

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Library Philosophy and Practice (e-journal)

Libraries at University of Nebraska-Lincoln

October 2006

Experimenting with a Model Digital Library of ETDs for Indian Universities Using D-Space

J. K. Vijayakumar

American University of Antigua, vijayakumarjk@yahoo.com

TAV Murthy

Information and Library Network Centre, Ahmedabad, India, tav@inflibnet.ac.in

M. T. M. Khan

Bundelkhand University, Jhansi, India, mtmkhan@rediffmail.com

Follow this and additional works at: <https://digitalcommons.unl.edu/libphilprac>



Part of the [Library and Information Science Commons](#)

Vijayakumar, J. K. ; Murthy, TAV; and Khan, M. T. M., "Experimenting with a Model Digital Library of ETDs for Indian Universities Using D-Space" (2006). *Library Philosophy and Practice (e-journal)*. 105.
<https://digitalcommons.unl.edu/libphilprac/105>

Experimenting with a Model Digital Library of ETDs for Indian Universities Using D-Space

J K Vijayakumar

*Assistant Director
Health Sciences Library
American University of Antigua
PO Box W-1451
St Johns, Antigua, West Indies*

T. A. V. Murthy

*Director
INFLIBNET Centre
PO Box No 4116, Navrangpura
Ahmedabad- 380009 India*

M. T. M. Khan

*Professor and Head
Institute of Library and Information Science
Bundelkhand University
Jhansi - 284128 India*

ETD - Electronic Theses and Dissertations

Electronic Theses and Dissertations (ETD) can take a variety of forms, from a Word or PDF version of a printed thesis, to a truly digital publication that includes audio and visual material and may be organized quite differently from a printed thesis [1].

The broader benefits of ETDs have been described as follows [2&3];

- Broader exposure of university research;
- New forms of creative scholarship through interactive elements, multimedia, hyperlinks, etc.;
- Ability to hyperlink to the thesis/dissertation on homepages and electronic CVs;
- Professional development for researchers who learn the basic skills of scholarly publishing in an electronic format;
- Conservation of paper, library storage space, and staff time;
- Faster access, with publication near the point of submission rather than three to four months later;
- Theses or dissertations accessible to any reader at any time.

ETD System: a Model

Formats for Main Files

When constructing the ETD system, there will be a need to convert files in different formats to a unified format, which can preserve the contents, format and layout of original documents, created by various programs/processors. Still there is no standard electronic format accepted for all kinds of documents, the PDF format is the most popular and adopted in most current ETD systems. It is always better if the text-based portion of the thesis or dissertation is in PDF, which allows documents created through word processing like MS Word, to be made available on the Web in an effective way. PDF retains the appearance of the print version across platforms and browsers [2&3].

Formats for Additional Files

Multimedia supplements can be mounted on an ETD server to support the text, but most ETDs still resemble their print equivalents. As scholars begin to desire better information along with better access to information, then the use of multimedia files will increase. MP3 is the file format of choice for audio files. This nonproprietary file format requires relatively little storage space. Apple QuickTime and MPEG Movie Player can be used to incorporate video clips. The widespread availability of these applications will ensure future access to the contents of ETDs, if the institutions take steps to make the software available by maintaining freely distributed applications by bundling them with corresponding media files on the host server [2&3].

Software to Manage ETDs

In India, for long-term preservation, we have to find an economical way to save digital content for future generations by using a variety of open source archiving solutions. We should give consideration to a number of systems that have already been tested and adopted by respected institutions. These include [DSpace](#), [Eprints](#), [Virginia Tech University's ETD-db](#), [Greenstone](#), and several other software packages. Many open source packages show some degree of similarity, but the key factors in selection are Suitability, Functionality, Interoperability and Sustainability. Copeland and Penman suggest the following criteria for selecting software for ETD systems [4].

Suitability: Software should be easy to install on a range of hardware and operating systems, and should be available free and open source. The ease of customization and availability of upgrade are prime concerns.

Functionality: It should have an intuitive and appealing user interface for administrator and author, and it encourage authors to submit content. Persistent URLs are essential for preserving long-term access to content. Simple and advanced metadata searching will allow a variety of search methods, ideally with full-text searching. It should be possible to apply metadata that conforms to national or institutional schemes. The software should support any file format or file size.

Interoperability: The software system must comply with the latest version of the 'Open Archives Protocol for Metadata Harvesting' (OAI-PMH), as well as satisfying individual institutional policies for integrating ETDs with other material in electronic repositories. This is an important to ensure that the system will import and export information from one system to another.

Sustainability: Repositories are long-term commitments, and the institution should be confident that the software will offer continued support and development. This is especially important because much ETD, digital library, and institutional repository software is relatively new and untested. As is common with much open source software, once a user community is established, the knowledge base can help 'keep the ball rolling,' by offering support to new users.

There are a number of options to compare when choosing ETD software. Jones, of the UK's [Theses Alive](#) project, evaluated two open source packages to deliver E-theses via a web-based interface: [ETD-db](#) by Virginia Tech and [D-Space](#) from HP and MIT. [ETD-db](#) is specifically designed for ETDs and endorsed by the [NDLTD](#), but was not usable by institutions who wanted to create a repository and host ETD as a part of it. Jones found that [D-Space](#) has a comprehensive and flexible system for applying and storing metadata. Because the Dublin Core registry within [D-Space](#) is customizable, and because of the option to modify the submission interface, [D-Space](#) will be able to accommodate future changes to metadata schemas. The [D-Space](#) archive is perhaps more geared toward digital preservation and the level of configuration available within the [D-Space](#) administrative area is also excellent [5]. The capability of [D-Space](#) to handle multilingual content, even at the metadata level, using the globally-accepted UNICODE standard, are important issues in countries like India for selecting this solution [6].

