Electronic Theses and Dissertations’ (ETDs) Repositories: A Case Study of Maharshi Dayanand University, Rohtak, Haryana (India)

by

Sunder Singh
Information Scientist, Vivekananda Library, M.D. University, Rohtak, Haryana - 124001, India
OCRiD 0000-0003-3420-2930
Email: sunderstanwar@gmail.com
Contact Nos. +911262393002 (M) +91 9467820821
M.Lib.Sc., B.Lib.I.Sc., MCA, UGC-NET and Pursuing Ph.D. in Library and Information Science

and

Neeraj Tanwar
M. D. University, Rohtak- 124001 India
Abstract:

A steep escalation is observed in Electronic Thesis and Dissertations’ (ETDs) Repositories throughout the world during the last few years. This increase in ETDs is transforming the way of learning, research and scholarly communication in academic institutions and their scholars. ETDs underpins research and national development, hence, there was a need to setup an ETD at our institutional level. The transformation from print to electronic theses and dissertations has brought new attention to the researchers. Theses and dissertations are considered original research works in every university. The quality of research is also a key criterion while measuring the repute of any academic institution. This paper focuses on the practical problems faced during implementation of this project. We faced many challenges to implement this project but we took it as an opportunity. It highlights techniques to seek knowledge using various software packages. The authors highlight the experiences in developing ETDs and experiments with knowledge discovery software packages. Furthermore, the paper explores the extent to which academic libraries are grappling with the emerging genres of ETDs, for example, the use of linked data to enhance discoverability. The paper also make recommendations to implement the ETDs to enhance effective utilization of ETDs and knowledge discovery.

This paper skilfully explores the workflow of making inventory, selecting the theses for digitization, unbinding, scanning, OCRing, formatting, quality checking, converting to secured PDF-A format, binding of theses, uploading into IR(Institutional Repository) and Shodhganga. The success of this innovative project has helped in making the mandatory submission of the Ph.D. thesis possible to the Vivekananda Library at MDU, Rohtak in desired PDF or MS-Word format to protect this important research treasure of the University.

**Keywords:** Digitization, Electronic Theses and Dissertations, ETDs, Shodhganga, Vivekananda Library, Institutional Repository

**Introduction:**

Emergence of the digital libraries has rattled the traditions of library and information science. The key elements in reforming the traditional library customs are growth of internet, information and communication technologies and digitization of documents. Digital libraries, institutional repositories, digital archives and learning objects are developed as the results of cutting-edge research in the area of information communication and technology for advanced information storage and retrieval systems and can store different types of digital information. These are revolutionizing the concept of libraries in a way that any type of information available can be accessed from anywhere in the world at any point of time.

Digitization refers to the process of conversion of a physical object such as printed text, manuscript, image, sound or any other format to a digital format so that it can be stored and accessed digitally by using a computer system. An “electronic photograph” (Parekh, 2001) can be generated by using a scanner or a digital camera. Along with this digital photograph a significant amount of information is also added to ease its access and to optimize its usage. The idea behind digitization which addresses needs of libraries is preservation of documents in library and enhancing their discoverability, accessibility and reusability(Tuna, Zogo, & Demirelli, 2013).
Initially, digital libraries were developed to preserve the documents of colossal cultural importance for human civilizations such as rare books, documents, archival records, manuscripts, paintings etc. But now, along with providing a sheath to such rare documents in digital form, digital libraries embrace the “born digital” documents as well. Digital material can be categorized in two types, viz., Digitized and Born Digital. One type of material that is initially produced in the print format such as printed books, manuscripts etc. and then by using digitization process it is converted into digital format. Another is “born digital” i.e. the material which is produced in machine readable form such as e-books, e-journals, multimedia etc.

Need and Purpose of Digitization:

- First goal of theses digitization is to preserve the research outputs in digital format and easy access of the valuable contents. Digitization provides sheathing to the rare and expensive material which is probable to damage if accessed and used regularly. So, the restricted access material can be accessed by anyone.
- Digitization allows easy searching and access to the library material. Digital material is stored along with some additional bibliographic information for indexing which helps users in locating the material in library.
- It provides a multi-user environment in the libraries. A single copy of digital material can be accessed by multiple users at the same time without any interruption.
- It solves the problem of travelling to the library in search of material. Users can access digital material through web interface of the digital library without physically travelling to the library.
- It solves the problem of limited library hours in a day i.e. unlike a traditional library; a digital library can be open to users every time they need to consult it.

