CHAPTER 4

E-THESIS REPOSITORIES: AN OVERVIEW

4.1 Introduction:

An ETD or Electronic Thesis & Dissertation is similar to traditional theses but more appropriate for worldwide access and retrieval with the help of computer, laptop, tablet, mobile phone etc. Depending on the nature of accessibility provided, an ETD can be accessed anytime, anywhere. It helps in increasing the visibility of the individual researcher as well as the institution involved in the research work. The Networked Digital Library of Theses & Dissertations (NDLTD) which is an international organization promotes the adoption, creation, use, dissemination and preservation of electronic theses and dissertations in various subjects and languages accessible through its union catalog. Since its inception, ETD activities have always been associated with enhancing the quality of education.

4.2 History of Electronic Thesis & Dissertation (ETD):

As mentioned by Edward Fox, Gail McMillan & V.Srinivasan¹, the history of ETD dates back to a Workshop in 1987 in Ann Arbor, Michigan, led by Nick Altair of University Microfilms International (UMI), who was involved in the Electronic Manuscript Project. The idea of marking up dissertations according to Standard Generalized Mark-up Language Document Type Definition (SGML DTD) was pursued by Virginia Tech working with SoftQuad, and led to perhaps the first ETD in 1988. Virginia Tech’s demonstration effort was initially funded by Gary Hooper of the Research Division. Still, making ETDs a common practice had to wait for improved technology support, among other things.

The project could not be developed for few years, until 1991, when Virginia Tech’s Dean Gary Hooper financed further critical development. Before the
release of the first version of PDF in 1992 and Adobe's Acrobat tool in 1993, a Virginia Tech team comprising of John Eaton (Graduate School), Edward Fox (Computer Science, Computing Center), and Gail McMillan (University Libraries) began discussions about ETDs with Adobe and worked with a pre-release version of their software. The Coalition for Networked Information (CNI) launched a discovery project guided by the Council of Graduate Schools, UMI, and Virginia Tech to explore ETDs. Ten universities in the United States and Canada sent representatives from their graduate programs, libraries, and computing/IT groups to an October 11, 1992, Design Meeting in Washington, D.C. This group agreed that working toward ETDs would be a reasonable initiative, so discussions continued at sessions of the CNI 1993 Spring Meeting held the following March in San Francisco. At the July 1993 meeting of the Monticello Electronic Library Project in Atlanta, participants from the Southeastern United States also discussed the role of ETDs in education and libraries.

Subsequently, the Virginia Tech library drafted a workflow description, developed and tested ETD submission scripts, and prepared a demonstration website in 1995. Following the faculty's recommendation from the Degree Requirements, Standards, Criteria, and Academic Policies Committee in 1996, graduate students were given the option to submit their theses and dissertations online. The library added incentives by waiving the Archiving Fee (previously the Binding Fee) for all who would submit ETDs in 1996. On January 1, 1997, ETDs became a requirement at Virginia Tech; West Virginia University followed in 1998 with its own requirement.

ETD initiatives have benefited from several funding sources. When the Southeastern University Research Association (SURA) sold SURAnet, some of the funds went to support a 1996-1997 project led by Eaton, Fox, and McMillan to launch ETD activities throughout the Southeastern United States (SURA 1997). Then the US Department of Education funded a 1996-1999 project led by these investigators to extend ETD efforts more broadly. The resulting National Digital Library of Theses and Dissertations quickly expanded into an international initiative, and was renamed the Networked Digital Library of Theses and Dissertations (NDLTD).
In 1998, the first ever ETD Conference was organised in Tennessee. Virginia Tech organised the second ETD Conference in 1999. The trend is still continued every year by different institutions hosting the conferences and focussing on a particular theme of ETD. The series of these conferences help in disseminating ideas and best practices in ETD.

4.3 Definition of Electronic Thesis & Dissertation (ETD):

ETD has been defined in various ways by national and international organisations, universities, institutions etc. explaining its characteristics, features, benefits and use.

1) Networked Digital Library of Theses & Dissertations (NDLTD) define ETD as:

"Electronic theses and dissertation (ETD) can be accessed on the web in full or partial-text."^2

2) According to The George Washington University (Washington DC),

"An ETD is a thesis/ dissertation created as an electronic document (or set of electronic documents). The electronic documents that make up an ETD can be created using any popular word processing software program. One can also use advanced software programs to produce multimedia animations and sounds for use in an electronic version of a thesis/dissertation."^3

3) Harrod’s Librarians’ Glossary and Reference Books defines ETD as,

"the digital version of a thesis or dissertation whose initiation is generally credited to Virginia Polytechnic Institute (Virginia Tech) who started creating ETDs in 1994."^4

4) The most comprehensive definition of ETD is given by Virginia Tech Digital Library and Archives where they define ETD by explaining its content, importance in candidate’s academic career, benefit of getting transformed from print to electronic format and it also highlights the aim of Virginia Tech
to make all the research work produced in the institute openly accessible to the public all over the world via Internet. According to Virginia Tech,

"A dissertation or thesis is a work submitted in support of candidature for a doctorate or master's degree, respectively which presents the author's research and findings. Electronic versions of theses and dissertations are called ETDs. An ETD 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, research methods used, the results, interpretation and discussion of the results, and a summary with conclusions. It provides a technologically advanced medium to express author's ideas. They are prepared using almost any word processor or document preparation system." \(^5\)

By taking into consideration various benefits of ETD, the researcher defines ETD as-

"ETD is electronic version of print thesis or dissertation that may include various file formats and can be produced using any popular word processing software. It facilitates dissemination of research through World Wide Web in an effective, expeditious and economic manner."

4.4 Benefits of ETD:

A thesis or dissertation is the document submitted by a candidate for an academic degree and it presents the research work performed by the researcher. Submitting thesis or dissertation in electronic format has further enhanced the usability of the document since it becomes readily available for dissemination through Internet. Some of the specific benefits of ETD\(^6,7\) are:

1) *Improved presentation of work:*

ETD can be created using any popular word processing software and it also enables the inclusion of audio-video files, images, spreadsheets, graphs, animations etc. This helps the researcher to express his/her ideas in a better way. Such facilities are not available with the print version.
2) Improved Accessibility and Increase in Citation Index:
ETD facilitates prompt dissemination of research work across the world through World Wide Web without the limitations of time and space. This helps in increasing the citation count of the researcher/author and increases prestige of the institution.

3) Increase control:
After submitting the ETD, the researcher can control the availability of his/her thesis or dissertation by choosing:
   a) To embargo (restriction in access for a specific period of time)
   b) To make it accessible full-text under open access.

4) Improves library services:
ETDs provide an economical solution for archiving and preservation of theses and dissertations which are one of the most important types of grey literature. Being in electronic format, it reduces the cost involved in the process of preservation, reduces need of library storage space thereby improving library services.

5) Lower Printing Costs:
Researcher can save a huge amount required for printing purposes if thesis or dissertations have to be submitted in electronic format.

4.5 Risks of ETD:

Jordan Lee and Brian Yandell (1999) of University of Wisconsin-Madison composed a ‘Working Draft on Risks and Benefits of Electronic Dissertations’. The risks of electronically submitting thesis or dissertations are-
1) This is a new system, and issues related to this are yet to be resolved regarding formats. Changes as standards are settled may affect submission and future access.
2) Immediate access can give competitive colleagues an edge. This can be overcome to some degree by delayed access.
3) Intellectual property rights: Electronic publishing might interfere with publication of chapters in some journals. This varies from journal to journal and discipline to discipline, and is in a great state of flux at this time. We believe this will not be an issue for electronic dissertations in the long run, but at present this is uncertain.