Prototype

[D-Space](#) is a digital library system for preserving faculty research [7]. Universities require a system that is strong in interoperability in order to sustain these digital repositories. Repositories must be able to provide access to many different kinds of digital object. [D-Space](#) has the flexibility and sustainability that is needed[5]. As an open source system, [D-Space](#) is freely available to run as-is, or to modify to meet local needs.

Setting up [D-Space](#) Server

[D-Space](#) was downloaded from <http://dspace.org/>. It was installed in a test-bed to experiment with its capabilities and performance. After testing, [D-Space](#) was customized and installed on server with a Linux (RedHat 9) platform. Then a request went to the Corporation for National Research Initiatives (CNRI) site to provide Persistent Identifiers (CNRI Handles), which promotes interoperability among open archives through the Open Archives Initiative Protocol for Metadata Harvesting, (OAI-PMH). After the necessary registrations, the [D-Space](#) Server went live (Figure 1).



Figure 1- [D-Space](#) Home Page

Creation of Community and Collection

After the server is set up, the next step is to create a Community, a group that will contribute content to [D-Space](#). Communities have Collections, which contain the content items, or files. Communities determine their own content and user access guidelines. A system administrator creates workflows for content to be accepted, edited, tagged with metadata, etc. The head of a community makes policy decisions where the content is created. Communities manage their own metadata and can also customize the look and feel of their pages in [D-Space](#). Figure 2 shows the creation of a community, ETD@INDIA.

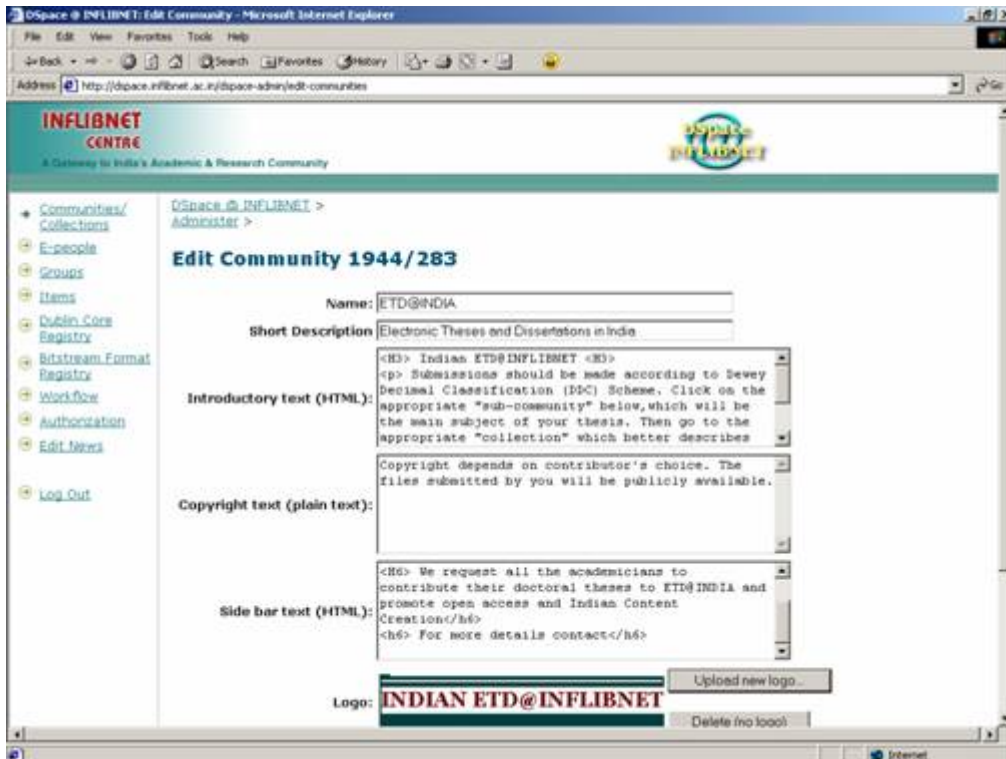


Figure 2 - Creation of Community ETD@INDIA

The next step is creating subcommunities, which are again divided into collections where researchers can submit their theses. Any number of subcommunities and collections can be created, and the following were made according to the Dewey Decimal Classification (DDC) Scheme.

- 000:Computer Science, Information and General Works
- 100:Philosophy & psychology
- 200:Religion
- 300:Social Sciences
- 400:Language
- 500:Science
- 600:Technology
- 700:Arts & Recreation
- 800:Literature
- 900:History & Geography

