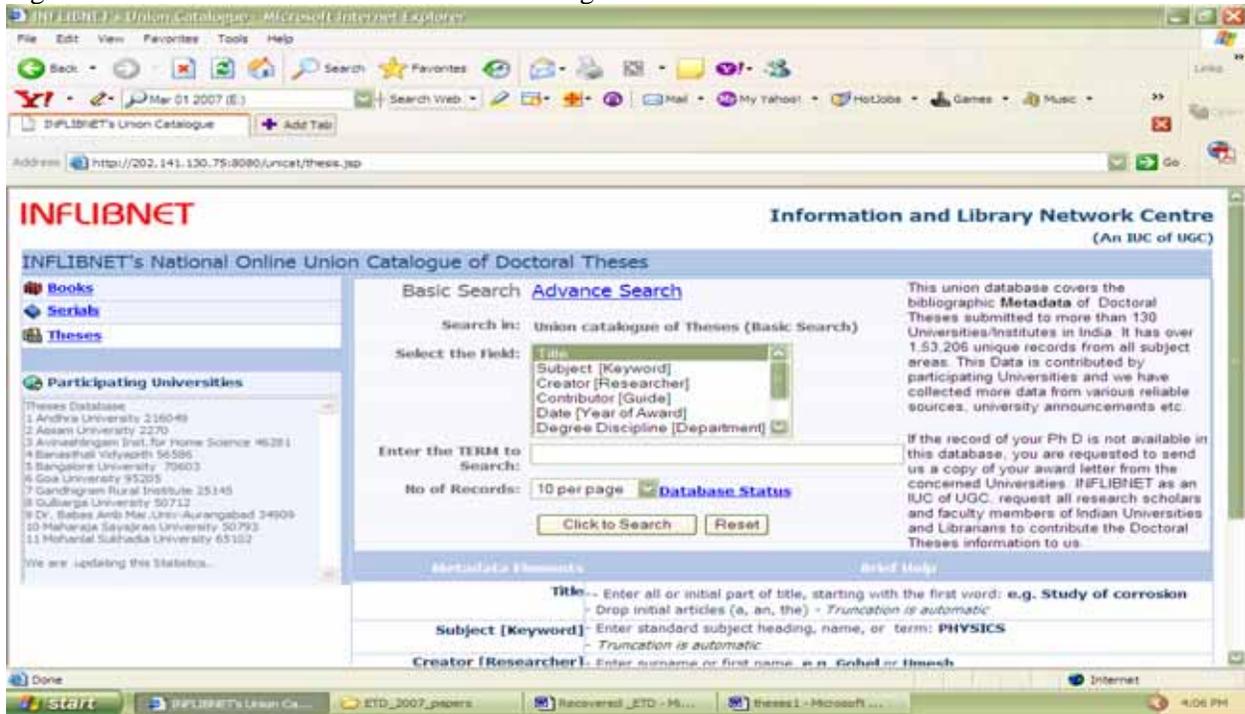


5.2 INFLIBNET’s National online union catalogue of Doctoral theses

Murthy (2005) mentioned “As per UGC-INFLIBNET Program, 142 Universities are funded for creation of databases of doctoral theses ...” A memorandum of understanding(MOU) is signed between universities and INFLIBNET in this regard. It is further reported that more than 200 universities /institutions from all over India have contributed data/records in electronic format (ISO). The metadata of Indian PhD theses are available since 1905 and more than 1,50,000 records of doctoral dissertations can be accessed as a single collection. The portal provides users with a simple and intuitive interface for searching and browsing through a merged metadata collection of theses. This online database has the provision to search from the following access points as Title, Researcher, Guide(s), Department,

University, Place, Year, subject(s), and Boolean Search etc. It has been proposed that UGC Infonet will have a data centre with large server capacity and e-theses can be maintained in this server.

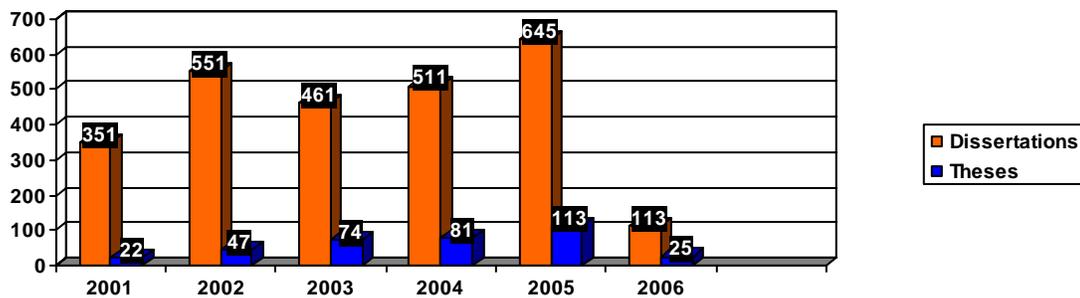
Figure: 4 INFLIBNET theses online union catalogue



