

**DESIGN AND DEVELOPMENT OF ELECTRONIC
THESIS AND DISSERTATION DATABASE FOR
DEPARTMENT OF LIBRARY AND INFORMATION
SCIENCE, PONDICHERRY UNIVERSITY USING
GREENSTONE DIGITAL LIBRARY SOFTWARE**

*Dissertation submitted to the Pondicherry University in partial fulfilment of
the requirements for the Degree of
Master of Library and Information Science*

By

ANNA POORANI V

(18390003)

Supervised by

Dr. M. LEELADHARAN



Department of Library and Information Science

School of Media and Communication

PONDICHERRY UNIVERSITY

2020

DECLARATION

I, hereby declare that the dissertation entitled "**DESIGN AND DEVELOPMENT OF ELECTRONIC THESIS AND DISSERTATION DATABASE FOR DLIS, PU USING GREENSTONE DIGITAL LIBRARY SOFTWARE**" submitted for the award of the degree of **Master of Library and Information Science**, Pondicherry University, is a record of original work done by me under the supervision and guidance of **Dr. M. LEELADHARAN**, Assistant Professor of the Department of Library and Information Science, School of Media and Communication, Pondicherry University, Puducherry.

ANNA POORANI V

Place: Puducherry,

Date:

CERTIFICATE

This is to certify that the dissertation entitled "**DESIGN AND DEVELOPMENT OF ELECTRONIC THESIS AND DISSERTATION DATABASE FOR D.L.I.S., P.U. USING GREENSTONE DIGITAL LIBRARY SOFTWARE**" is the bonafide work of **ANNA POORANI V (Reg.No.18390003)**. It has been carried out at the Department of Library and Information Science, Pondicherry University, under my supervision and guidance for the partial fulfilment of the requirements for the award of the Degree of **Master of Library and Information Science**. This work has not been submitted for the award of any other degree of this university or any other universities.

Dr. M. LEELADHARAN

Assistant Professor

Department of Library and Information Science

Pondicherry University

Puducherry

Place: Puducherry

Date:

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the Almighty for giving me strength and direction to do this project

I express my whole-hearted gratitude to my guide, my mentor **Dr. M. Leeladharan, Assistant Professor**, Department of Library and Information Science, Pondicherry University, Pondicherry for having offered valuable insights, encouragement on the topic and for his inspiration, supervision, guidance, advice, commitment, enthusiasm, understanding, support, above all for his consistent patient and constant motivation in various ways at every stage of this project. It has been an honor and a privilege to be his student.

I extend my heartfelt thanks to all the faculty members in the Department of Library and Information Science, Pondicherry University including the Dean **Prof. Dr. C.K. Ramaiah** for selflessly sharing his vast exposure and exuberant knowledge, the HOD **Dr. Rekha Rani Varghese** for her radical inspiration, her eager and even passionate encouragement, my tutor **Dr. R. Sevukan** for his radiant enthusiasm and constructive guidance, and **Dr. Mangkhollen Singson**, one who got me introduced to greenstone, for sharing his cordial and unfailing intelligence. **Dr. J.P.S.Kumeravel** for his benevolent support and valuable suggestions, in the entire course of the degree as well as to complete this project and report. I am deeply owed the vibrant professors for their love, knowledge and optimism they have given in me.

I am deeply thankful for my parents and my brother who had been there for me and encouraged me throughout my life as well as in the time of this project work. This thesis is heartily dedicated to my mother who has been a constant source of inspiration all my life.

I am also genuinely thankful to my husband for his constant encouragement, constructive criticism and suggestions. He got my back when I needed the most and his challenges brought this work towards a completion. It is with his support that this work came into existence.

Furthermore, I would like to share my sincere thanks to my parents in law for encouraging me to pursue my studies and their extended support.

At this juncture, I am so grateful for Kadiresan Nagarattinam, Somipam R Shimray Research scholars, Department of Library and Information Science, Pondicherry University for generously sharing their Knowledge and Contribution. And a special mention to Pavithra Shanmugam for helping me in the digitizing process.

Finally, I wish to thank all my friends and my classmates who made this entire course of time joyful but always efficient.

And to many others who had not been mentioned but none forgotten.

ANNA POORANI V

ABSTRACT

Digital libraries are quintessential tools to manipulate a non-commercial mechanism for disseminating knowledge originating from the research output of Academic institutions. Research and other activities of academic institutes generate valuable information consistently. Libraries collect, organize and store such information for easy retrieval. Digital Libraries foster the ability for these academic libraries to build digital information collections locally.

The Greenstone digital library software is an extensive platform for constructing and disseminating digital library collections. It is a flexible tool that caters a way of organizing and publishing information based on metadata on both online and offline. Being “opensource” software, Greenstone can be customized to meet the needed functionalities.

This project brought forth the edges of developing a digital library collection of Electronic Thesis and Dissertations using greenstone digital library software for the Department of Library and Information Science, Pondicherry University. This illustrates how greenstone is promoting open educational resources (OER) movement and demonstrates the customization of GSDL in for a specific collection.

Keywords: Digital libraries, Institutional repositories, Digital archives, Collection management, ETD, Greenstone Software.

Table of Contents

Contents	Page No
Declaration -----	i
Certificate -----	ii
Acknowledgements -----	iii
Abstract -----	iv
Table of Contents -----	vi
List of Figures-----	xi
List of Tables -----	xiii
1.0 Introduction -----	01
1.1 Digital Library -----	02
1.1.1 Definitions of Digital Library -----	03
1.2 Archives and Institutional Repositories -----	04
1.2.1 Definitions of Institutional Repositories -----	05
1.2.2 Characteristics of Institutional Repositories -----	06
1.2.3 Functions of Institutional Repositories -----	07
1.3 Electronic Thesis and Dissertation (ETD) -----	07
1.4 Open Source Software -----	09
1.5 DSpace -----	11
1.6 EPrints -----	12
1.7 Greenstone Digital Library Software -----	12
1.8 AIM -----	13
1.9 Purpose of Study -----	13

1.10	Objectives -----	13
1.11	Prospectus -----	13
1.12	Chapterization -----	14
2.0	Literature Review -----	15
2.1	Introduction -----	15
2.2	Digital library and IR systems -----	15
2.3	Electronic Theses and Dissertation (ETD) -----	16
2.4	Workflow -----	18
2.5	Standards -----	21
2.6	Open Source Software -----	24
2.7	Greenstone Digital Library (GSDL) -----	24
2.8	Summary -----	26
2.9	References -----	26
3.0	Methodology -----	31
3.1	Introduction -----	31
3.2	Design and Development Methodology -----	31
3.3	ADDIE Model -----	34
3.3.1	Analysis -----	35
3.3.2	Design -----	37
3.3.2.1	Workflow Creation -----	38
3.3.2.2	Selection of Digital Library Platform -----	39
3.3.2.3	Evaluation Design -----	41
3.3.3	Development -----	42

3.3.3.1 Prototype Creation -----	42
3.3.3.2 Extraction of Bibliographic data -----	43
3.3.3.3 Valuation of Prototype -----	43
3.3.4 Implementation -----	43
3.3.4.1 Creation of Database -----	44
3.3.4.2 Customization of the User Interface -----	44
3.3.5 Evaluation -----	45
3.4 Summary -----	46
4.0 Design and Development -----	47
4.1 Introduction -----	47
4.2 Digitizing Collection -----	47
4.2.1 Tools for Digitization -----	50
4.2.2 Digitizing Process -----	51
4.3 Gathering Metadata -----	53
4.4 Development of ETD -----	58
4.4.1 Software Requirement and System Configuration -----	58
4.4.2 GSDL Installation -----	58
4.4.3 Interface Design-----	60
4.4.4 Creation of New Collection -----	61
4.4.4.1 Gather -----	62
4.4.4.2 Enrich -----	63
4.4.4.3 Design -----	65
4.4.4.4 Format -----	68

4.4.4.5 Create -----	74
4.4.5 Customization of Home Page -----	74
4.4.6 Moving the Collection to CD-ROM -----	76
4.4.7 ETD Database Design Preview -----	76
4.5 Summary -----	82
5.0 Evaluation -----	83
5.1 Introduction -----	83
5.2 Design Evaluation -----	83
5.2.1 Formative Evaluation -----	85
5.2.2 Summative Evaluation -----	86
5.2.2.1 Heuristic Evaluation -----	86
5.3 Summary -----	88
6.0 Conclusion -----	89
6.1 Introduction -----	89
6.2 Summary -----	91
6.2.1 Trends in Institutional Repositories -----	91
6.2.2 Design and Development of ETD Database -----	91
6.2.2.1 Analysis -----	91
6.2.2.2 Tools and Technologies -----	92
6.2.2.3 Digitization of Physical Collection -----	92
6.2.2.4 Development of ETD Database -----	92
6.2.2.5 Evaluation of ETD Database -----	93
6.3 Research Contributions -----	94

6.4	Research Gaps -----	95
6.5	Recommendation for future research -----	95
6.6	Conclusion -----	96
 Bibliography -----		098
Appendix- A: List of Abbreviations -----		105
Appendix- B: Librarian Interface Guide-----		107
Appendix- C: User Guide -----		109

List of Figures

Figure 1.1 Hierarchy of ERM -----	08
Figure 2.1 High-level ‘life cycle’ model -----	19
Figure 2.2 The Workflow of Producing Digital Product -----	19
Figure 2.3 System Architecture of the Workflow Reference Model -----	20
Figure 2.4 Digitization of Ph.D. Theses at IIT Delhi -----	20
Figure 2.5 News Clipping Workflow -----	21
Figure 3.1 ADDIE Model -----	34
Figure 3.2 Courses of Analysis Phase -----	35
Figure 3.3 The Customized Workflow Layout in the Development of ETD DB ----	38
Figure 4.1 The Konica Minolta PS 7000 book scanner -----	49
Figure 4.2 DIY Scanner Setup -----	49
Figure 4.3 Process of Digitization -----	52
Figure 4.4 Excel sheet -----	53
Figure 4.5 Starting a new collection -----	61
Figure 4.6 Adding information about collection -----	62
Figure 4.7 Gather Panel -----	63
Figure 4.8 Enrich Panel -----	64
Figure 4.9 Document Plugin -----	66
Figure 4.10 Search Index -----	66
Figure 4.11 Browsing Classifiers -----	68
Figure 4.12 General Macro -----	69
Figure 4.13 Format Features -----	70

Figure 4.14 HTML Code for Title Classifier -----	70
Figure 4.15 HTML Code for Subject Classifier -----	71
Figure 4.16 HTML Code for Guide Classifier -----	71
Figure 4.17 HTML Code for Year Classifier -----	72
Figure 4.18 HTML Code for Search Index -----	72
Figure 4.19 HTML Code for Collection Specific Macro -----	73
Figure 4.20 Create Panel -----	74
Figure 4.21 Home Page -----	75
Figure 4.22 HTML Code for Home Page -----	75
Figure 4.23 Moving Collection to CD-ROM -----	76
Figure 4.24 User Interface Design Before Customization -----	76
Figure 4.25 User Interface Design After Customization -----	77
Figure 4.26 Search Index -----	77
Figure 4.27 Title Classifier -----	78
Figure 4.28 Subject Classifier -----	78
Figure 4.29 Subject Classifier -----	79
Figure 4.30 Guide Classifier -----	79
Figure 4.31 Guide Classifier -----	80
Figure 4.32 Year Classifier -----	80
Figure 4.33 Year Classifier -----	81
Figure 4.34 Help Window -----	81
Figure 4.35 Settings Window -----	82

List of Tables

Table 1.1 Comparison of Digital Library Software -----	10
Table 3.1 Comparison Matrix of Different Digital Library Software -----	39
Table 4.1 Mapped Dublin Core Metadata -----	56
Table 5.1 Documentation of Formative Evaluation -----	86
Table 5.2 Documentation of Heuristic Evaluation -----	87
Table 6.1 Navigation Indexes and Classifiers -----	93

Chapter 1

Introduction

CHAPTER 1

INTRODUCTION

1. Introduction:

The Digital age has brought a spectacular change in the methodology of storage and retrieval. It shall be stated by three distinct features: abundance, currency and easy access to information. This has tremendously shifted the concept of libraries, their collection, and services. In turn, it becomes vital for an information professional to redefine their role in disseminating information to their patrons.

In this age of exponential growth of information, a digital library can be referred to as an organized collection of digital documents, in any digital form, that would permit ease of access to the information users. Digital libraries do replace the concept of a library as a warehouse of knowledge with the notion of it as a provider of information services to users in print as well as digital form, both physical and digital libraries are quintessential and can work in unison, supplementing each other. Digitizing the bibliographic data, organizing by systematic collections and making the materials accessible to users is the foremost application of digital libraries.

Institutional Repository is a novel concept in digital era. Digital information generated by novel works of scholarly Community shall be collected, managed, disseminated and preserved. To manage the Institutional Repositories several software have been developed in the past and Greenstone is one among them. Greenstone is an open-source software for building digital repositories. The Greenstone Digital Library Software (GSDL) is one open-source software that offers tools to systematically build and distribute digital collections. It aids to publish digital collections in different digital forms.

1.1 Digital library

Digital Libraries are redefining the role of libraries. A digital library serves as a rich resource for its patrons with the collection of structured system of electronic information.

"a focused collection of digital objects, including text, video, and audio, along with methods for access and retrieval, and for selection, organization, and maintenance of the collection" - **Ian Witten and David Bainbridge**

Growth and befitting use of digital libraries can generally be attributed to the emergence of computing tools in information storage, access, and delivery. Availability of radical yet simple and intuitive UI and the recent developments in cloud storage made storage of large corpus of digitized information possible.

Digital library enables the electronic access to large collection of electronic documents as well as digital access to material that already exists within physical library collections to the patrons.

The digital libraries are regarded as "logical extension and augmentation of physical libraries in the electronic information society (Marchionini, 1998)". The digital library extends and augment their physical counterparts by extending existing resources, services and enable the development of new possibilities for information access and retrieval (Fox, 1998).

The digitization of the library has focused profoundly on the creation of foster records of printed physical form of documents available in a library or for providing computerized digital services through secondary databases held locally on CD ROM escalating scope for digital libraries.

1.1.1 Definitions of Digital Library

Michael Lesk defined "Digital libraries are organized collections of digital information. They combine the structuring and gathering of information, which libraries and archives have always done, with the digital representation that computers have made possible."

Borgman (1992) highlights digital libraries as not only a point of access to digital information but as a combination of

- A service;
- An architecture;
- A set of information resource, and
- A set of tools and capabilities to locate, retrieve and utilize the information resources available.

William Arms stated "An informal definition of a digital library is a managed collection of information, with associated services, where the information is stored in digital formats and accessible over a network. A crucial part of this definition is that the information is managed. A stream of data sent to earth from a satellite is not a library. The same data, when organized systematically, becomes a digital library collection."

Terence R. Smith (1997), articulated digital libraries as "controlled collections of information-bearing objects (IBOs) that are in digital form and that may be organized, accessed, evaluated and used through a heterogeneous and extensible set of distributed services that are supported by digital technology".

Harter defined "A digital library contains digital representations of the objects found in it - most understanding of the digital library probably also assumes that it will be accessible via the

Internet, though not necessarily to everyone. But the idea of digitization is perhaps the only characteristic of a digital library on which there is universal agreement."

Clifford Lynch (1995), termed the digital library as "a system providing a community of users with coherent access to a large, organized repository of digital information and knowledge. The digital library is not just one entity, but multiple sources seamlessly integrated.

1.2 Archives and Institutional Repositories

An *archive* is a collection of historical records, Archives contain primary source documents which have accumulated throughout an individual or organization's lifetime. Archival records are generally unpublished and almost always unique, unlike books or magazines for which many identical copies exist.

While, *Institutional repositories* are digital collections capturing and preserving the intellectual output of a single or multi-university community (Crow, 2002).

Institutional Repository shall be store information in different forms. It can in the form of an archive, a museum, a storehouse that shall store for future use and safe keeping of an institutions records or artifacts with a scope outlining the institutional repository. Institutional Repository is a cognitive content that shall refer to storage and preservation of an institution's digital information and knowledge resource.

However, an Institutional Repository has taken connotation that refers to the storage and preservation of an organization's digital information or knowledge assets.

Institutional Repository can also be defined as "a Web-based archive of scholarly materials produced by members of the institution. Since the Repository relies on input from members of the

institutional commitment and participation of contributors, users and managers are crucial” (Johnson 2002).

The prime purpose of the Institutional Repository is to organize systematically and store the intellectual output of a scholarly community, it provides tools that help the scholarly community to exhibit their Intellectual productivity to its spectators.

Institutional Repository supports as a complement to traditional forms of publication. They enable information seekers to find the works of the scholarly community in comfortably organized way and indexed in a way that can be easily accessed, making it more visible to users like employers, colleagues and fund providers. The incentives and commitments also act as a motivational factor to change the process of scholarly communication.

The impact of information flare-up, dwindling library budget, space constraints, exalted demand for information have forced the libraries to seek ways to store and disseminate information among the users, outcome of which the concept of IR's emerged among the academic institutions. Now the repositories are built by individual academic institutions by themselves, which act as micro repositories.

1.2.1 Definitions of Institutional Repositories:

Lynch (2003) defined Institutional Repository as “a university-based Institutional Repository is a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members”.

Foster and Gibbons (2004) defined as “an electronic system that captures preserves and provides access to the digital work products of a community”.

According to Raym Crow (2004) Institutional Repository as a “Digital Archive of an intellectual product created by the faculty, research staff, and students of an institution and accessible to end-users both within and outside the institution, with few, if any barrier to access. The content is institutionally defined, scholarly, cumulative and perpetual, open and interoperable.”

Chang (2003) explained Institutional repository with scholarly works created in digital form. Based on the new method, capturing, collecting, managing, disseminating and preserving in digital form.

Heery & Anderson (2005) articulates Institutional Repositories as:

- Contains content, deposited by the owner, creator, or third party;
- Repository architecture manages content as well as metadata;
- The repository offers a minimum set of basic services, e.g. put, get, search, access control;
- The repository must be sustainable and trusted, well-supported and well managed;
- If an Open Access Repository, it must also;
- Provide Open access to its content (notwithstanding legal constraints);
- Provide Open Access to its metadata for harvesting.

1.2.2 Characteristics of Institutional Repositories:

Gibbons enlists the five core features of an Institutional Repository: -

- It is *institutionally defined* — the Institutional Repository captures only the intellectual output of its parent institution.
- Content can be either *specifically academic* or may contain administrative, teaching and research materials, both published and unpublished.

- It is *cumulative and perpetual*. The intellectual assets cannot be withdrawn, with it along with term obligation to the parent institution to preserve Institutional Repository content.
- It is *open and interoperable*.
- In gathering, storing and disseminating information it contributes to the process of *scholarly communication*.

1.2.3 Functions of Institutional Repositories:

Gibbons list the six core functions of Institutional Repository: -

- Material submission
- Metadata application
- Access control
- Discovery support
- Distribution
- Preservation

1.3 Electronic Thesis and Dissertation (ETD)

An Electronic Thesis and Dissertation (ETD) is self-explanatory by its name, ETD is prepared as a result of research work carried out by students of post-graduate course or research degree. It provides a technologically advanced medium for presenting one's notions.

ETD is structured in a system simultaneously suitable for machine archives and worldwide retrieval. It can be prepared using any word processor or document preparation system, incorporating relative multimedia objects resulting in less expensive, virtual library space that can be accessed at patron's convenience.

There are many ETDs, the prominent few to name include:

The Networked Digital Library of Theses and Dissertations (NDLTD) is a collaborative effort of universities around the world to promote creating, archiving, distributing and accessing Electronic Theses and Dissertations (ETDs). Since its inception in 1996, many universities have joined the initiative, underscoring the importance institutions place on training their graduates in the emerging forms of digital publishing and information access. The outreach and training mission of NDLTD is an ongoing project. Recent research has focused on creating a union database that will provide a means to search and retrieve ETDs from the combined collections of NDLTD member institutions. In response to the need for a focused and accessible catalog with a low barrier to participation, NDLTD has adopted a solution that uses the Open Archives Initiative's Metadata Harvesting Protocol (OAI-PMH) to gather metadata in the ETDMs format and then to make it accessible at a central portal. NDLTD project has international members from over a dozen countries sharing electronic theses and dissertations.

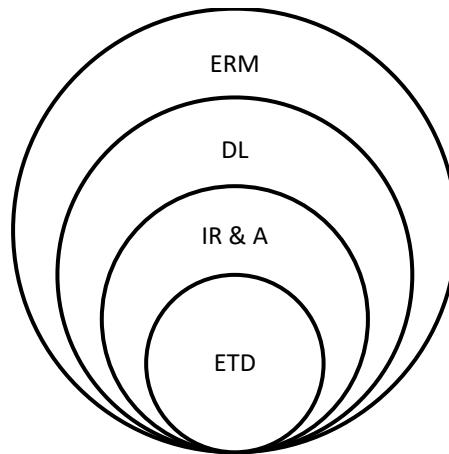


Figure 1.1 Hierarchy of ERM

1.4 Open Source Software

Open-source software has been a vague term to any software that is free and is often mistaken with freeware and shareware. The Open Source Initiative (OSI)¹ has, therefore, become a certification body for open source software under a commonly agreed-upon definition for “open source”, a method of software development that harnesses the power of distributed peer review and transparency of progress. This technique delivers better software’s that are economic still holds higher reliability, flexibility, and eliminate the traditional vendor lock-in. The source code and rights that were normally reserved for copyright holders are now distributed under a free software license that permits users to study, change, and improve and to redistribute the software.

There is many prominent open-source software used to create repositories some of them are EPrints (Southampton University), DSpace (MIT), CDSWare (CERN), and FEDORA (Cornell and U. of Virginia), Greenstone (University of Waikato), i-Tor and MyCoRe.etc. In addition to this, there is a few more software, which is being used for archiving. There are a handful of Open-source packages for creating and maintaining such archives. They are distributed either under GNU public license or the BSD license. Each of the software has a multitude of features, unique facilities and excellent capabilities, which the users are open to explore and experiment with it.

The three leading software packages, Greenstone, DSpace, and EPrint are available free under Open source licenses,

¹ www.opensource.org

Table 1.1: Comparison of Digital Library Software

Features	DSpace	EPrints	Greenstone
Year of creation	2002	2000	1997
Product Type	Software	Software	Software
Update cost	Free	Free	Free
Resource Identifier	CNRI Handles	No	OAI Identifier
OAI-PMH	Yes	Yes	Yes
Supported Item Types (Storage and rendition)	Can store and manage all types of content	Can store and manage all types of content	Can store and manage all types of content
Metadata formats	Dublin Core, Qualified DC, METS	Dublin Core, METS	Dublin Core, Qualified DC, METS, NZGLS (New Zealand Government Locator Service), AGLS (Australian Government Locator Service)
User interface functions	End-user depositions, Multilingual support	End-user depositions, Multilingual support	End-user deposition, Multilingual support
Thumbnail Preview	Images	Images, Audio, Video	Images, Audio, Video
Searching Capabilities	Field Specific, Boolean Logic, Sorting options	Field Specific, Sorting options	Field Specific, Boolean Logic,
Browsing options	By Author, Title, Subject and	Browsing can be done using	Browsing can be done
Syndication	RSS, ATOM	RSS, ATOM	---
User Authentication	LDAP Authentication,	LDAP Authentication	User Groups

	Shibboleth Authentication		
Statistical reporting	Count of Full Records	Count of Full records	Count of Full records
Software Platforms	Linux or Unix, Solaris, Windows	Linux, Unix, Windows,	Linux, Unix, Windows, Mac-OS
Databases	Oracle, PostgreSQL	MySQL, Oracle, PostgreSQL, Cloud.	Its Own
Programming Language	Java and JSP	Perl	C++, Perl, Java
Web Server	Apache and Tomcat	Apache	Apache/IIS
Associated Software	Java, Apache, PostgreSQL, or Oracle	Linux or Unix, Apache, Perl	Apache, PERL, GNU C++ Compiler, JAVA, GNU Database manager
Machine-to- Machine Interoperability.	OAI-MHP, OAI- ORE, SWORD, SWAP	OAI-MHP, OAI- ORE, SWORD, SWAP, RDF	Z39.50, OAI-MHP
License	GNU	BSD	GNU

1.5 DSpace

DSpace² is a platform from MIT that captures items in any format — in text, video, audio, and data. It distributes it over the web. It indexes the entity, so users can search and retrieve them. It preserves the digital manifestations of the original work over the long term. DSpace provides a way to manage research output and publications in an organized repository to furnish them with greater visibility and accessibility.