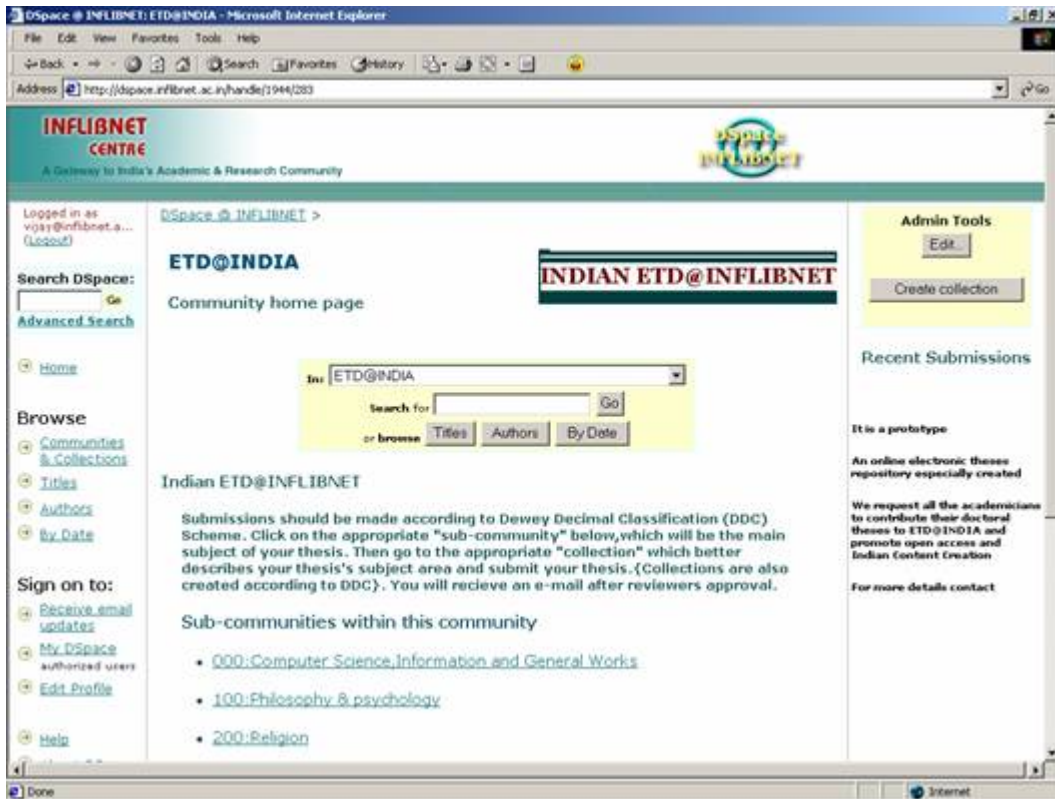


Figure 3 - ETD@INDIA Community Home Page

Collections were also created according to DDC. For example, the subcommunity "000:Computer Science, Information and General Works" was divided into collections which can be seen in Figure 4.

- 010 Bibliography
- 020 Library & information sciences
- 030 General encyclopedic works
- 050 General serial publications
- 060 General organizations & museology
- 070 News media, journalism, publishing
- 080 General collections
- 090 Manuscripts & rare books

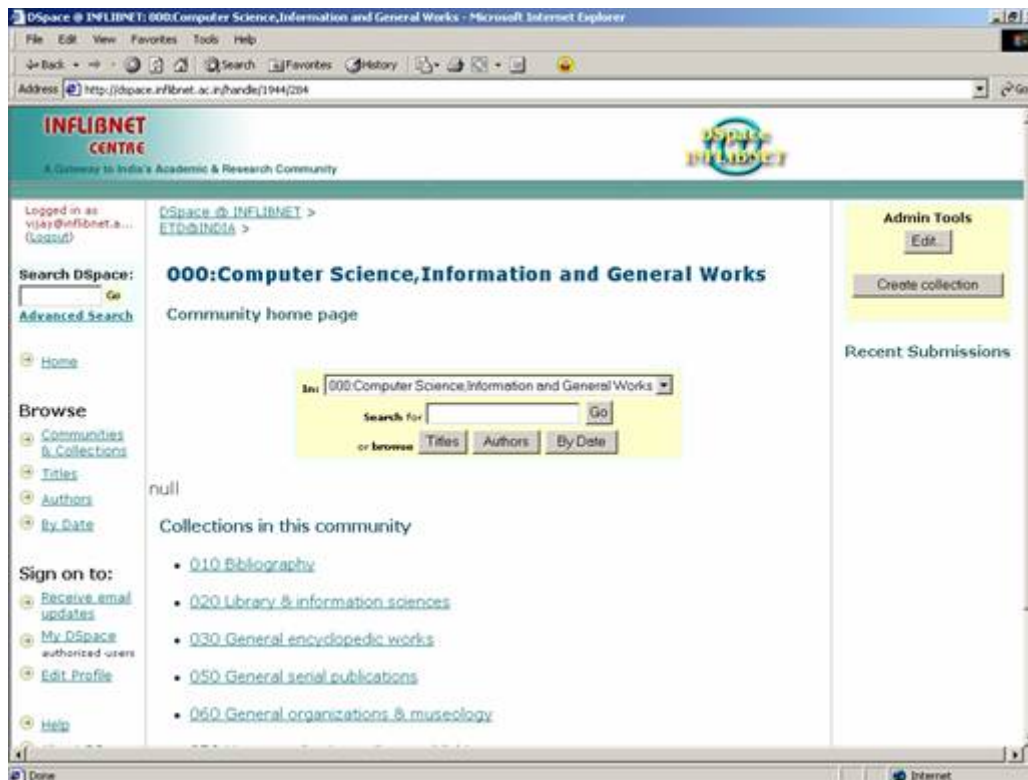


Figure 4 - ETD@INDIA Collections Home Page

Submitting the Item

To submit a thesis, researchers access the Submitting page by clicking on Communities and Collections. They can select the collection that contains their subject. Submitting a thesis has these steps, which are described in the following figures.

- Selecting the Collection - Figure-5
- Describe the item (1) - Figure-6
- Describe the item (2) - Figure-7
- Describe the item (3) - Figure-8
- Upload the item - Figure-9
- Verify the item (1) - Figure-10
- Verify the item (2) - Figure-11
- Licensee Check - Figure 12
- Finishing Message - Figure 13