4) Electronic publishing can interfere with publication of a whole thesis as a book. Some publishing houses will not consider works that have appeared in any form on the web, even if it might undergo substantial revision before book publication.

5) Copyright issues about web-published material are not well understood.

6) Plagiarism on the Internet is a growing concern. Under widely accepted "fair use" principles, anyone can copy part of your dissertation as long as it is for noncommercial purposes. Ethically it is important to properly cite such work. However, it is possible with cut-and-paste technology to repackage large chunks of work as "new".

The apex body of higher education in India, University Grants Commission (UGC), in “Electronic Thesis Online (India) UGC (Submission of Metadata and Full-text of Doctoral Theses in Electronic Format) Regulations 2005” mention risk of plagiarism as one of the most important concerns for students and faculty. The regulations also mention how the electronic format of theses and dissertations has made plagiarism an easy task. The regulations also explain the working of plagiarism detecting software and how the use of such software packages is useful in detecting plagiarism in submitted ETDs.

4.6 ETD-MS: Interoperability Metadata Standard for Electronic Theses and Dissertations:

Metadata is the term used to describe data about data. According to NDLTD, for items in open access repositories, metadata usually consists of a full bibliographic reference, abstract, keywords and similar information about the related digital object(s) i.e. ETDs. UGC-ETD Guidelines mention that the following three types of metadata are associated with the digital objects:
• Descriptive Metadata: Include content or bibliographic description consisting of keywords and subject descriptors.

• Administrative or technical Metadata: Incorporates details on original source, date of creation, version of digital object, file format used, compression technology used, object relationship, etc. Administrative data may reside within or outside the digital object and is required for long-term collection management to ensure longevity of digital collection.

• Structural Metadata: Elements within digital objects that facilitate navigation, e.g. table of contents, index at issue level or volume level, page turning in an electronic book, etc.

Institutions dealing with electronic theses and dissertations have developed their own standards or adapted the existing metadata standards. NDLTD, the world’s largest consortium of ETD Universities and which maintains an international union catalogue of ETDs from around the world has defined a ‘Standard Set of Metadata Elements’ used to describe an electronic thesis or dissertation which is not a replacement for metadata scheme already in use for any university but should be used as a guideline to develop a cross-walk between local metadata standards and a single standard used for sharing information about ETDs. The set is based on the work of Dublin Core Metadata Initiative and the proposed set of qualifiers put forward by the DC Date Working Group.

4.7 Standards for Metadata Harvesting:

Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)- Harrod’s Librarian’s Glossary & Reference Books mentions, “OAI-PMH is the protocol which provides an application-independent interoperability framework for the Open Archives Initiative to enable the automatic gathering of metadata across multiple repositories.”

The OAI-PMH is a protocol devised to make machine-readable metadata widely available for use. The development of the OAI-PMH protocol took root in a
meeting that was convened in late 1999 at Santa Fe, New Mexico to address problems of the e-print world. As disciplinary e-print servers became more common, it was difficult to support searching across multiple repositories. Repositories needed greater capabilities to automatically identify and access papers that had been deposited in other repositories.

The Open Archives Initiative Protocol for Metadata Harvesting\textsuperscript{13} provides an application-independent interoperability framework based on metadata harvesting. There are two classes of participants in the OAI-PMH framework:

- \textit{Data Providers} administer systems that support the OAI-PMH as a means of exposing metadata; and
- \textit{Service Providers} use metadata harvested via the OAI-PMH as a basis for building value-added services.

The metadata stored in the data providers' database is transferred in bulk to the metadata database of the service providers. The transfer of metadata is done in a series of requests and responses between the data provider and the service provider/harvester. The OAI-PMH Protocol depends upon the HTTP-transaction framework for communication between a harvester and a repository. Requests may be made using either the HTTP GET or POST methods. All successful replies are encoded in XML, and all exception and flow-control replies are indicated by HTTP status codes.

### 4.8 File Formats for ETDs:

Virginia Tech\textsuperscript{14}, the first University to create ETDs have recommended a list of file formats for electronic theses and dissertations are-

- \textbf{Full-text of ETDs}

  1) \textbf{PDF(.pdf)-}

    PDF is an acronym of Portable Document Format and is most commonly used format for submitting full-text of ETDs. PDF is used to present documents in a manner independent of application software, hardware and operating
systems. PDF is best used to store vector-based graphics (i.e. graphics drawn using lines and curves rather than pixels).

- Additional texts/notices
  2) PPT(.ppt)-
  PPT or Power-Point Presentation may also be included in ETDs.
  3) JPEG(.jpeg or .jpg)-
  JPEG or Joint Photographic Experts Group is a commonly used method for glossy compression for digital images particularly produced by digital photography. The degree of compression can be adjusted taking into consideration the storage size and image quality.

- Images
  4) GIF(.gif)-
  Graphics Interchange Format or GIF is a bitmap image format that was introduced by CompuServ in 1987, has since come into widespread use on the WWW due to its wide support and portability. It is best used to store screen-quality images that do not contain many colours. GIF files are typically very small, but cannot reproduce the range of colours necessary to reproduce photographic images which JPEG can do.
  5) TIFF(.tif)-
  TIFF, an acronym for Tagged Information File Format is an archival format, meaning that it does not sacrifice image quality in order to reduce file sizes. The TIFF format is widely supported by image-manipulation applications, by publishing and page layout applications, and by scanning, faxing, word processing, optical character recognition and other applications. TIFF images are excellent for storing detailed, high quality images.
  6) PNG(.png)-
  Portable Network Graphics (PNG) is a bitmap file format that supports lossless data compression. It was created as an improved, non-patented replacement for GIF and is the most used loss less image compression format on the internet. Since it can store high color images, it can also be used for storing photographic content.
7) MPEG (.mpeg)-
Moving Picture Experts Group or MPEG is the oldest and most widely used supported format for movies. The MPEG format is most commonly used as an output format from UNIX utilities that generate video content.

8) QuickTime (.mov, .qt)-
QuickTime is an extensible multimedia framework developed by Apple Inc., capable of handling various formats of digital video, picture, sound, panoramic images, and interactivity. Support for QuickTime is good on Macintosh and Windows but not much on Unix.

• Video

9) AVI (.avi)-
Audio Video Interleave (AVI) files contain both audio and video data in a file container that allows synchronous audio-with-video playback. It is more of a Windows supported format and is not as well supported on other formats.

10) WAV (.wav)-
Waveform Audio File Format or WAVE is a default standard for Windows Sound Files. It is also supported by other platforms. It is mainly used for raw and uncompressed audio files.

11) AIFF (.aif)-
Audio Interchange File Format (AIFF) is an audio file format standard used for storing sound data for personal computers and other electronic audio devices. The AIFF format is a Macintosh-specific equivalent of the WAV format. It is not as well supported on all platforms as the WAV format.

• Audio

12) MPEG-3 (.mp3)-
MPEG-3 (or MP3) format eliminates sound data which is not as strongly perceived by the human ear and brain, and as such creates files of reasonable quality that are as much as 10 times smaller than the raw data itself. MP3 files are good for storing long passages of sound content where high quality is not required.
• Other file formats

13) Zip Files-
It is an archive file format that supports lossless data compression. A .zip file may contain one or more files or folders.