¹ <https://duraspace.org/dspace/>

1.6 EPrints

EPrints software³ is developed as a part of a digital library project at the University of Southampton, UK. It is available free under the term of GNU General public license. It runs under the Linux and creates online archive libraries of electronic prints. The default configuration creates a research paper archive but could not be modified and use for other purposes. The document can be stored in any format and each research paper (e-print) can be stored in more than one format. The print can be submitted through a powerful www-based interface. The data integrity checks are performed.

1.7 Greenstone Digital Library Software

The New Zealand Digital library Project at the University of Waikato produces greenstone digital library software. This project is a research program aiming to develop the underlying technology for digital libraries and make it available publicly so that others can use it to create their collections. The main architects of the software are Roger McNab and Stefan Boddie. The Greenstone Digital Library Software is a suite of software for creating, maintaining, presenting and distributing collections of digital resources. It organizes the collection digitally for publishing it on the internet or CD-ROM. This software is developed and distributed in cooperation with UNESCO and the human info NGO. It is open-source software available from <http://www.Greenstone.org> under the term of the GNU General public license

Greenstone Digital Library Software (GSDL)⁴. Greenstone collections offer effective full-text searching and metadata-based browsing facilities that are attractive and easy to use, and a user-friendly interface called *The Collector* makes it easy for people to assemble their library

² <https://www.eprints.org/uk/>

⁴ <http://www.greenstone.org/>

collections from disparate source documents. To address the exceptionally broad demands of digital libraries, the system is public and extensible: It is open-source software and it is distributed under the terms of the GNU General Public License.

1.8 AIM

The aim of this project is to design and develop an “Electronic Thesis and Dissertation Database for the Department of Library and Information Science, Pondicherry University”.

1.9 Purpose of the study

The purpose of this project is to study the process of non-commercial mechanism for digitization and management of intellectual output of the Department of Library and Information Science, Pondicherry University.

1.10 Objectives

The Specific Objectives that guided the study are as follows.

1. To digitize the physical collection of Thesis and Dissertation of DLIS, PU.
2. To categorize the intellectual materials that are digitized.
3. To devise E-Thesis and Dissertation (ETD) database.

1.11 Prospectus

The web-based ETD systems have various advantages, and a few have been listed below:

- Providing Bibliographic searching for ETD Collection retrieval
- Global access to ETD Collection in real-time by remote users
- Browsing facility through various metadata fields

- Instant access to ETD Collection over the web through a common user interface.
- Up-to-the date updated access to ETD Collection.
- Dynamic updating of the digitized collection from any location

1.12 Chapterization

There are Six chapters in this Dissertation.

The Introduction chapter provides general information about Digital Libraries, IR, ETD and GSDL; The aim, purpose, objectives and prospects of the study are discussed in this chapter.

The Second chapter is the literature review of different trials conducted in the specific field of studies.

The Third chapter discusses the methodologies employed in designing, the ADDIE model and its aspects in the prospect of development of the ETD Database.

The Four Chapter details the digitization of Thesis and dissertation collection and the design and development of ETD using GSDL Software.

The Fifth chapter covers the evaluation of the ETD Database, formative and summative evaluations.

The Sixth chapter deliberates discussion for future developments on ETD Database and the Conclusion.

Chapter 2

Literature Review

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

To understand the themes of Digital libraries, IR, ETD, GSSDL better, before delving further into the research, it is essential to review some published works related to the topic. This review includes a holistic analysis of the latest trends, collected from various sources including publications in journals and web. Thus, a literature review gives a better perspective of the topic under research, enriching from gathered sources of knowledge. The review endeavor to provide a better explanation on scope of the research. The literature review has been arranged thematically under the various subthemes.

Having understood the concept of ETD, the topic has been divided it into five major components: Digital library and IR systems, workflow, standards, OSS, GSSDL. *ETD* deals with the management electronic versions of thesis and dissertation while *Digital library and IR systems* relates with organized collection of other sources of information, *Workflow* represents the life-cycle of the process involved, *standards* are the compliances that implies to the development of ETD systems, and *Open Source Software (OSS)* tools for management of e-resources were analyzed for the selection of appropriate tool for the present context. Each component has been individually reviewed in the standard literature review.

2.2 Digital library and IR systems

Self-Archiving is conceived out of glitches with the current scholarly communication model developed by commercial publishers and vendors. Digital libraries assist human *Design and Development of ETD DB for DLIS,PU using GSDL*

development by providing a non-commercial mechanism for channelizing humanitarian information and can reverse the negative impact of information technology on developing countries (Witten et al., 2002). The establishment of self-archiving in such countries helps their national research becomes mainstream and contributes on an equal footing to the global knowledge pool. (Anuradha, 2005)

Institutional Repositories provide seamless access to abundant information and knowledge in a digital world to serve its users' information needs (Gohain, R.R., 2011). It is a method for capturing, collecting, managing, disseminating, and preserving scholarly works created in digital form by the constituent members of an institution (Chang, 2003).

Anuradha in her paper "Design and development of institutional repositories: A case study" emphasis self- archiving as extremely sporadic, for instance institutional repository of IISc is compiled from several identified, authentic sources by extracting metadata by constructing an appropriate search strategy. The repository, named "PRABHAVI", is web-enabled using Greenstone Digital Library software (Anuradha, 2005).

2.3 Electronic Theses and Dissertation (ETD)

Theses and dissertations were an early target for electronic archiving and distribution. ETDs are the electronic archives that mainly focus on the productivities of the latest researches by scholars during their process of cognition, exploration and analysis. The ETD repository can contain published journal articles, papers presented at scientific conferences, master's theses, doctoral dissertations, presentations, etc. (Dhamdhere & Smet, 2014)

Gupta (2014) made an analytical evaluation of the Indian digital repositories especially ETD and the government guidelines for acceptance of ETDs, found that most of the Indian

universities and higher-level institutions concentrate more on Government funded retro-conversion of the timeworn PhD theses rather than acquisition the current ETD and ETD creating facilities for the researchers.

Mittal & Mahesh (2007) identified and evaluated the collections within digital libraries and institutional repositories in India available in the public domain and found the extensive use of open-source software in the creation of institutional repositories. Apart from which, major digital library initiatives such as the Digital Library of India used custom-made software. where the collection size is found to be in a few hundred.

Webley and Suleman (2011) in their paper “Creating a National Electronic Thesis and Dissertation Portal in South Africa” discusses their experience on development of a custom software solution, using a multitiered simple architecture of complex components where a tiered architecture is tightly integrated into a commonly-used application/operating system framework.

Anuradha et al., (2011) tried to enable full-text search features in open-source library automation package Koha, through integrating it with Greenstone Digital Library Software (GSDL) by passing and search and retrieval by URL (SRU) request to GSDL from Koha. The text documents are indexed individually in Koha and in GSDL.

Reeves, Hagen and Jewell (2006) argue that ETDs have matured enough to capture the attention of the higher education communities worldwide. With more organizational interest in the growing Institutional Repository movement, the links between ETDs and institutional repositories can be build, well positioned as open access and archival preservation techniques become common in electronic and scholarly publishing.

2.4 Workflow

The Council of Library and Information Resources report in 2004 recommended the move from print to electronic journals would result in time savings for Technical Services staff and would help them take on new responsibilities. Thus, the entire process of management of Thesis and Dissertation gets renovated more proficiently, with lesser steps and fewer staff involvement. ETDs contemporaneous many logistical issues. Submission, authentication, distribution, and preservation are major progressions that require a careful planning in maintain the integrity of the intellectual output (Lubas, 2009).

Boock & Kunda (2009) in their paper “Cataloging & Classification Quarterly Electronic Thesis and Dissertation Metadata Workflow at Oregon State University Libraries” discussed four major negative aspects of employing Cataloging in ETD Systems.

1. Cataloging of Thesis and Dissertation are time-consuming and costly
2. The Core subject content of Thesis and Dissertation are often narrow and specialized, making it difficult to assign subject headings accurately
3. Thesis and Dissertation often contain cutting-edge research and LCSH has not always caught up with newfound knowledge
4. Keyword and full-text searching are alternatives to LCSH

Here are Few of the workflow model that has major influence in designing the workflow for the creation of this ETD database, that include;

The High-level “lifecycle” model from The Ohio State University Libraries (2016) laying out the general steps in the workflow to build digital collections.(Shaw et al., 2018) The comprehensive workflow of producing digital product,

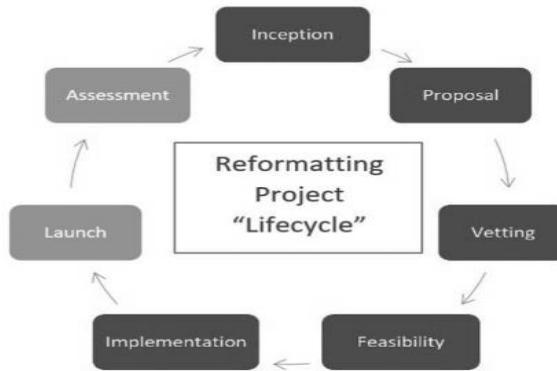


Figure 2.1 High-level 'life cycle' model from the Ohio State University Libraries (2016)

The workflow in producing digital product from planning of project to site maintenance was discussed in detail in University of Michigan Library blog by Kate Hagedorn⁵ is as follows

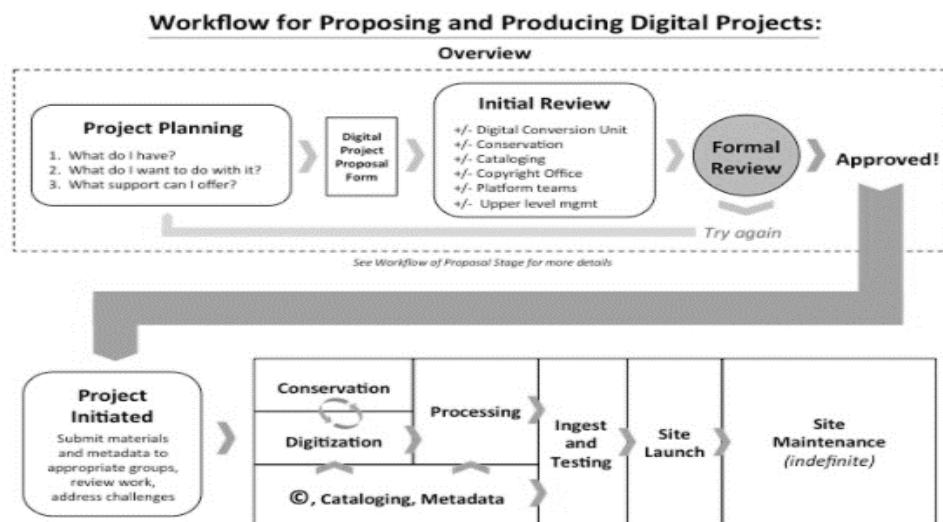


Figure 2.2 The Workflow of Producing Digital Product in University Michigan Library

Another workflow model that details the overall system architecture in the establishment of Digital archives/Digital Library define, the following recommendations of the Workflow Reference Model, has been a commonly accepted framework by Places et al., (2016)

⁵ <https://www.lib.umich.edu/blogs/library-tech-talk/how-create-and-keep-creating-digitization-workflow>

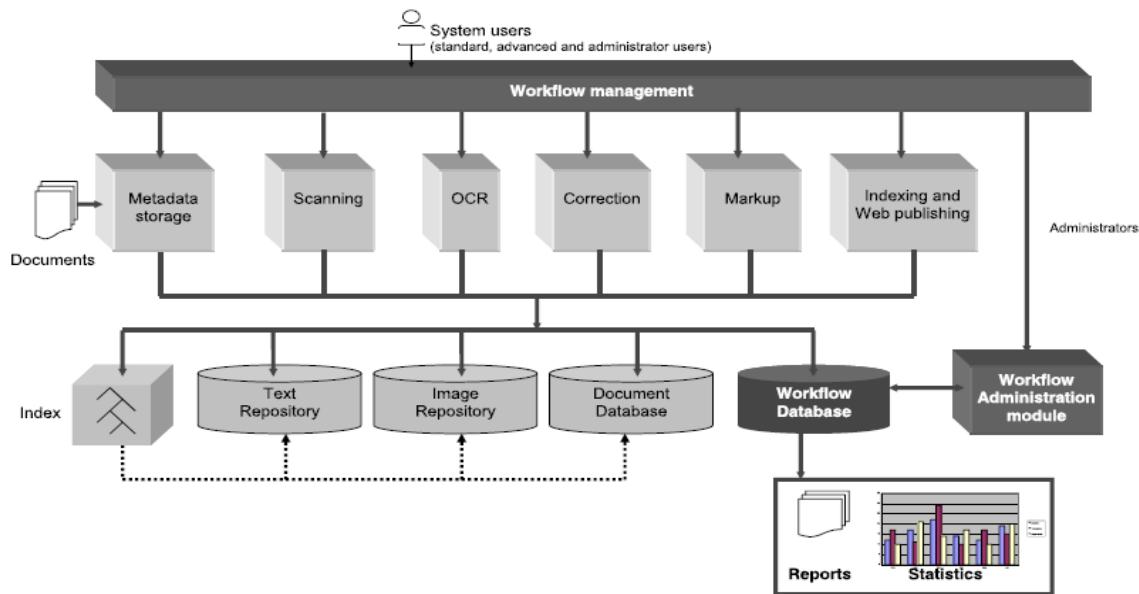


Figure 2.3 System Architecture of the Workflow Reference Model

Arora, (2013) has exemplified the process of Digitization of Ph.D. theses at the IIT Delhi in his paper Network-enabled digitized collection at the Central library, IIT Delhi.

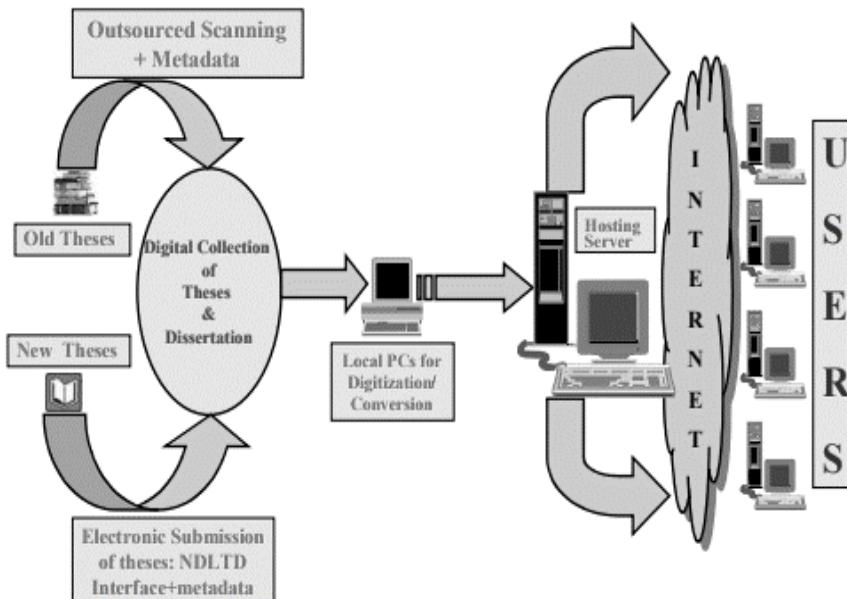


Figure 2.4 Digitization of Ph.D. Theses at IIT Delhi

Sonkar et al., (2005) build a ‘digital library of newspaper clipping’ through a simple workflow that illustrates the development of digital library and implementation of greenstone digital library software.

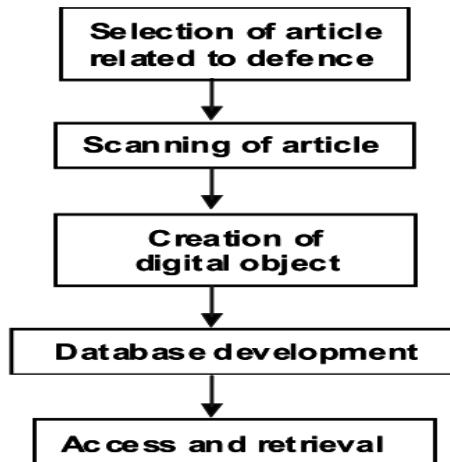


Figure 2.5 News Clipping Workflow

2.5 Standards

Once the workflow has been assessed, the key focus in design of any digital repository is conceiving metadata standards. There are different types of metadata standards tailored for different purposes, being practiced among research community, which have been discussed by various authors as follows.

Gilliland-Swetland (1998) interprets “metadata also documents how that objects behaves, its functions and use, relationship to other objects and how it should be managed”.

Among the different metadata standards, the most commonly practiced one’s are the Dublin Core and MARC 21 metadata standards.

In 2nd International World Wide Web Conference, October 1994, Chicago, Yuri Rubinsky (SoftQuad) Stuart Weibel and Eric Miller (OCLC) Terry Noreault, (Director - OCLC Office of Research) and Joseph Hardin (Director - National Center for Supercomputing Applications)

discussed on “semantics on Web and the difficulty of finding resources”. Enunciated the instigation of Dublin Core Metadata standards. The name refers to Dublin, Ohio, U.S., where the standard originated from the OCLC/NCSA Metadata Workshop hosted in 1995 by OCLC (NCSA is the National Center for Supercomputing Applications. The Dublin Core Metadata Initiative (DCMI) incorporated as an independent entity (separating from OCLC) in 2008 that provides an open forum for the development of interoperable online metadata standards.

Janick and McLaughlin (2004) referred minimalist Dublin Core as a limitation as a cataloger had to parallel catalogue the paper copy in MARC their local catalogue and OCLC. MARC 21 is an ISO 2709-based metadata communications protocol maintained by the Library of Congress. It has two-character encoding MARC 8 (ASCII, ANSEL, selected ISO, EACC) and Unicode (limited to equivalents of MARC 8 repertoire) which no has been extended to XML expression

El-Sherbini and Klim (2004) documented the OhioLINK ETD Center using enhanced MARC from the author-submitted metadata records in OCLC and the OhioLINK catalogue. Catalogers during the transition period maintained both library-collected paper and electronic versions, the record for the paper version was enhanced and linked to the electronic version with the provisions for presenting metadata in MARC, basic Dublin Core (DC), the DC-based Electronic Theses and Dissertation Metadata Standard (ETD-MS), and Html.

On the criticism that Dublin Core had “not many elements” The Texas Digital Library started using the Metadata Object Description Schema (MODS) paving way for alternative Metadata schemes (Surratt, 2006).

Metadata Object Description Schema (MODS) is essentially MARC 21 recast in an XML-native framework maintained by Library of Congress. MODS support Text-based tags rather than

numeric ones, selected clusters of related MARC 21 attributes condensed into single MODS element. MARC 21 readily converts to MODS, but you can't do a lossless reverse conversion of MODS to MARC 21.

The lack of interoperability of different metadata schema became a barrier to cross-searching. To enable cross-institutional electronic theses and dissertations search, Hussein Suleman proposed the approach of applying the fundamental concepts of the OAI protocol to the task of building DLs by communicating using XML-coded unqualified Dublin Core as the common metadata language in the data interchange mechanism (Agosti Maristella & Costantino Thanos, 2002)

Kent State library harvests metadata from the central OhioLINK catalogue using a Perl script with the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) along with author-generated metadata for full level MARC cataloguing, giving them the advantage of instantaneous discovery and hand-built MARC (McCutcheon et al., 2008).

Kansas State University provides verified author-generated metadata complementing DSpace Dublin Core elements through a MARCit style sheet in building the library catalogue, the MARC records are uploaded into OCLC for WorldCat (Lubas, 2009).

Despite this pervasive criticism, unqualified DC also continued to proliferate via OAI-PMH (Jordan & Shearer, 2006). Many ETDs employed the qualified Dublin Core fields tailored for theses and dissertations (Jones, 2004). Dublin Core has later emerged as a practical approach to metadata language switching. over time, blended use of qualified Dublin Core with harvesting and crosswalks, and author-generated metadata have proved useful (Lubas, 2009).

2.6 Open Source Software (OSS)

Software now-a-days have become the lifeline of modern-day administrations. Libraries also need software for a parallel digital library with features which we may not find in a traditional library. There is several open source software available to create a digital library.

A general-purpose digital library like greenstone is a useful tool to provide information services in our libraries. Absence of knowledge on how to use it should not come in the way of exploiting the advantages it offers. This documentation may be used as a tool to bring in more people to the growing constituency of greenstone users. Open source software requires a greater degree of computing responsibility than commercial software.

Jancik and McLaughlin (2004) mentions a discovery need, criticizing DSpace for not making it easy to bring together a list of all the theses and dissertations completed by a certain department. The comparative table drawn by help the professionals who are planning to create a digital library.

2.7 Greenstone Digital Library (GSDL)

Greenstone is a software tool for building and distributing digital library collections. It is not a digital library by itself but a tool for building digital libraries (Witten, 1972).

Greenstone digital library software provides a way of organizing information and publishing it on the Web in the form of a fully searchable, metadata-driven digital resource. It can also be run in a non-networked environment (standalone), operating from removable media such as a USB Flash Drive and DVD (Tapkir, 2017).

Witten (2004) explains customization of digital library collections by adding documents and metadata to collections, creating new collections whose structure emulates existing ones, and building systematic collections for users to view through an end-user interface.

Sonkar, S.K. (2005) emphasis the issues related to the development of digital library of newspaper clippings using Greenstone Digital Library software.

Randhawa, S. (2012) highlights the comparison, features, function and usability of OSS like Greenstone Digital Library, DSpace.

Anuradha, K.T. (2009) enable users for full text search feature in a widely used open source library automation package viz., Koha, by integrating it with an open source digital library software viz., Greenstone Digital Library Software (GSDL), by making use of Search and Retrieval by URL (SRU) feature available in both Koha and GSDL.

K Rajasekharan (2007) has constructed a Self-Instructional Guide for Beginner's to Build up a Digital Library with Greenstone.

Verma & Kumar (2018) compares the different digital library software documents and respective technical manuals. Based on the results of the comparison, the implementation of Digital Library Management Software at DESIDOC has also been dealt in detail.

Places et.al., (2016) proposed a set of workflow strategies and system architecture that support the control and coordination of the people and tasks involved in the process of automating activities that are prone to errors and optimizes the performances.

Deb (2006) describe the setting up of the Energy and Resources Institute (TERI) Digital library, provides better single window access for researchers to access the structured information.

Alam (2010) hypothesizes a mechanism for the development of a union catalogue for geoscience theses based on open source software (GSDL, Apache, PERL, JRE), that can be used

both in offline mode (i.e. self-installable CD-ROM) and online distributed digital library for geoscience theses through the web serving.

Doctor (2007) highlights the phases involved in the creation of the digital repository. Steps involved in the Collection development using the GLI tool, “Greenstone Librarian Interface”, such as Gathering of documents into the repository, enriching them with metadata, setting up of appropriate search indexes were given emphasis for efficient retrieval of information.

2.8 Summary

In this chapter the trends in the literature of different facets of the dissertations: DL, IR, ETD, Workflow, Standards, OSS, GSDL were deliberated in brief. Different trials in the design and development process were outlooked in the perspective of resources available, existing potential, hopes, expectations, limitations justifications made. The upcoming chapter methodologies deals with the models adopted, in the process of the design development of the ETD Database.