Indian Scenario:

Librarians are participating in introducing new advancements from Information Technology because of its influence in society and systematic nature. Use of information technology in our field of library and information science has pioneered new challenges for libraries. Technology has a major influence on changing the nature and architecture of libraries and their status in the society. Libraries of today have evolved from traditional libraries to a global face of resource providers via advanced information and communication technologies.

The systematic approach of information technology in libraries have resulted an enormous quantitative growth in information which is referred by term “Information explosion”. We cannot deny the advantages of this systematic technological approach in Indian libraries also. This approach has contributed a lot in the fields of information processing and digitization. But where the systematic approach accentuates on quantity of information, the Indian traditional meditative approach to knowledge emphasizes on quality. We Indians have been slow to introduce or develop a program to implement this approach completely, by preserving our rich cultural heritage and traditional values at the same time (Neelankavil, 2003).

India is a country with huge cultural diversity where a multicultural society encircles people with different religious, ethnic, regional, democratic and language backgrounds. It has witnessed the origin of many famous ancient literatures which contribute in the rich cultural heritage collection of the country. From time to time many cultural institutions have created
the repositories to embrace this cultural heritage. As a result of these initiatives various kinds of literature have been produced. In a country like India with such a unique rich cultural and traditional methods in library science, there are still opportunities and obligations of development of digital libraries (Das, 2012).

**Major digital library initiatives in India:**

Digital library initiatives in India have contributed in preserving the rare cultural documentary resources. Several steps have been taken by the Government of India to institute efficient digital libraries. Some of these initiatives are discussed here below:

- **Digital Library of India (DLI)** collectively implemented by Indian Institute of Science (IISc), Bengaluru; IIIT Hyderabad and C-DAC. This project is funded by Ministry of Communication and Information Technology (MCIT).
- **Kalasampada: Digital Library- Resource for Indian Cultural Heritage (DL-RICH)** hosted by Indira Gandhi National Centre for Arts (IGNCA) and funded by Ministry of Communication and Information Technology (MCIT).
- **Traditional Knowledge Digital Library (TKDL)** implemented by National Institute of Science Communication and Information Resources (NISCAIR). This project is funded by Department of Indian Systems of Medicine and Homoeopathy (ISM&H).
- **Mobile e-library** developed by C-DAC Noida under the Ministry of Communication and Information Technology (MCIT).
- **Nalanda Digital Library** developed by National Institute of Technology Calicut and funded by All India Council for Technical Education (AICTE).
- **Archives of Indian Labour: Integrated Labour History Research Programme** hosted by V.V.Giri National Labour Institute and Association of Indian Labour Historians.
- **National Science Digital Library (NSDL)** developed by National Institute of Science Communication and Information Resources (NISCAIR).
- **Down the Memory Lane** an initiative of National Library of India under the Ministry of Culture.
- **Digitization of Manuscripts initiative** of National Mission for Manuscripts under the Ministry of Culture.
- **Digitization, electronic archiving, indexing and retrieval system of the Indian Journal of Medical Research (IJMR)** an initiative of Indian Council of Medical Research (ICMR).

**About the MDU, Rohtak: a succinct summary**

The Maharshi Dayanand University (MDU), Rohtak, Haryana (India) was established in 1976, through an act of Haryana State Legislative Assembly, as a residential University, with the objective to promote inter-disciplinary higher education and research with special emphasis on studies of environmental, ecological and life sciences.

**Affiliated Colleges, Departments and Programmes offered:**

In a span of 35 years, the University has progressed leaps and bounds to emerge as a leading centre of higher education not only in the state of Haryana, but also in the entire Northern Region of India. It has since acquired the status of a teaching-cum-affiliating University with 518 colleges presently affiliated to it, including 298 Colleges of Education offering B.Ed. and
M.Ed. programmes, 96 Degree Colleges, 80 Engineering Institutions, 34 Management Institutes, and 6 Law Colleges.

At present, the University offers courses of study through 11 faculties, 37 post-graduate teaching departments (UTD), one autonomous constituent institute known as University Institute of Engineering and Technology (UIET), a University Institute of Law and Management Studies at Gurugram, and a Directorate of Distance Education (DDE) that operates from the main campus.

Courses of Study:

Maharshi Dayanand University is a fast growing hub of wide range of courses being offered in the traditional, professional, and job-oriented streams. The nascent initiatives include establishment of a Centre for Bioinformatics, a Centre for Medical Biotechnology, Department of Food Technology, and Department of Library and Information Science. The University runs 74 programmes of study through its 38 PG departments on the campus and three centres, besides offering Ph.D. programmes in almost all subjects.