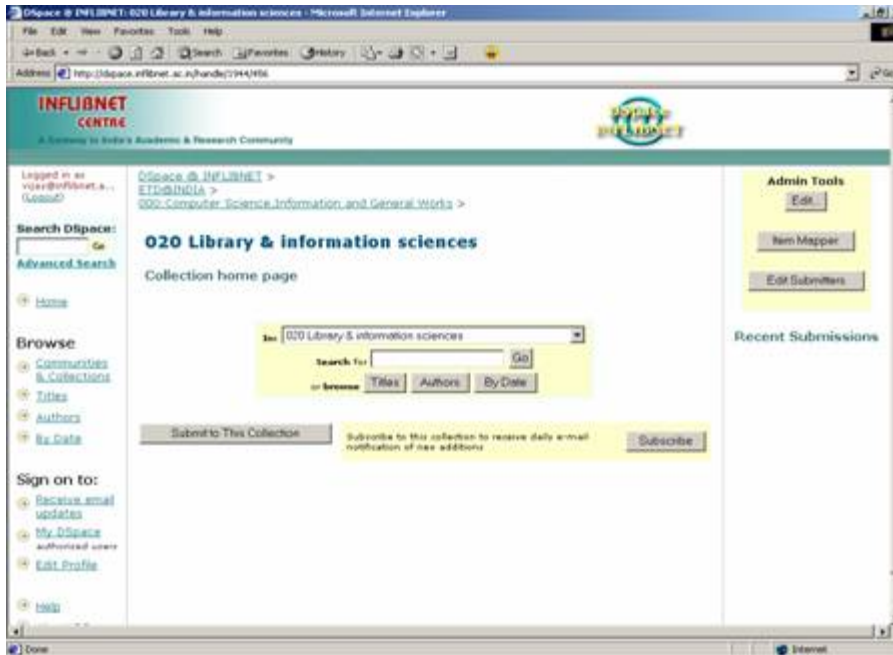


Figure 5 - ETD@INDIA Selecting Collection

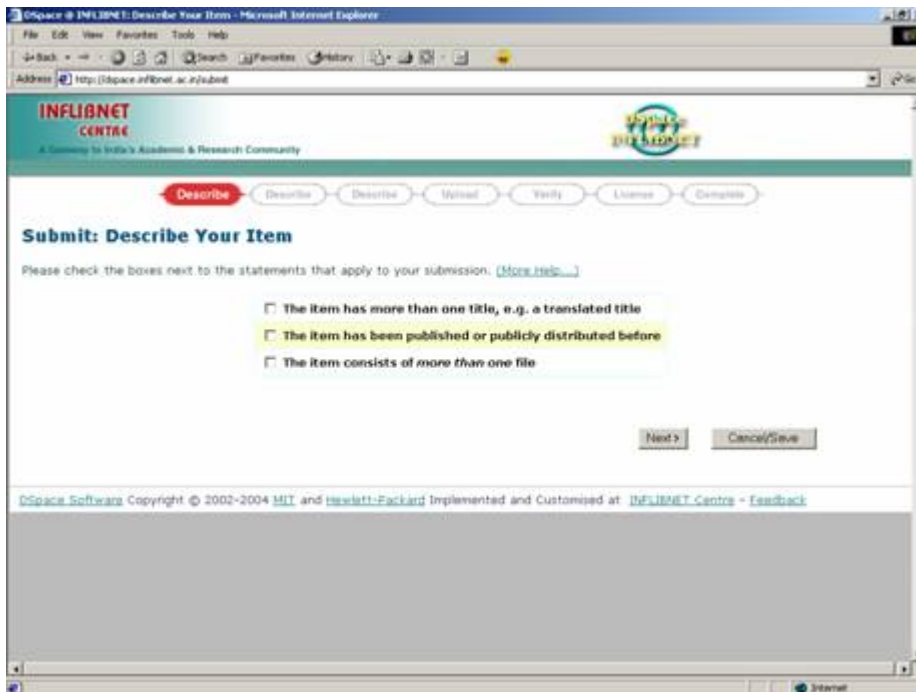


Figure 6 - ETD@INDIA Describe the item (1)

Submit: Describe Your Item

Please fill in the requested information about your submission below. In most browsers, you can use the tab key to move the cursor to the next input box or button, to save you having to use the mouse each time. [\(More help...\)](#)

Enter the names of the authors of this item below.
 Last name First name(s) + "Jr"
 e.g. Smith e.g. Donald Jr

Authors: Vjyayakumar JK Remove This Author
 Murthy T.A.V.
 Khan M.T.M.
 Add More

Enter the main title of the item.
Title: Importance of Ph.D. Theses in Furthering Research - a Model

Enter the series and number assigned to this item by your community.
 Series Name Report or Paper No.
Series/Report No.: Add More

If the item has any identification numbers or codes associated with it, please enter the types and the actual numbers or codes below.
Identifiers: URI Add More

Select the type(s) of content you are submitting. To select more than one value in the list, you may have to hold down the "CTRL" or "SHIFT" key.
Type: Recording musical, Recording oral, Software, Technical Report, Thesis, Video

Figure 7- ETD@INDIA Describe the item (2)

Submit: Describe Your Item

Please fill further information about your submission below. [\(More help...\)](#)

Enter appropriate subject keywords or phrases below.
Subject Keywords: Digital Library ETD Add More

Enter the abstract of the item below.
Abstract: As a research resource, the Indian university theses have a limited distribution and may be under used outside the host institutions. The importance of having Theses and Dissertations in electronic form, which has to be made available for easy search and access, is well established through the Literature Search and Survey of Indian Universities conducted as the part of this

Enter the names of any sponsors and/or funding codes in the box below.
Sponsors:

Enter any other description or comments in this box.
Description:

< Previous Next > Cancel/Save

Figure 8 ETD@INDIA Describe the item (3)