2.9 References

- Ahmed, A., Alreyaae, S. and Rahman, A. (2014), "Theses and dissertations in institutional repositories: an Asian perspective", New Library World, Vol. 115 No. 9/10, pp. 438-451.
<https://doi.org/10.1108/NLW-04-2014-0035>
- Alam, N. (2010). *GeoTheses : development of union catalogue of Indian geoscience theses using GSDL*. <https://doi.org/10.1108/02640471211252175>
- Anuradha, K. T. (2005). Design and development of institutional repositories: A case study. *The International Information & Library Review*, 37(3), 169–178.

<https://doi.org/10.1016/j.ijlr.2005.07.006>

Anuradha, K. T., Sivakaminathan, R., & Kumar, P. A. (2011). *Open-source tools for enhancing full-text searching of OPACs Use of Koha , Greenstone and Fedora.* 45(2), 231–239.

<https://doi.org/10.1108/00330331111129750>

Anuradha, K.T., Sivakaminathan, R., (2011). 2009 International Symposium on Computing, Communication, and Control (ISCCC 2009), Singapore, pp.330-333

Boock, M., & Kunda, S. (2009). *Cataloging & Classification Quarterly Electronic Thesis and Dissertation Metadata Workflow at Oregon State University Libraries.*

<https://doi.org/10.1080/01639370902737323>

Chang, S.-H. (2003). Institutional repositories: The library's new role. *OCLC Systems and Services*, 19(3), 77–79.

Deb, S. (2006). *TERI Integrated Digital Library Initiative.*

<https://doi.org/10.1108/02640470610671213>

Dhamdhere, S. N., & Smet, E. De. (2014). *ABCD open source software for managing ETD repositories source software.* 35(4), 387–397. <https://doi.org/10.1108/LM-08-2013-0072>

Doctor, G. (2007). *Knowledge sharing : developing the digital repository of SIPS.*

<https://doi.org/10.1108/03055720710742043>

El-Sherbini, M., & Klim, G. (2004). Metadata and cataloging practices. *The Electronic Library*, 22(3), 238–248.

Gupta, D. K. (2014). *Analytical study of the ETD repositories and government initiatives for depositing ETDs in India.* 35(4), 308–319. <https://doi.org/10.1108/LM-09-2013-0092>

- Janick, S., & McLaughlin, T. (2004). Using DSpace to administer an ETD program: The Drexel experience. Retrieved March 26, 2009, from <http://www.uky.edu/ETD/ETD2004/janick/etd062004.ppt>
- Jones, R. (2004). Using DSpace for e-theses at Edinburgh University Library. Retrieved March 26, 2009, from <http://www.thesesalive.ac.uk/archive/DSpaceUGMeeting.ppt>
- Jordan, M., & Shearer, K. (2006). A metadata application profile for Canadian repositories. Retrieved March 26, 2009, from <http://www6.bibl.ulaval.ca:8080/etd2006/pages/papers/SP6 Mark Jordan.ppt>
- Lubas, R. L. (2009). Defining best practices in electronic thesis and dissertation metadata. *Journal of Library Metadata*, 9(3–4), 252–263.
<https://doi.org/10.1080/19386380903405165>
- McCutcheon, S., Kreyche, M., Maurer, M., & Nickerson, J. (2008). Morphing metadata: Maximizing access to electronic theses and dissertations. *Library Hi Tech*, 26(1), 41–57.
<https://doi.org/10.1108/07378830810857799>
- Mittal, R., & Mahesh, G. (2007). *Digital libraries and repositories in India : an evaluative study*. 2004. <https://doi.org/10.1108/00330330810892695>
- Places, Á. S., Fariña, A., Luaces, M. R., Pedreira, Ó., & Seco, D. (2016). A workflow management system to feed digital libraries: Proposal and case study. *Multimedia Tools and Applications*, 75(7), 3843–3877. <https://doi.org/10.1007/s11042-014-2155-3>
- Rajasekharan, K., & Nafla, K. M. (2007). *Building up a Digital Library with Greenstone A Self-Instructional Guide for Beginner 's Kerala Institute of Local Administration*. 22.
- Randhawa, S. (2012). Open Source Software for Creation of Digital Library : A Comparative Study of Greenstone Digital Library Software & DSpace. *Indian Journal of Library and Information Science*, 18(1), 1–10.

- Information Science*, 6(3), 45–52.
- Reeves, S., Hagen, J. and Jewell, C. (2006), "Unlocking Scholarly Access: ETDs, Institutional Repositories and Creators: Highlights of ETD 2006, the 9th International Symposium on Electronic Theses and Dissertations", *Library Hi Tech News*, Vol. 23 No. 7, pp. 12-15.
<https://doi.org/10.1108/07419050610704303>
- Sonkar, S. K., Makhija, V., Ashok Kumar, A. K., & Singh, M. (2005). Application of Greenstone Digital Library (GSDL) Software in Newspapers Clippings. *DESIDOC Bulletin of Information Technology*, 25(3), 9–17. <https://doi.org/10.14429/dbit.25.3.3655>
- Surratt, B. (2006). MODS meets Manakin: Innovations in the Texas Digital Library's thesis and dissertation collection. Retrieved March 26, 2009, from http://www6.bibl.ulaval.ca:8080/etd2006/pages/papers/SP6_Brian_SURRATT.pdf
- Tapkir, P. V. (2017). *APPLICATION OF GREENSTONE DIGITAL LIBRARY SOFTWARE (GSDL) IN COLLECTION*. 3(04), 721–729.
- Thanos, C. (Ed.). (2002). Research and Advanced Technology for Digital Libraries: 6th European Conference, ECDL 2002, Rome, Italy, September 16–18, 2002, Proceedings (Vol. 2458). Springer Science & Business Media.
- Verma, L., & Kumar, N. (2018). Comparative analysis of open source digital library softwares: A case study. *DESIDOC Journal of Library and Information Technology*, 38(5), 362–369.
<https://doi.org/10.14429/djlit.38.5.12425>
- Webley, L., Chipeperekwa, T., & Suleman, H. (2011). *Creating a National Electronic Thesis and Dissertation Portal in South Africa*.
- Witten, I. H. (1972). The Development and Usage of the Greenstone Digital Library Software. *Bristol Medico Chirurgical Journal*, 87(324), 53–56.

- Witten, I. H., Trujillo, M. F., & Bainbridge, D. (2002). *The promise of digital libraries in developing countries*. 20(1). <https://doi.org/10.1108/02640470210418218>
- Witten, I. H. (2004). Creating and customizing digital library collections with the Greenstone Librarian Interface. *Proc. International Symposium on Digital Libraries and Knowledge Communities in Networked Information Society, DLKC. 2004*, 97–104.
<http://researchcommons.waikato.ac.nz/handle/10289/1802>

Chapter 3

Methodology

CHAPTER 3

METHODOLOGY

3.1 Introduction

In this chapter we are about to discuss in detail the methodologies implemented to accomplish the objectives of the project. There are several methods incorporated widely across the world in developing a product or a system. This chapter deliberates adoption a specific methodology for particular needs in the development of the database.

A Methodology is a theoretical model of systematic process and quantitative or qualitative techniques employed in subsequent phases of development of product or a system. In other terms it is a procedure, or a set of procedures used in a particular area of study or a specific process of development.

3.2 Design and Development Methodology

The design and development of Electronic Thesis and Dissertation Database is can be contemplated as a software-based product development project. Where the substantial output of the project is a customized database for simple search and retrieval of metadata from the E-Thesis and Dissertation collection of the department.

The Database design methodology is a systematic process that helps in creating a database and developing simple, effective, appealing, and efficient search techniques to access within metadata collection. The components of the design methodology include analysis of patrons' problems and needs, designing of design model include search patterns and metadata.

Development of objective goals is done in order to meet the patrons' need; implementation includes evaluating the search patterns and the archetype exhibit of resultant metadata.

A Structural approach is necessary in the design and development of the database because the user requirements and the system complexity increase in the process of project developments. An appropriate development model may be used, taking into account to meet the users' expectations and needs. There are various design models that are used by the designers and the developers of Digital library.

The major system development models along with brief overview is given below:

1. Cleanroom Model
 2. Component Based Model
 3. Hacking Model
 4. Incremental Model
 5. Prototyping Model
 6. Spiral Model
 7. Waterfall Model
 8. ADDIE Model
1. *Cleanroom Model*: Cleanroom model is based on error prevention rather than error correction.
 2. *Component-Based Model*: Component-Based Model is used in object-oriented technologies. This emphasis on creation of software classes and the classes can be reused across different applications.
 3. *Hacking Model*: Hacking model is a non-symmetric approach to software development. This model does not have specifications and leads to high cost in later stages

4. *Incremental Model:* Incremental model combines the elements of waterfall model, which are then applied in an iterative manner. It delivers a series of releases called “increments” that provides progressively more functionality for the client as each increment is delivered.
5. *Prototyping Model:* The prototyping model is used when client states the requirements in general nature. This type of system is used when the developer is unsure of the stability of the system feature, user interface, programming environment etc. The various steps in this model include, requirement gathering, develop “Quick Design”, prototype and evaluation.
6. *Spiral Model:* The spiral model is same as that of Incremental model, with more emphasis placed on risk analysis. The spiral model has four phases: Planning, Risk Analysis, Engineering, and Evaluation.
7. *Waterfall Model:* This model is also called “Classic Life Cycle” or “Waterfall Model”. Waterfall model is sequential design process often used in software development processes, in which progress is seen steadily downward (like a waterfall) through the phase of conception, initiation, analysis, design, construction, testing and maintenance.
8. *ADDIE Model:* The ADDIE model is same as that of spiral model or iteration model with succession of iteration of the five phase - Analysis, Design, Development, Implementation, and Evaluation. Where the process repeatedly passes through these phases in iteration. Requirements are gathered during analysis phase, in design and development phase the workflow, and the prototype are produced, the system is developed in the implementation phase. The evaluation phase allows the user to evaluate the output of the system through continuous iterations.

3.3 ADDIE Model

In the design and development of this database, the model employed is the ADDIE model in order to structure all the components of the design.

The ADDIE Model is a product development concept enunciated by Florida State University in 1970s'. The abbreviation of ADDIE stands for Analysis, Design, Development, Implementation, and Evaluation. It is a generic process that is widely used in instructional design.

In the design and development of ETD Database, the basic framework of ADDIE model and its concept of succession of Iteration of stages has been adopted to construct a system that meets users' needs and facilitates digital experience in accessing thesis and Dissertation Collection of the department.

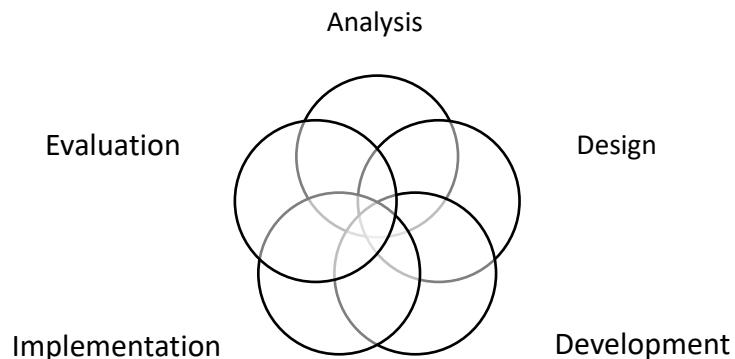


Figure 3.1 ADDIE Model

The ADDIE model incorporates the five phases —Analysis, Design, Development, Implementation, and Evaluation—represent a dynamic, flexible guideline for building effective database. Each phase has an outcome that feeds into the subsequent phase. Each phase in ADDIE model is related to and interacts with each other.

Analysis > Design > Development > Implementation > Evaluation

The ADDIE model, along with the use of succession of Iteration and rapid prototyping i.e. the concept of receiving continual or formative feedback while the database is being created, makes this model is economic as well as time saving by predicting problems while they are still easy to fix.

3.3.1 Analysis

Analysis is the key phase in this process. When the analysis phase is done before creating the plan, developing, or even implementing, they save considerable amount of effort, and time. In the analysis phase, the research problem is clarified, the design goals and objectives are established and the digital platform and user's existing skills, objectives and needs in accessing e-resources are identified.

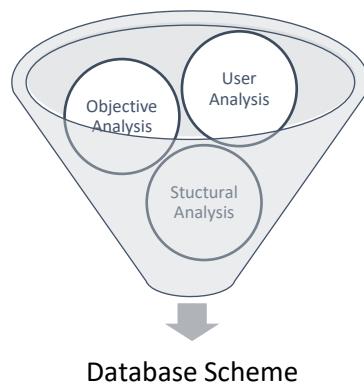


Figure 3.2 Courses of Analysis Phase

In *User Analysis*, the prospective users of the database, and their prior knowledge about the digital facilities to access the database must be identified, so that the plan can build on what they know, what information the audience needs and how much. Their objectives need and problems of the users must be known in advance. In order to identify that, the pre-assessments of the audiences must be done to collect data about them.

In this instance, as the database is accessed in offline mode, the potential users of the database can be the professors, students and the administrative staffs of the department of Library and Information Science, Pondicherry University. The sorted user group, in general have a good knowledge on metadata, and a decent basic skill on digital library platforms, and information retrieval. From the observations made, some of their objectives and needs in accessing the thesis and dissertation collection in the department are

1. To get an overview on patterns in the researches made over the years.
2. To retrieve different metadata components for specific topics of interest.
3. To access Index and Abstract content of the specific topics of interest.
4. To access Bibliographic flow in the research made over particular subjects.
5. To avoid repetition of research work that has already been done.

And the problems that the users come across in accessing the physical collection of thesis and dissertation in the department are

1. The Huge Size of the collection demands substantial effort to access a particular content.
2. It's physically hard to sort and filter for specific subject.
3. The real time physical search can lead to human error.
4. The whole process is time consuming.
5. The administrative staffs find it intricate to carefully preserve and avoid mishandling of the primary manuscripts.

In *Objective Analysis*, the evident goals or objectives for the specific ETD Database are identified. The clearly identified specific goals, will save much in time and sources. In fact, the specific objectives in constructing the database is discussed in the Chapter 1: Introduction.

In *Structure Analysis*, once the specific goals are identified, all the important steps and all the details that need to achieve the goals are formalized. This phase needs to edit back and forth.

As the specific objectives are identified, the in-depth probe on constructing the ETD Database is seen as a streaming flow of physical collection to digitized metadata for structured storage and controlled retrieval. Through digitizing the physical collection of the intellectual output of the department and categorizing the intellectual materials which are digitized by extracting the index, abstract, bibliographic details and other metadata in devising the electronic thesis and dissertation (ETD) database and assessing benefits and challenges in organization and management of the intellectual output.

3.3.2 Design

Design phase is the integral part of developing a system. Design is really about applying the instruction. This phase focus on how design can really be effective in ways that facilitate user and their interaction with the digital materials that are created and provided access. Design phase helps in preparing the system design specification document from the user's requirement. Furthermore, in design phase the designer evolves and emphases on designing a prototype, reviewing prototype, resolving defects, renovating modules and succession of Iteration of steps, in order to make the system more effective. The common design phase modules include,

- Creating Workflow.
- Selection of Digital Library Platform.
- Evaluation Design.

3.3.2.1 Workflow Creation

A workflow is a sequence of steps toward the accomplishment of a pre-defined task in achieving specific goals. Workflow is in general used to manage repetitive processes. It is expounded as a tool for communicating with internal stakeholders, the administrators, patrons who are not directly involved in reformatting collections, as well with designers outside of the organization, the sequence of steps and their specific roles in the process. It also helps identifying patterns, gaps, and bottlenecks to enhance efficiency and improve work dynamics.

The basic flow of Work in the process of digitization, digital content creation, creation of digital libraries, digital archives, Digital Content delivery, database creation, evolves over period through numerous tailored customizations for particular needs and objectives of specific systems.

Based on the certain workflow models that was earlier discussed in the Chapter 2: Literature Review, here is a new workflow layout that has been specifically customized for the design and development of The ETD Database.

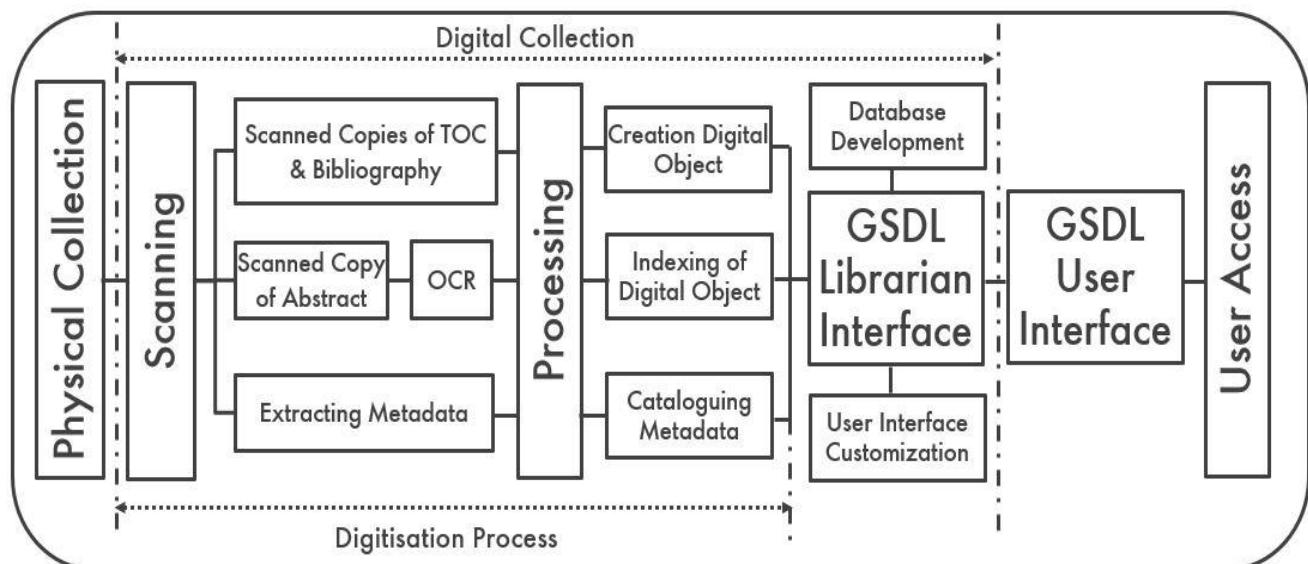


Figure 3.3 The Customized Workflow Layout in the Development of ETD Database

3.3.2.2 Selection of Digital Library Platform

A digital archive is an organized collection of digital documents, in a format such as image, text that would permit easy access by information users and proper maintenance by librarians. Digital libraries do not replace the concept of libraries as a repository of knowledge with a notion of it as provider of information services to users.

The Greenstone Digital library software offers exciting ways to build and distribute digital document collections. It helps us to publish digital collections on the internet or CD-ROM. GSDL comes with tools that build full-text search indexes and browsing classifiers for any collection of digital documents. Once initiated, the collection building process will take place mechanically, running into several hours or days for a very large collection. Creating digital objects, organizing them into focused collections and making the materials accessible to others can be a prime application area of digital libraries.

Justification for Selection of GSDL

Greenstone is used for building the content repository after doing a comparative analysis of three open source digital library software's DSpace, EPrints, and Greenstone as in table

Table 3.1: Comparison Matrix of Different Digital Library Software

Features	DSpace	Greenstone	EPrints
Open Source	Yes	Yes	Yes
Unicode Compliant	Yes	Yes	Yes
Language Interface (Existing)	No	Yes	No
Built in HTML viewers	No	Yes	No
Size	6 MB	69.7 MB	2 MB

Server	UNIX/Windows	UNIX/Windows	UNIX
Offline Viewing/Browsing	No	Yes	No

The features supported by greenstone that were found useful are as follows⁶:

1. *Support for Digital formats and content types:* Plugins distributed with Greenstone processes plain text, HTML, Word and PDF documents, Power Point presentations, images, spreadsheets, Email Messages, etc. New plugins can be customized for different document types.
2. *Access and Navigation:* The Users access the library via a future rich iconic interface. The collections are listed on one interface and the user can navigate easily through the collections. The user has also the option of changing the language, search and presentation preference through web-based interface. The documents are uploaded through a librarian interface GUI that allows the librarian to assign metadata, choose plugins, assign classifiers for browsing indexes, etc. The Documents can be grouped by titles, authors, date, organization, etc.
3. *Search and Retrieval:* Greenstone uses Managing Gigabytes (MG) for compression, indexing the textual information in the collections. It has a facility for “cross-collection searching” which allows several collections to be searched at once, with the results combined behind the scenes as though you were searching a single unified collection. Any of the open collections to be included in the searches.
4. *Technology Platform:* The system is designed to run on UNIX, Windows, and Mac OS/X servers and comprises other open source middle ware and tools. The code is in the JAVA and C++ programming Language. All versions of greenstone use the GNU Database Manager, GDBM. It is supplied with all windows versions of greenstone and installed automatically

⁶ <http://www.dil.iitb.ac.in/docs/Digital%20Library.pdf>

during the installation comes in 2 flavors – the local library with a build in web server that can be used for sharing the library on the intranet and the web library version for sharing on the internet.

5. *Compliance to standards and protocols:* Unicode is used throughout greenstone. This allows any language to be processed and displayed in a consistent manner. Collections have been built containing Arabic, Chinese, English, French, Maori, Hindi, Spanish, etc. Besides, there are standard templates available in languages like Hindi, Kannada, Guajarati, etc. Greenstone also follows the OAI Protocol for metadata harvesting and the documents are tagged with metadata that follow the Dublin core standard. It also adheres to the Z39.50 Information Retrieval Protocols.

The library can be saved on the computer and viewed even when not connected to the internet and distributed in CDs. Various formats are supposed(such as Adobe pdf, MS Power point, MS Word etc.) and can be viewed in the library even without having the applications installed on the computer (i.e. Acrobat reader, MS Word etc.) and can be viewed in the library even without having the applications installed on the computer(i.e. Acrobat Reader, MS Office are not necessary).

3.3.2.3 Evaluation Design

Testing and user evaluation are quality control mechanisms to verify that the final product confirms to the software requirements. Evaluating a system is very essential but before, that it is important to know how you will evaluate this system when working on the design of the system. To do the evaluation effectively, the data that had been already collected from the former stage has to be used, which is Analysis phase. Along with that, the evaluation must be strongly related to the

content and context on which the system is build. Here the final product of the system must confirm the user requirements and also meets their expectations. The methodology adopted for evaluation design has been discussed in detail in the fore coming section ‘Evaluation’ from this same chapter.

3.3.3 Development

This Development phase is based on the first two phases, which are the analysis and the design phase. That implies, if these phases were appropriately formulated, the development shall be economic as well as effortless. In this third phase, the designers integrate the technology with the theoretical setting and process. Moreover, the designer starts to develop and create a good quality prototype for the design, the metadata, and run through of the digital library software.

3.3.3.1 Prototype Creation

In this step samples of the system Is created in order make sure that system is executed in the right track. In this project, a prototype is developed with our workflow that was created in the previous phase (the design phase).

This way, the feedback about the progress of the system development and working to improve the prototype can be achieved. This step is important as the feedback from the prototype of the system that the design is created for, the implementation can be done smoothly, and the shortcomings can be identified without having to redo all the process. Since there are some distinct strength and weakness in every system. This step helps us identify these strengths and weakness that need to be revised.

3.3.3.2 Extraction of Bibliographic data

On one hand the prototype has been developed and design is satisfactory, on the other hand it is as important to start building competitive digital corpus of the metadata extraction. Which would be included the system along with the prototype that was made earlier. The process of digitization of the content along with extraction of bibliographic data from physical collection of thesis and dissertation is discussed in detail in the fore coming chapter 4: Design and Development.

3.3.3.3 Valuation of Prototype

After we are done from developing digital corpus of the metadata extraction. Valuation of the prototype is ready to start, which is a run through of the design, like a practice run or a pilot test. This step is the actual time to do rehearsal for database using all materials that had been created. The user has to be actively involved and use the exact same conditions of the real-time as if use the database. In this phase furthermore, attention must be paid to, the assessment made from feedback and suggestions of the user in order to find out weaknesses and improve the system.