The University has diploma level courses in French and Spanish to cultivate in its students linguistic skills in select foreign languages to create better mobility and greater employability in a wider arena.

Collaborations with Prominent Institutions:

The University sincerely embraces its mission of providing quality education. To further this goal, it has inked MoUs with reputed educational and research organizations both at national and international level. The University teaching departments have entered into research collaborations with 48 national and 15 international institutions. It has inked MoUs with 13 international and 3 national organisations, and 1 industry for the promotion of academic and research activities. More prominent institutions are the National Institute of Malaria Research (N.I.M.R.), American Institute of Indian Studies, Institute of Humanity and Nature (Japan), and Central Electronics Engineering Research Institute (CEERI) have proved to be highly delivering. Having successfully completed a 5-year collaborative research project with the Korean Institute of Energy Research, the Chemistry Department is participating in this project with Universities of Valencia and Madrid in Spain, University of Minho in Portugal, and three others (one each from Poland, Chile, and Mexico).

Embracing Technology:

The University pays special care to the use of latest technological advancements in academic governance. This has resulted in massive computerization of all important day-to-day functions of the University. A dynamic Campus Wide Area Network already stands provided. In fact, ours is the first varsity which went into becoming a Wi-Fi campus in the region. All the teaching departments have been provided modern teaching aids and efforts are being made to initiate, adopt, and popularize the emerging IT-enabled teaching processes. Accordingly, each member of the faculty has been provided with a PC or laptop to further the cause of effective teaching delivery mechanism. Taking “networking” to the doorstep of every teacher and the ongoing drive to extend the facility to all hostels has been the top-most priority of the University administration during the last two years, with a good measure of progress already made in achieving the targets.

University Library Services:
The University has carved out a highly conducive and enabling academic environ with its Vivekananda Library, named after the great social reformer Swami Vivekananda. Its rich knowledgebase encompasses 3,50,632 print books, 20,146 eBooks, 50,800 bound volumes of Journals, 442 Indian and 108 Foreign Journals are subscribed in print form, more than 13,000 electronic journals and 5 databases are accessible through e-ShodhSindhu – a MHRD, Govt. of India consortium. Housed in a modern spacious building with 84,000 sq.ft. carpet area, the library provides seating capacity for 963 library users, with another 358 seats in its five offshoots. It has an air-conditioned Internet Lab with 80 PCs, 1 Gbps bandwidth internet connectivity. The university library system is indeed fully automated in true sense, which others would envy to emulate. The innovative Radio Frequency Identification Technology (RFID) has made the task of check-out check-in, self check-in and tracking of books and other material easier and more efficient.

**Need for Digitization of Theses:**

The Vivekananda Library has a separate Archives Section where 18,089 Ph.D. and M.Phil. theses produced in the University are stacked with restricted access to information seekers. An air-conditioned Reading Hall with a seating capacity for 250 readers remains open 24X7 throughout the year, with a large number of readers thronging around all the time in search of a seat inside. Laced with the latest technologies, Vivekananda Library of Rohtak is contributing its part well and good in the digitization. The digitization project of Vivekananda Library aims to provide an open access environment to the academic community world-wide by online availability of electronic theses and dissertations through centrally managed repository. It can provide easy access and archiving of the theses and dissertations of Maharshi Dayanand University to help raising the standard and quality of research among academic community.

**Implementation Plan:**

For providing digital resources such as electronic theses and dissertations, Vivekananda Library has collaborated with the Shodhganga. It is a digital repository of Indian Electronic Theses and Dissertations set-up by the INFLIBNET Centre set-up using an open source digital repository development software D-Space developed by MIT (Massachusetts Institute of Technology)in contribution with Hewlett- Packard (HP). To execute and implement the project in a transparent way a detailed tender consisting work and the technical requirements was prepared meticulously to avoid any possible shortcomings in the work.

**Execution of digitization of theses:**

Brief execution details for digitization work are as under:

- Ph.D. Theses segregated from Post-Graduation and M.Phil Theses
- Made inventory of all Ph.D. theses
- Searched for duplicate theses and identified unique theses for scanning
- Selected theses were unbounded for scanning
- Scanned theses using Image Scanner fi-7260 in .TIFF file format on 300 dpi (dot per inch) resolutions with [Bit depths: Bit, not less than 8-bit Grayscale, or 24-bit Color.
- Done basic image enhancement, Raster cleaning, De-skew, De-Speckle Cropping & hole removal etc. on each images for optimum images clarity without compromising the quality of text, photographs etc.
- Checked files for integration and noted missing pages and repeated the process of scanning, cleaning and converting files in to PDF/PDFA file
- Created chapter-wise PDF/PDF-A files in a systematic way
• Captured and created Metadata for each and every thesis and created .xls file using MS-Excel.
• Rebinding done after final check of the soft-copies of the PDF/PDF-A files
• Upload the Digitized PDF on Document Management Systems (DMS), that is, our own Institutional Repositoty and Shodhganga
• Metadata of digitized theses uploaded to our Library Management System Libsys which is searchable through Web OPAC.
• Digital copy of entire digitized content is backed up in Portable Hard Disk.