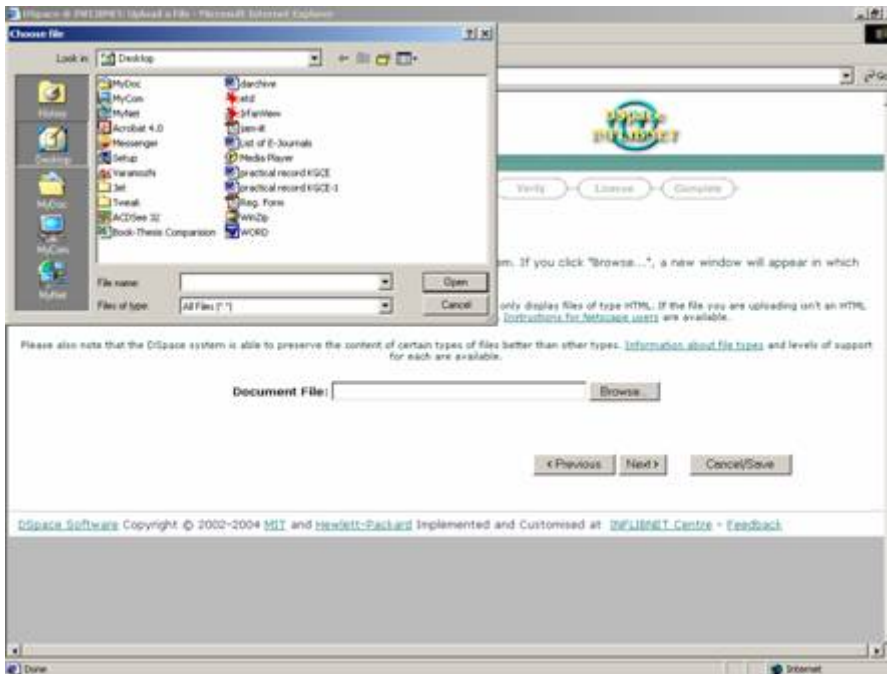


Figure 9 - ETD@INDIA Upload the item



Figure 10 ETD@INDIA Verify the item (1)

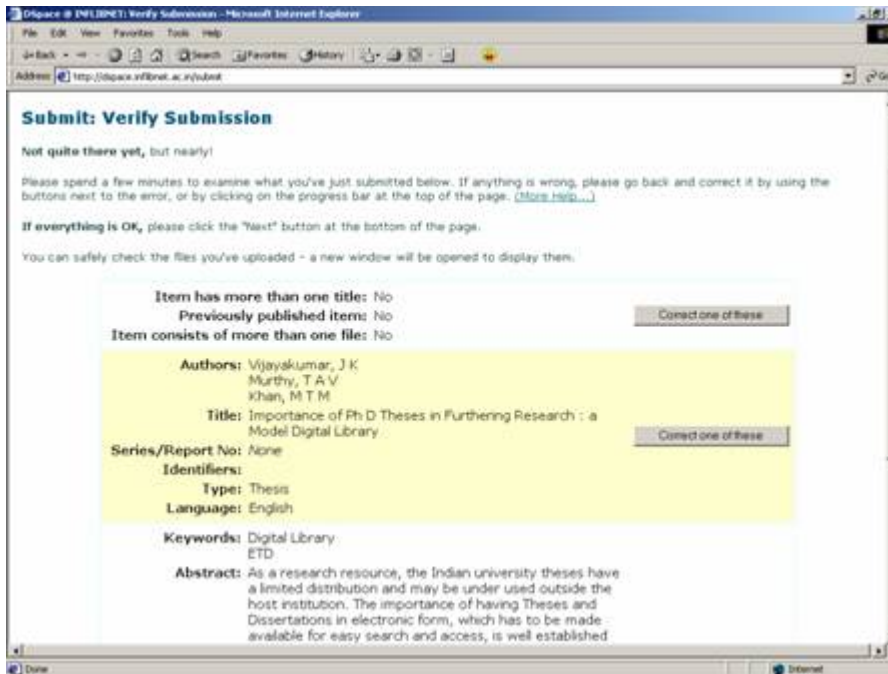


Figure 11 - ETD@INDIA Verify the item (2)

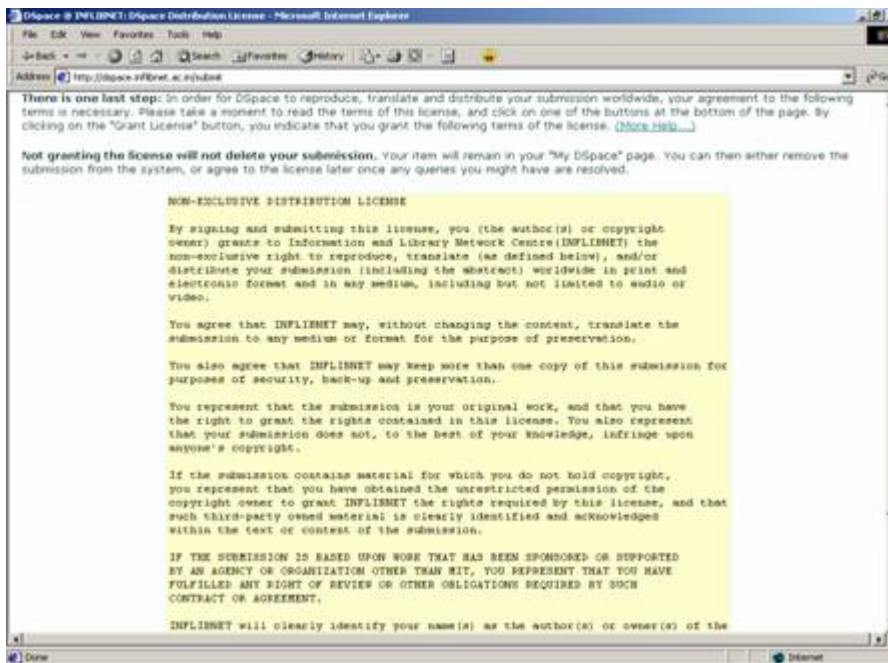


Figure 12- ETD@INDIA Licensee Check



Figure 13 - ETD@INDIA Finishing Message

Search for an Item

D-Space allows users to find content in a number of ways. Users have high expectations from search engines, so one goal for D-Space is to supply as many search features as possible. Browsing is an important mechanism for discovery in D-Space. The browse features allows the user to view a particular index, such as the title index. The browse subsystem allows users to specify an index and a subsection of that index. Browsable indexes are title, issue date, and authors. The browse can be limited to items within a particular collection or community [8]. The following figures illustrate searching from the Home Page of ETD@INDIA.