5.3 ETD initiatives at Central library, IIT Bombay

The ETD digital library project was initiated at central library, IIT Bombay in 1999. It has bibliographic records as well as full text; online submission is mandatory. Statistics of ETDs submitted by IIT Bombay students from 2001 to early 2006 is furnished below:

Figure: 5 Number of Theses/ dissertations submitted to IIT Bombay



[Adapted from Jadhav, 2006]

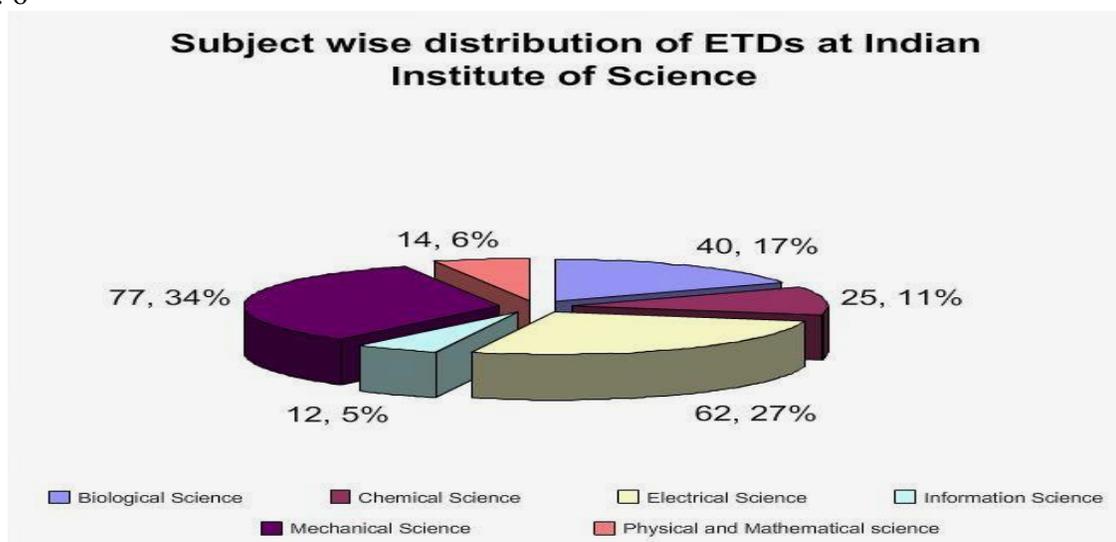
5.4 PhD theses at central library, IIT Delhi

IIT Delhi provides web-based Access to bibliographic database of PhD theses (<http://indest.iitd.ac.in/isisdb/phd-index.html>). The bibliographical list is available since 1966 till 2006 includes nearly 2000 records of PhD theses. The institutional repository eprints@IIT Delhi is registered as e-theses repository in ROAR. The Dspace e-theses repository is accessible within the campus and thirty items are available in the repository; 21 theses are from computer science and four theses are from Chemistry. The central library offers mediated submission service to ease the load of academic community.

5.5 ETD digital library at Indian Institute of Science

The Indian Institute of Science (IISc) is a premier institution of research and higher learning, offers postgraduate and doctoral research programmes. The ETD digital repository is registered to harvesting services like ARC, OAIster etc to facilitate institute’s research output to international community. It is compliant to OAI PMH V-2.0 and ETD –MS V1.01 metadata standard for theses and dissertations.

Figure: 6



Following are the unique features of etd@IISc identified by Jobish [et.al.](2006)

- ✧ Reflection of departments and divisions in Communities and collection
- ✧ Compliance to ETD–MS metadata standard for theses and dissertations
- ✧ Metadata and full text quality assessment by Library staff
- ✧ Support for assignment of subject category
- ✧ Automatic collection assignment of student during registration
- ✧ Common fields and pre fields are made hidden to save the time of the submitter
- ✧ Automatic metadata assignment and validation during submission

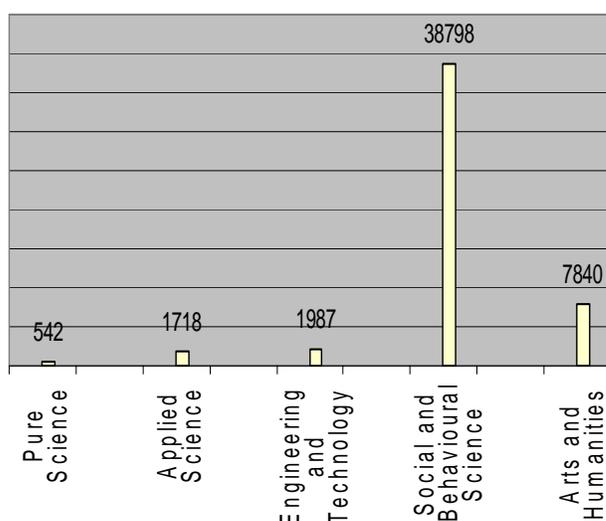
5.6 ‘Vidyanidhi’ digital library

As evident from the website and mentioned in the published literature “Vidyanidhi is supported by the Ford Foundation and Microsoft India, hosts more than 500 full-text and 50,000 bibliographic records of theses submitted to the universities in India. Vidyanidhi also has the strategic support from UGC