14) HTML-
Hyper Text Mark-up language (HTML) is the language with which Web pages are designed. This standard has been defined by the World Wide Web Consortium (W3C). HTML allows web documents to be created with ease. HTML is a plain text file and any text editor as simple as Notepad can be used to create HTML documents.

15) XML-
eXtensible Mark-up Language (XML) provides a structured representation of data that can be implemented broadly and is easy to deploy. XML is a subset of SGML (Standard Generalized Mark-up Language), modified and optimized for delivery over the Web. This standard has been defined by the World Wide Web Consortium (W3C). XML can be used to format and transfer data in an easy and consistent way. XML is also a plain text format.

4.9 Digital Preservation Strategies of ETDs:
A digital preservation strategy is a well-considered and documented approach to the preservation of digital objects.

4.9.1 Definition of Digital Preservation:
a) Harrod’s Librarian’s Glossary & Reference Book explain Digital Preservation as “method of keeping digital materials ‘alive’ so that they remain usable as technological advances render original hardware and software specifications obsolete.”

b) The Working Group on Defining Digital Preservation- Association of Library Collections & Technical Services (a Division of ALA) states “Digital preservation combines policies, strategies and actions to ensure the accurate rendering of authenticated contentover time, regardless of the challenges of media failure and technological change. Digital preservation applies to bothborn digital and reformatted content.”
4.9.2 Need of Digital Preservation:

Digital Preservation Europe Project\textsuperscript{17} mentioned the reasons for taking care of digital objects and need for digital preservation:

a) \textit{Storage media/data carrier problem:}

Digital objects are much more 'fragile' than traditional analogue documents such as books or other hard copy mediums. They are also much more venerable to physical damage. Digital objects require pro-active intervention to remain accessible.

b) \textit{Hardware obsolescence:}

After five years if a digital object and the disk are in perfect condition, also we have the software that can open the file, but if that file is on a disc your computer doesn't have a drive for, you will not be able to access it.

c) \textit{Software and format obsolescence problem:}

The software or file format can become obsolete for a number of reasons. Without the intervention of digital preservation techniques the information contained will no longer be accessible.

4.9.3 What does 'Long-Term' Digital Preservation mean?

Consultative Committee for Space Data Systems (CCSDS, 2002)\textsuperscript{18} explain the meaning of 'Long-term' in digital preservation as "a period of time long enough for there to be concern about the impacts of changing technologies, including support for new media and data formats, and of a changing user community, on the information being held in a repository. This period extends into the indefinite future."

4.9.4 Criteria for Selecting a Digital Preservation Strategy:

Various theories have been developed to preserve digital material which range from preserving the original technology on which the archival object
runs to preserving only significant properties of the object which are defined independently of any hardware or software platform.

Thibodeau\textsuperscript{19} suggests that a digital archive should take the following four criteria into consideration when selecting a preservation strategy:

a) \textit{Feasibility}: possession of hardware and software capable of implementing the chosen method.

b) \textit{Sustainability}: the method should be capable of being applied indefinitely into the future; or there should be another path which will offer a sequel to the method if it ceases being sustainable.

c) \textit{Practicality}: implementation should be within reasonable limits of difficulty and expense.

d) \textit{Appropriateness}: The chosen approach should be appropriate for the particular types of digital objects to be preserved and the objectives of their preservation.

\subsection*{4.9.5 Digital Preservation Strategies:}

Many digital preservation technologies have been proposed but a single technology is not suitable for all data types and institutions.

Tristram (2002)\textsuperscript{20} discussed following digital preservation techniques-

1) \textit{Bitstream Copying}:

It is more commonly known as "backing up your data" and refers to the process of making an exact duplicate of a digital object. It is often combined with remote storage so that the original and the copy are not subject to the same disastrous event. Bitstream copying should be considered the minimum maintenance strategy for even the most lightly valued, ephemeral data.
2) **Refreshing:**

Refreshing is the process of copying digital information from one long-term storage medium to another of the same type, with no change in the bitstream (e.g. from an older CD-RW to a new CD-RW). Refreshing is a necessary component of any successful digital preservation program, but is not itself a complete program.

3) **Durable/Persistent Media:**

Durable media has no impact on any other potential source of loss, including catastrophic physical loss, media obsolescence, as well as obsolescence of encoding and formatting schemes. It has the potential for endangering content by providing a false sense of security.

4) **Technology Preservation:**

This type of preservation is based on preserving the technical environment that runs the system, including operating systems, original application software, media drives, and the like. It is sometimes called the "computer museum" solution. This is not a strategy that an individual institution can implement. Maintaining obsolete technology in usable form requires a considerable investment in equipment and personnel.

5) **Digital Archaeology:**

Digital Archaeology includes methods and procedures to rescue content from damaged media or from obsolete or damaged hardware and software environments. It is explicitly an emergency recovery strategy and usually involves specialized techniques to recover bitstreams from media that has been rendered unreadable, either due to physical damage or hardware failure such as head crashes or magnetic tape crinkling.

6) **Analog Backups:**

It combines the conversion of digital objects into analog form with the use of durable analog media. An analog copy of a digital object can, in some respects, preserve its content and protect it from obsolescence, while
sacrificing any digital qualities, including sharability and lossless transferability.

7) **Migration:**
Migration is to copy data, or convert data, from one technology to another, whether hardware or software, preserving the essential characteristics of the data. Migration includes refreshing as a means of digital preservation but differs from it in the sense that it is not always possible to make an exact digital copy or replica of a data base or other information object as hardware and software change and still maintain the compatibility of the object with the new generation of technology.

8) **Replication:**
Replication is a term used to mean multiple things. Bitstream copying is a form of replication. Open Archival Information System (OAIS) considers replication to be a form of migration. LOCKSS (Lots of Copies Keeps Stuff Safe) is a consortial form of replication, while peer-to-peer data trading is an open, free-market form of replication.

9) **Reliance on Standards:**
It seeks a way to "harden" the encoding and formatting of digital objects by adhering to well-recognized standards and favouring such standards over more hard and less well-supported ones. It assumes in part that such standards will endure and that problems of compatibility resulting from the evolution of the computing environment (applications software, operating systems) will be handled by the continuing need to accommodate the standard within the new environment.

10) **Normalization:**
It is a formalized implementation of reliance on standards. Within an archival repository, all digital objects of a particular type (e.g., colour images, structured text) are converted into a single chosen file format that is thought to embody the best overall compromise amongst characteristics
such as functionality, longevity, and preservability. The advantages and disadvantages of reliance on standards also apply to normalization.

11) Canonicalization:
It is a technique designed to allow determination of whether the essential characteristics of a document have remained intact through a conversion from one format to another. Once created, this form could be used to algorithmically verify that a converted file has not lost any of its essence. Canonicalization has been postulated as an aid to integrity testing of file migration, but it has not been implemented.

12) Emulation:
It combines software and hardware to reproduce in all essential characteristics the performance of another computer of a different design, allowing programs or media designed for a particular environment to operate in a different, usually newer environment. Emulation requires the creation of emulators, programs that translate code and instructions from one computing environment so it can be properly executed in another.

13) Encapsulation:
It may be seen as a technique of grouping together a digital object and metadata necessary to provide access to that object. Appropriate types of metadata to encapsulate with a digital object include reference, representation, provenance, fixity and context information. Encapsulation is considered a key element of emulation.