3.3.4 Implementation

This phase is about transforming the prototype into active system. In order to go through this phase, two major steps must consider, which are Creation of the complete database and Customization of the User Interface Design.

3.3.4.1 Creation of the Database

In the implementation phase, the real time digital corpus of the metadata extraction is built into an active database through the system designed and prototype tested using, Greenstone's "librarian" interface (GLI) of Greenstone digital library software package.

This allows the system to collect sets of documents, import or assign metadata, and build them into a Greenstone collection. This includes four basic activities, which can be interleaved but are nominally undertaken in this order:

1. Copy documents along with their digital corpus of the metadata from the computer's file space, into the new collection.
2. Enrich the documents by adding further metadata to individual documents or groups of documents.
3. Design the collection by determining its appearance and the access facilities that it will support.
4. Build the collection using Greenstone.
5. Preview the newly created collection.

The librarian interface also allows to add "external" metadata to documents, metadata that pertains to the document as a whole. But documents need to be structured into sections and subsections, and "internal" metadata might be associated with each part.

3.3.4.2 Customization of the User Interface

Customization of the collection by determining its appearance and the access facilities is determined in this phase. The Design panel in the Greenstone's "librarian" interface (GLI) allows to specify the structure, organization, and presentation of the collection being created. This step

involves a series of separate build in, each dealing with one aspect of the collection design. In effect, it serves as a graphical equivalent to the usual process of editing the configuration file manually.

With these two steps ETD Database can be displayed in very active and authentic ways to achieve the implementation phase. The detailed manifestations of implementation of the design using Greenstone Digital Library Software is described in the fore coming chapter 4: Design and Development.

3.3.5 Evaluation

The final phase in ADDIE model is Evaluation phase. Evaluation is to find out what the user wants and what problems they experience while using (Jenny et. al., 1998). Evaluation is the process by which the interface is tested against the needs and practices of user (Faulkner, 1998). It is very important to evaluate each step-in order to make sure that we achieve our goals using the system design and extracted metadata to meet the user needs. In evaluation, the attitude of the users towards the system are usually measured based on their feedback.

Additionally, when we talk about evaluation, we talk about two types of evaluation, which are formative evaluation and summative evaluation. First, formative evaluation is a continuing process that we do as we are working on our system design in each phase on ADDIE model. In the summative evaluation, effectiveness of all the components of the design is evaluated, after the implementation of design.

Dix et. al. (1998) identified three main reasons for doing evaluation:

1. To access the extent of the system's Functionality.
2. To identify specific problems with the system.

3. To access the effect of the interface on user.

In this project the evaluation of the offline ETD Database is subjected to be conducted in laboratory condition as a combination of formative and summative evaluations. The process is discussed in detail in the chapter 5: Evaluation.

3.4 Summary

This Chapter covered the various methodologies available in the design, development and evaluation of a system in general and the framework of ADDIE model that had been specifically adopted for the design and development along with rapid prototyping model for the evaluation of ETD Database. The next chapter illustrates the implementation of the models utilized in the design and development discussed in this chapter.

Chapter 4

Design and Development

CHAPTER 4

DESIGN AND DEVELOPMENT

4.1 Introduction

This Chapter discusses the design and development of the Electronic Thesis and Dissertation Database. This ETD is built with bibliographic data collected in the form of Text, Images, and PDFs that are digitized from the physical hardcopies of Thesis and Dissertation collection of the department.

The process of digitizing materials involves different paces depending upon material, technology and requirement. Various technical requirements, like hardware and software, file formats and file compression and then the post processing requirements for making the digitized copies of bibliographic data are discussed here. Later the bibliographic pile is compiled into an accessible through User friendly Digital Library Interface using Greenstone Digital Library Software.

4.2. Digitizing Collection

The Core base of the ETD is the digitized collection of the bibliographic data upon which the database is build. “Digitization is the process of renovating the content in the physical media to the digital media.”

It is a method of translating a piece of information such as book, journal articles, sound recordings, pictures, audio tapes or video recordings, etc. into bits. While Bits are the fundamental units of information in a computer system.

From the printed material an image of the physical media is captured using a scanner or digital camera and altered into a digital format that can be electronically stored and accessed using

computer or mobile devices. In general, hard copies can be converted into digital form in three different means:

- 1) Keying
- 2) Scanning and capturing
- 3) OCR

The keying costs way more than scanning and capturing as image files in terms of time and energy. However, if we could convert them into OCR that would be comparatively be cost-effective.

Scanning technology has improved significantly over the years in terms of speed and resolution. There are different types of scanning devices. In general, these scanning machines are based on either of Charge Couple Device (CCD) technology, in low end devices uses Contact Image sensor (CIS) technology, in some high-end devices Photo Multiplier Tube (PMT) technology or CMOS (Complementary Metal Oxide Semiconductor) sensing technology is used.

The scanners operate by illuminating the document and directing the reflected light through a series of mirrors and lenses onto photo sensitive element. The photo sensitive element could be CCD, CIS, PMT or CMOS based technology depending on the type of the scanners. Light sensitive photo sites arrayed along the photosensitive element are converted into electronic signals which finally processed into digital image.

In a broad-spectrum, the book scanners, the face-up scanning system are used for scanning large-sized books, ledgers and other bound materials in libraries and archives. They are sophisticated techno device that almost exclusively uses Charge Couple Device (CCD) sensors for capturing images along with Adobe Acrobat for post processing and are usually pricey.



Figure 4.1: The Konica Minolta PS 7000 book scanner

As an alternative for these high end book scanner, a relatively economic DIY book scanner was build using hand held 13-megapixel f/2.2 aperture mobile camera that enfolds the combination of the low-noise performance of Charge Couple Device (CCD) and high quality CMOS sensors mounted on a tripod for stability and coupled with Adobe Scan Software along with Adobe Acrobat through adobe document cloud for post processing. This setup was used in digitizing the bibliographic content of thesis and the dissertation collection.

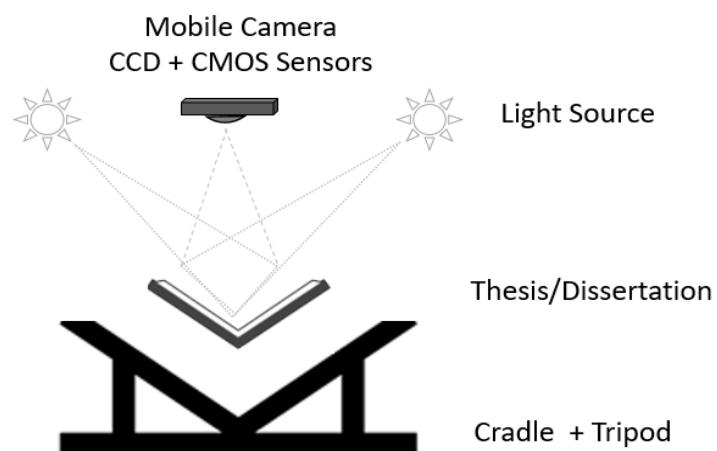


Figure 4.2: DIY Scanner Setup

4.2.1 Tools for Digitization

Hardware Requirements

- Handheld mobile camera with CCD and CMOS sensors
- External Light Source
- Tripod
- Book Holder or Cradle

Software Requirements

- Adobe Scan
- Adobe Document Cloud
- Adobe Acrobat Reader

Adobe Scan

Adobe Scan is an on-the-go scanner, manufactured by Adobe Systems. Adobe Scan is a free software that essentially transforms any standard device such as a mobile phone or a tablet into a portable scanner. It automatically uploads the images scanned into a PDF format. These can then be stored in a separate folder or sent to another recipient through Adobe Document Cloud. Its user-friendly design has automatic functions such as image sharpening and the detection of borders will be handled by algorithms as opposed to be the responsibility of the end user. The content can be further attuned with the help of Adobe Acrobat Reader.

Adobe Document Cloud

Adobe Document Cloud is Adobe's Industry Standard document suite for Full Cloud integration. It consists of all the major desktop applications, including Adobe Scan and Acrobat Reader. It includes Lots of improved Cloud related features. The applications can be simultaneously accessed on your desktop as well as handheld device. There are also number of

new cloud related features that have been added to the suite including file syncing, feedback, fonts, settings, and updates. The added cloud dimension supports the extensive collaboration in a distributed environment.

Adobe Acrobat Reader

The basic Acrobat Reader is freeware, the main function of Adobe Acrobat is creating, editing and viewing PDF documents. It also supports printing and annotating of PDF files. The commercial proprietary Acrobat, can create, edit, convert, digitally sign, encrypt, export and publish PDF files. It can import document and image formats, it is also possible to import a scanner's output, a website, or the contents of the Windows clipboard and save them as PDF.

Once a PDF document is created, its natural organization and flow cannot be meaningfully modified. In other words, Adobe Acrobat can modify the contents of paragraphs and images but doing so does not repaginate the whole document to accommodate for a longer or shorter document. Acrobat can crop PDF pages, change their order, manipulate hyperlinks, digitally sign a PDF file, add comments, redact certain parts of the PDF file, and ensure its adherence to such standards as PDF/A. It available for both desktop and mobile platforms.

4.2.2 Digitizing Process

In the process of digitization, the primary step is to capture the documents available in print for conversion into digital form. In the case of print-based material, it is the hard copy of the document which needs to be scanned and digitized. The hard copy in here is a paper-based document i.e. the physical copy of Thesis and Dissertation. Capturing devices for print based material requires scanners and digital cameras attached with a computer.

In the current study, DIY Scanner Setup is deployed as the hardware tool for scanning and digitizing the documents. A simple mobile camera that holds the specific scanner sensors was engaged in capturing the images of thesis and the dissertation collection.

While software verso of the setup was carried by Adobe scan software. The Adobe Scan software, captures and converts documents into high-quality Adobe PDFs. The scanned copies of physical collection that are stored in the form Adobe PDFs can be accessed from any computing platform that holds access to Internet and Adobe Acrobat Reader software through Adobe Document Cloud. The documents scanned using Adobe Scan Software are simultaneously processed by editing and indexing the document with Adobe Acrobat Reader software.

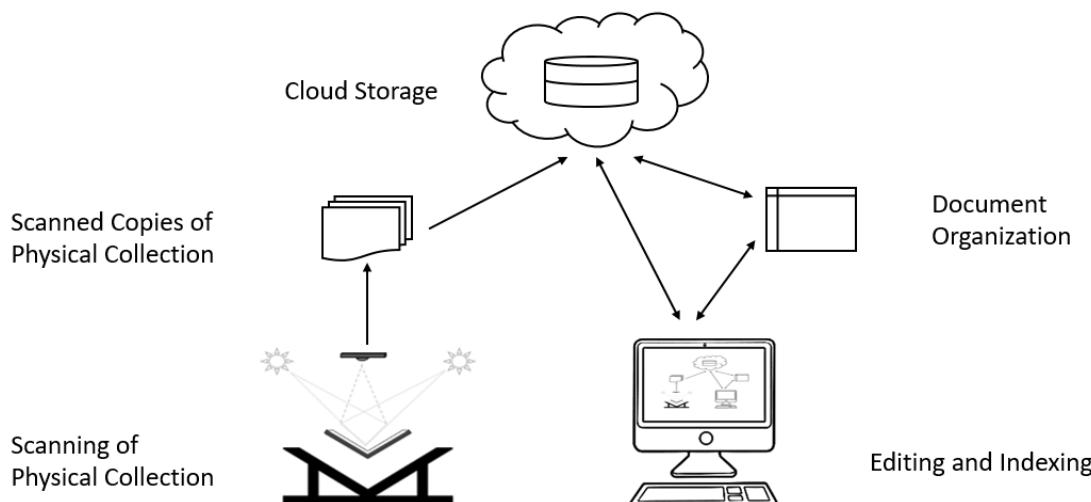


Figure 4.3: Process of Digitization

The Adobe Acrobat Reader treats the scanned PDFs by sharpening the edges, removing noise like shadows and glare. As the Greenstone Digital Library software versions accepts PDF only of versions up to 1.4, the scanned pdf documents gets delimited back to version 1.4 using Adobe Acrobat Reader software and then the title individual PDF documents gets split into Title, Abstract, TOC, Bibliography PDF compilations for the ease of access to the patrons. Later the

compilation of PDFs are indexed at the document level as well as folder level for a well-defined organization the document while feeding it into the Greenstone digital library collection.

The Adobe Document Cloud stands as an online dais to store and share the scanned PDF files in the cloud and to access them on cross computing platforms of scanning as well as devices performs editing and organizing documents.

4.3 Gathering Metadata

The Dublin Core Metadata components are extracted from the physical collection of Thesis and Dissertation through manual scanning of the documents and recording it in spreadsheet.

Record	Registration No.	Author	Title	Area of Study	Supervisor/Guide
2	1	1076101 ARUNAGIRI,S.	Government Arts and Science College Libraries in Pondicherry - An Analytical study	Academic Libraries	Dr. M. Bavakutty
3	2	1076103 JAYABARATHAN,A.	A Study of Pondicherry University Library	Academic Libraries	Dr. M. Bavakutty
4	3	1076104 KARUJINKAR,N.	Collection development: A case study of Pondicherry University Library.	Library Management	Dr. M. Bavakutty
5	4	1076105 RAVI,K.	Digital infrastructure facility in the Medical college libraries in Pondicherry	ICT	Dr. M. Bavakutty
6	5	1076106 SUSANTA NARZARY	Use of Electronic Information Resources by the Science Doctoral Students in Pondicherry University: A study	E-Resource Management	Dr. M. Bavakutty
7	6	1086101 ARIVOLIJ.	Research productivity of Faculties in Pondicherry University: An Empirical Analysis	Infometrics	Dr. R. Sekukan
8	7	1086102 NARAYAN BASUMATARY	Open Access Institutional Repositories in Asia: A Webometric Analysis	Webometrics	Dr. Rekha Rani Varghese
9	8	1086103 NAYAN KUMAR BORO	Computerized Information Service in Ananda Rangbillsi Library, Pondicherry University: A Users Study	Library Services	Dr. K. Anas
10	9	1086104 RAJESH	Awareness of Wi-Fi Infrastructure: A Case of Student Community in Pondicherry University	ICT	Dr. Mangkhollen Singson
11	10	1088105 RAU BRAHMA	Scholarly Publications in India Journal of Pharmacology during 2006 to 2008: A Citation Analysis	Infometrics	Dr. M. Leeladharan
12	11	1088106 SANJEEB KUMAR BRAHMA	Digital Information Literacy among PG students and scholars of Pondicherry University: An analytical study	Information Literacy	Dr. R. Sekukan
13	12	1088107 SANSUMWI BORO	Open Access Journals in Social Sciences: A Webometric Analysis	Webometrics	Dr. Rekha Rani Varghese
14	13	1088108 SANTHI,N.	Users Perception on the Utilization of E-Resources in Pondicherry University Library: A Case Study	User Studies	
15	14	1096101 ANITHA,N.	Research trends of biology in selected Central University of India: A Scientometrics	Scientometrics	Dr. R. Sekukan
16	15	1096102 SUJITHA,B.	Public Library Systems and Services in Pondicherry: A State-of-the-art Report	Public library	Dr. K. Anas
17	16	1096103 HIMABINDU DRAKSHARAM	Evaluation of E-Book enterprise	E-Resource Management/Bibliometric	Dr. Chennupatti K. Ramala
18	17	1096104 JOICY, A. J.	The websites of Research and Development institutions in India: A webometric study	Webometrics	Dr. Rekha Rani Varghese
19	18	1096105 SHEJINA SREENIVAS	Open Access Journals in Medium and allied subject	Information Sources	Dr. M. Leeladharan
20	19	1096106 SURESH,G.	Information Search Pattern of Research Scholars in Pondicherry University	Information Retrieval	Dr. Rekha Rani Varghese
21	20	1096107 SUNKARA RAMARAJNA	Online social networking in higher education: A survey on Pondicherry University students' experience	Social Networking	Dr. Mangkhollen Singson
22	21	1096108 VINODA BHARATHA	ICT literacy among Library Professional in higher academic institution in Pondicherry: A study	Information Literacy	Dr. Mangkhollen Singson
23	22	1106101 BHUVANESH,N.	Pondicherry University Library Portal: An Influence of Image on user perception	Web Technology	Dr. Mangkhollen Singson
24	23	1106102 BREMALATHA,M.	E-Reading Habits of the Student of Pondicherry University: A Survey	User Studies	Dr. Chennupatti K. Ramala
25	24	1106103 BREMNATH,V.	Use of E-Resources by Medical Science Students in Pondicherry Institute of Medical Sciences (PIMS)	User Studies	Dr. Rekha Rani Varghese
26	25	1106104 DEEPTHI,S.	Library Automation in Higher Educational Institute in Southern Kerala: A Survey	Library Automation	Dr. R. Sekukan
27	26	1106105 GNANASELVI JOY,M.	Scholarly Communication attitude and Behavior of Faculties in Pondicherry University	User Studies	Dr. Mangkhollen Singson
28	27	1106106 ILLAMMARAN,A.	Open source Library Management Software Literacy among the Library Professional of Professional Colleges in Pond	Library Professionals	Dr. M. Leeladharan
29	28	1106107 JEEVA,G.	Research Trends in Information Literacy : A Study Based on LISTA	Infometrics	Dr. Rekha Rani Varghese
30	29	1106108 VAIJUJANAN,V.	Awareness and Acceptance of E-resources among Social Science Researchers in Pondicherry University		Dr. K. Anas

Figure 4.4 Excel sheet

The extracted metadata components are categorized into year wise tables to facilitate piling metadata in a hierarchy that defines year wise contribution of the intellectual output made by the Department of Library and Information Science, Pondicherry University.

In the conversion of Excel list to Greenstone, the attributes of each object become metadata in categories used for access or classifying the current collection and future acquisitions. Further,

when Greenstone's metadata sets are used in building the collection, these classifications will be recognized and can be accessible to user interface of greenstone.

To enable this conversion from Excel list to Greenstone, a set of systematized protocol is followed⁷. They are

1. Create a Standard Excel Database
2. Import to CSV format
3. Review Standard Greenstone Metadata Categories
4. Adding Metatags to Greenstone
5. Importing the Database

Create a Standard Excel Database

For a standard list to be used as a "database", All the entities, in the list should follow some standard formats.

1. Each line is a single Record with data on the one person, item or other entity on that line.
2. Each column is headed by a descriptor that is termed a "Field" and all the data in that column has similar characteristics.

Import to CSV format

The purpose of this process is to create hundreds of (null/empty) records, with each of their data elements a new metatag, ready to be joined to a document, photo or other digital item. i.e. The "Title" field created in the excel list can be converted into a key index in the gr. For which the database must be "exploded" in the greenstone librarian interface. An Excel file cannot be

⁷ <http://frazernash-usa.com/digital/Greenstone-Imports.htm>

"exploded", but such a file can easily save to a "comma-delimited" (csv) file. This can be done by importing the Excel file, by "saving as" the "csv" format.

Review Standard Greenstone Metadata Categories

"Dublin Core Metadata Standard", which is the basic classification scheme used in Greenstone and widely recognized by digital libraries. With Dublin core dc. is prefixed with name of the metadata element such as Title or Publisher etc. There are 15 metadata elements in Dublin Core Metadata Elements Set (DCMES),

The Dublin Core basically consists of these elements:

- Title
- Creator
- Subject
- Description
- Publisher
- Contributor
- Date
- Type
- Format
- Identifier
- Source
- Language
- Relation
- Coverage
- Right

Subsequent to the specification of the original 15 elements, 55 exemplary terms extending or refining the DCMES in the Qualified Dublin Core Metadata Elements Set (QDCMES).

In the ETD Database “Qualified Dublin Core Metadata Elements Set” was used as the classification scheme.

Adding Metatags to Greenstone

In Greenstone, each Dublin Core element is prefixed with "dc.", so they appear as dc.Title, dc.Creator, etc. Because these elements are widely accepted and recognized, it is a good

idea to match your field names to the Dublin Core elements, insofar as that is possible. For the ETD Database, the mapping scheme employed is as follows:

Table 4.1: Mapped Dublin Core Metadata

Archive File	Dublin Core Metadata
Title	dc.Title
Area of Study	dc.Subject
Abstract	dc.Description
Year	dc.Date
Registration Number	dc.Type
Record Number	dc.Identifier
Source	dc.Source
Dissertation/Thesis	dc.Format
Author	dc.Creator
Guide	dc.Contributor
	dc.Publisher
	dc.Language
	dc.Relation
	dc.Coverage
	dc.Rights

It is not all Dublin Core metatags are mapped from the file scheduled to be imported; unused metatags can be added after the import, directly in Greenstone. The "mapping" as done by

renaming the column headings to the relevant Dublin Core element or a new metatag that has to be added to Greenstone.

Importing the Database

The records in the excel list, with the data elements were the new metatags, to be tagged thesis and dissertation PDF copies. Once the database was "exploded" i.e. the data elements from the excel list were imported to the greenstone digital library software as null/empty records, The "Source metadata element" were merged with the "Target metadata set" which is "Dublin Core", "Creator" or "Date", by selecting a "Target metadata element" and Merge.

Greenstone continue to go through each field name and new folders in the Collection panel, with names derived from the name of the imported file. With a restriction that each folder is limited 100 (or less) records. Hence the records were sorted in a year wise folder whose record count does not exceed 35-40 each. In the "Enrich" panel, the metadata for the record in each metadata element were manually enhanced. The new folders created by the import process were renamed (or create new one), such as "2019", "2018", etc. and the records were moved into the new or renamed folders. The newly scanned documents/PDF copies of thesis and dissertation were dragged into the appropriate folder and each matched with its correct metadata.

The metadata can either be selected from the existing imported metadata for the scanned documents or the "null" record can be replaced with the actual document. In either the Gather or Enrich panels, just by right clicking the "null" record, and browsing for the actual document. The document can be merged with the metatags.

4.4 Development of ETD

4.4.1 Software Requirement and System Configuration

Software Requirements to run Greenstone Digital Library:

- Windows/Linux Operating System
- Apache Webserver
- Java Run-time Environment (JRE)
- Web Browser
- Greenstone Digital Library Software 2.87

System Configuration to run Greenstone Digital Library:

Install Java Run-time Environment (JRE) version 1.4 or above before installing GSDL as JRE is required for Greenstone Librarian Interface (GLI). Apache Web Server Recommended. PERL gets installed automatically. C++ compiler, to compile the source code for Customization.

Storage Capacity:

- 50MB for a binary installation
- 155MB for compiling Greenstone from source code
- 200MB for optional Greenstone demonstration collections
- 5MB for documentation
- 24MB for Greenstone's "CD exporting" function

4.4.2 GSDL Installation

Alternative choices for getting the GSDL software:

1) UNESCO CD-ROM (version 2.70) or FAO IMARK CD-ROM, (but this is an earlier version 2.51) which contain the Greenstone software, plus documented example collections, four

language interfaces (English, French, Spanish, Russian), the Export to CD-ROM package, the ImageMagick graphics package, the Java runtime environment, and an installer that installs all of these.

2) IITE Digital Libraries in Education CD-ROM, or a Greenstone workshop CD-ROM. This CD-ROMs contains the tutorial exercises and a set of sample files to be used for these exercises apart from the requisite software listed above.

3) Download directly from <http://www.greenstone.org> that contains the latest version of Greenstone. Downloading directly needs Java⁸ to run Greenstone. To work with image collections, ImageMagick⁹ is needed.

The following steps need to be carried out to install Greenstone:

- 1) Install the Java 2 Runtime Environment (latest version).
- 2) After installing J2RE, go for GSDL folder choose setup gsdl 2.70.
- 3) Choose setup Language. English (US) is the default. We choose English
- 4) Welcome to the Install Shield Wizard for the Greenstone Digital Library Software. Click <Next>
- 5) License Agreement. Accept the agreement and then click <Next>
- 6) Choose location to install Greenstone. Leave at the default and click <Next>
- 7) Setup Type. Leave at the default (Local Library) and click <Next>
- 8) (For older installers you must now select collections. Leave at the default, Documented Example Collections, and click <Next>)
- 9) Set admin password. Choose a suitable password and click <Next> (If your Computer will not be serving collections online; the password doesn't matter)

⁸ <http://java.sun.com>

⁹ from <http://www.imagemagick.org>

- 10) Click <Install> to complete the installation
- 11) Files are copied across and Installation is complete.