(University Grants Commission) and a membership of the Networked Digital Library of Theses and Dissertations (NDLTD). The Vidyanidhi Digital Library has two layers viz. metadata database and full text of theses. Indian Theses database (<http://210.212.200.226/home/firstpage.htm>) has records of bibliographic descriptions, and Indian ETD collection has Dspace based full text theses digital library (<http://dspace.vidyanidhi.org.in:8080/dspace>)

Figure: 7

Subject wise distribution of ETDs in Vidyanidhi



Vijayakumar (2004) lamented “Even though, ‘Vidyanidhi’ aims to evolve as a national depository, but no provision is made for mandatory submission”.

It is evident from the chart that Vidyanidhi has large collection (i.e. nearly thirty nine thousand records) in Social and Behavioural Science and moderate collection (i.e. less than eight thousand) in the area of Arts and Humanities. The number in each subject area of applied science and Engineering and technology is less than 2000. Pure Sciences has the lowest number i.e. below thousand.

The user interface, search and browse features of selected digital libraries of ETDs is summarized in the following table.

Table 2: Web site analysis of ETD repositories

Particulars	IITK	IITB	IISc	Vidya nidhi	IIA	NIO	IITD	NCL
Submission guidelines/help	Y	Y	Y	Y	Y	Y	Y	Y
<u>Browse</u>								
Author/Title	Y	Y	Y	Y	Y	Y	Y	Y
Supervisor/guide	Y	N	N	Y	N	N	N	N
Date	Y	Y	Y	Y	Y	Y	Y	Y
Subject	Y	Y	Y	Y	Y	Y	Y	Y
<u>Search option</u>								
Keyword search	Y	Y	Y	Y	Y	Y	Y	Y
Boolean Search	Y	Y	Y	Y	Y	Y	Y	Y
Truncation	Y	Y	N	Y	N	N	N	N
Proximity	Y	N	N	N	N	N	N	N

Recent Submission list	Y	Y	Y	Y	Y	Y	Y	Y
Theses style/Template	N	N	Y	N	N	N	N	N
Usage statistics	Y	N	N	N	Y	N	N	N
Links to other ETD services	Y	N	Y	N	N	N	N	N
Copyright information	N	N	N	N	Y	Y	N	N

Findings:

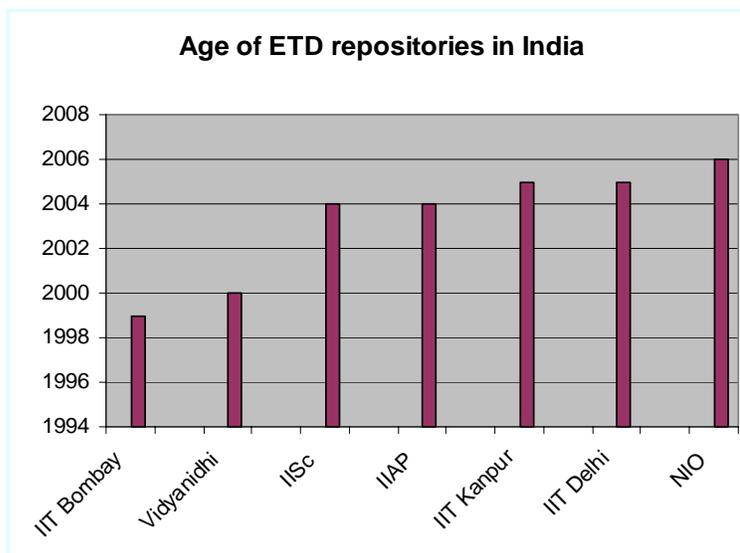
- There are mainly three search options viz. simple search, advance search and browse
- Users are offered the choice to browse by Author, Title, and date
- IITK and IISc have provided additional search option i.e. supervisor or guide
- In Vidyanidhi search can be done by name of the institutions, degree, language
- Only IISc have included theses style/ template
- Monthly usage statistics are available at IIA and IITK web site.
- All the sites have included recent submission list
- All the digital libraries of ETDs have Boolean search facility
- Copy right information is provided in NIO and IIA websites
- Links to other ETD services are available in IITK and IISc web sites

6. Data analysis and discussions

In this analysis, the author has presented data, which is largely the result of questionnaire Survey, face to face/ telephonic interview and content analysis of ETD web sites/bibliographical databases. In preparing this paper, a brief questionnaire was sent to twelve ETD digital libraries to solicit information on their Theses and Dissertations as well as information on ETD projects.