14) Universal Virtual Computer:
It is a form of emulation. It requires the development of a computer program independent of any existing hardware or software that could simulate the basic architecture of every computer since the beginning, including memory, a sequence of registers, and rules for how to move information among them. Users could create and save digital files using the application software of their choice, but all files would also be backed up in a way that could be read by the universal computer.
4.10 Copyright & ETDs:

Copyright is one of the most important factors of ETD. Infringement of Copyright Act is one of the major barriers in submission of ETDs in repository. In spite of the benefits of ETD, due to the fear of getting their intellectual work plagiarised, researchers hesitate to submit their ETDs and contribute in the scholarly output of the institution through open access.

Harrod's Librarian's Glossary & Reference Book explains Copyright as “a procedure whereby the originator of a piece of intellectual property (book, article, piece of music etc.) acquires a series of rights over the work created, including copying, publishing, performing, broadcasting and adaptation.”

Ramaiah mentioned that:
1) The author of the ETD automatically holds the copyright. There is no need to transfer copyright.
2) Registering of Copyright is optional. It can be done yourself or through ProQuest (a fee is required in both cases)
3) The author gets cultural and legal benefits.
4) Use of a copyright statement in your ETD is not required, but is highly recommended.

According to the Office of Graduate & Professional Studies, Theses Office, Texas A&M University, any submitted thesis or dissertation is automatically copyrighted under U.S. Federal Law (United States Code, Title 17) as soon as the respective student saves a draft of it on paper or computer.

Stanford University Libraries & Academic Information Resources mention that “in copyright law, copying is known as ‘reproduction’, and it’s one of the exclusive rights of copyright owner. The right to publicly display a work is an exclusive right of copyright owner. While there is no general exemption for preservation activities in copyright law, there are exemptions that can help individuals and especially libraries and archives legally preserve expressive
works for the future. Furthermore, in the absence of a specific exemption, one can always consider fair use as a defense when making a preservation copy.”

For the present research study, the researcher became member of NDLTD Listserv member in order to get information and participate in discussions related to various topics of ETD. In order to get the authentic information regarding copyright policies for ETDs, the researcher posted the following query to the ETD mailing group (etd@ndltd.org) on 29th May 2014-

Dear Professionals,

Once a student submits a Doctoral Thesis or Master's Thesis, with whom does the copyright of the thesis remain and why?

a) with the researcher?

b) with the University?

Fig. 4.1: Researcher posted query to NDLTD Mailing Group

Three Repository administrators replied to the query.

i) Max Read (Associate Director, Student Academic Services, Office of the Dean and Vice-Provost, Graduate and Postdoctoral Studies, The University of British Columbia, Vancouver Campus) informed of researchers retaining the Copyright after submitting the ETD.

(Max Read, personal communication, May 29, 2014)
ii) **Rhonda Marker** (Head, Scholarly Communication Center, Repository Collection Librarian, Rutgers University Libraries, Rutgers, The State University of New Jersey) mentioned that “university's copyright policy says that the author holds copyright in his/her thesis or dissertation.”

Marker also provided the Rutgers University Copyright Policy\(^{25}\) which reaffirms the faculty’s rights to retain copyright ownership to the scholarly and artistic works they create, such as books, monographs, journal articles, musical compositions, and artwork, in whatever format they are created, print or electronic, without regard to the extent of university resources involved in the creation of these works. Students typically will own the copyright to works created as a requirement of their coursework, degree, or certificate program. The university, however, retains the right to use student works for pedagogical, scholarly, and administrative purposes.

(Rhonda Marker, personal communication, May 29, 2014)
Fig. 4.3: Response to researcher’s query by Rhonda Marker & Mohammed Butt

iii) **Mohammed Butt** (American University Library) mentions that “as soon a person creates a thing, in our case as soon as student creates a document, by law they have the copyright to that thing/doc. The copyright trademark symbol is not needed, though it is certainly helpful. ProQuest simply distributes the documents. It might be worth registering with the copyright office (copyright.gov) for legal benefits.

The responses received from the repository administrators clarify that once the ETD is submitted to the repository, the author of the ETD retains the copyright.

(Mohammed Butt, personal communication, June 6, 2014)

### 4.11 Types of Access Provided for ETDs:

Theses and dissertations are one of the most important forms of grey literature that provide access to the intellectual thoughts of the researcher. It is one of the major sources of scholarly communication. Print formats of theses and
dissertations are kept in closed access in the respective institution/university’s library due to which the research done by the researcher remains locked and unavailable for use. Submission of theses and dissertations in electronic format has made the dissemination of information possible outside the boundaries of institution/university libraries with the help of World Wide Web. However, in order to keep a check in the purpose and amount of use of ETDs, researchers/institutions/universities have developed different types of access categories of ETDs. The types of access as mentioned by UGC, India for ETD are as follows:

1) **Worldwide Access:** This access level allows immediate worldwide access to all ETDs, as soon as they are submitted. Onsite and offsite users can read and download the ETDs.

2) **Campus-only Access:** The access in this category is restricted only to a particular group of users’ on-campus. The access is controlled either by IP addresses, a valid ID, or through a proxy server.

3) **Temporary Restricted Access:** The ETD could be restricted for one, two or three years and then it is released for worldwide access, unless the University/Department authorities give an approval for delayed release.

4) **Mixed Access:** The student has the opportunity to release parts of the ETD but not the entire thesis or to publish two versions of a chapter, each with different access levels.

Peter Suber, one of the most famous supporters of open access mentions that ‘ETDs are the most invisible form of useful literature and the most useful form of invisible literature.’ Suber identified nine major reasons for providing mandatory open access to ETDs:

1. Nowadays most theses and dissertations are born digital. They are already ETDs even if the university only wants to deal with printouts.
2. ETDs are Phase One, royalty-free works of research literature. Their authors lose no revenue by consenting to OA.
3. ETDs are not formally published. Hence there are no publishers in the picture to resist or oppose OA. There are no publisher fears of lost revenue to answer. There are no publisher permissions to seek. There are no publisher negotiations to delay or deter OA archiving.

4. Mandates work and exhortations do not. This is the universal lesson from OA mandates to date, whether at funding agencies or universities.

5. OA solves the invisibility problem for ETDs. Without OA, there is almost no access, visibility, or indexing for dissertations. They are hard to retrieve even if discovered, and they are hard to discover. When they are OA, ETDs are not only searchable by cross archive search tools that index the ETD repositories; they are also indexed (in growing numbers but jerky stages) by Google, Yahoo, and Microsoft. Scirus already indexes the ETDs held by the Networked Digital Library of Theses and Dissertations (NDLTD).

6. Universities are in a good position to mandate OA. They can make it a simple condition of submission and acceptance.

7. Mandating OA for ETDs will educate the next generation of scholars about OA, when they do not already know about it.

8. An OA mandate will elicit better work.

9. Finally, an OA mandate shows that the university takes the dissertation seriously.

Suber strongly believed that OA for ETDs is one of the best ways to increase visibility and impact. He knew about the students fears that OA for ETDs will prevent future publication are understandable but groundless. According to Suber "there are easy work-arounds for ETDs describing patentable discoveries or containing chapters for which copyright has already been transferred to publishers."
4.12 ETD Initiatives at International Level:

4.12.1 Networked Digital Library of Theses and Dissertations (www.ndltd.org)\(^{28}\)

The Networked Digital Library of Theses and Dissertations (NDLTD) is an international organization dedicated to promoting the adoption, creation, use, dissemination, and preservation of electronic theses and dissertations (ETDs). They support electronic publishing and open access to scholarship in order to enhance the sharing of knowledge worldwide. The website includes resources for university administrators, librarians, faculty, students, and the general public and include topics like how to find, create, and preserve ETDs; how to set up an ETD program; legal and technical questions; and the latest news and research in the ETD community.