If the installation is from a CD-ROM, the installer will offer to install ImageMagick, and Java.

To invoke the Greenstone Reader's interface, go to the Greenstone Digital Library Software item under Programs on the Windows Start menu and select Greenstone Digital Library. To invoke the Librarian interface, go to the same item and select Greenstone Librarian Interface.

4.4.3 Interface Design

GSDL comprises two interfaces, the Librarians Interface and the Website which serves as the user interface. The “*Librarian Interface*” in GSDL is for creation, management and updating collections. It is programmed in JAVA highly based on creation of the necessary commands. The *Website* is served by internal www-server or Apache. Webpages are created by Perl and Java Servlets which is customizable via CSS and text-files.

Customization of GSDL

The first step towards building this ETD Database is to customize the greenstone software according to requirements.

The Greenstone Librarian Interface is a tool for collecting and marking up documents, then building digital library collections. It provides access to GSDL software functionality from graphical point. The macros are needed to be modified to create a customized digital library. A JAVA-PERL applet¹⁰ provides an interactive graphical interface for the Greenstone Librarian Interface with the following main macros/functions

¹⁰ gliserver.pl

- Download
- Gather
- Enrich
- Design
- Format
- Create

4.4.4 Creation of New Collection

To create a new collection, open the file menu and choose new option. Several fields are needed to fill out for new collection. The following screens will appear as follows. While creating the collection the default directory can be changed, while the collections can be stored. After creating the collection, these steps have to be followed to create the ETD Database.

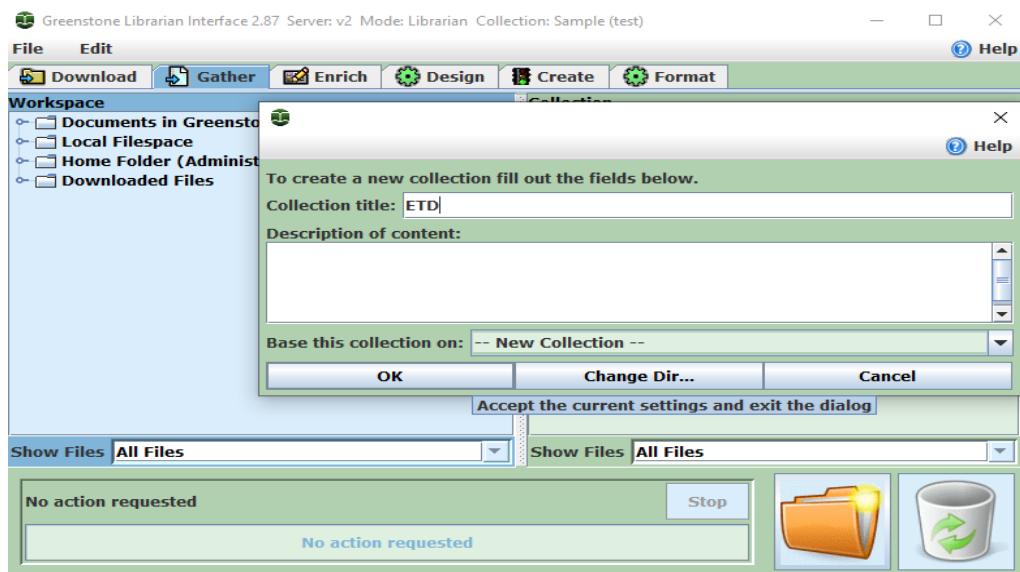


Figure 4.5 Starting a new collection

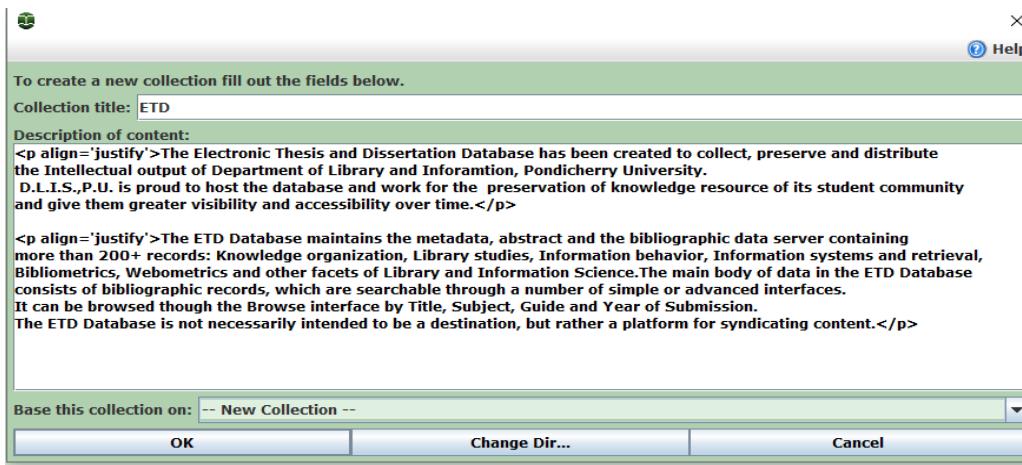


Figure 4.6 Adding information about collection

4.4.4.1 Gather

The Gather panel, allows the user to explore the local file space and existing collections, gathering up selected documents for the new collection. The panel is divided into two sections, the left for browsing existing structures and the right for the documents in the collection.

This enables user to collect the required documents by exploring the entire computer. Then Selecting the files or directories from ‘local file space’ by browsing the folders in the computer or Local Network or downloading using protocols viz. WWW, OAI (Open Archives Initiative), Z39.50, SRW (Search and Retrieve Web service), MediaWiki. Drag and drop them into the right-hand panel of the mouse. The documents can be dragged individually or as a set of documents in folders/subfolders.

Now the scanned pdf files compiled in year wise sub-folders in dissertation hierarchy are required which were collected before. By clicking on the gather tab Dissertation files can be included to the collection created. Basically “Gather” is to collect the data. Files can be moved to the collection by dragging and dropping. The figure below represents the Gather panel with

hierarchy of records. In this figure the null records were replaced with the pdf files of the corresponding records.

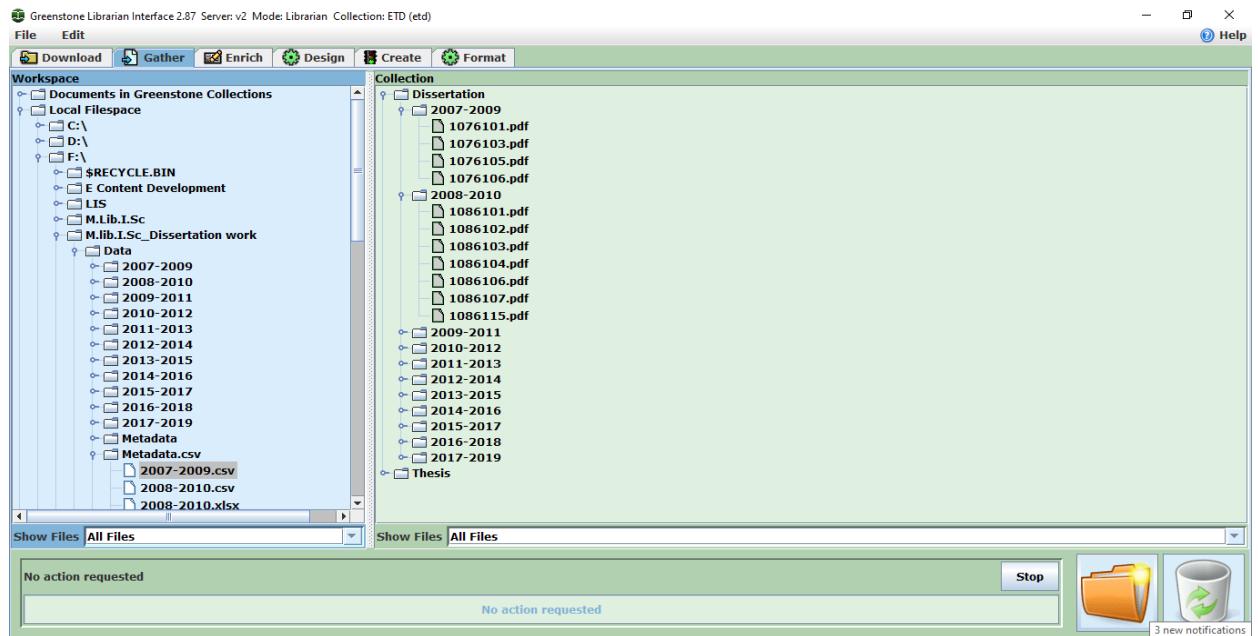


Figure 4.7 Gather Panel

4.4.4.2 Enrich

Having gathered all the scanned PDF files into the collection, the next phase in collection building is to enrich the documents by adding metadata. This section explains how metadata is created, edited, assigned and retrieved and how to use external metadata sources.

The Enrich panel allows Cataloguing with metadata, i.e. assign values to metadata-fields-Dublin Core. Metadata editor allows creating/changing sets and assigning values, automatic inheriting for lower levels, multiple values, pick lists or hierarchical structure.

Documents that are copied during the first step come with any applicable metadata attached. If a document is part of a Greenstone collection, previously defined metadata is carried over to the new collection. If the element's values have a hierarchical structure, the hierarchy can be extended in the same way. Of course, this new collection may have a different metadata and set new values can be added to the set of existing values for an element, or perhaps just a subset of

the defined metadata, and only metadata that pertains to the new collection's set is carried over. Resolution of such conflicts may require user intervention via a supplementary dialog. Any choices made are remembered for subsequent file copies.

In the existing collection of ETD Database, Name of the metadata can be different with respect to my metadata naming. So, names in the metadata are modified by clicking on "Manage Metadata Set". Then select the "Qualified Dublin Core Metadata Elements Set" and click on Edit. There are 55 exemplary terms in the Qualified Dublin Core Metadata Elements Set (QDCMES) and the ETD Database contains 20 elements only. The entire elements names are modified according to requirement.

The Enrich tab, left side of the panel under the collection tab shows the files. The right side, of the panel allows adding metadata for each document on each metadata field in the value box against the Element.

Once the names of all the elements are changed and then assign the metadata information of PDF copies of thesis and dissertation to respective elements from the null/empty records with the metatags imported from the excel sheet. New metadata values can be added or modified to the set of existing values for an element in the right side enrich panel.

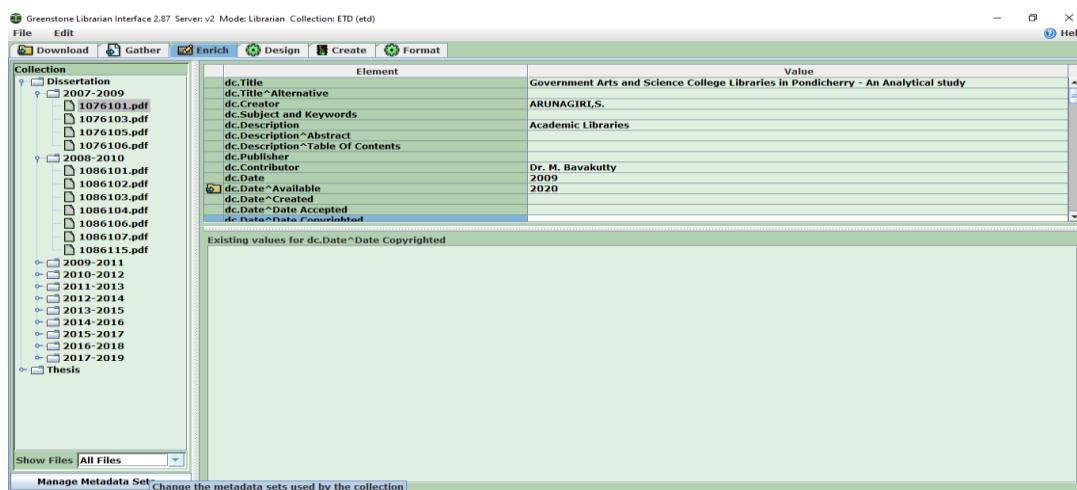


Figure 4.8 Enrich Panel

4.4.4.3 Design

Once the files are marked up with metadata, the appearance of the user end is decided. Then the collection was designed by choosing the features needed from design menu. Collection design consists of many facets as given in the left side panel. What kind of information is searchable? What ways are provided to browse through the documents? What languages are supported? Where do the buttons appear on the page? And other user interface facets can be customized in the design tab.

Design section involves selection of plugins (e.g. GA, TEXT, PPT, Word, PDF, RTF, e-mail, XLS, Fox, DB, as well as ISIS, DSpace, MARC, ProCite...), defining Search index, Partitioning of sub-collections and setting Browsing classifiers, hierarchical or A-Z. ‘plug-ins’ (filters), Indexing the documents and providing preview facility for direct access to webpage with search-interface produced by GLI is done at this stage. Once build is successful then the collection needs to be linked to previewing.

1. Document Plugin

Click on the document plug-ins to add the required plug-ins needed to convert the document into the document format (Greenstone archive format) required for greenstone. All plug-ins, needed for handling common documents, will be loaded by the default at the time of installation. The proper plug-ins have to be loaded at the time of installation. Kindly note that if proper Plug-in is not loaded; the software cannot build the digital library collection.

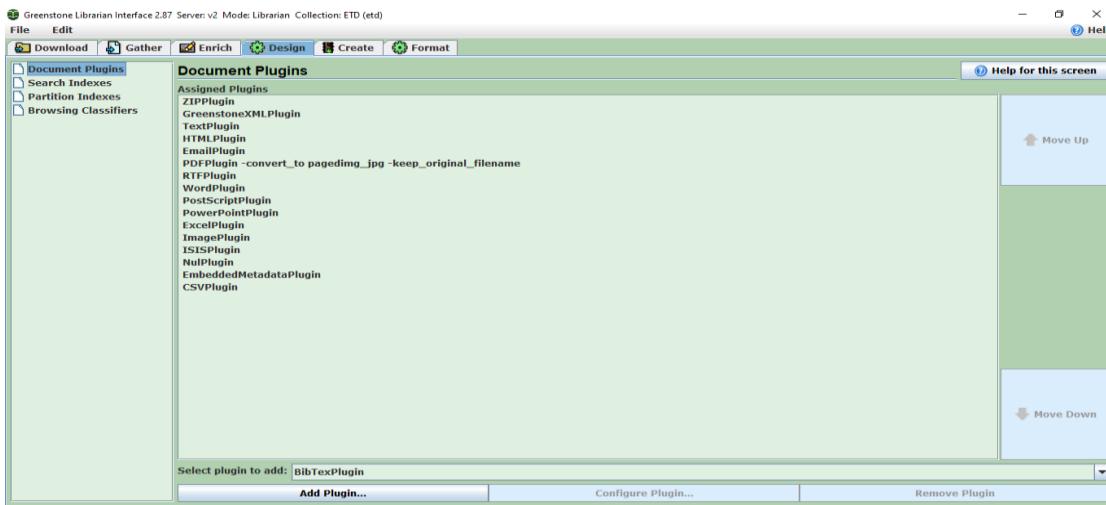


Figure 4.9 Document Plugin

The plugin for PDF files is required to be added. Click on the “Add Plugin” to add PDF plugin in greenstone.

2. Search Indexes

Choose the search Indexes, shown next below on left plane, for creating *Search Indexes*.

Search Indexes determine whether to combine the search to title name, author name, guide name, year. Here search the names of the fields on which the indexes are to be built can be added. The process of adding indexes for searching is given in fig*

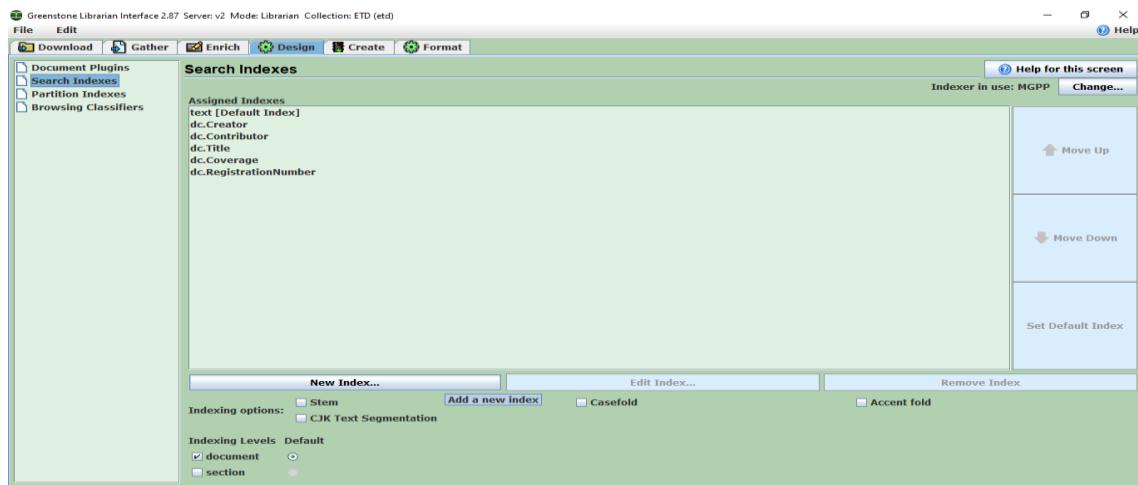


Figure 4.10 Search Index

Remove Default Indexes

Remove the default indexes for *ex.Title* and *ex.Source* by selecting the index description under Assigned indexes and then by clicking on Remote index button. Do not remove the search index for text [Default Index]. If search index for the text isn't provided the entire text of the document cannot be searched.

Adding New Indexes

Click on *New Index* button. Select *dc.Title*, *dc.Subject*, *dc.Guide* and *dc.Year* by tick marking on the check box one by one, and then one by one by clicking on the Add Index button. That means select *dc.Title* first add it, then select *dc.Author* and add it and so on.

1. Browsing Classifiers

To browse on Metadata elements, browsing Classifiers has to be set up, independently of creating an index on this metadata element. Browsing Classifiers such as Title, Subject, Guide, and Year help browse the collection. Select *Browsing Classifier* by clicking on it and then right click on *Remove Classifier* button for removing them one by one. Remove the default browsing classifiers for *Title* and *Source*.

By clicking *Add Classifier* button, the window for choosing Browsing Classifier pops up. Select the browsing classifier for Title by choosing the metadata (*dc.Title*) likewise Subject, Guide, Year classifiers for *dc.Description*, *dc.Contributer*, *dc.Date* metadata were chosen.

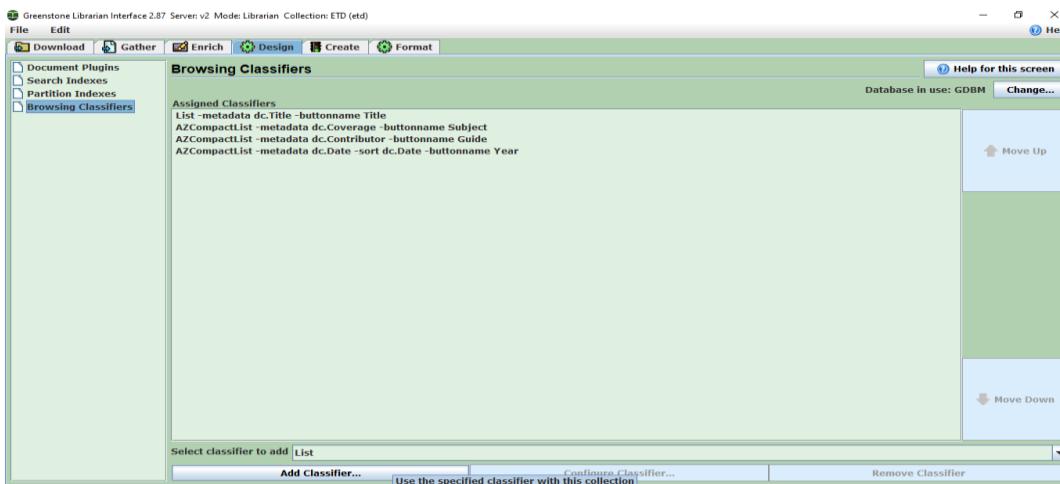


Figure 4.11 Browsing Classifiers

4.4.4.4 Format

The page display of the resultant digital library including the display page that appears on clicking the browsing classifiers or on making a search, are governed by the features provided in the *Format* tab.

1. General

Choose the format tab and select general to provide the general information about the collection

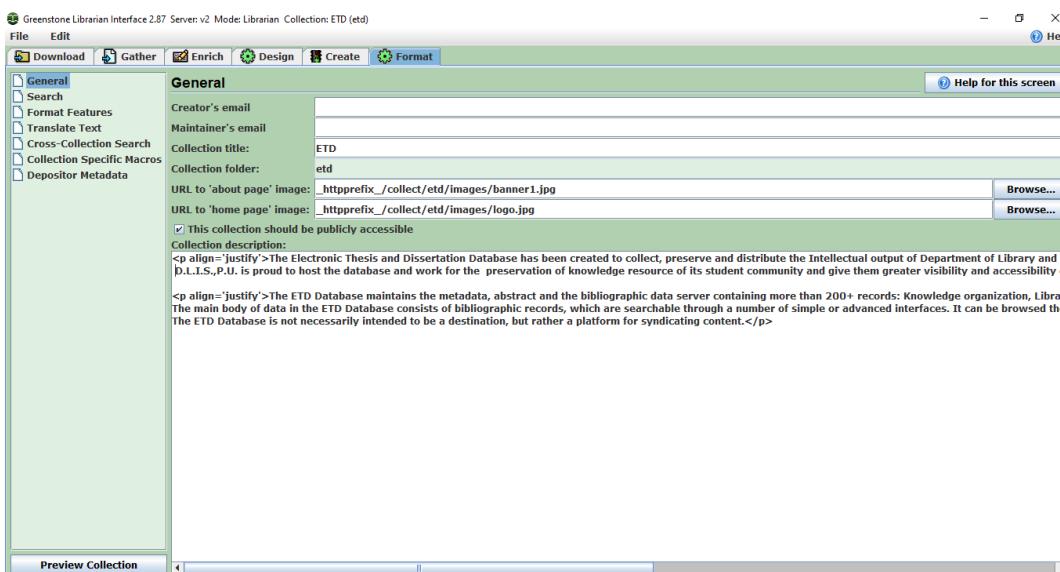


Figure 4.12 General Macro

Put a picture as a collection icon: Choose a small picture of *100*100* pixel that needs to appear as an icon of the collection in the home page, by clicking on *Browse* button on the middle right and select the picture. On choosing the picture, the full path of the picture will appear in the address box lying against the *Browse* button. Same or a different picture can be provided as the image for the ‘about page’ also. About page is the first page about an individual digital library collection that contains a short description about the collection.

2. Format Features

While changing the format feature, the appropriate one (for e.g. CL1 AZList-metadata *dc.Title*) from choose feature pull down list is to be selected and the format string to the format features is to be added by clicking on the “*Add Format*” button. The CL1 Browsing classifier for title can be added to the format features.

Similarly, any other string from choose feature box and click Add format button for customizing it. By editing the HTML string under the Format features box by selecting it and editing them in HTML Format String box.

Images of the titles were kept with the PDF copies of Table of Content, Bibliography. The Title of the images were given similar name as of PDF copies. After built, images will come on the respective PDF files on the page. This will help users to access the PDF files.