- General Search [by Author Name] - Figure 14
- First Level of Display - Figure 15
- Second Level Display - Figure 16
- Full Text Display - Figure 17

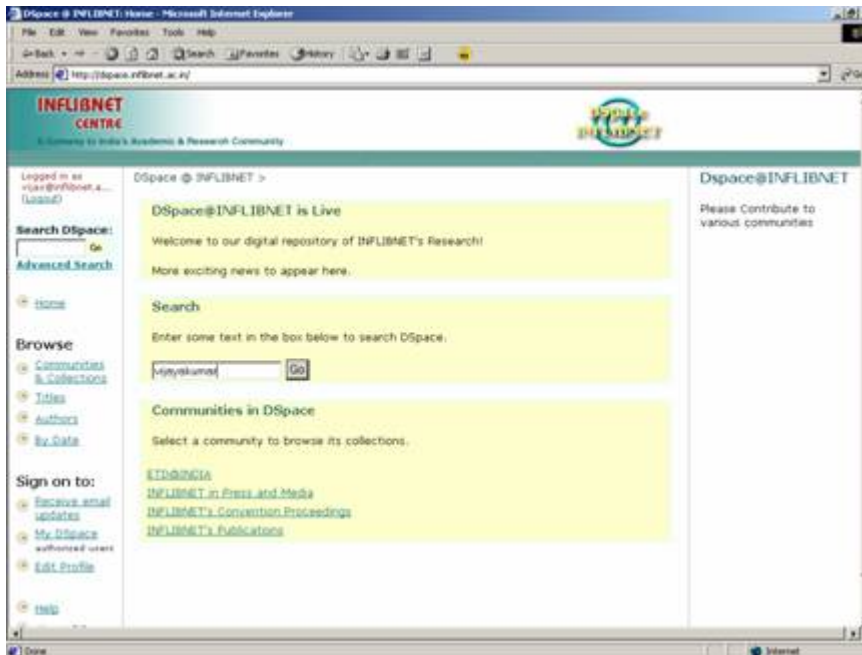


Figure 14 - ETD@INDIA General Search [by Author Name]