The National Digital Library of Theses and Dissertations was established in 1996, directed by an informal steering committee. As its scope became international, the organization kept the acronym NDLTD, but changed its name to the Networked Digital Library of Theses and Dissertations.

4.12.2 DART-Europe E-theses Portal (http://www.dart-europe.eu/)\(^{39}\)

DART-Europe is a partnership of research libraries and library consortia who are working together to improve global access to European research theses. DART-Europe is endorsed by LIBER (Ligue des Bibliothèques Européennes de Recherche), and it is the European Working Group of the Networked Digital Library of Theses and Dissertations (NDLTD).

The DART-Europe partners help to provide researchers with a single European Portal for the discovery of Electronic Theses and Dissertations (ETDs), and they participate in advocacy to influence future European e-theses developments. DART-Europe offers partners a European networking forum on ETD issues, and may provide the opportunity to submit collaborative funding applications to achieve DART-Europe's vision for ETDs.
4.12.3 DIVA (http://www.diva-portal.org/)

DiVA - Academic Archive On-line, is a publishing system for research and student theses and a digital archive for long-term preservation of publications. DiVA began its development in the year 2000 at Uppsala University Library. Today the technical development is driven by the EPC in co-operation with the participating universities and colleges. All universities and publicly financed research institutions both in Sweden and abroad are welcome to join DiVA in its co-operative effort. DiVA portal is a finding tool for research publications and student theses written at the following 34 universities and colleges of higher education. Publications found in DiVA portal can also be found in the local databases using a local interface.

4.12.4 PQDT OPEN-ProQuest Digital Theses (http://pqdtopen.proquest.com/)

It provides free of charge access to full text of open access theses and dissertations in PDF format.

4.13 ETD Initiatives: Country Specific

4.13.1 EThOS-Electronic Theses Online Service (http://ethos.bl.uk/)

EThOS is the United Kingdom’s national thesis service which aims to maximise the visibility and availability of the UK’s doctoral research theses. It demonstrates the quality of UK research, and supports the UK Government’s open access principle that publications resulting from publicly-funded research should be made freely available for all researchers, providing opportunities for further research. EThOS helps institutions to meet the expectation of the UK Research Councils according to whom, PhDs supported by a Research Council Training Grant should be made freely available in an open access repository. There are approximately 350,000 records relating to theses awarded by over 120 institutions. Around 120,000
of these also provide access to the full text thesis, either via download from the EThOS database or via links to the institution's own repository. Of the remaining 250,000 records dating back to at least 1800, three quarters are available to be ordered for scanning through the EThOS digitisation-on-demand facility.

4.13.2 Theses Canada ([http://www.collectionscanada.gc.ca/](http://www.collectionscanada.gc.ca/))

Canadian universities participate in the program voluntarily by submitting approved theses and dissertation to Theses Canada. Library and Archives Canada (LAC) staff provide advice to Canadian universities that are establishing electronic theses and dissertations (ETD) programs and welcome new electronic participants. Library & Archives Canada (LAC) collection includes two formats, electronic theses and theses on microfiches. The electronic theses and dissertations on this site are for the personal use of students, scholars and the public. Any commercial use, publication or lending of them in libraries is strictly prohibited.


This program began in 1998 as a project funded by the ARC, transferred to Council of Australian University Librarians (CAUL) in July, 2001, became fully supported by CAUL members in 2003 and CONZUL members from 2006. The ADT officially became the Australasian Digital Theses Program in September, 2005, when CAUL adopted the revised business plan which includes the New Zealand Universities as full members. The Australasian Digital Theses (ADT) Program ceased operation on 28 March, 2011. The database server has been decommissioned, and the content of that database is accessible from the National Library of Australia’s Trove service. Each participating university will continue to host their own digital theses and house their own print and other non-digital theses.
4.13.4 DissOnline & Online Dissertations at the German National Library
(http://www.dnb.de/EN/Wir/Kooperation/dissonline/dissonline_node.html)\textsuperscript{35}

The German National Library has the largest national collection of online dissertations in Europe. Since 1998, it holds the German National Library online dissertations and post-doctoral theses; the stock has since grown to more than 120,000 documents.

4.13.5 NARCIS (http://www.narcis.nl/)\textsuperscript{36}

National Academic Research and Collaborations Information System (NARCIS) is the main national portal for those looking for information about researchers and their work. Besides researchers, NARCIS is also used by students, journalists and people working in educational and government institutions as well as the business sector. NARCIS provides access to scientific information, including (open access) publications from the repositories of all the Dutch universities, KNAW, NWO and a number of research institutes, datasets from some data archives as well as descriptions of research projects, researchers and research institutes.

4.13.6 National ETD Portal-South African theses and dissertations
(http://www.netd.ac.za/)\textsuperscript{37}

This site is run by the University of Cape Town UCT-CS Digital Libraries Laboratory on behalf of the NDLTD. Metadata for all theses and dissertations produced internationally are collected and made accessible, as well as disseminated further, from this site.

4.13.7 China Doctoral/Master Dissertation Database (http://www.cnki.net/)\textsuperscript{38}

The database has collected dissertations covering basic science, engineering technologies, agriculture, medical science, philosophy, humanity, social sciences, and so on. By June 2011, the database had collected more than 1,50,000 excellent domestic doctoral dissertations from 397 institutions and 11,50,000 excellent domestic master dissertations from 598 institutions.
4.13.8 DATAD- Database of African Theses & Dissertations (www.aau.org)³⁹

African research results are rarely indexed in major international databases, a problem that is further exacerbated by the inaccessibility of theses and dissertations completed in the region, many of which contain local empirical data that is not available in international literature. The Association of African Universities (AAU) found it necessary to initiate and support efforts towards putting Africa’s research output onto the mainstream of world knowledge. The idea of DATAD was born out of a project in 2000 as a result of a positive recommendation of a feasibility study carried out for a pilot project to index, abstract, and distribute theses and dissertations completed in African universities. The Database of African Theses and Dissertations (DATAD) is a programme to improve management and access to African scholarly work. Theses and dissertations represent a significant proportion of Africa’s research activity.

4.14 ETD Initiatives in India (Full-text):

4.14.1 Shodhganga (http://shodhganga.inflibnet.ac.in)⁴⁰

"Shodhganga" is the name coined to denote digital repository of Indian Electronic Theses and Dissertations set-up by the INFLIBNET Centre. The word "Shodh" originates from Sanskrit and stands for research and discovery. The "Ganga" is the holiest, largest and longest of all rivers in Indian subcontinent. The Ganga is the symbol of India’s age-long culture and civilisation, everchanging, ever-flowing, ever-loved and revered by its people, and has held India’s heart captive and drawn uncounted millions to her banks since the dawn of history. Shodhganga stands for the reservoir of Indian intellectual output stored in a repository hosted and maintained by the INFLIBNET Centre.
The Shodhganga@INFLIBNET is set-up using open source digital repository software called DSpace. The DSpace uses internationally recognized protocols and interoperability standards. Shodhganga provides a platform for research scholars to deposit their Ph.D. theses and make it available to the entire scholarly community in open access. The repository has the ability to capture, index, store, disseminate and preserve ETDs (Electronic Theses and Dissertations) submitted by the researchers. At present 182 Universities have signed MoU with Shodhganga and 18143 theses are uploaded.

ShodhGangotri is the Repository of Research in Progress in India in various universities in variety of subjects. Shodhganga is an extension to ShodhGangotri.