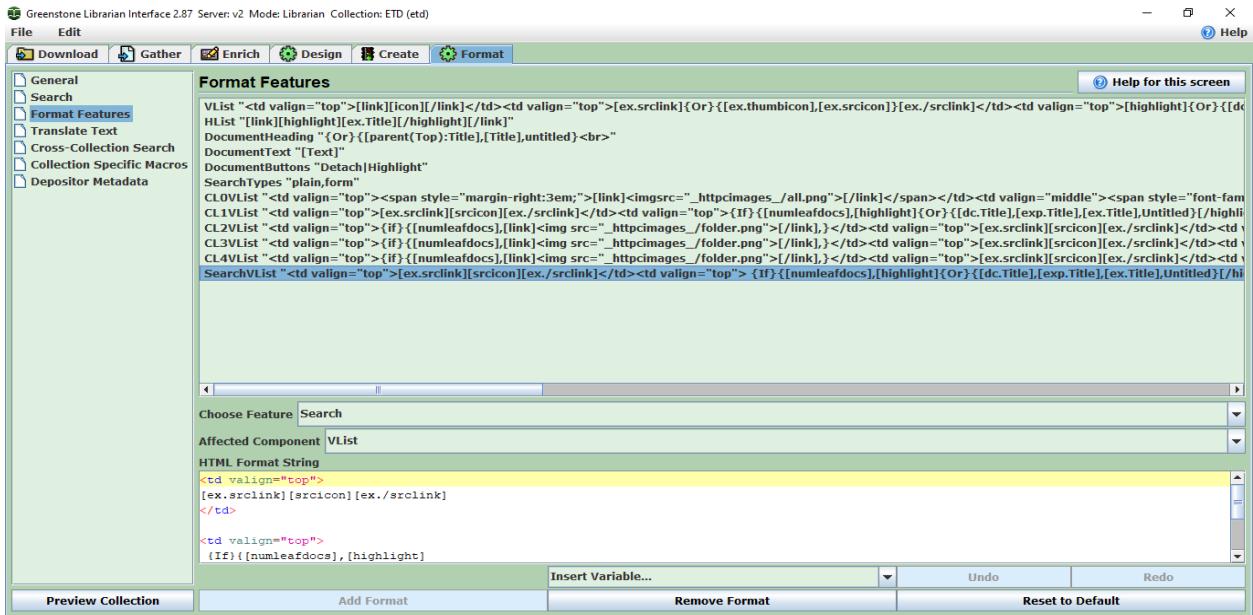


Figure 4.13 Format Features

The browse feature can also be added in the choose feature option. This will help to access the library by category. The HTML code are modified to add more feature to this option. By editing this, the digital library can be browsed by title and can browse the whole collection at a time etc.

The HTML Code for different classifiers and search index are as follows are as follows,

```
#Title (CL1 : Metadata dc.Title)
<td valign="top">
[ex.srclink][srcicon][ex./srclink]
</td>

<td valign="top">
{If}{[numleafdocs],[highlight]}
{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}
[/highlight]&nbsp;<i>([numleafdocs])</i>,
[dc.Title]<br><b> Author : </b><i>[dc.Creator]</i><br>
<b> Registration Number : </b><i>[dc.Identifier]</i><br>
<b> Guide : </b><i>[dc.Contributor]</i><br>
<b> Area of Study : </b><i>[dc.Description]</i><br>
<b> Year : </b><i>[dc.Date]</i><br>
}
</td>
```

Figure 4.14 HTML Code for Title Classifier

```
#Subject (CL2 : Metadata dc.Description)
<td valign="top">
{if}{[numleafdocs],[link][/link],}
</td>

<td valign="top">
[ex.srclink][srcicon][ex./srclink]
</td>

<td valign="top">
{If}{[numleafdocs],[highlight]
{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}
[/highlight]&ampnbsp<i>([numleafdocs])</i>,
[dc.Title]<br><b> Author : </b><i>[dc.Creator]</i><br>
<b> Registration Number : </b><i>[dc.Identifier]</i><br>
<b> Guide : </b><i>[dc.Contributor]</i><br>
<b> Area of Study : </b><i>[dc.Description]</i>
}
</td>
```

Figure 4.15 HTML Code for Subject Classifier

```
#Guide (CL3 : Metadata dc.Contributor)
<td valign="top">
{if}{[numleafdocs],[link][/link],}
</td>

<td valign="top">
[ex.srclink][srcicon][ex./srclink]
</td>

<td valign="top">
{If}{[numleafdocs],[highlight]
{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}
[/highlight]&ampnbsp<i>([numleafdocs])
</i>,
[dc.Title]<br><b> Author : </b><i>[dc.Creator]</i><br>
<b> Area of Study : </b><i>[dc.Description]</i>
}
</td>
```

Figure 4.16 HTML Code for Guide Classifier

```
#Year (CL4 : Metadata dc.Date)
<td valign="top">
{if}{[numleafdocs],[link][/link],}
</td>

<td valign="top">
[ex.srclink][srcicon][ex./srclink]
</td>

<td valign="top">
{If}{[numleafdocs],[highlight]
{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}
[/highlight]&nbsp;<i>([numleafdocs])
</i>,
[dc.Title]<br><b> Author : </b><i>[dc.Creator]<i><br>
<b> Registration Number : </b><i>[dc.Identifier]<i><br>
<b> Guide : </b><i>[dc.Contributor]</i><br>
<b> Area of Study : </b><i>[dc.Description]</i>
}
</td>
```

Figure 4.17 HTML Code for Year Classifier

```
#Search
<td valign="top">
[ex.srclink][srcicon][ex./srclink]
</td>

<td valign="top">
{If}{[numleafdocs],[highlight]
{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}
[/highlight]&nbsp;<i>([numleafdocs])</i>,
[dc.Title]<br><b> Author : </b><i>[dc.Creator]<i><br>
}
</td>
```

Figure 4.18 HTML Code for Search Index

3. Collection Specific Macros

Specific Macros can be edited to manage the library in a better way. They help in changing the default background color and outlook of the Greenstone. The data option from the page can be removed. Coding for this macro is as follows

```

package Global

_linktextHOME_{Home}
_linktextHELP_{Help}
_linktextPREFERENCES_{settings}

package Style
_collectionspecificstyle_
{
<style type="text/css">
    body.bgimage \{background-image:none;\}
    p.bannertitle \{background-image:none;font-family:Cambria;color:black;text-transform:capitalize;\}
    body\{font-family:Cambria;font-size:18px;color:grey;margin-left:0.5%;\}
    a.navlink\
    {
        border:1px solid #ccc; padding:1px 20px 6px;color:black;font-family:Cambria;text-decoration:none;font-weight:bold;
        text-transform:capitalize;font-size:16px;line-height:30px;background:#dadb8d;border-radius:12px;-moz-border-radius:12px;
        box-shadow:1px 1px 2px rgba(0,0,0,.5);-webkit-box-shadow:1px 1px 2px rgba(0,0,0,.5);-moz-box-shadow:1px 1px 2px rgba(0,0,0,.5);
        text-shadow:#fff 0px 1px 1px;
        background:-webkit-gradient(linear,left top,left bottom,from(#eeeeee),to(#cccccc));background:-moz-linear-gradient(top,#eeeeee, #cccccc);\}
    }

    a.navlink_sel\
    {
        border:1px solid #ccc; padding:1px 20px 6px;font-family:Cambria;text-decoration:none;text-transform:capitalize; font-size:16px;line-height:30px;
        background:#border-radius:12px;-moz-border-radius:12px;box-shadow:1px 1px 2px rgba(0,0,0,.5);-webkit-box-shadow:1px 1px 2px rgba(0,0,0,.5);
        -moz-box-shadow:1px 1px 2px rgba(0,0,0,.5);text-shadow:#fff 0px 1px 1px;
        background:-webkit-gradient(linear,left top,left bottom,from(#eeeeee),to(#cccccc));background:-moz-linear-gradient(top,#eeeeee, #cccccc);\}
    }

    a.navlink:hover\{background-image:none;background-color:#cce6ff;\}
    div.divbar\{background-image:none;background-color:none;\}
    div.navbar\{background-image:none;background-color:none;\}
</style>
}

package about

_titleabout_={}
_content_={}
_navigationbar_={}
_textabout_={}

_textabout_={}
<div id="content" style="width:600px">
    <div id="spacer" style="padding:2em;"></div>
    <div id="text" style="height:300px; width:1000px; float:'center'; margin-right:3em;">
        <p align='justify'>The Electronic Thesis and Dissertation Database has been created to collect,
        preserve and distribute the Intellectual output of Department of Library and Information, Pondicherry University.
        D.L.I.S., P.U. is proud to host the database and work for the preservation of knowledge resource of its student community and
        give them greater visibility and accessibility over time.</p>

        <p align='justify'>The ETD Database maintains the metadata, abstract and the bibliographic data server containing more than 200+ records:
        Knowledge organization, Library studies, Information behavior, Information systems and retrieval, Bibliometrics, Webometrics and other
        facets of Library and Information Science.
        The main body of data in the ETD Database consists of bibliographic records, which are searchable through a number of simple or advanced interfaces.
        It can be browsed through the Browse interface by Title, Subject, Guide and Year of Submission.
        The ETD Database is not necessarily intended to be a destination, but rather a platform for syndicating content.</p>
    </div>
</div>
<div id="footer" style="clear:both;text-align:center;font-size:small;font-style:italic;">
<hr>This Collection contains _about:numdocs_ documents. Last update: _about:builddate_ days ago.
</div>
}

```

Figure 4.19 HTML Code for Collection Specific Macro

4.4.4.5 Create

Build the Collection

Through the *Create* panel, the collection is built by clicking on the *Build Collection* button and the process bar displays the progress of building the collection.

Preview the Collection

Once the collection is built completely, the preview collection button gives a trial of user experience on how the user interface appears and the ways the collection could be accessed.

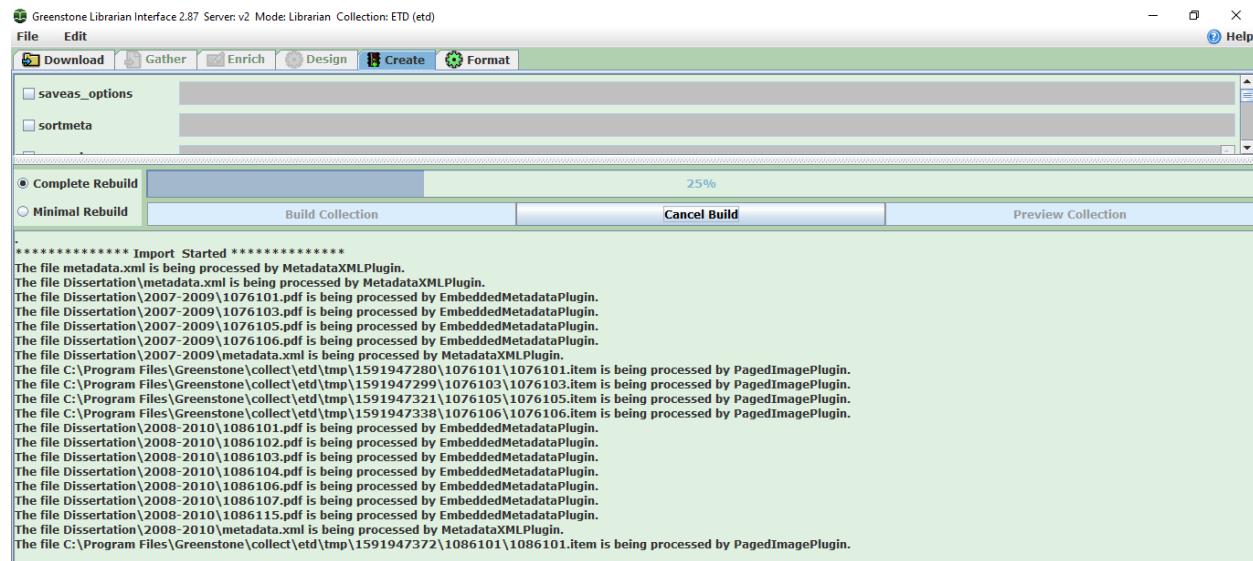
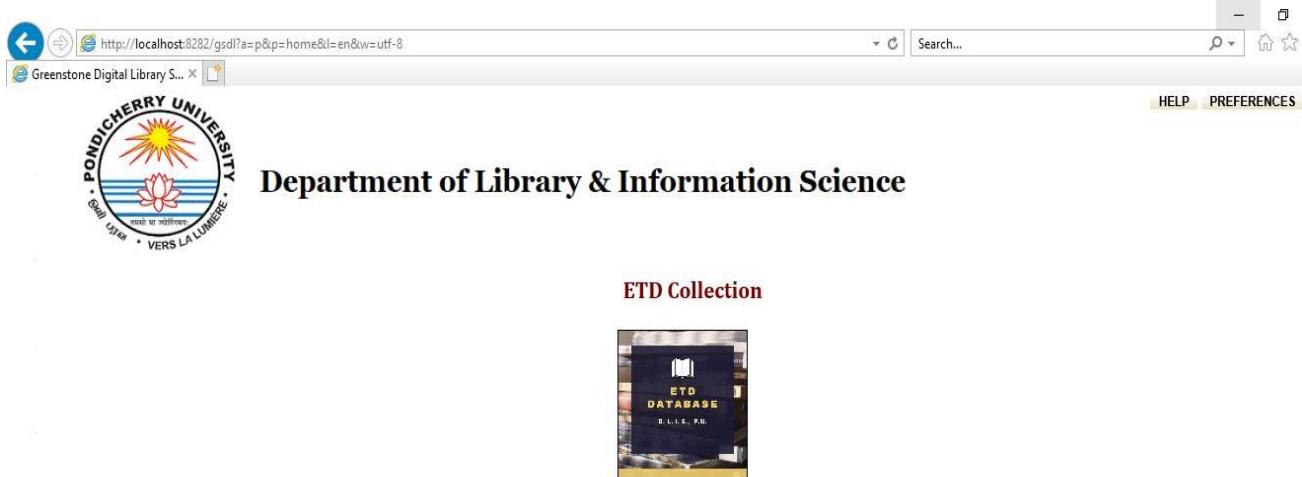


Figure 4.20 Create Panel

4.4.5 Customization of Home Page

The Customization of the home page was achieved by altering the HTML Code of the home macro from greenstone home.dm file, package used for the build of the home page.

Home page gives an overview of the department of library and information science and the link to access the ETD Database.



About Department

The Department of Library and Information Science was started during the academic year 2007-08. The Department offers a professional Master's Degree programme (MLIS) and a Ph.D. Degree Programme in Library and Information Science. In addition, an Add on Course - Post Graduate Diploma in Library Automation and Networking (PGDLAN), which is of one year duration with two semesters, is also offered in the evenings. MLIS course duration is of two years and comprises four semesters.

Infrastructure Facilities

The Department has modern classrooms and well established computer lab with sufficient computers installed with the required software for teaching and practice. Internet facility with higher bandwidth is also available and it may be highlighted that the campus is WiFi enabled.

Figure 4.21: Home Page

```

content_ {


<b>ETD Collection</b>



_homeextra_


|


_textaboutgreenstone_


}

_textaboutgreenstone_ {


<b>About Department</b><br><br>The Department of Library and Information Science was started during the academic year 2007-08. The Department offers a professional Master's Degree programme (MLIS) and a Ph.D. Degree Programme in Library and Information Science. In addition, an Add on Course - Post Graduate Diploma in Library Automation and Networking (PGDLAN), which is of one year duration with two semesters, is also offered in the evenings. MLIS course duration is of two years and comprises four semesters.</p>



<b>Infrastructure Facilities</b><br><br>The Department has modern classrooms and well established computer lab with sufficient computers installed with the required software for teaching and practice. Internet facility with higher bandwidth is also available and it may be highlighted that the campus is WiFi enabled.</p>


}

```

Figure 4.22: HTML Code for Home Page customization

4.4.6 Moving the Collection to CD-ROM

The Collection can be exported to a CD-ROM. The collection of the greenstone can also be converted into an installable CD-ROM for distribution among wider audience. After choosing “write CD/DVD Image” option, the library exported to be selected. The output file is to be renamed. It was created under the temp folder of the greenstone. Now the collection is copied to the CD writer.

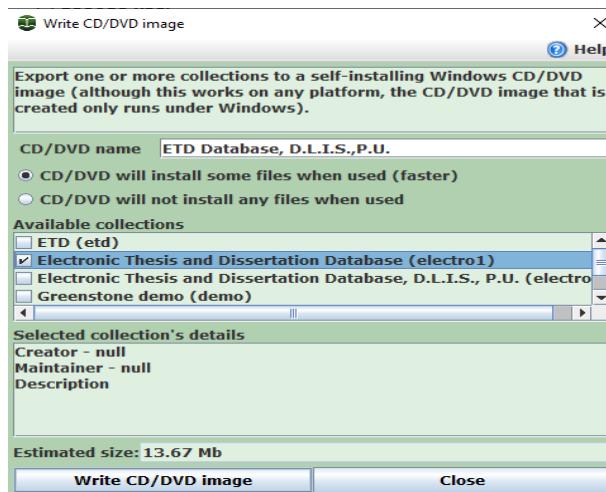


Figure 4.23 Moving Collection to CD-ROM

4.4.7 ETD Database Design Preview

The following are the quick cognitive walkthrough on user experience of the user interface.

1. User Interface Design Before and After Customization

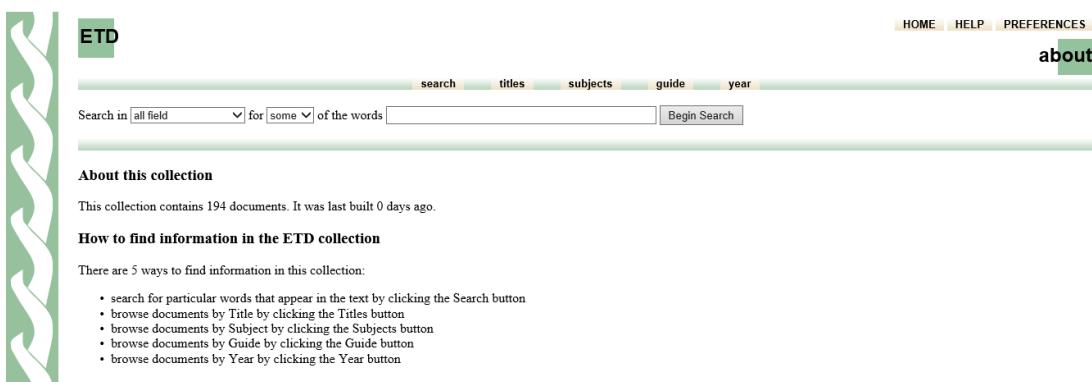
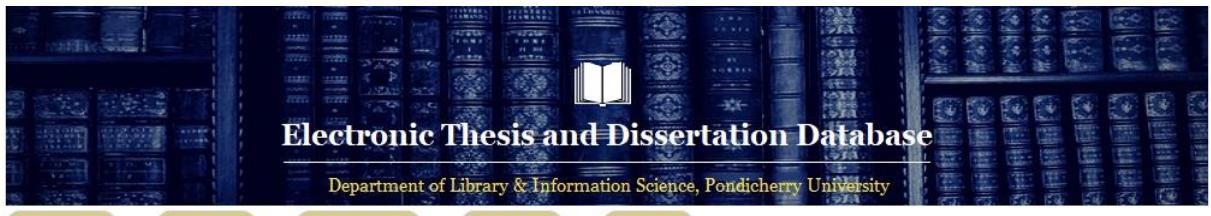


Figure 4.24 User Interface Design Before Customization



The Electronic Thesis and Dissertation Database has been created to collect, preserve and distribute the Intellectual output of Department of Library and Inforamtion, Pondicherry University. D.L.I.S.,P.U. is proud to host the database and work for the preservation of knowledge resource of its student community and give them greater visibility and accessibility over time.

The ETD Database maintains the metadata, abstract and the bibliographic data server containing more than 200+ records: Knowledge organization, Library studies, Information behavior, Information systems and retrieval, Bibliometrics, Webometrics and other facets of Library and Information Science. The main body of data in the ETD Database consists of bibliographic records, which are searchable through a number of simple or advanced interfaces. It can be browsed though the Browse interface by Title, Subject, Guide and Year of Submission. The ETD Database is not necessarily intended to be a destination, but rather a platform for syndicating content.

This Collection contains 194 documents. Last update: 1 days ago.

Figure 4.25 User Interface Design After Customization

3. Search Index

Figure 4.26 Search Index

4. Title Classifier

The screenshot shows the homepage of the 'Electronic Thesis and Dissertation Database' from the Department of Library & Information Science, Pondicherry University. The page features a decorative background of bookshelves. At the top right is a search bar with a magnifying glass icon. Below it are five buttons: 'Search' (yellow), 'Titles' (green), 'Subjects' (blue), 'Guide' (yellow), and 'Year' (yellow). A navigation menu at the top left includes letters A through W. The main content area displays three thesis titles with their details:

- A Bibliometric study of Ph.D Thesis in Tamil Language and Literature
Author : VINAYAGAMURTHY,A.
Registration Number : 15390021
Guide : Dr. Mangkollen Singson
Area of Study : Bibliometrics
Year : 2017
- A Study of Pondicherry University Library
Author : JAYABARATHAN,A.
Registration Number : 1076103
Guide : Dr. M. Bavakutty
Area of Study : Academic Libraries
Year : 2009
- An Evaluative Study of Photo Archive in the French Institute of Pondicherry
Author : DHANABAGYASELVI,I.
Registration Number : 13390002
Guide : Dr. Chennupati K. Ramaiah
Area of Study : Information Systems and Services
Year : 2015

Figure 4.27 Title Classifier

5. Subject Classifier

The screenshot shows the 'Subjects' section of the database. It has the same header and navigation as the previous screen. The main content area lists subjects with their counts, each preceded by a yellow folder icon:

- Academic Libraries (6)
- Acquisition (2)
- Bibliometrics (11)
- Citation Analysis (4)
- Digital Library (3)
- E-Resource Management (1)

Figure 4.28 Subject Classifier

The screenshot shows the homepage of the Electronic Thesis and Dissertation Database. The header features a blue background with bookshelves and the text "Electronic Thesis and Dissertation Database" and "Department of Library & Information Science, Pondicherry University". Below the header are five buttons: "Search", "Titles", "Subjects" (which is highlighted in green), "Guide", and "Year". A sidebar on the left contains a folder icon followed by the text "Academic Libraries (6)". The main content area lists six thesis titles under this category, each with a small PDF icon, author, registration number, guide, and area of study.

- Perception of the use on the service of Aarupadai Veedu Medical College Library in Puducherry: a Study
Author : BANUPRIYA,S.
Registration Number : 11390004
Guide : Dr. R. Sevukan
Area of Study : Academic Libraries
- Reading Habit of higher secondary school students in Pondicherry
Author : BHARATHI,R.
Registration Number : 11390005
Guide : Dr. Rekha Rani Varghese
Area of Study : Academic Libraries
- Informal Learning Spaces in Pondicherry University Campus: An Impact Analysis
Author : NENAVATH LAXMAN
Registration Number : 16390011
Guide : Dr. R. Sevukan
Area of Study : Academic Libraries
- Government Arts and Science College Libraries in Pondicherry - An Analytical study
Author : ARUNAGIRI,S.
Registration Number : 1076101
Guide : Dr. M. Bavakutty

Figure 4.29 Subject Classifier

6. Guide Classifier

The screenshot shows the homepage of the Electronic Thesis and Dissertation Database. The header features a blue background with bookshelves and the text "Electronic Thesis and Dissertation Database" and "Department of Library & Information Science, Pondicherry University". Below the header are five buttons: "Search", "Titles", "Subjects", "Guide" (which is highlighted in green), and "Year". A sidebar on the left contains a folder icon followed by the text "DR. A. Manoharan (2)", "Dr. Chennupati K. Ramaiah (34)", "DR. J.P.S. Kumaravel (3)", "Dr. K. Anas (9)", "Dr. Mangkhollen Singson (31)", "Dr. M. Bavakutty (4)", "Dr. M. Leeladharan (32)", "Dr. Rekha Rani Varghese (35)", "Dr. R. Sevukan (34)", and "DR. V. Pulla Reddy (6)".