The institutional repository in Vivekananda Library is developed using D-Space and a backend of Ubuntu version 14.04 LTS (A Linux based operating system). Hardware and Software specifications for digitization project by Vivekananda Library are as follows:

• **Scanner** – FUJITSU Image Scanner fi-7260 – Duplex scanning speeds of 120 images per minute at 300 dpi in Color, Grayscale, and Monochrome, data captured with smart ultrasonic technology, and reduce rescans by cleaning up documents automatically in a single pass.

• **ABBYY FineReader OCR**: It enables you to convert different types of documents, such as scanned paper documents, PDF files or images captured by a digital camera into editable and searchable data.

• **14.04 LTS version of Ubuntu - Linux Operating System for Backend of Repository**
• **D-Space for Frontend of Repository. Lifecycle is as under:**

Classification of total 18089 Theses in MDU: There is a large collection of 18089 theses in Vivekananda Library of Maharshi Dayanand University, Rohtak, Haryana (India). We created an inventory of all these theses and found that the number of M.Phil and Post
Graduate theses and dissertations are 80.3% (14,524) and 19.7% (3,565) theses are doctoral theses.

**Format of Doctoral Theses:** Since the inception of Library, total 3,565 Ph.D. were submitted to the library. Out of this, 73% theses (2,593) were received in traditional print format. During the last six years, we received 27% theses (972) in digital format. All the digitized and born digital theses have been uploaded on Shodhganga and 2550 (out of 2593) theses digitized under Shodhganga ETD project have been uploaded on our Institutional Repository.

**Linguistic bifurcation of digitized theses:** We digitized total 2550 doctoral theses. We observed that out of these 2550 theses 74.94% (1944) theses were written in English language and remaining 25.6% (639) theses were written either Hindi or in Sanskrit.

**Faculty-wise distribution of digitized theses:**

<table>
<thead>
<tr>
<th>Name of Faculty</th>
<th>Theses Digitized</th>
<th>%age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commerce</td>
<td>157</td>
<td>6.2</td>
</tr>
<tr>
<td>Education</td>
<td>253</td>
<td>9.9</td>
</tr>
<tr>
<td>Engineering &amp; Technology</td>
<td>71</td>
<td>2.8</td>
</tr>
<tr>
<td>Humanities</td>
<td>592</td>
<td>23.2</td>
</tr>
<tr>
<td>Law</td>
<td>164</td>
<td>6.4</td>
</tr>
<tr>
<td>Life Sciences</td>
<td>219</td>
<td>8.6</td>
</tr>
<tr>
<td>Management Sciences</td>
<td>132</td>
<td>5.2</td>
</tr>
<tr>
<td>Performing &amp; Visual Arts</td>
<td>75</td>
<td>2.9</td>
</tr>
<tr>
<td>Pharmaceutical Sciences</td>
<td>16</td>
<td>0.9</td>
</tr>
<tr>
<td>Physical Sciences</td>
<td>351</td>
<td>13.8</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----</td>
<td>------</td>
</tr>
<tr>
<td>Social Sciences</td>
<td>512</td>
<td>20.1</td>
</tr>
</tbody>
</table>

**Findings of Study**

- Out of total digitized theses, 73% (2,593) were received in traditional print format.
- 74.94% theses were written in English language and 25.6% theses were written either in Hindi or in Sanskrit.
- University Faculty of Humanities contributed highest 23.2% theses while Faculty of Pharmaceutical Sciences contributed lowest 0.9% theses to the library collection.

**Suggestions:**

- The university library has a large collection of M.Phil. dissertations which needs to be digitized and preserved for the use of future generations.
- Migration to Future Storage Formats and Technology: Computer technologies are very dynamic in nature, therefore, the university must prepare a future plan to decide what digital storage formats and technologies shall constitute acceptable archival storage and migrate the Collection to those formats and technologies as and when required.
- The digital copy should be kept on cloud by the library for purpose of future use and record.
References