4.14.2 Vidyanidhi Digital Library of Indian Doctoral Dissertations
(http://dspace.vidyanidhi.org.in:8080/dspace)\textsuperscript{41}

Vidyanidhi (Meaning 'Treasure of Knowledge' in Sanskrit) is India's premier Digital library initiative to facilitate the creation, archiving and accessing of doctoral theses. This project was initiated in the year 2000 at the Department of Library & Information Science, University of Mysore, Manasagangotri, Mysore sponsored by the National Information System for Science and technology (NISSAT), Department of Scientific & Industrial Research, Ministry of Science & technology, Government of India. The initiative is supported and founded by The Ford Foundation and Microsoft Foundation. It also has a strategic support from the University Grants Commission (UGC).

4.14.3 CSIR Explorations (http://eprints.csirexplorations.com/)\textsuperscript{42}

CSIR Explorations is a digital library of electronic dissertations, theses and research reports of the fellowships and projects supported by CSIR. URDIP is working in close association with the Human Resource Development Group of CSIR and National Information System for Science and Technology (NISSAT) of DSIR to put this database of basic research in the public domain on the World Wide Web It is developed using EPrints.
e-Thesis is a database of doctoral theses submitted by CSIR fellows. It is the only one central authoritative source for information about CSIR doctoral dissertations. The potential users of this database include doctoral candidates who would like to make sure that their proposed thesis topics have not already been written about. Students, faculty and other researchers may search it for titles related to their scholarly interests. The database represents the work of authors from over 250 national laboratories, universities and research institutes in India.

4.14.4 **KrishiPrabha** ([http://krishikosh.egranth.ac.in/](http://krishikosh.egranth.ac.in/))

KrishiPrabha is a full-text electronic database of Indian Agricultural Doctoral Dissertations submitted by research scholars to the 36 State/Deemed Agricultural Universities during the period from 1.1.2000 to 31.12.2006. This database has been created by Nehru library, Ch. Charan Singh Haryana Agricultural University, Hisar with financial support from Indian Council of Agricultural Research, New Delhi under its National Agricultural Innovation Project.

4.14.5 **Librarians’ Digital Library** ([http://drtc.isibang.ac.in/](http://drtc.isibang.ac.in/))

Librarians’ Digital Library is a repository where any digital resource related to Library and Information Science domain can be archived by anyone across the world. The site is not updated anymore.

4.14.6 **DELNET Database of Theses and Dissertations**

A database of Theses and Dissertations submitted to Indian Universities has been started, which covers various subjects. The database has 70,293 records. Access is provided only to DELNET members.
4.14.7 NASSDOC Database of Ph.D Dissertations

National Social Science Documentation Centre (NASSDOC) was established in 1969 as a division of Indian Council of Social Science Research (ICSSR). It provides library and information support services to researchers in social science. It provides access to bibliographic details of 5000 dissertations in Social Sciences.

4.15 ETD Directories/ Registries:

4.15.1 ROAR-Registry of Open Access Repositories (http://roar.eprints.org)

The aim of ROAR is to promote the development of open access by providing timely information about the growth and status of repositories throughout the world. Open access to research maximises research access and thereby also research impact, making research more productive and effective. ROAR is hosted at the University of Southampton, UK and is made possible by funding from the JISC. ROAR is part of the EPrints.org network. Presently it holds information about 3793 repositories across the globe in 12 types of content. It provides information about 284 repositories containing ETDs.

4.15.2 OpenDOAR (www.opendoar.org)

The Directory of Open Access Repositories (OpenDOAR) is an authoritative directory of academic open access repositories. OpenDOAR is one of the SHERPA Services including RoMEO and JULIET, run by the Centre for Research Communications (CRC). Current development work is currently funded by JISC, with contributions from the CRC host organisation, the University of Nottingham. Repositories need to be categorised with clear information on their policies regarding tagging peer-reviewed/non peer-reviewed material, their subject coverage, the constituency they draw on for content, their collection and preservation etc. Presently it provides
information about 2700 repositories from the world in 12 types of content. It provides information about 1474 repositories containing ETDs.

4.15.3 OAIster (www.oaister.org)

OAIster began at the University of Michigan in 2002 funded by a grant from the Andrew W. Mellon Foundation and with the purpose of establishing a retrieval service for publicly available digital library resources provided by the research library community. During its tenure at the University of Michigan, OAIster grew to become one of the largest aggregations of records pointing to open access collections in the world.

In 2009, OCLC formed a partnership with the University of Michigan to provide continued access to open access collections aggregated in OAIster. Since OCLC began managing OAIster, it has grown to include over 30 million records contributed by over 1,500 organizations. OCLC is evolving OAIster to a model of self-service contribution for all open access digital repositories to ensure the long-term sustainability of this rich collection of open access materials. OAIster data is harvested from Open Archives Initiative (OAI)-compliant digital libraries, institutional repositories and online journals using the self-service WorldCat Digital Collection Gateway. OAIster.worldcat.orgis a freely accessible OCLC site for searching the millions of OAIster records alone.

4.15.4 BASE-Bielefeld Academic Search Engine (http://www.base-search.net/)

BASE is one of the world's most voluminous search engines especially for academic open access web resources. BASE is operated by Bielefeld University Library. As the open access movement grows and prospers, more and more repository servers come into being which use the "Open Archives Initiative Protocol for Metadata Harvesting" (OAI-PMH) for providing their contents. BASE collects, normalises, and indexes these data. BASE provides more than 60 million documents from more than 3,000 sources. It provides access to the full texts of about 70% of the indexed documents. The index
is continuously enhanced by integrating further OAI sources as well as local sources. Our OAI-PMH Blog communicates information related to harvesting and aggregating activities performed for BASE. BASE is a registered OAI service provider and contributed to the European project "Digital Repository Infrastructure Vision for European Research" (DRIVER). Database managers can integrate the BASE index into your own local infrastructure (e.g. meta search engines, library catalogues) via an interface.

4.15.5 ScientificCommons (http://en.scientificcommons.org/)

ScientificCommons is a project of the University of St. Gallen Institute for Media and Communications Management. The major aim of the project is to develop the world’s largest archive of scientific knowledge with full-texts freely accessible to the public. It includes a search engine for publications and author profiles. It also allows the user to turn searches into customized RSS Feeds of new publications. ScientificCommons also provides a full-text caching service for researchers. Since the beginning of 2013, ScientificCommons has been inaccessible.

4.16 UNESCO’s Contribution towards ETD:

The UNESCO Guide to Electronic Theses and Dissertation promotes the sharing of knowledge available in universities across the world. The Guide is specifically designed by academic researchers. The Guide is an international, "living" document, written by ETD scholars throughout the world. It is published and mainly funded by UNESCO. It was prepared by an international team of faculty and staff; coordinated by Shalini Urs and edited by Joseph Moxley. This work is a living document that will continue to be updated in connection with the work of the Networked Digital Library of Theses and Dissertations; It was born as a result of the support provided by UNESCO in grants given to Virginia Tech, USA and the University of Montreal.
Participants in the project are from the Australian Digital Theses Project, the Humboldt-University Berlin, Germany, the Ibero-American Science Technology Education Consortium (ISTEC), the National Library of Portugal, the Digital Library of Indian Electronic Theses (VIDYANIDHI), the Virginia Tech University, USA; the University of Chile, the University of Lyon, France, and the University of Montreal, Canada. At present it is available in English and Spanish. French, Russian and German versions are under preparation due to presence of various translators funded in part by UNESCO, and others volunteering their assistance.