Figure 4.30 Guide Classifier

The screenshot shows the 'Electronic Thesis and Dissertation Database' homepage with a blue header featuring bookshelves. Below the header, there are five navigation buttons: 'Search', 'Titles', 'Subjects', 'Guide' (which is highlighted in green), and 'Year'. Under the 'Guide' button, a section titled 'Dr. M. Bavakutty (4)' is displayed. It lists four thesis titles with their authors and areas of study:

- Digital infrastructure facility in the Medical college libraries in Pondicherry.
Author : RAVI,K
Area of Study : ICT
- Use of Electronic Information Resources by the Science Doctoral Students in Pondicherry University: A study
Author : SUSANTA NARZARY
Area of Study : E-Resource Management
- Government Arts and Science College Libraries in Pondicherry - An Analytical study
Author : ARUNAGIRIS.
Area of Study : Academic Libraries
- A Study of Pondicherry University Library
Author : JAYABARATHAN,A.
Area of Study : Academic Libraries

Figure 4.31 Guide Classifier

7. Year Classifier

The screenshot shows the same database interface as Figure 4.31, but the 'Year' classifier is selected. A list of years from 2009 to 2019 is shown, each with a folder icon and the count of entries in parentheses.

Year	Count
2009	(4)
2010	(7)
2011	(6)
2012	(22)
2013	(23)
2014	(22)
2015	(27)
2016	(17)
2017	(22)
2018	(17)
2019	(26)

Figure 4.32 Year Classifier

2010 (7)

Computerized Information Service in Ananda Rangapillai Library, Pondicherry University: A Users Study
Author : NAYAN KUMAR BORO
Registration Number : 1086103
Guide : Dr. K. Anas
Area of Study : Library Services

Scholarly Publications in India Journal of Pharmacology during 2006 to 2008: A Citation Analysis
Author : RAJU BRAHMA
Registration Number : 1086115
Guide : Dr. M. Leeladharan
Area of Study : Infometrics

Digital Information Literacy among PG students and scholars of Pondicherry University: An analytical study
Author : SANJEEB KUMAR BRAHMA
Registration Number : 1086106
Guide : Dr. R. Sevukan
Area of Study : Information Literacy

Open Access Institutional Repositories in Asia: A Webometric Analysis
Author : NARAYAN BASUMATARY
Registration Number : 1086102

Figure 4.33 Year Classifier

8. Help

Topics

- [How to find information in the ETD collection](#)
- [How to read the documents](#)
- [How to search for particular words](#)
 - [Search terms](#)
 - [Query type](#)
 - [Scope of queries](#)
 - [Advanced searching using the MGPP search engine](#)
 - [Fielded searching](#)
- [Changing your preferences](#)
 - [Collection preferences](#)
 - [Language preferences](#)
 - [Presentation preferences](#)
 - [Search preferences](#)

How to find information in the ETD collection

There are 5 ways to find information in this collection:

- search for particular words that appear in the text by clicking the Search button
- browse documents by Title by clicking the Titles button
- browse documents by Subject by clicking the Subjects button
- browse documents by Guide by clicking the Guide button
- browse documents by Year by clicking the Year button

Figure 4.34 Help Window

9. Settings

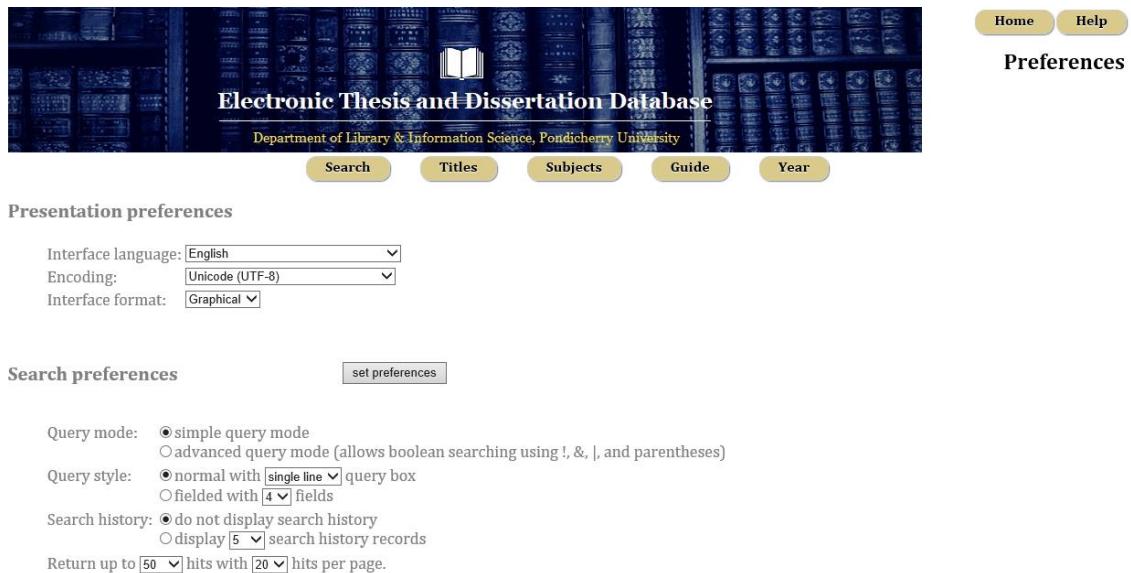


Figure 4.35 Settings Window

4.5 Summary

The design and development chapter detail the way the collection had been developed through different phases of digitization, gathering metadata, collection building, interface customization, and finally exporting the database to CD-ROM. The consecutive evaluation chapter explicates the process of evaluation of the ETD Database.

Chapter 5

Evaluation

CHAPTER 5

EVALUATION

5.1 Introduction

Evaluation is the systematic and impartial assessment of the design, implementation and the resultant output, primarily for the purposes of error detection and error deduction. Evaluation deliberates value, merit, worth, significance or quality (Scriven, 1991).

Evaluation helps to relate what works, for whom, in what respects, to what extent, in what contexts, and how (Pawson and Tilley, 2004). Evaluation scrutinizes expected and achieved indicators, to realize accomplishments and the lacks. Evaluation emphasis on relevance, accessibility, comprehensiveness, integration, fulfillment of objectives, effectiveness, impact, cost, efficiency, and sustainability (Patton, 1997; OECD, 2010). Conclusions made in evaluations encompass both an empirical aspect and a normative aspect (Fournier, 2005).

An evaluation has to be methodical, and provide information that is credible, reliable, and useful to enable the incorporation of feedback into cyclic reiteration of prototyping. In this chapter we about to discuss in detail the methodology implemented in the evaluation of the design.

5.2 Design Evaluation

The General approaches in evaluating a design are:

1. *Cognitive Walkthrough approach:* The cognitive walkthrough approach is a technique for evaluating the design of a user interface, with special attention to how well the interface supports “exploratory learning”, i.e., first-time use without formal training. The strengths and limitations of the walkthrough method are considered, and it is placed into context of more complete design approach.

2. *Heuristic evaluation:* A heuristic evaluation is the systematic inspection of a user interface design, based on interface design guidelines and principles. It specifically involves evaluators examining the interface and judging its compliance with recognized usability principles (the heuristics).
3. *Review based evaluation:* review based evaluation is more efficient if the reviewer selects carefully, noting the experiment design chosen, the population of subject used, the analysis performed, and the assumptions made.
4. *Questionnaire:* Questionnaire-based user survey provides a vast amount of information and is good measure of user attitudes. There are several styles of question, they are:
 - i. General
 - ii. Open-ended
 - iii. Closed-ended
 - iv. Scalar
 - v. Multi Choice
 - vi. Ranked
5. *Observation of task performance:* In the observation technique, the users are watched carefully as they perform task interacting with the system. In the ideal situation, the users ignore the presence of evaluator and out loud what they are doing, as well as problems faced. In another situation, people may act differently in the presence of the evaluator.
6. *Co-operative evaluation:* This is a cost-effective technique for identifying the usability problems in the prototype product and processes. This technique encourages design teams and users to collaborate in order to identify usability issues and their solutions. Users work with prototype and carry out tasks as set by design team. During this procedure users

explain what they are doing by talking or thinking aloud. An observer records unexpected user behavior and user comments regarding the system. The observer also actively participates by questioning the users with respect to their intentions and expectations. This provides qualitative information concerning any difficulties that the user experience and the features or interface elements, which give rise to those problems. The method is applicable in the early stages of design cycle when the prototype is available for evaluation or at later stage when an existing product is available for evaluation or at a later stage when an existing product is to receive further development.

In this project the evaluation of the offline ETD Database is subjected to be conducted in laboratory condition as a combination of formative evaluation through rapid prototyping alongside with summative Heuristic evaluation.

5.2.1 Formative Evaluation

Formative evaluation, formative assessment, formative feedback, including diagnostic testing, is a range of formal and informal assessment procedures conducted during the development process in order to modify prototype to attain robustness. Their aim is to provide both the designer with a gauge of where their level of prototype is at the current moment and enable the designer to adjust accordingly to meet the emerging needs from the user perspective.

The Errors detected in each reiteration of prototype were documented for purpose of debugging errors and error prevention in future prototype design.

Table 5.1: Documentation of Formative Evaluation

S. No	Lapse/Error	Response
1	Simple Search Index	Advanced Sorting and Filtering Options
2	Exhaustive Listing of Records	Classified Browsing Index
3	Listing of Classified Documents	Supplemented with Number of Records
4	Visual Noise	Personalized Home and Browsing pages
5	Fixed Navigation Bar	Customization of Navigation bar
6	Default Banner	Redesigned Banner for ETD Database
7	Conventional Greenstone Icons	Custom Icons

5.2.2 Summative Evaluation

Summative evaluations come at the end of development, to know to what extent the objectives previously fixed have been accomplished. They are intended to provide a package of results used to assess whether a program works or not. In other words, it is the evaluation of output at the end of a development of the model. The main objective of the summative evaluation is to assign grades to the design. It indicates the degree to which the design has achieved its objectives. It helps to judge the appropriateness of objectives. Summative evaluation is generally the work of standardized tests. The Heuristic evaluation has been adopted as the standard test procedure to assess the design.

5.2.2.1 Heuristic Evaluation

The Heuristic evaluation involves experts having examined the interface and judge its compliance with recognized usability principles (the "heuristics"). Heuristic evaluation (Nielsen and Molich, 1990; Nielsen 1994) is a usability engineering method for finding the usability

problems in a user interface design so that they can be attended to as part of an iterative prototyping process.

In general, the Heuristic evaluation is performed by 2 or more usability experts (Nielsen and Molich, 1990). Instead, The ETD Database' Heuristic evaluation was directed with Heuristic guideline to achieve same extent of robustness.

The Heuristic Evaluation was performed in 2 steps

1. Choosing the Heuristic Guideline
2. Define a Ranking System

The widely used Jakob Nielsen's 10 general principles for interaction design (the "heuristics") has been adopted as *the Heuristic Guideline* for evaluating the database along with binary ranking system.

Table 5.2: Documentation of Heuristic Evaluation

S. No	Usability Heuristics for User Interface Design	Binary Ranking
1	Visibility of system status	No
2	Match between system and the real world	Yes
3	User control and freedom	Yes
4	Consistency and standards	Yes
5	Error prevention	No
6	Recognition rather than recall	Yes
7	Flexibility and efficiency of use	No
8	Aesthetic and minimalist design	Yes
9	Help users recognize, diagnose, and recover from errors	No
10	Help and documentation	Yes

The ETD Database using greenstone digital library software had established six of ten Usability Heuristics designed by Jakob Nielsen for User Interface Design. The Interface design hence does have minimum usability experiences leaving leaves to improve.

5.3 Summary

During the formation of prototypes, a more realistic pattern of development, is for designers to switch back and forth through the various stages as the task proceeds. The Rapid prototyping model coupled with feedback loop stances the formative evaluation ensuring the robustness of the database design while summative heuristic evaluation substantiates the worth of the end database interaction design. This Chapter analyzed systematically the performance of the database with the standards contributing value to the offline ETD Database. The final chapter six briefs the summary of the proceedings of the project, research contributions, research gaps, recommendation for further research and finally the concluding phase of the dissertation.

Chapter 6

Conclusion

CHAPTER 6

CONCLUSION

6.1 Introduction

In the paradigm shift of the digital libraries, from an experimentation stage to institutionalization, the central role of the library professional remains the same, but then methods and tools for information storage and retrieval continue to grow and change radically. This has been brought forth with a remarkable progression of the computing technology, IT and ICT's and the substantial growth in internet, network, cloud computing and open source software, etc... Made in its supporting technologies both hardware and software has established a comprehensive range of services in the Library.

The institutional repositories and ETD are a few among many such library services that takes a shape in meeting the information needs of digital age patrons with the collections which were previously maintained in-house. The exponential growth of e-resources and the complexity in managing those e-resources has posed new challenges for the library professionals. To cope up with the information needs and the e-resources with speed, relative accuracy and reliability, the digital library platforms has emerged as an imperative and reliable recourse and the knowledge discovery turns up to be its crucial aspect.

The development of digital library, by collecting, compiling, digitizing, cataloging and preserving the intellectual resources of an institution is a herculean task. Greenstone offers scores of collections and represents the cutting edge of digital library research using greenstone as a vehicle for dissemination. It helps to create different types of collection like audio, video, image text and multimedia collection. The digital library arena offers a unique challenge to emerging breed of digital librarians, to combine principles, practices and tools of information management

to create new information product and service. The GSDL offers a dynamic platform for knowledge organization and retrieval.

This project is aimed to design and develop GSDL based Electronic Thesis and Dissertation Database for the Department of Library and Information Science, Pondicherry University. The ETD Database was developed to share the digital compilation of abstract, bibliography, and other metadata of the physical collection of thesis and dissertation in the department, with the fellow students and staffs for an active and an ethical exploitation of intellectual resources of the department.

The work was attempted to build a system by proposing and testing a component framework to support the construction of a digital repository in an exhaustive and repeatable fashion. In this work, the research is being carried out on exploring the ways to digitize the physical collection, the physical collection of Thesis and the dissertation in the department are digitized before being fed into customized database. A Generic Digital library system like Greenstone must be customized for specific purpose to build a system that is more interactive and user friendly.

Once the metadata corpus of the ETD Database is constructed, different components of Greenstone Digital Library were explored. Few macros of Greenstone Digital Library were modified based on the understandings obtained from analysis of the user needs, objectives, and requirements. The E-Thesis and Dissertation Database gets devised in the Greenstone platform is continuously evaluated for glitches and undergoes iterative revisions before being finally exported to a CD-ROM, which will be helpful for preservation and distribution among potential users.

6.2 Summary

6.2.1 Trends in Institutional Repositories

The literature review surfaced the Self Archiving archetypes and their past research and processes in the design and development of Digital Libraries, Institutional Repositories, E-Thesis and Dissertation (ETD) Collection, which includes different open source digital library software packages and projection of the justification on use of Greenstone digital library software platform in the development of ETD Database.

6.2.2 Design and Development of ETD Database

In the Design and Development of ETD Database the ADDIE model had been deployed as the methodology in the development process. It includes 5 major key phases before the database gets developed. They are Analysis, Design, Development, Implementation, and Evaluation. They stood as the standard framework/guideline in the development of this database.

6.2.2.1 Analysis

Initially the user requirements in personalizing user interface was totally unknown. In the analysis phase three key analysis were made in building the structure of ETD Database. The basic user requirements were broadly taken into consideration. The prospective users of the database, their problems, skills, and needs were taken into consideration. The evident goals and objectives in building the database was made clear and concise. Structural proposal of the system had been developed.

6.2.2.2 Tools and Technologies for developing the ETD Database

Appropriate tools and technologies in creating customizable collection of E-Thesis and Dissertation were identified. As the collection had been planned to be distributed in the offline mode Greenstone Digital Library software had been chosen as the digital library platform to build the ETD Database. Manifestations of modules were redesigned based on the objectives and user experience envisaged in the prelude.

6.2.2.3 Digitization of Physical Collection

In the early phase of the project, the physical collection of Thesis and Dissertation was initially digitized for the archival preservation of the research data of Department of Library and Information Science, P.U. the collection had been particularly scanned for Abstract, Table of Content, Bibliography and the Title page using the custom made scanner that had been built temporarily for the specific purpose. The DIY Scanner was used alongside with the software backend of Adobe scan software. The Adobe scan software, Adobe document cloud, and Adobe Reader software were used in tandem for simultaneous post processing the scanned documents.

6.2.2.4 Development of ETD Database

The Electronic Thesis and Dissertation Database was developed using Opensource Greenstone Digital Library Software. The metadata manually gathered from the physical collection of thesis and dissertation were exploded into the greenstone collection etd from excel sheet. The null records with the metadata were replaced with the corresponding abstract and bibliography PDFs that were digitized and indexed earlier. Total of 194 such records were built within the database.

As the Collection once get build from the gather and enrich panel, the navigation of the user end was customized with design and format macros.

Table 6.1: Navigation Indexes and Classifiers

Search Index	Title, Subject, Author, Guide, Registration Number, Key words, Full text, all fields.
Browsing Classifiers	Title, Subject, Guide, Year.

The search index and the browsing classifiers assigned in the design macro were customized in the format macro. The user interface of the ETD Database was personalized with custom specific macros. The HTML code for the customization at different levels of format macro were altered for specific personalized user experience. The collection then gets built and preview in the create panel. After multiple reiterations of prototyping and error fixing the current model was finalized, evaluated and exported to CD-ROM.

6.2.2.5 Evaluation of ETD Database

The E-Thesis and Dissertation (ETD) database has been an offline repository and the evaluation of the ETD Database was subjected to be conducted in laboratory condition as a combination of formative evaluation through rapid prototyping alongside with summative Heuristic evaluation.

In the formative evaluation the prototypes were tested for bugs at every single phase in the development of multiple subsequent prototypes and the errors were fixed, apposite features were improvised, and the prototypes were revised every now and then through rapid prototyping.

The summative evaluation was conducted at the end of the development cycle, once the model was fully developed. Heuristic evaluation model was chosen for the summative evaluation of the database and found the ETD Database design satisfying six of ten Jakob Nielsen's heuristic guideline.

6.3 Research Contributions

It is anticipated that the result of the project will alter the people's presumption on digital libraries, its services, and the way digital libraries are built. The core contribution of the project is to create an Electronic Thesis and Dissertation Database. The subventions of the present dissertation work include

- 1 *Digitalized the physical collection of Thesis and Dissertation.* In aiding cordial exploitation of the intellectual output of the department to its fuller potential, simultaneously evading wither and tear of the primary manuscripts.
- 2 *Gathered Metadata of Thesis and Dissertation Collection.* To foster a reliable database with sound indexing and classification.
- 3 *Finding useful features of Open Source Digital Library software.* To bet the best possible software solution in developing the digital database.
- 4 *Explored different components of Greenstone Digital library software.* In amending the design of the ETD Database.
- 5 *Customized Greenstone Macros* to create an affable user interface in accessing offline ETD Database.

6.4 Research Gaps

The most important constrain during the development of the collection is the digitization of Thesis due to the current Covid-19 situation accessing the physical copies of the thesis has been restrained. Hence the process digitizing the bibliographic contents of thesis had been held back, limiting the collection to digitized form of dissertations.

Another significant limitation of the work is Standardization of Nomenclature for subject headings. As the dissertation work has been a time bounded one. There had been a tight rein on assigning a standard nomenclature for 200+ records.

Again, due to the prevailing covid situation Real time evaluation of the database had been impeded. Facilitating the target patrons, the access to the database, evaluating their user experience, getting feedback of the prototypes and reconstructing the prototypes made based on the recommendation had been saddled with the access restrain.

After all the attempts were made to incorporate individual PDF copies of Abstract, TOC, Bibliography under same record. the idea was mete out due to time constrains.

The process of ETD Database development had to put a ceiling on Greenstone Digital Library Software Version 3.09 as the version emanates with higher level HTML programing and comes with their own set of UI design packages restraining the customization of greenstone macros and ultimately the end user interface.

6.5 Recommendation for future research

There is always a room for further research in creating a reliable Greenstone digital library and user-friendly interface. Some of this corresponds to development of resources, while others refer to specific details of implementing and improving current model.

In this work, Greenstone digital library software program is executed to build a collection of record containing metadata of dissertation submitted in the department till date. Therefore, the collection of the thesis and dissertation metadata can be extended with years to come.

Further scope of the work is to explore Greenstone digital library software more and more. The greenstone macros for developing collection of records with multiple document types can still be explored. There is still scope for further study on Greenstone Digital Library Software Version 3.09.

The User experience on the database can be studied, and their feedback can be used for improving the mobilization and customizing the macros to attain robustness in the ETD collection and user friendliness in accessing the ETD Database.

The Bibliometric analysis on the metadata gathered from Thesis and Dissertation collection of the department gives a complete portrait of Intellectual output of the department: Trends, growth and Obsolescence of different facets of library and information science, relationship between facets, structure of subjects, bibliographic coupling, co-citation, citation patterns, citation behavior and the list goes on,

6.6 Conclusion

Overall, there are positive factors influencing to adopt ETD Database. A big incentive for adopting electronic collection of thesis and dissertation is that patrons are now expecting to be provided with e-services. The second reason for adopting digital services is that it is also more economic to provide e- resources rather than to use primary hardcopy materials. The main advantage of ETD Database is its accessibility for all students and staff.

On the other hand, some negative factors that prevent libraries from adopting e-services are worth considering. Funding is a major problem that constrains libraries from adopting new services, it is necessary to investigate their users' information needs and to decide resources and services need to be provided. The ETD Database based on open source greenstone digital library software is paradigm revelation in such a case.

Some other obstacles in implementing digital services were whether some resources are suitable to appear online was the first issue to be considered. It is also necessary to keep a reasonable balance between paper-based resources and online resources based on users' information needs. The ETD Database provides its potential user only the metadata, abstract and bibliography of the Thesis and Dissertation preserving the integrity of the scholarly work.

The field of Digital Libraries experiences constant and rapid changes. Looking ahead, ETD will affect libraries' work processes and librarians' jobs. Libraries have a responsibility to accommodate all users' needs and should provide the necessary help to support users' information needs. Academic library must also ensure that the intellectual output of the institution is well preserved. At the same time, the content of the intellectual output must be exploited to its fuller potential. This ETD Database can serve as a handy tool in preserving as well exploiting the scholarly output of the Department of Library and Information Science, Pondicherry University.

Bibliography

Bibliography

- Aalberg, T., Papatheodorou, C., Dobreva, M., & Tsakonas, G. (2004). Research and Advanced Technology for Digital Libraries. *8th European Conference, ECDL 2004, 8092 LNCS*, 156–167. <https://doi.org/10.1007/978-3-642-40501-3>
- Ahammad, N. (2019). Open source digital library on open educational resources. *Electronic Library*, 37(6), 1022–1039. <https://doi.org/10.1108/EL-11-2018-0225>
- Alam, N., & Pandey, P. (2012). GeoTheses: Development of union catalogue of Indian geoscience theses using GSDL. *The Electronic Library*, 30(4), 456–468. <https://doi.org/10.1108/02640471211252175>
- Aldoobie, N. (2015). ADDIE Model. *American International Journal of Contemporary Research*, 5(6), 68–72.
- Anuradha, K. T. (2005). Design and development of institutional repositories: A case study. *The International Information & Library Review*, 37(3), 169–178. <https://doi.org/10.1016/j.iilr.2005.07.006>
- Anuradha, K. T., Sivakaminathan, R., & Kumar, P. A. (2011). Open-source tools for enhancing full-text searching of OPACs: Use of Koha, Greenstone and Fedora. *Program*, 45(2), 231–239. <https://doi.org/10.1108/00330331111129750>
- Arora, J. (2004). Network-enabled digitized collection at the Central library, IIT Delhi. *International Information and Library Review*, 36(1), 1–11. <https://doi.org/10.1080/10572317.2004.10762618>
- Bainbridge, D., Thompson, J., & Witten, I. H. (2003). Assembling and enriching digital library collections. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, 2003-Janua*, 323–334. <https://doi.org/10.1109/JCDL.2003.1204885>
- Biswas, G., & Paul, D. (2010). An evaluative study on the open source digital library softwares for institutional repository : Special reference to Dspace and greenstone digital library. *Journal of Library and Information Science*, 2(1), 1–10.
- Boddie, S., & Thompson, J. (2008). Coping with very large digital collections using greenstone. ... on Very Large Digital http://www.researchgate.net/publication/228765296_Coping_with_very_large_digital_collections_using_greenstone/file/9fcfd506ca05229a9d.pdf
- Boock, M., & Kunda, S. (2009). Electronic thesis and dissertation metadata workflow at Oregon state university libraries. *Cataloging and Classification Quarterly*, 47(3–4), 297–308. <https://doi.org/10.1080/01639370902737323>
- Chang, S. H. (2003). Institutional repositories: the library's new role. *Oclc Systems & Services: International Digital Library Perspectives*, 19(3), 77–79. <https://doi.org/10.1108/10650750310733199>
- Choi, N., & Pruett, J. A. (2019). The context and state of open source software adoption in US academic libraries. *Library Hi Tech*, 37(4), 641–659. <https://doi.org/10.1108/LHT-02-2019-0001>

0042

Creating Digital Libraries Using Gsdl. (2000).