Figure: 8

Interviews were semi-structured and project coordinators were asked to comment on their experiences on digitization, choice of software, staffing and infrastructure, funding, submission policies, promotion, user feedback, institutional support and future plans. The author visited two IIT libraries other than her own institution and interacted with the staff working on ETD projects. Few ETD digital libraries did not respond to the questionnaire, therefore described in brief. It is observed that premier institutes viz. Indian Institute of Science and IITs have extensive,



sophisticated, and well-developed system of Electronic resources. All the libraries are using open source information repository software namely DSpace, except central library, IIT Bombay, which has developed its digital library by using Greenstone. The high level of involvement and/or collaboration of all stakeholders are one of the prerequisite for successful ETD projects. The author has particular interest in finding information about the academics involvement in ETD repository development. It is found that faculty involvement is high in two IIT libraries viz. IIT Bombay and IIT Kanpur, where ETD project coordinator/advisors are faculty members. The repository of IIT Delhi, Indian Institute of Astrophysics, Indian Institute of Oceanography and Vidyanidhi are developed and managed by librarians.

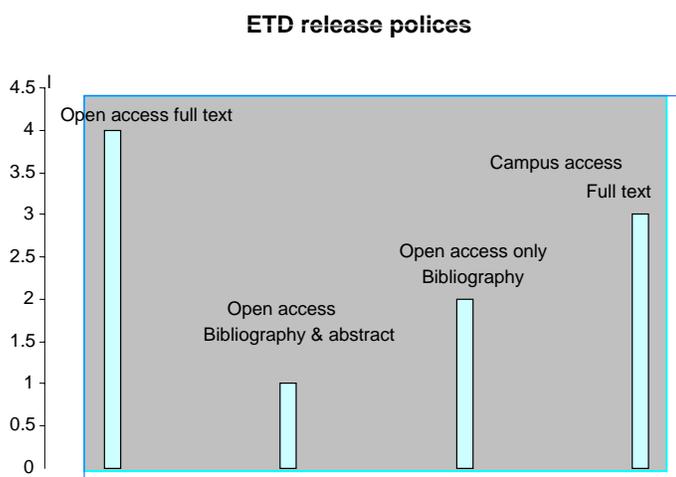
Table 3 Collection size and coverage

ETD repository/ database	Number of items/records	Subject coverage
Dspace@IITK (2005) http://172.28.64.70:8080/dspace http://www.lib.iitk.ac.in:8080/examples/thesis/index.html	9333 PhD theses and M.Tech dissertations Full text	Engineering sciences, Bio Sciences, Pure Sciences, Humanities and Social sciences
etd@IITB(1999) http://etd.library.iitb.ac.in http://dspace.library.iitb.ac.in/dspace/	3700 PhD theses and M.Tech dissertations Full text	Engineering sciences, Bio and Pure Sciences, Humanities and Social sciences
etd@IISc (http://etd.ncsi.iisc.ernet.in/)	233 PhD theses Full text	Pure sciences, Bio sciences, Engineering sciences
IIA (2004) http://prints.iiap.res.in/	83 Full text	Astrophysics, Astronomy and related physics
ETD @Central Library, IIT Delhi (2005) http://eprint.iitd.ac.in/dspace/ http://indest.iitd.ac.in/isisdb/phd-index.html	30 PhD theses (full text) 2000 bibliographic record	Engineering sciences, Pure Sciences
Digital repository service at NIO drs@NIO (2006) http://drs.nio.org/drs/index.jsp	Theses- 498 Dissertations-21	Oceanography and allied subject
Vidyanidhi(2000) http://210.212.200.226/index.html http://dspace.vidyanidhi.org.in:8080/dspace/	95,000 bibliographical record and 4000 ⁺ full text open access repository	All subjects
DSpace @ NCL, Pune http://dspace.ncl.res.in/dspace/index.jsp	208 PhD theses full text	Chemistry and Chemical engineering
INFLIBNET(1994) Theses online database http://202.141.130.75:8080/unicat/thesis.jsp	PhD theses 1,50,000 bibliographical records	All subjects

It is clear that ETD repositories have already achieved some status and developed infrastructure hence, can form the basis of a national repository project that presupposes the country wide availability of ETDs. An increasing number of universities are accepting and encouraging the submission of theses and dissertations in electronic format. At this moment it is considered that ETD awareness is high among Indian academia. Institutions' ETD release policies have to meet the requirements of intellectual property law and balance the needs of multiple stakeholders' viz. graduate students, faculty, researchers, librarians and publishers.

Despite functioning in similar infrastructural and cultural environments, analysis reveals that institutions have implemented a diversity of release and submission policies.

Figure: 9



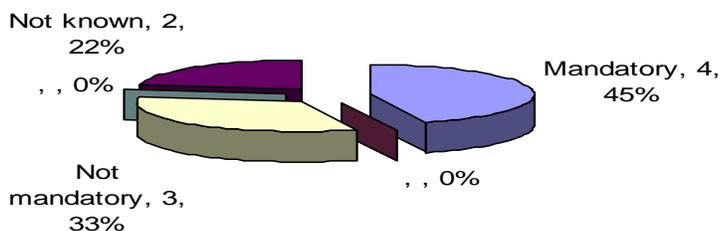
Indian Institute of Astrophysics, Indian Institute of Science, National Institute of Oceanography, Goa institutions have adopted open access policies. On the other hand IIT Kanpur, IIT Bombay, IIT Delhi have restricted access, i.e Full text are available online within the campus; even authors do not have access to their ETDs when they are outside of their campus.