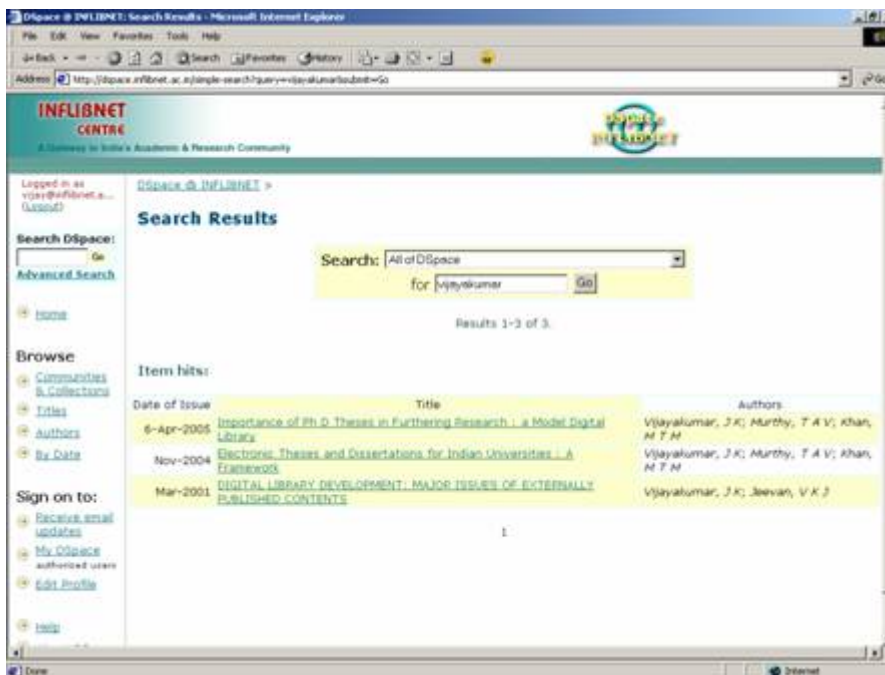


Figure 15- ETD@INDIA First Level of Display

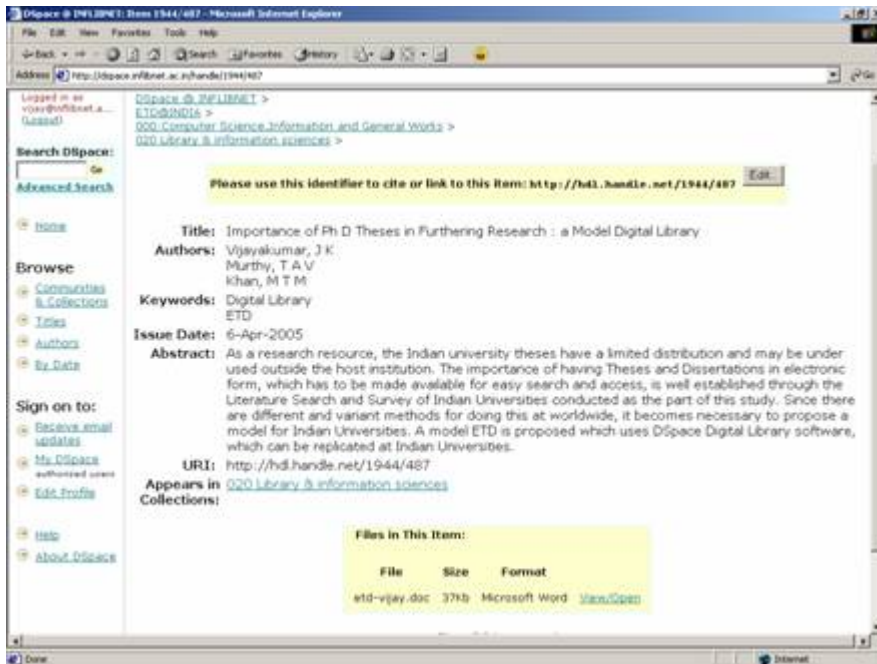


Figure 16 - ETD@INDIA Second Level Display

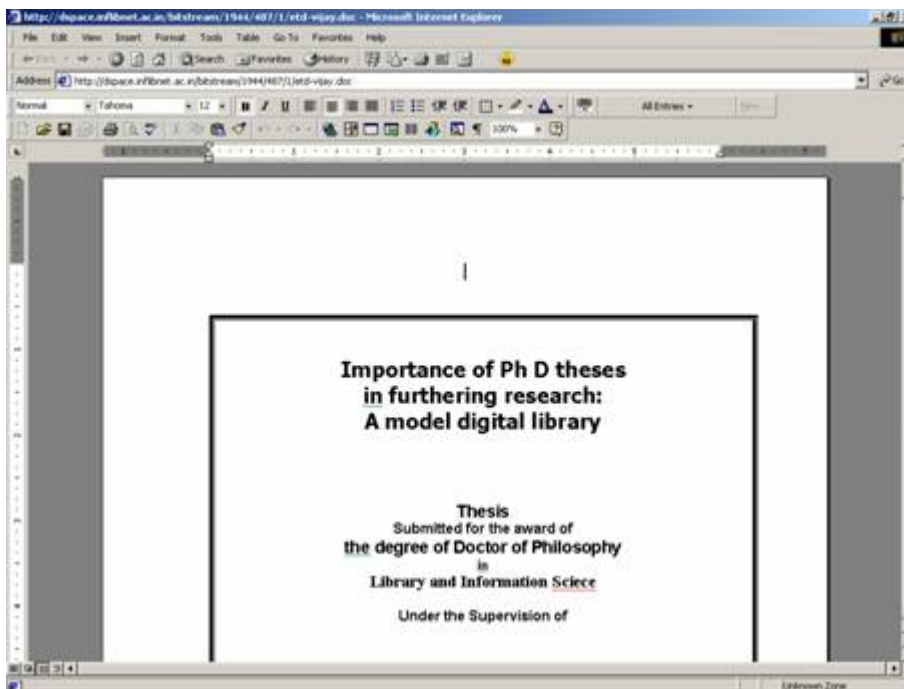
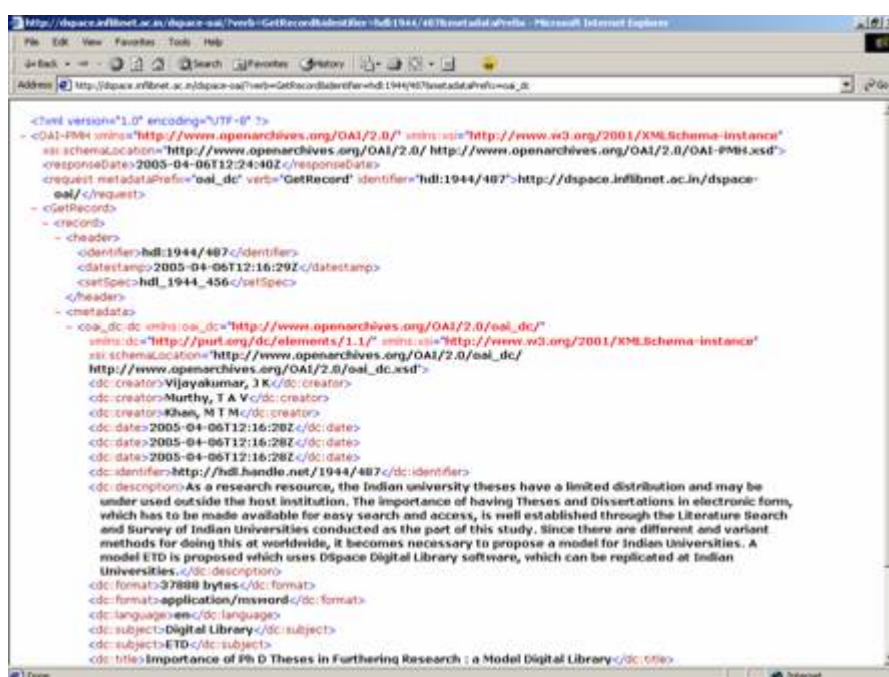


Figure 17- ETD@INDIA Full Text Display

Metadata for ETD@India

D-Space holds three types of metadata about archived content: descriptive metadata, administrative metadata, and structural metadata. Descriptive metadata contains a qualified Dublin Core record. Administrative metadata includes preservation information, provenance, and authorization policy data, most of which is held within D-Space's relation DBMS schema, where provenance information is stored in Dublin Core. The structural metadata includes has information on the relationship between parts of an item, and determines its presentation to users [8].