The UNESCO ETDs Guide aims to help all those interested in projects and programmes involving ETDs. Its goal is of aiding students of every university in creating electronic documents and using digital libraries. It has particular focus on the emerging genre of ETDs, which should enhance the quality, content, form, and impact of scholarly communication that involves students engaged in research. The guidelines given intend to help universities to develop their local infrastructure, especially regarding electronic publishing and digital libraries, which in turn build upon networking, computing, multimedia, and related technologies. The various sections of the Guides are aimed to address the needs of universities (including administrators and faculty), students (including those who wish to create ETDs as well as those who wish to make use of already-created works), and those involved in training or setting up ETD projects or programmes.

4.17 How to Set Up an ETD Program

NDLTD has provided the most comprehensive and Model Guidelines to Set up an ETD Program. 53

Step 1: Assemble the team
Set up a project team with representatives from the library, the IT department, the Faculty of Graduate Studies, senior administrators from
the university, and the Graduate Student Association. It is important to have cross-institutional representation in order to develop a balanced initiative.

**Step 2: Prepare the Proposal**

Prepare a pilot project proposal for consideration by the appropriate university administrators. This is the stage during which you should work out policies specific to your university.

You may want to include information about ETDs providing increased exposure to the university and graduate students' research and scholarship, publication potential, intellectual property and rights management, plagiarism, orientation and training, standards, costs, restrictions on access, archiving and preservation.

**Step 3: Decide on Software and Metadata**

This is also the time to decide which technical infrastructure you plan to use. These days most universities are implementing institutional repositories (IRs), which include ETD collections. There are any number of choices of institutional repository software available, both open source and proprietary. Some examples of open source systems include: Archimede, DSpace, Eprints, Fedora, and ETD-db. The latter is specifically designed for theses. Examples of proprietary systems include: Digital Commons, CONTENTdm, DigiTool, Open Repository, and Vital. Smaller institutions that have no IR may choose a remote-host option, such as Open Repository through BioMed Central. Digital Commons provides a total beginning-to-end submission package. Other non-IR options include: the ETD Administrator from ProQuest/UMI and VALET for ETDs from VTLS. The NDLTD highly recommends use of ETD-MS, the metadata standard specifically for electronic theses and dissertations. This standard includes minimal basic descriptive information related to an ETD. Other popular metadata schemes include Dublin Core
and MARC.

Step 4: Create an ETD Website
Once the project is approved, set up a website for ETDs. This can be done by the IT staff at your university or by the university library or graduate school. Information on the websites should include an overview of your ETD program, submission guidelines, ETD procedures, policies, information on tutorials etc.

Step 5: Train Graduate Students
Implement an orientation and training program for graduate students. They may need training on all aspects of ETDs, including creating the word document, application of styles, use of templates, conversion to PDF, and submitting online to the institutional repository. Some universities offer in-person training sessions and/or online training resources. Ohio State University has a tutorial on Creating a PDF. West Virginia University also has a Convert to PDF tutorial. The University of Victoria, Canada, has developed the online tutorial How to Submit an ETD.

Step 6: Set Up the ETD Workflow
Establish the ETD workflow, deciding who is responsible for the various stages of the process from submission to approval, preservation, and access. In the most common scenario students upload their ETD files to the Graduate School office where they are reviewed, approved, and released to the Library for preservation and access. The metadata is then made available for harvesting by other organizations, such as Google Scholar or the NDLTD.

Step 7: Establish a Pilot Project
Run a pilot project either with a limited number of students or with one or two specific departments for a semester or another limited time period. This will allow you to fine-tune your procedures and workflow. As few as
20 or 30 ETDs are sufficient to run a pilot project. At the end of their pilot projects some universities adopt a voluntary e-theses submission model for a set time period before moving to mandatory submission of electronic theses and dissertations. If possible, the best practice is to recommend mandating ETDs at your institution as soon as you begin your ETD submission program (by decree by the Provost or Graduate School Dean, on recommendation of faculty governance).

**Step 8: Prepare for OAI Harvesting**

Once your ETD submission program is established, have your IT staff prepare your ETD collection for Open Access Initiative (OAI) harvesting. The IT staff should read the Technical Requirements in order to implement the IR as an OAI data repository. Some systems have a built-in OAI interface. An open source version called OAIcat from OCLC is available online.

**Step 9: Evaluate and Enhance the Program**

Like any program, your institution’s ETD submission program should be periodically evaluated and enhanced.

- Software should be upgraded.
- Graduate School staff should always be aware of policy changes and how copyright impacts the program.
- IT staff must monitor appropriate listservs in order to become aware of upgrades, enhancements, and security issues.
- Library staff can implement enhancements to the program, such as creating catalogue records automatically from the ETD-MS metadata.
- Administrators should be aware of changing standards (e.g., PDF/A) and legislation (e.g., copyright, privacy issues, etc.).
- To stay informed about ETDs, your institution may want to become a member of NDLTD, subscribe to the ETD-L and network with NDLTD in Facebook.
4.18 ETD Terms and Definitions:

Some of the most commonly used terms and their definitions related to electronic thesis and dissertations are given below:

1) Access Type:
The manner in which an ETD is made available (or not) to the university community and/or public.

2) Archiving:
The application of processes, techniques and protocols to preserve the scholarly record over time.

3) Availability:
Level of access to which an ETD is available to the general public or in the institutional repository.

4) Born Digital:
An item is born-digital if it has been generated entirely electronically by using a word-processor and/or electronic composition software. For electronic files which have been produced by scanning.

5) Copyright:
A form of protection grounded in the U.S. Constitution and granted by law for original works of authorship fixed in a tangible medium of expression. Copyright covers both published and unpublished works. As intellectual property law, copyright protects original works of authorship including literary, dramatic, musical, and artistic works, such as poetry, novels, movies, songs, computer software, and architecture.

6) Closed Access:
The full text and sometimes the metadata of closed access ETDs are only available to authorized members of University staff and external examiners for administrative purposes. This is also known as an
7) **Creative Commons:**
A non-profit corporation dedicated to making it easier for people to share and build upon the work of others, consistent with the rules of copyright. They provide free licenses and other legal tools to mark creative work with the freedom the creator wants it to carry, so others can share, remix, use commercially, or any combination there of.

8) **Digital Library:**
A managed collection of information, with associated services, where the information is stored in digital formats and accessible over a network.

9) **Digital Preservation:**
The management process of ensuring digital objects and information are accessible over the long term. Development of standards, format compatibility, format migration and systems interoperability are important aspects of this process. Digital preservation systems are under development to provide appropriate digital preservation techniques.

10) **Dissertation:**
A research document written by a post-Master, Doctoral-level student. The term ‘dissertation’ may be used in some countries or universities to mean a post-baccalaureate, Master-level research project.

11) **Dissemination:**
Methods of distributing electronic documents on the Web.

12) **Embargo:**
Restricting access to an electronic document for a specific period of time. Also, called Publication Delay.
13) **EPrint:**

Any version of a work available online which has been either submitted for formal publication or has been accepted after formal review. The term encompasses both preprint and post-print.