The submission of soft copy of the approved theses/ dissertations is mandatory for IIT Kanpur, IIT Bombay, Indian Institute of Astrophysics, National Chemical

Laboratory, Pune.

Figure: 10

State of ETD Submission policy



It is understood that submission of bibliographic information is mandatory for universities who have signed a MOU with INFLIBNET. ETD submission is not mandatory for Indian Institute of Technology, Delhi, Indian Institute of Oceanography, Goa and Vidyanidhi. The non availability of information on state of submission policy in two institutes is termed as not known. These institutes are now better prepared

to share their information resources through improved management and provision of access to their collections of theses and dissertations.

7. Need for developing a national ETD repository

The vast bulk of research work (theses and dissertations) produced by Indian universities or other tertiary level institutions remains as grey literature or accumulates dust on the shelves of libraries and is thus effectively lost. One of the major limitations of INFLIBNET database is incomplete coverage and non availability of full text or even abstract. Vidyanidhi's coverage is far from being comprehensive therefore not successful in achieving the status of a national e-theses repository. We need to rethink and develop a model which will involve all the stakeholder and work in collaboration. Developing a comprehensive web site (an e-portal of theses and dissertations) is required to facilitate a national ETD project. A unified repository is likely to be better indexed and generally ranked higher by search engines than many separate repositories. The local ETD repositories generally not visited by internet robots. Since the primary purpose of ETD repositories is accessibility, a central reservoir with unifying access is the need of the hour.

7.1 Proposed model

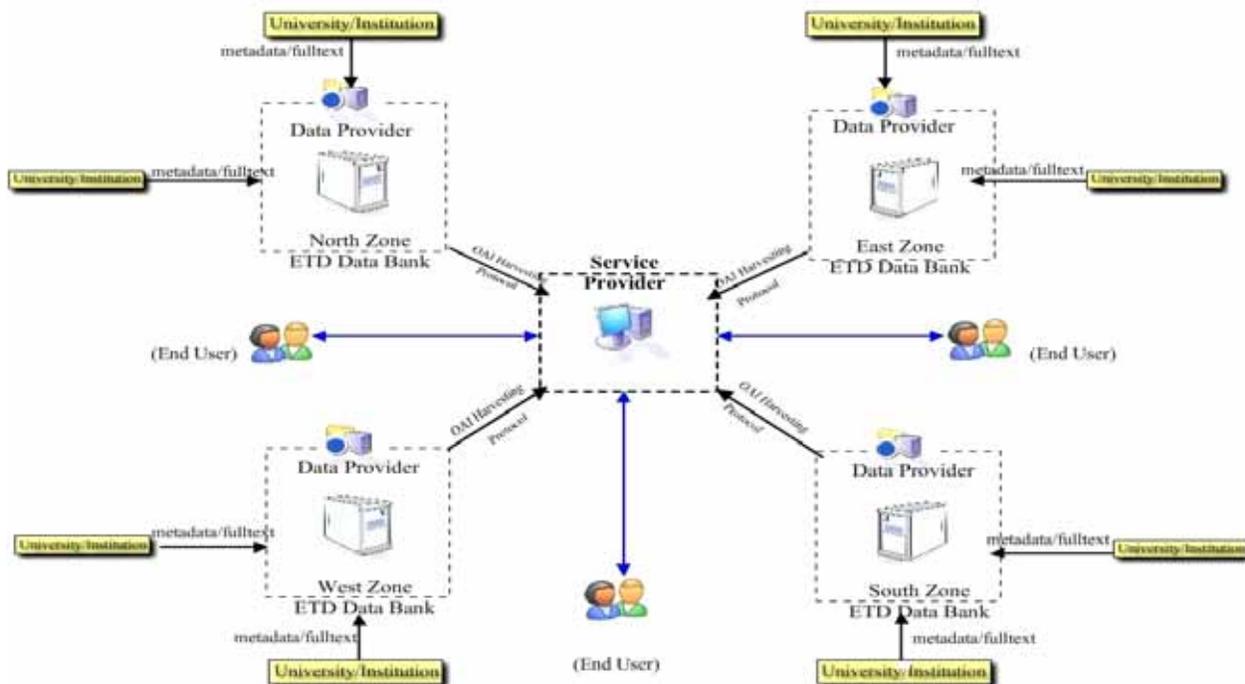
It is evident that in a vast country like India, successful management of the ETD archives and distribution requires participation of all stakeholders. It looks impossible for a single agency to develop and maintain a national repository, as there is large number of PhD granting institutions spreading across the country with variety in institutional settings. A cost effective model is necessary in order to sustain the service and infrastructure, and fund the digitization of theses currently available only in hardcopy at the participating institution. It is envisaged that in order to maintain a viable service, the national ETD repository would have the following architecture:

- ✧ Four zonal data banks as data provider and a central agency as service provider; the service will not be for-profit.
- ✧ Participating universities/ institutions will be offered a choice of two participation options: data provider and ordinary member

The model has been proposed to keep in mind the UGC regulations 2005 of establishing OAI compliant e-theses repositories at each University and metadata harvesting services at national level; each university would send the metadata to a centralized agency to be named by the UGC.