OAI-PMH is an interoperability framework based on metadata harvesting. Any [D-Space](#) server can process OAI-PMH. The Items in the repository will have a unique identifier, e.g., CNRI Handles. An example generated by OAI-PMH, based on the Dublin Core schema, is shown in Figure 18 below.



```
<?xml version="1.0" encoding="UTF-8" ?>
<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
  <responseDate>2005-04-06T12:24:40Z</responseDate>
  <request metadataPrefix="oai_dc" verb="GetRecord" identifier="hdl:1944/487">http://dspace.inflibnet.ac.in/dspace-
  oai/</request>
  <GetRecord>
    <records>
      <header>
        <identifier>hdl:1944/487</identifier>
        <timestamp>2005-04-06T12:16:29Z</timestamp>
        <setSpec>hdl_1944_456</setSpec>
      </header>
      <metadata>
        <dc:dc xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
          xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
          xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
            http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
          <dc:creator>Vijayakumar, J K</dc:creator>
          <dc:creator>Murthy, T A V</dc:creator>
          <dc:creator>Khan, M T M</dc:creator>
          <dc:date>2005-04-06T12:16:28Z</dc:date>
          <dc:date>2005-04-06T12:16:28Z</dc:date>
          <dc:date>2005-04-06T12:16:28Z</dc:date>
          <dc:identifier>http://hdl.handle.net/1944/487</dc:identifier>
          <dc:description>As a research resource, the Indian university theses have a limited distribution and may be
            under used outside the host institution. The importance of having Theses and Dissertations in electronic form,
            which has to be made available for easy search and access, is well established through the Literature Search
            and Survey of Indian Universities conducted as the part of this study. Since there are different and variant
            methods for doing this at worldwide, it becomes necessary to propose a model for Indian Universities. A
            model ETD is proposed which uses DSpace Digital Library software, which can be replicated at Indian
            Universities.</dc:description>
          <dc:format>37888 bytes</dc:format>
          <dc:format>application/msword</dc:format>
          <dc:language>en</dc:language>
          <dc:subject>Digital Library</dc:subject>
          <dc:subject>ETD</dc:subject>
          <dc:title>Importance of Ph D Theses in Furthering Research : a Model Digital Library</dc:title>
        </dc:dc>
      </metadata>
    </records>
  </GetRecord>
</OAI-PMH>
```

Figure 18 - ETD@INDIA OAI-PMH Metadata Display

Metadata can be created at the point of submission. The template provided should have tagged elements for all metadata to be collected, except for rights administration, which must be verified by staff, who also may add supplementary metadata, keywords, and classifications. The default submission procedure of [D-Space](#) does not support automatic metadata extraction. After the license agreement type is selected and recorded in the metadata, the submission process is complete.

The administrator of the system can review submissions and approve or reject them. If an item is not accepted, the administrator's comments can be sent to the author by email. The author can revise the material and resubmit, or review any submitted version with the administrator. Additional duties of the administrator of the ETD System include designating groups who can approve user submissions, and management of user accounts. Individual user

accounts can use the local authentication scheme, and [D-Space](#) supports customizable authentication systems. After final approval, a message goes to researchers and registered users about the submission of a new ETD. Policies on access to ETDs can also be implemented by the administrator through the [DSpace](#) server, for Intranet or Internet access [9].

Limitations of the Model

The proposed model uses Dublin Core metadata for describing the ETDs. The [ETD-MS](#) Metadata standard especially developed by NDLTD for ETDs can be implemented in this model [10]. The full version of an ETD system can be developed by creating add-ons to [D-Space](#), according to the requirements or practices of a university. This will be a fully automated system, similar to existing practices for writing a dissertation, including registration of a researcher, submitting the proposal, approval by the committee, interaction with the advisor, sending interim reports, submitting the thesis to committee members, comments and revisions by members and advisor, final submission, and award of degree. [TAPIR \(Theses Alive Plug-in for Institutional Repositories\)](#) as an add-on to [D-Space](#) is a good example of this. Similar things can be developed for Indian conditions [11].

Conclusion

The proposed model can be used by Indian universities to create their ETDs and provide access either on their Intranet or the Internet. Well-equipped computer labs must be put in place to provide workstations, software, and technical support staff for students writing ETDs. Standards are needed for the presentation of dissertation research [12]. This model is derived by keeping in mind the embryonic nature of this idea in India and, at the same time, the changing technological environment in Indian campuses. Worldwide ETD initiatives should consider being part of OAI for global access to scholarship, so that Indian universities can also benefit from sharing and contributing. At the same time we must consider several strategies for acquiring the latest technological developments, including the availability of free solutions for a better and effective implementation of ICT in Indian campuses.

References

1. Robert Gordon University ETD Site, <<http://www2.rgu.ac.uk/library/etds/what.html>>
2. University of Pittsburgh ETD Project site, <<http://www.pitt.edu/AFShome/g/r/graduate/public/html/etd/copyright.html>>
3. Oregon State University ETD FAQ, <http://oregonstate.edu/dept/grad_school/current/thesis_faq.html>
4. Copeland, Susan and Penman, Andrew (2004). The Development and Promotion of Electronic Theses and Dissertations (ETDs) within the UK, *The New Review of Information Networking*, 10 (1), 2004: Pp 19-32.
5. Jones, Richard (2004). [DSpace vs. ETD-db: Choosing software to manage electronic theses and dissertations](#), *Ariadne*, Issue 38, January 2004. <<http://www.ariadne.ac.uk/issue38/jones/intro.html>>
6. Patel, Yatrik, Vijayakumar, JK and Murthy, TAV. Institutional Digital Repositories / E-Archives: INFLIBNET's Initiative in India , *Proceedings of the 7th MANLIBNET Annual National Convention* , May 5-7, 2005, Calicut, India. Pp. 312-318.
7. [D-space Website](#). <<http://dspace.org>>
8. [D-Space Technology](#) <<http://dspace.org/technology/system-docs/functional.html>>
9. Jones, Richard (2004). [The Tapir: Adding E-Theses Functionality to DSpace](#), *Ariadne*, Issue 41, October 2004 <<http://www.ariadne.ac.uk/issue41/jones/>>

10. Atkins, Antony, et al. (2004). ETD-MS: an Interoperability Metadata Standard for Electronic Theses and Dissertations. 1.00, 2.
<<http://www.ndltd.org/standards/metadata/current.html>>
11. TAPIR Website <http://www.thesesalive.ac.uk/dsp_home.shtml>
12. Edminster, Judith (2002). The diffusion of New Media Scholarship: Power Innovation and Resistance in Academe, Dissertation, University of South Florida.
<<http://etd.fcla.edu/SF/SFE0000035/etd.pdf>>