Deb, S. (2006). TERI Integrated Digital Library initiative. *Electronic Library*, 24(3), 366–379. <https://doi.org/10.1108/02640470610671213>

Dhamdhere, S. N., Smet, E. De, Naikal, A., Paynter, G., Choi, N., Pruett, J. A., Library, P. H., Angelo, S., Boddie, S. J., Thompson, J., Parichi, R., Nisha, F., سعید، ف., Witten, I. H., Ahammad, N., Gul, S., Ganaie, S. A., Asia, S., Doctor, G., ... Dzandza, P. E. (2013). Developing an Institutional Repository Using Greenstone Digital Library (Gsdl) Software. *Bristol Medico Chirurgical Journal*, 25(4), 53–56. <https://doi.org/10.1007/s11042-014-2155-3>

Dhanabalan A., P. R. (2012). Design and Development of Institutional Repositories at Annamalai University. *Journal of Advances in Librar Nad Information Science*, 1(August), 32.

Doctor, G. (2007). Knowledge sharing: Developing the digital repository of SIPS. *Vine*, 37(1), 64–73. <https://doi.org/10.1108/03055720710742043>

Dublin Core Metadata Initiative. (n.d.). Retrieved May 23, 2020, from <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>

Dzandza, P. E. (2019). Digitizing the intellectual output of Ghanaian universities. *Collection and Curation*, November. <https://doi.org/10.1108/CC-05-2019-0012>

Eden, B., Mccutcheon, S., Kreyche, M., Beecher Maurer, M., & Nickerson, J. (2008). Morphing metadata: Maximizing access to electronic theses and dissertations. *Library Hi Tech*, 26(1), 41–57. <https://doi.org/10.1108/07378830810857799>

Fox, E., Mcmillan, G., Eaton, J., & Tech, V. (n.d.). UNESCO International Guide for the Creation of ETDs by. *Society*.

Fuhr, N., Tsakonas, G., Aalberg, T., Agosti, M., Hansen, P., Kapidakis, S., Klas, C. P., Kovács, L., Landoni, M., Micsik, A., Papatheodorou, C., Peters, C., & Sølvberg, I. (2007). Evaluation of digital libraries. *International Journal on Digital Libraries*, 8(1), 21–38. <https://doi.org/10.1007/s00799-007-0011-z>

Ghosh, M. (2009). E-theses and Indian academia: A case study of nine ETD digital libraries and formulation of policies for a national service. *International Information and Library Review*, 41(1), 21–33. <https://doi.org/10.1080/10572317.2009.10762794>

Goh, D. H. L., Chua, A., Khoo, D. A., Khoo, E. B. H., Mak, E. B. T., & Ng, M. W. M. (2006). A checklist for evaluating open source digital library software. *Online Information Review*, 30(4), 360–379. <https://doi.org/10.1108/14684520610686283>

Gohnim, R. R. (2011). CURRENT TREND AND DEVELOPMENT OF INSTITUTIONAL REPOSITORIES IN INDIA. *International Journal of Information Research*, 1(1).

Greenstone Blog. (n.d.). Retrieved January 23, 2020, from <http://www.greenstone.org/blog>

Greenstone Digital Library Software. (n.d.). Retrieved January 23, 2020, from

- <http://www.greenstone.org/>
- Greenstone Wiki.* (n.d.). Retrieved January 23, 2020, from <http://wiki.greenstone.org/doku.php>
- Gul, S., Bashir, S., & Ganaie, S. A. (2019). Evaluation of institutional repositories of South Asia. *Online Information Review*, 44(1), 192–212. <https://doi.org/10.1108/OIR-03-2019-0087>
- Hartmanis, J., & Leeuwen, J. Van. (n.d.). *Research and Advanced Technology for Digital Libraries*.
- Harvell, T. A. (2005). Electronic Resources Management Systems. *The Serials Librarian*, 47(4), 125–136. https://doi.org/10.1300/j123v47n04_12
- Hirwade, M. A. (2011). Digitization of Theses and Dissertations : Status Quo India. *Proceedings of the 14th International Symposium on Electronic Theses and Dissertations, September*, 13–17.
- Importing Files Into Greenstone.* (n.d.). Retrieved January 15, 2020, from <http://frazernash-usa.com/digital/Greenstone-Imports.htm>
- Institutional Digital Repository, IIT Kharagpur.* (n.d.). Retrieved December 31, 2019, from www.idr.iitkgp.ac.in
- Institutional Repository, Indian Institute of Management, Ahmedabad.* (n.d.). Retrieved December 31, 2020, from <http://vslr.iima.ac.in:8080/xmlui/>
- K. Gupta, D., & Gupta, N. (2014). Analytical study of the ETD repositories and government initiatives for depositing ETDs in India. *Library Management*, 35(4/5), 308–319. <https://doi.org/10.1108/lm-09-2013-0092>
- Kamila, K. (2011). Institutional Repository Building through GSDL : A Case Study of Burdwan University. *International Conference on Digital Libraries & Knowledge Organization*.
- Kemp, K., & Beard, K. (2014). Digital Library. In *Encyclopedia of Geographic Information Science*. <https://doi.org/10.4135/9781412953962.n48>
- Khan, S. Z., & Tiwari, R. K. (2017). Electronic thesis and dissertation: their importance and need for research work". *ISST Journal of Advances In Librarianship*, 6(2), 43–51.
- Khaparde, V., & Ambedkar, B. (2014). Growth and Development of Electronic Theses and Dissertation (ETDs) in India. *Journal of Library and Information Sciences*, 2(1), 99–116.
- Koganuramath, M. M., & Angadi, M. (2003). Design and Development of Digital Library : an initiative at TISS. *5th National MANLIBNET Convention, Jamshedpur (India)*, 8313. <http://eprints.rclis.org/4978/>
- Kökörčený, M., & Bodnárová, A. (2010). Comparison of digital libraries systems. *International Conference on Data Networks, Communications, Computers - Proceedings*, 97–100.
- Koulouris, A., Kyriaki-Manessi, D., Giannakopoulos, G., & Zervos, S. (2013). Institutional Repository Policies: Best Practices for Encouraging Self-Archiving. *Procedia - Social and Behavioral Sciences*, 73(February), 769–776. <https://doi.org/10.1016/j.sbspro.2013.02.117>
- Kruk, S. R., & McDaniel, B. (2009). Semantic digital libraries. *Semantic Digital Libraries, Design and Development of ETD DB for DLIS,PU using GSDL*

- August 2016, 1–245. <https://doi.org/10.1007/978-3-540-85434-0>
- Larivière, V., Zuccala, A., & Archambault, É. (2008). The declining scientific impact of theses: Implications for electronic thesis and dissertation repositories and graduate studies. *Scientometrics*, 74(1), 109–121. <https://doi.org/10.1007/s11192-008-0106-3>
- Lihitkar, S. R. (2013). Designing a prototype virtual learning environment for library and information science students. *Library Hi Tech News*, 30(4), 13–15. <https://doi.org/10.1108/LHTN-04-2013-0019>
- Lubas, R. L. (2009). Defining best practices in electronic thesis and dissertation metadata. *Journal of Library Metadata*, 9(3–4), 252–263. <https://doi.org/10.1080/19386380903405165>
- Manjunatha, K. (2014). *History and growth of institutional repositories 3.0 introduction*. Derek 2007, 52–97. shodhganga.inflibnet.ac.in/bitstream/10603/14547/8/08_chapter 3.pdf
- Mcmillan, G. (1987). *E-THESIS REPOSITORIES : AN OVERVIEW* (pp. 127–177).
- Mishra, R., Vijaianand, S. K., & Shukla, G. (2007). *Development of ETD Repository at IITK Library using DSpace*. July, 249–259.
- Mittal, R., & Mahesh, G. (2008). Digital libraries and repositories in India: An evaluative study. *Program*, 42(3), 286–302. <https://doi.org/10.1108/00330330810892695>
- Naikal, A. (2014). *Design and Development of Digital Library : A Conceptual Framework for ATEC Central Library Abstract : April 2004*.
- Namdev Dhamdhere, S., De Smet, E., & Lihitkar, R. (2014). ABCD open source software for managing ETD repositories. *Library Management*, 35(4/5), 387–397. <https://doi.org/10.1108/lm-08-2013-0072>
- Nielsen, Jakob. (n.d.). *10 Heuristics for User Interface Design*. Retrieved May 16, 2020, from www.nngroup.com
- Nielsen, Jukob, Molich, R., & Bitnet Denmark, J. (1990). *HEURISTIC EVALUATION OF USER INTERFACES*.
- Nisha, F., & Parichi, R. (2015). Greenstone Digital Library Management System – a functional review based on selected criteria. *Library Hi Tech News*, 32(10), 16–21. <https://doi.org/10.1108/LHTN-06-2015-0041>
- Osborn, W., Fox, S., Bainbridge, D., & H.Witte, I. (2011). Automating the Maintenance of Greenstone Collections. *Digital Libraries - Methods and Applications*. <https://doi.org/10.5772/14121>
- Paknejad, A. (2013). Technical assessment of Greenstone toward development of digital libraries in Iran. *Electronic Library*, 31(5), 593–607. <https://doi.org/10.1108/EL-01-2011-0010>
- Paynter, G. (2002). Greenstone: Collection management for digital works. ... *Source Projects in* <http://greenstonesupport.iimk.ac.in/greenstone2010/pdf/articles/GWP-IHW-DB-SB-Greenstone.pdf>

- Places, Á. S., Fariña, A., Luaces, M. R., Pedreira, Ó., & Seco, D. (2016). A workflow management system to feed digital libraries: Proposal and case study. *Multimedia Tools and Applications*, 75(7), 3843–3877. <https://doi.org/10.1007/s11042-014-2155-3>
- Raghavan, K., Kumar, K., Professor, R., Kashyap, M., Viswanathan, T., Yanthan, Z., Sharma, J., Talwar Kanungo, N., Talwar Kanungo Uma Kanjilal, N., Soni Vijendra, S., & Sethi, R. (2014). *Digitisation and Digital Libraries–DSpace and GSDL*. <http://www.ignou.ac.in>
- Rajasekharan, K., & Nafala, K. (2007). Easy Ways to Customize the Display of Greenstone Pages, 2007. *Kerala Institute of Local Administration.(Published)[Guide/Manual]*, 1–12.
- Rajasekharan, K., & Nafla, K. M. (2007). *Building up a Digital Library with Greenstone A Self-Instructional Guide for Beginner's*.
- Randhawa, S. (2012). Open Source Software for Creation of Digital Library : A Comparative Study of Greenstone Digital Library Software & DSpace. *Indian Journal of Library and Information Science*, 6(3), 45–52.
- Rani, K. (2011). Brief Information about Institutional Repository. In *Library & Information Science Network*. <http://www.lisbdnet.com/brief-information-institutional-repository/>
- Reinbold, S. (2013). Using the ADDIE Model in Designing Library Instruction. *Medical Reference Services Quarterly*, 32(3), 244–256. <https://doi.org/10.1080/02763869.2013.806859>
- Riddle, K. (2015). Creating policies for library publishing in an institutional repository: Exploring purpose, scope, and the library's role. *OCLC Systems and Services*, 31(2), 59–68. <https://doi.org/10.1108/OCLC-02-2014-0007>
- Saracevic, T. (2004). Evaluation of digital libraries : An overview Tefko Saracevic School of Communication , Information and Library Studies. *Methodology, September*.
- Satyanarayana, K. V., & Babu, B. R. (2007). *Trends in the development of E-Theses in India: issues, constraints and solutions*. <http://epc.ub.uu.se/ETD2007/files/papers/paper-17.pdf>
- Seshaiah. (2009). *Developing an Institutional Repository Using Greenstone Digital Library (Gsdl) Software*.
- Shaw, E., Garcia-Spitz, C., Bragg, M., Hagedorn, K., & Porter, E. (2018). *Finding the balance: Modelling successful workflows for digital library collections*. 6(3), 295–311.
- Sheeja, N. K. (2012). Knowledge management and open access e-theses: Indian initiatives. *Library Review*, 61(6), 418–427. <https://doi.org/10.1108/00242531211284339>
- Smithsonian/NASA Astrophysics Data System (ADS)*. (n.d.). Retrieved December 31, 2020, from <https://science.nasa.gov/astrophysics/astrophysics-data-centers/smithsonian-nasa-astrophysics-data-system-ads>
- Sonkar, S. K., Makhija, V., Ashok Kumar, A. K., & Singh, M. (2005). Application of Greenstone Digital Library (GSDL) Software in Newspapers Clippings. *DESIDOC Bulletin of Information Technology*, 25(3), 9–17. <https://doi.org/10.14429/dbit.25.3.3655>
- Srivastava, P., Parabhoi, L., & Sonkar, S. . (2016). E- Resources Collection Development Policy.

- International Journal of Library and Information Studies*, 6(3), 69–76.
http://ijlis.org/img/2016_Vol_6_Issue_3/69-76.pdf
- Sumner, T., & Marlino, M. (2004). Digital libraries and educational practice: A case for new models. *Proceedings of the ACM IEEE International Conference on Digital Libraries, JCDL 2004*, 1, 170–178.
- Sutcliffe, A. (2001). Heuristic evaluation of website attractiveness and usability. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2220 LNCS, 183–198. https://doi.org/10.1007/3-540-45522-1_11
- Tapkir, P. V. (2017). *Application of Greenstone Digital Library Software (Gsdl) in Collection*. 3(04), 721–729.
- Tramboo, S., Humma, H., M Shafi, S., & Gul, S. (2012). A study on the Open Source Digital Library Software's: Special Reference to DSpace, EPrints and Greenstone. *International Journal of Computer Applications*, 59(16), 1–9. <https://doi.org/10.5120/9629-4272>
- Tsakonas, G., & Papatheodorou, C. (2008). Exploring usefulness and usability in the evaluation of open access digital libraries. *Information Processing and Management*, 44(3), 1234–1250. <https://doi.org/10.1016/j.ipm.2007.07.008>
- Tsakonas, G., Papatheodorou, C., Saracevic, T., Fuhr, N., Tsakonas, G., Aalberg, T., Agosti, M., Hansen, P., Kapidakis, S., Klas, C. P., Kovács, L., Landoni, M., Micsik, A., Papatheodorou, C., Peters, C., Sølvberg, I., Sutcliffe, A., Rajasekharan, K., Nafala, K., ... Paynter, G. G. (2007). Open Source Software for Creation of Digital Library : A Comparative Study of Greenstone Digital Library Software & DSpace. *Online Information Review*, 6(1), 288–298. <https://doi.org/10.1080/19386380903405165>
- Tutorials Point*. (n.d.). <https://www.tutorialspoint.com/>
- UNESCO*. (n.d.). Retrieved December 31, 2020, from <https://en.unesco.org/>
- Unwin, T. (2009). The technologies: identifying appropriate solutions for development needs. *ICT4D - Information and Communication Technology for Development*, 76–124.
- Verma, L., & Kumar, N. (2018). Comparative analysis of open source digital library softwares: A case study. *DESIDOC Journal of Library and Information Technology*, 38(5), 362–369. <https://doi.org/10.14429/djlit.38.5.12425>
- W 3 School*. (n.d.). Retrieved April 5, 2020, from www.w3schools.com
- W3.org*. (n.d.). Retrieved April 5, 2020, from <https://www.w3.org/TR/REC-html40/struct/links.html>
- Webley, L., & Suleman, H. (2011). *Creating a National Electronic Thesis and Dissertation Portal in South Africa* (Issue September).
- WebsiteSetup*. (n.d.). Retrieved April 5, 2020, from <https://websitesetup.org/>
- Witten, I. H. (1997). *Customizing digital library interfaces with Greenstone Customization in Greenstone*. 1–6.

- Witten, I. H. (2004). Creating and customizing digital library collections with the Greenstone Librarian Interface. *Proc. International Symposium on Digital Libraries and Knowledge Communities in Networked Information Society, DLKC. 2004*, 97–104.
<http://researchcommons.waikato.ac.nz/handle/10289/1802>
- Witten, I. H. (2009). The Development and Usage of the Greenstone Digital Library Software. *Bulletin of the American Society for Information Science and Technology*, 36(2), 31–38.
- Witten, I. H., & Bainbridge, D. (2005). Creating digital library collections with Greenstone. *Library Hi Tech*, 23(4), 541–560. <https://doi.org/10.1108/07378830510636337>
- Witten, I. H., Bainbridge, D., & Boddie, S. J. (2001). Greenstone: Open-source digital library software with end-user collection building. *Online Information Review*, 25(5), 288–298. <https://doi.org/10.1108/14684520110410490>
- Witten, I. H., Bainbridge, D., Paynter, G., & Boddie, S. (2002). Importing documents and metadata into digital libraries: Requirements analysis and an extensible architecture. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 2458). https://doi.org/10.1007/3-540-45747-x_29
- Witten, I. H., Bainbridge, D., Tansley, R., Huang, C. Y., & Don, K. J. (2005). StoneD: A bridge between Greenstone and DSpace. *D-Lib Magazine*, 11(9). <https://doi.org/10.1045/september2005-witten>
- Witten, I. H., Loots, M., Trujillo, M. F., & Bainbridge, D. (2002). The promise of digital libraries in developing countries. *The Electronic Library*, 20(1), 7–13. <https://doi.org/10.1108/02640470210418218>
- Witten, I. H., McNab, R. J., Boddie, S. J., & Bainbridge, D. (2000). Greenstone: A comprehensive open-source digital library software system. *Proceedings of the ACM International Conference on Digital Libraries*, 113–121.
- Zhang, B. A. B. (2008). *Customizing Greenstone User Interface Greenstone User Interface Overview*. 1–20.

Appendices

Appendix A

List of Abbreviation

Abbreviation	Expansion
ANSEL	American National Standard for Extended Latin
ASCII	American Standard Code for Information Interchange
CCD	Charge Coupled Device
CD-ROM	Compact Disc Read-Only Memory
CIS	Contact Image Sensor
CMOS	Complementary Metal-Oxide Semiconductor Sensor
CSS	Cascading Style Sheets
CSV	Comma Separated Value
dc.	Dublin Core
DCMDS	Dublin Core Metadata System
DL	Digital library
DVD	Digital Versatile Disk
EACC	East Asian Character Code
ERM	Electronic Resource Management
ETD	Electronic Thesis and Dissertation
ETD-MS	Electronic Theses and Dissertation Metadata Standard
GSDL	Greenstone Digital Library
HTML	Hyper Text Markup Language
ICT	Information Communication Technology
IR	Institutional Repository
IRS	Information Retrieval System
ISO	International Organization for Standardization
JRE	Java Runtime Environment
MARC	Machine Readable Cataloging
MODS	Metadata Object Description Schema

NDLTD	Networked Digital Library of Thesis and Dissertation
OAI-PMH	Open Access Initiative Protocol for Metadata Harvesting
OCR	Optical Character Recognition
OKI	Open Knowledge Initiative
OSM	Open Source Movement
OSS	Open Source Software
PDF	Portable Document Format
PERL	Practical Extraction and Report Language
PMT	Photo Multiplier Tube
QDCMDS	Qualified Dublin Core Metadata System
SEM	Search Engine Marketing
SEO	Search Engine Optimization
SWOT	Strength Weakness Opportunity Threat
UI	User Interface
UX	User Experience
XML	Extensible Markup Language

Appendix B

Librarian Interface Guide

To base the periodic updation of the collection on the same structure as the existing ETD Database.

1. Install GSDL Software.
2. Run the GSDL Librarian Interface in the default Librarian mode.
3. Move Source collection of the ETD Database from CD-ROM to GSDL Librarian Interface,
 - a. choose "Open" from the "File" menu to get the Open Collection prompt.
 - b. click "Browse" for a file system browsing dialog.
 - c. Select ETD Database collection to see its description, and
 - d. click "Open" to load ETD Database.
4. The Gather panel from Librarian Interface, allows to explore the local file space and gather up new thesis/dissertations for updating the collection.
5. Enrich the documents by adding new metadata values to the set of existing values. The chronological hierarchy of the records can be extended in the same way.
6. As ETD Database Collection's design and indexing configurations were built within the package, once records are marked up with metadata, rebuild the collection by clicking "Build Collection" from the Create tab.
7. When error prompts appear during collection building process. switch the GLI into expert mode, set the build option "verbosity" to 5, and rebuild, check for error messages and then reboot them with developer's guide¹¹.

¹¹ http://www.greenstone.org/manuals/gsdl2/en/html/Develop_en_index.html

8. Once the collection gets fully build, review the database collection by clicking "Preview Collection" from the Create tab.
9. The updated collection can to be written back to the CD-ROM,
 - a. Open the "File" menu and choose "Write CD/DVD Image".
 - b. Tick the check boxes of the ETD Database collection to export.
 - c. Enter ETD Database in the "CD/DVD name" box:
10. Upon completion, Greenstone will show the name of a folder containing the exported collections. Use the CD/DVD writer to copy its contents to a blank CD/DVD.

Appendix C

User Guide

To Access the Collection:

The ETD Database CD-ROM contains all the source files of the Electronic Thesis and Dissertation Database of Department of Library and Information Science, Pondicherry University.

To access the ETD Database, click on the Greenstone server icon directing to Home Page of the Department. In home page click on the ETD Database Collection icon to access the Database.

The about page of the ETD Database gives an overview about the collection, the banner serves as a common link to the about page from any other page. In addition to that the about page also includes Home, Help, Preferences tabs to ease the navigation of the user.

To Locate a File in the Collection:

Any Thesis or Dissertation metadata can be located from the database either by searching with keywords or by browsing with classifiers. For easy access, a separate tab is provided in the user interface for each classifier and search can be narrowed down with search delimiters.

By using *Search tab*, ***Search in*** category: Title, Author, Guide, Subject, Registration Number can be selected and by entering a keyword of required information in the search box, User will get Thesis/dissertation metadata as result against the keyword entered.

In the Browsing Classifiers, By Clicking on *Title tab*, the list of titles available in the collection were enlisted aligned in alphabetical order. In addition to Title, the other metadata components of Author, Registration Number, Guide, Area of Study and Year details for each record shall be displayed.

On *Subjects tab*, the records are consolidated into individual folders based on area of study and the number within the closed braces represents the number of records/studies conducted within each subject. Individual records compiled within the subject folders provides with other metadata details corresponding to the record.

In the same way, on *Guide tab* and *Year tab*, the records are compiled based on guide who supervised the project and year the project has been completed respectively.

To Read the Documents:

Once the desired level of document is reached either through searching or browsing, each record gives the complete metadata details corresponding to the record, by clicking on the respective PDF icon one can read the title page, abstract, keywords, table of Content, bibliography corresponding the Thesis or dissertation.

To set preferences:

The settings button at the top right corner of the page, will facilitate limited changes in features of the interface to suit user's self-requirements such as Language, Presentation and Search preferences.

Note: For further assistance, user can always click on Help button at the top right corner of the corresponding page.