It is further proposed that central agency would work closely with the zonal data banks and other stakeholders to develop standards and tools to support local developers. To maintain good quality collection only elite institutions will be encouraged to contribute as data provider. A minimum subscription fees can be charged if necessary from the libraries requesting documents and have not contributed any data. The data providers or members of the national networks and consortium may avail services at free of cost. The core services of content upload to the Central Hub/ metadata harvesting will be open to all higher educational institutions. The central server will serve as a common database built with metadata gathered from local e-theses server or zonal databank. Small libraries without digital library infrastructure can directly upload theses into the server of the zonal databank. In order to achieve the interoperability between four zonal systems and central server an intense effort of standardization is essential. The adoption of a common metadata set is a fundamental step to achieve interoperability between Zonal ETD data banks. The central server will offer ETD services tailored to the requirements of the institutions without local repositories.

Figure: 11 proposed model



The centralized activities will include: Standardization of record formats, quality control monitoring, distribution /dissemination of full text, data integrity for the central server, access authorization to document delivery, management and maintenance of the central server. On the other hand decentralized activities have: Data entry and editing, Quality control, Storage, archiving and preservation, Copyright and IPR control for full text etc.

7.2 Barriers to setting up ETD repository

The growth in digital libraries of ETDs brings along many new problems and challenges including those of copyright, digitization problems, long term preservation of e-theses etc. There are numerous challenges which relate to the ability of organizations to integrate the management of ETDs into their organizational structure. The following major issues are being faced in many developing countries:

1. Lack of Expertise/ Awareness
2. Lack of support from Faculty/academic staff
3. Leadership problem- The lack of initiatives on the part of parent institutions and the absence of action plans or priorities.
4. Lack of Funding/ infrastructure
5. Access and security
6. Copyright and Intellectual Property Right

IPR is one of the key issues and a significant barrier often confronted in institutions which have embarked on ETDs projects. In most institutions, libraries are one of the key role players in the management of theses and dissertations. With the growing diversity of media and technologies for the production of theses and dissertations, there are a wide range of copyright and licensing issues that need to be taken into consideration.

8. Recommendations:

ETDs is a new generation of theses and dissertations that can include color diagrams, color images, hypertext links, audio, video, animations, spreadsheets, databases, simulations, and virtual reality worlds. Greater collaboration is required to improve ETD sharing and, ultimately, to develop a nationwide digital library of theses and dissertations. The first and main focus at the moment is on setting up the infrastructure and getting the content into the institutional repositories. Further, there is the question of which organizations should build cross-institutional services; this seems less problematic in nations with organizations, such as INFLIBNET, Vidyanidhi and INDEST; each of them have already developed infrastructure and can act as zonal centre. Continuing the natural development of these isolated efforts of ETD digital libraries, effective handling of ETDs automatically requires some collaboration among the stakeholders in institutions. Countries such as Korea, China, Brazil Australia, have implemented, or are in the process of implementing, national policies and developed national ETD repositories, India should follow their footsteps, in order to reduce the information gap.

Following are the recommendations to be considered for a nationwide ETD repository

- ✧ There is no system of theses and dissertations legal deposit in India and universities/institutions are not under the subject of the archiving organizations, so the collecting work is not very smooth and some theses and dissertations are simply lost or accumulate dust. There is need for concerted effort by all ETD stakeholders to integrate theses and dissertations produced in our country in a unified database to be coordinated by a central agency of Government of India.
- ✧ A project for developing a complete national union catalogue of ETDs or amalgamation of existing bibliographic databases already in possession under INFLIBNET, DELNET or other library networks need to be initiated in order to avoid any duplication of work.
- ✧ E-portal and interface- a good user interface is essential for visibility and maximizing access and act as a guide for the stakeholders on IPR issues in particular. A comprehensive web site (e-theses portal) need to be developed for providing guidance to researchers, postgraduate students and supervisors on ETD submission format, metadata creation etc.
- ✧ Personalization and value added services- incorporating personalization into ETDs, will be extremely helpful for users; researchers should be encouraged to submit multimedia theses, which can be easily integrated with other academic applications.
- ✧ ETD Awareness program/ promotional activities- an improved understanding of the issues and options can encourage cooperation among institutions that are planning or have implemented ETD programs. There is need to sensitize all the stakeholders about the need and issues involved in the creation of ETDs. Workshops and training programs for both students and faculty need to be conducted on issues of copyright, choice of access, and encourage researcher to submit theses online. Indian information professionals are setting their minds to take up the challenges and actively getting involved in creation and hosting e-theses, need vigorous training. Training is also very important, in terms of tools, processes, and standards.