14) **Fair Use:**

One of the more important copyright limitations is the doctrine of ‘fair use.’ The doctrine of fair use has developed through a substantial number of court decisions over the years and has been codified in section 107 of the copyright law. Section 107 contains a list of the various purposes for which the reproduction of a particular work may be considered fair, such as criticism, comment, news reporting, teaching, scholarship, and research. Section 107 also sets out four factors to be considered in determining whether or not a particular use is fair:

1. The purpose and character of the use, including whether such use is of commercial nature or is for non-profit educational purposes
2. The nature of the copyrighted work
3. The amount and substantiality of the portion used in relation to the copyrighted work as a whole
4. The effect of the use upon the potential market for, or value of, the copyrighted work

The distinction between fair use and infringement may be unclear and not easily defined. There is no specific number of words, lines, or notes that may safely be taken without permission. Acknowledging the source of the copyrighted material does not substitute for obtaining permission.

15) **Infringement:**

Misuse of copyrighted material.

16) **Institutional Repository:**

An online database that provides access to digital collections such as
theses and dissertations for online viewing and provides the associated metadata regarding the documents (e.g. student and university name, year of graduation, document title, abstract, keywords). A type of digital repository designed to collect the work of a particular institution.

17) Intellectual Property Rights:
Intellectual property (IP) refers to creations of the mind: inventions, literary and artistic works, and symbols, names, images, and designs used in commerce. IP is divided into two categories: Industrial property, which includes inventions (patents), trademarks, industrial designs, and geographic indications of source; and Copyright, which includes literary and artistic works such as novels, poems and plays, films, musical works, artistic works such as drawings, paintings, photographs and sculptures, and architectural designs. Rights related to copyright include those of performing artists in their performances, producers of phonograms in their recordings, and those of broadcasters in their radio and television programs.

18) Keywords:
Search terms or phrases relating to the item/body of work. Keywords are often words or phrases from the document title and/or abstract, are discipline-specific and which provide topic/subject search terms for online discovery.

19) LOCKSS:
Lots of Copies Keeps Stuff Safe. Originally developed the Stanford University Libraries, this digital preservation methodology allows inter-institutional sharing of archival collections. The NDLTD participates in the MetaArchive cooperative, a consortium of LOCKSS institutional partners.
20) **Mandate:**
Required ETD submission policy. Often the University President, Provost, or Graduate School Dean may implement such a policy to launch an ETD program at their institution.

21) **Metadata:**
Data that describes other data. For items in open access repositories, this usually consists of a full bibliographic reference, abstract, keywords, and similar information about the related digital object(s) (i.e. ETDs).

22) **Open Access:**
Information readily available on the Web at no cost and without access restrictions.

23) **Open Access Initiative:**
The OAI develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content. Its major contribution is the OAI Protocol for Metadata Harvesting (OAI-PMH), a set of guidelines that enable repositories to expose the metadata describing their content to service providers who harvest the metadata into large aggregations.

24) **OAI-PMH:**

25) **OpenDOAR:**
The Directory of Open Access Repositories. Allows users to search for specific repositories, or search within the contents of all the repositories in the directory.
26) Plagiarism:
The use or close imitation of the language and thoughts of another author and the representation of them as one's own original work. Within academia, plagiarism by students, professors, or researchers is considered academic dishonesty or academic fraud, and offenders are subject to academic censure, up to and including expulsion. The development of the Internet, where articles appear as electronic text, has made the physical act of copying the work of others as well as the detection of plagiarism much easier.

27) Preservation:
The management process of ensuring printed and digital objects and information are accessible over the long term.

28) Print-only Access:
Theses and dissertations are in paper and are placed on library shelves. Secure method to assure few people will ever view theses and dissertations. ETD is accessible only from library shelves or Interlibrary Loan. May also apply to some ETD programs where an indefinite campus restriction is allowed. Although ETD submission may be required, access may be limited to login for authorized university community or by interlibrary loan request by print distribution.

29) Restricted Access:
For ETDs this generally signifies that the complete work or aspects of the work will not be accessible to the public for a specific period of time. During this time the ETD may or may not be available to the university community, although the metadata is generally available to the public. This term may used to refer to ETDs that are available to a limited population as well as ETDs where access is embargoed.
30) **RoMEO Project:**
A project that defined the archiving policies of publishers. Now part of SHERPA. You will see publishers defined as having a Romeo color of white, yellow, blue, and green, which mean:
- White: archiving not formally supported
- Yellow: can archive preprint (i.e., re-refereeing)
- Blue: can archive post-print (i.e., final draft post-refereeing)
- Green: can archive preprint and post-print

31) **Scholarly Communications:**
The creation, transformation, dissemination and preservation of knowledge related to teaching, research and scholarly endeavours. Among the many scholarly communications issues include author rights, the economics of scholarly resources, new models of publishing including open access, institutional repositories, rights and access to federally funded research, and preservation of intellectual assets.

32) **Self-archiving:**
Placing a copy of your work in a digital/institutional repository or professional or departmental website.

33) **Sequestered:**
The act of removing, separating or seizing anything from the possession of its owner under process of law for the benefit of creditors or the state. For ETDs, this is usually an ‘embargoed’ or ‘No Access’ condition, typically requested for security purposes.

34) **SPARC:**
The Scholarly Publishing and Academic Resources Coalition is an international alliance of academic and research libraries working to correct imbalances in the scholarly publishing system.
35) **SWORD:**

A repository deposit protocol; in other words, a way to get items into the repository. Funded and supported by the Joint Information Systems Committee (JISC), SWORD allows you deposit resources (like electronic theses and dissertations) into repositories powered by platforms such as DSpace, Eprints, Fedora, IntraLibrary, and Zentity.

36) **University-only Access:**

The full text of university-only ETDs are only available to authorized members of University students, faculty and staff, or in some situations access may be restricted within a State or consortium arrangement of member schools via login or IP restriction. Many universities allow interlibrary loan service to provide print and/or electronic versions as requested from other academic libraries.

37) **Withheld:**

Theses and dissertations are withheld from the public for a certain period of time or indefinitely.

**4.19 Summary:**

The present chapter gives the overview of Electronic theses and Dissertations by explaining various concepts related to ETDs.

The chapter starts with the sneak preview of historical events which led to the production of first E-thesis in 1992 by Virginia Tech. The researcher discussed various definitions of ETDs given by various international organizations and reference source. A comprehensive definition of ETD is also stated by the researcher.

The chapter then proceeds to benefits and risks involved in submitting theses or dissertation in electronic format. The researcher mentions NDLTDs
‘Standard Set of Metadata Elements’ which can be used as a standard for sharing information about ETDs by the institutions.

Various file formats included in electronic format of theses and dissertations is also discussed. PDF is the exclusively used format for submitting full-text of ETDs. But there are various other formats used for submitting additional files. The next section explains very important concept for ETDs related to Preservation and Copyright of ETDs. The researcher discussed definition of digital preservation, need, concept of long term preservation, criteria for selecting most appropriate digital preservation strategy for an institution, benefits and various traditional and modern strategies of preservation used for archiving and preserving the ETDs. Views of other authors and Copyright Act are presented in this section.

The discussion now proceeds towards brief introduction of some fully operational ETD Projects at International Level, Specific Country related ETDs, Full-text ETD Projects in India and various Directories/Registries which provide list of E-theses Repositories.

The concluding section presents NDLTDs Guide to Set up ETD Program in an Institution. Also, there are numerous technical terms involved in ETD Program. The researcher listed some of the most commonly used terms and provided definitions. The terms are selected from USETDA and NDLTDs ETD Terms and Definitions. Being pioneering international organization, NDLTD explained ETD from various angles.

The next chapter analyses the data collected with the help of responses received to the web survey questionnaire sent through e-mail to various E-theses repositories across the globe.
References:


44. Librarians’ Digital Library. Retrieved July 18, 2014 from http://drtc.isibang.ac.in