9. End note

ETD building is developing very fast in India and ETD projects are resulted in a new and improved inter institutional relationship. Since first started at IIT Bombay in 1999 till date there are nearly a dozen ETD repositories have been developed in India, but the repositories size is small and scattered throughout the country. Majority of the ETD repositories have restricted full text access and open access to bibliographic databases. The proactive participation of college and university libraries/ research institutions is required for national e-theses repository to succeed owing to its cooperative nature. Recently National Knowledge Commission has submitted a report to the Prime minister of India, (December, 2006) for establishing a national level institute namely Indian Institute of Library and Information Science (ILIS) in pattern of Indian Institute of Technology (IITs) in India. It is expected that this proposal will be implemented soon. In that case a national e-theses repository under the aegis of ILIS will be an ideal solution to information poverty in a vast country like India where around three hundred Universities and hundred R&D laboratories are producing a sizable number of research theses and dissertations. A national-level mechanism is essential to promote ETD movement and improve awareness.

Acknowledgements:

The author is grateful to Professor Sanjay G. Dhande, Director, and Professor Somenath Biswas, Dean of Faculty Affairs, Indian Institute of Technology, Kanpur for their continual support, encouragement and permission to attend ETD symposium 2007 at Uppsala library, Sweden.

The author would like to express appreciation to the following persons for providing data/information in writing this article:

Mr. Umed Singh, Assistant Librarian, IIT Kanpur

Mr. Mahendra Jadhav, Assistant Librarian, IIT Bombay

Mr. Ajay Kumar Sharma, Central Library, IIT Delhi

Mr. Prem Chand and Rajesh Chandrakar INFLIBNET, Ahmedabad

Mr. Murari P Tapaswi, librarian, NICMAS at NIO, Goa

Mr. P. Prabakar, IIA, Bangalore

References:

Alamsyah, Oling Pushpa (2003) Digital Preservation of Scholarly Electronic Publications, (MSc dissertation, Information studies, Unpublished), School of Communication and Information, Nanyang Technological University, Singapore. Pg 92

Arunachalam (2004) India's march towards open access

available at <http://www.scidev.net/content/opinions/eng/indias-march-towards-open-access.cfm> (23/3/07)

Caplan P and Thomas C (2006) DAITSS: Another Preservation Option For Electronic Theses & Dissertations In proceedings of the 9th International Symposium on Electronic Theses and Dissertations, June 7 – 10, 2006, Quebec City, Canada.

Available at: http://www6.bibl.ulaval.ca:8080/etd2006/pages/papers/SP10_Charles_Thomas.pdf (23/3/07)

Choudhury S and Müller Eva (2006) Added values to E-theses: A preliminary version, 9th International Symposium on Electronic Theses and Dissertations June, 2006, Quebec City, Canada

Available at: http://www6.bibl.ulaval.ca:8080/etd2006/pages/papers/SP4_Muller_Choudhury.pdf (23/3/07)

- Dobratz, Sussane (2005) Thinking the long term: the XML based publishing work flow for handling electronic theses and dissertations at Humboldt university, Berlin
Available at: <http://adt.caul.edu.au/etd2005/papers/075Dobratz.pdf> (3/5/07)
- Edminster Jude, and Moxley, Joe (2002) Graduate Education and the Evolving Genre of Electronic Theses and Dissertations. Computers and composition, no.19,pp 89-104
Available at: <http://personal.bgsu.edu/~jedmins/dissertation.pdf> (3/5/2007)
- Fox, Edward A(1998) Networked Digital Library of Theses and Dissertations: A Framework for East-West Collaboration Paper for Proceedings of East meets West An Asian Digital Library Workshop (5-7 August 1998) University of Hong Kong.
Available at: <http://docs.ndltd.org:8080/dspace/bitstream/2340/39/1/Asia98.htm> 13/4/07
- Fox, Edward. A (1999) The Evolving Genre of Electronic Theses and Dissertations In Proceedings of the 32nd Hawaii International Conference on System Sciences 1999.
Available at: <http://ieeexplore.ieee.org/iel5/6293/16782/00772646.pdf> (13/5/07)
- Fox, Edward, A et. al (ed.) Electronic Theses and Dissertations: a sourcebook for Educators, Students and Librarians, Marcel Dekker Inc, 2004, New York, pg 214.
- Fox, Edward. A et.al (2006) ETDs, NDLTD, and open access: a 5S perspective Ciencia da Informacao v.35 no. 2 May/Aug 2006 pp 75-90
- Jadhav, Mahendra N (2006) Enhancing Digital Repository of Scholarly Publications at Indian Institute of Technology, Bombay (IITB) India, 9th International Symposium on Electronic Theses and Dissertations IXe June 7 – 10 2006, Quebec City, Canada.
Presentation available at: <http://www6.bibl.ulaval.ca:8080/etd2006/pages/prog.jsf>] April 13, 2007
- Jin, Yi (2004) The development of the China Networked Digital Library of Theses and Dissertations, Online Information Review, V. 28 No 5, pp. 367-370
- Jobish,P et.al (2005) etd@IISc : a Dspace- based ETD-MS and OAI compliant theses repository service of Indian Institute of Science in 8th International symposium on electronic theses and dissertation,28-30 September, Sydney, Australia. Available at: <http://adt.caul.edu.au/etd2005/papers/049Minj.pdf> (3/5/07)
- Marcondes, C.H & Sayao, L.F (2003) Brazilian Digital Library of Theses and Dissertations, International information and library review, v. 35, pp 265-279
- Murthy, T. A.V et. al (2005) UGC-INFLIBNET Initiatives in E-journal consortia and digital library of Doctoral theses for Indian Universities in proceedings of National Conference on Digital Library and e-Theses (NCDLET 2005), January 7-8, 2005 at Jadavpur University, Kolkata, India
- Seng, Foo Tun(2004) Design and development of a digital library of bibliographic data of theses and dissertations, (MSc dissertation, Information Studies, unpublished) School of Communication and Information, Nanyang technological University, Singapore, pg.108
- Southwick, S.B (2006) The Brazilian electronic theses and dissertations digital library: providing open access for scholarly information Ci. Inf., Brasilia, v. 35, n. 2, pp 103-110, May/Aug.
- Southwick, S.B and Pavani, A M B (2004) Some Practical points on organizing ETD consortia, 7th international symposium on electronic theses and dissertation. Available at:

http://docs.ndltd.org:8080/dspace/bitstream/2340/211/1/ETD_04_consortia_pavani_southwick.pdf
(13/5/07)

Urs, Shalini R (2003) Vidyanidhi - the evolving Indian Digital Library of Electronic Theses Initiative
Available at: <http://edoc.hu-berlin.de/etd2003/urs-shalini/PDF/Urs.pdf>(13/5/07)

Vijayakumar, J K and Murthy, T A V (2001) Need of a Digital Library for Indian Theses and Dissertations: a model on par with the ETD initiatives at International Level.
In Urs, Shalini R, Rajashekhar, TB and Raghavan, KS(ed) Digital Libraries Conference Papers of the 4th International Conference of Asian Digital Libraries (ICADL-2001), Bangalore, India. 10-12 December, pp 384-390.

Vijaykumar, J. K et.al. (2004) Electronic theses and dissertations for Indian universities: a framework. In Murthy, T. A. V. (eds.) Proceedings 2nd PLANNER-2004, pp. 65-70, Imphal India. available at: <http://eprints.rclis.org/archive/00005656/>

Vijayakumar, J K et.al (2005) Introducing Electronic Theses and Dissertations in Universities: An Indian Perspective in International Conference on Grey Literature(GL-7), INIST/CNRS, Nancy, France, December. Available at: <http://eprints.rclis.org/archive/00005670/> 15/5/2007

Wolf, Martin (2006) Electronic Theses Online Service (EThOS) – A national UK Project in 9th International Symposium on Electronic Theses and Dissertations, June 7 – 10 2006, Quebec City, Canada
Available at: <http://www6.bibl.ulaval.ca:8080/etd2006/pages/prog.jsf> (3/5/07)

Zhang, Yin et.al (2001) Usage patterns of an electronic theses and dissertations system, Online Information Review, V. 25 No. 6, pp. 370-378

List of web sites consulted:

1. UNESCO ETD Guide website at <http://etdguide.org/> (3/5/07)
2. Electronic Thesis Online (India): UGC (Submission of Metadata and Full-text of Doctoral Theses in Electronic Format) Regulations, 2005: http://www.ugc.ac.in/new_initiatives/etd_hb.pdf
http://www.ugc.ac.in/new_initiatives/consultation_electronic.html retrieved on 3/5/2007
3. National Knowledge commission recommendation on Libraries
<http://knowledgecommission.gov.in/recommendations/libraries.asp> (13/5/07)
4. INFLIBNET's national online Union catalogue - <http://202.141.130.73/unicat/thesis.jsp> (3/5/07)
5. DELNET database of theses and dissertations- <http://delnet.nic.in/access-catalogs-databases.htm>
6. <http://www.ncstrl.org> (15/5/07)
7. Virginia Tech Electronic Thesis and Dissertation home page." <<http://etd.vt.edu/>> (9/5/07)
8. ETD Digital Library." <<http://www.theses.org/>> (9/5/07)
9. TUG Electronic Thesis Project." <<http://www.lib.uwaterloo.ca/TUG/ETD/>> 13/5/07
10. <http://library.caltech.edu/etd/index.html> (13/5/07)
11. <http://www.lib.iitk.ac.in:8080/examples/thesis/index.html> (13/5/07)
12. <http://dspace.vidyanidhi.org.in:8080/dspace/> (13/5/07)
13. <http://www.inflibnet.ac.in/> (13/5/07)

