# Unveiling the Growth and Usage Pattern of Electronic Theses and Dissertations Through National ETD Repository: A Case Study of Shodhganga

**Surbhi[1] and Rajan Kumar[2]**

[1]Scientist C (LS), INFLIBNET Centre, Gandhinagar, Gujarat
[2]Scientist B (LS), INFLIBNET Centre, Gandhinagar, Gujarat

**Abstract**

*Poor visibility and closed accessibility of printed copies of Theses available in the Libraries were the major issues among the research community in India before the year 2009. Electronic Theses and Dissertations (ETDs) which is the digitised version of traditional print theses and dissertations, have emerged as an important means of sharing, transferring and collaborating the research and publication output of any Institution. ETDs, not only, support the documents submission and archiving, but also, play an important role in collaboration and Interoperability among the Institutions in terms of research output i.e. PhD Theses. The objective of this study is to explore the growth and usage analytics of Shodhganga. The present study explored various methods to analyse the data from Shodhganga in terms of Theses submissions, usage pattern, accessibility, growth and usability, metadata formats and standards, workflow management, metadata creation, data validation, text mining, Citations export tool and a file format, etc.*

**Keywords:** Dublin Core, Electronic Theses and Dissertations (ETD), Institutional Repositories, Metadata Standards, Shodhganga, Usage Analytics

## 1. Introduction

Theses and Dissertations as a primary and rich source of information remains untapped and unpublished through a formal channel of publications which in turn leads to duplication of research and research topics and ultimately leads to plagiarism. Shodhganga is a National Digital Repository came into existence in 2010 based on the UGC Notification 2009. It is the largest National ETD Repository in India with 4,75,000+ records of full-text Theses, plays an important role to enhance academic integrity and ethics among academicians and researchers in the form of full text Theses and provides a platform to research scholars to deposit, re-use and share their research work. Shodhganga has shown a progressive growth in terms of Theses submissions as well Universities/Institutions voluntarily joined Shodhganga by signing the MoUs and started contributing the softcopies of the Theses in open access for academic and research community as a national research output to the worldwide without any restrictions/subscriptions. Compared to other types of digital libraries, a digital library of ETDs is unique in its collection, user groups, and goals. (Zhang, Lee, & You, 2001).

Corresponding Author: Dr. Surbhi, Email: surbhi@inflibnet.ac.in and Mr. Rajan Kumar, Email: rajankumar@inflibnet.ac.in

## 2. Historical Background of the Initiative

Shodhganga as national repository was first initiated in 2010 based on the UGC Notification, 2009 The UGC Notification (Minimum Standards & Procedure for Award of M.Phil. / Ph.D Degree, Regulation, 2009 amendment made on 2016) dated 5th May 2016 mandates submission of electronic version of theses and dissertations by the researchers in universities with an aim to facilitate open access to Indian theses and dissertations to the academic community world-wide. (http://shodhganga.inflibnet.ac.in/newmoredetails/about.html) .Based on the UGC Notification, the responsibility of maintaining and hosting the digital repository was assigned to the INFLIBNET Centre and the repository name was coined as "Shodhganga". Shodhganga is a word coined by combining two Indian regional language words "Shodh" which means research and "Ganga" the holiest and longest Indian river to represent the unseen and unpublished overall research output of the nation produced as PhD Theses and submitted to the Indian Universities. The UGC Notification 2009/2016 amendment made in 2022 and in addition to the recent Directives issued by the Ministry (MoE) dated 30 Nov. 2022 for CFTIs/INIs mandate for uploading the soft copies of the Theses into Shodhganga.

## 3. Review Literature

In order to establish the importance of the current study in terms of the studies already conducted and to find out the literature gap of the studies conducted on the same issues, the authors have thoroughly studied the available literature in the field and tried to find out the relevant studies related to the concept. Various published and available literature was studied using Books, journals, conference papers, online databases etc. using the Keywords like, Shodhganga, Institutional repositories (IRs), Online Theses and Dissertations, Usage Analytics, Metadata Standards, Dublin Core etc. Some of the relevant studies from the last five years 2019-2023 are thoroughly reviewed and found that the past studies have covered and discussed Shodhganga as an institutional repository and its growth, but no study was found covering growth and usage analytics, State Wise MoU signed and their Contribution to Shodhganga etc.

Kenning Arlitsch (2020) in their study demonstrates that aggregated data from the Repository Analytics and Metrics Portal (RAMP) have significant potential to analyze visibility and use of institutional repositories (IR) as well as potential factors affecting their use, including repository size, platform, content, device and global location. The RAMP dataset is unique and public. Shajitha & Majeed (2021) evaluate the institutional repositories (IRs) in South India in terms of policy and procedures, technology, content and contributors, promotion and assessment and personnel.

(Gul et. al.,2020) explored the status of institutional repositories (IRs) in the South Asian region. The various characteristic features of IRs are studied. India, Sri Lanka and Bangladesh lead other South Asian nations in terms of IRs count. Majority of the IRs are operational in nature with a higher number of operational IRs from India. In terms of record count, India leads the list. "Journal articles" outscore other content types and majority of the IRs have OAI-PMH as their base URL. DSpace is a prioritized software for content management

in IRs. Majority of the IRs have not defined their content management policies. English stays a prioritized language of the content dotting the South Asian IRs and majority of the IRs not providing usage statistics. A good score of IRs has incorporated Web 2.0 tools in them with RSS as the preferred Web 2.0 tool. A good count of the IRs has not customized their interface. Majority of the IRs have interfaces in two languages. Chapepa et. al, (2023) investigated the metadata creation practices in a functional academic institution repository in Malawi, with a specific focus on the Lilongwe University of Agriculture and Natural Resources (LUANAR) library. Paskali & Ivanoviæ, (2021) determined the digital library usage patterns as a means of improving the system, as well as the user experience, to give appropriate recognition to the most popular dissertations' authors and to measure the interest of non-academic users for dissertations defended at the University of Novi Sad (UNS). Brush & Jiras (2019) shared the knowledge and lessons learned about the process of developing an institutional repository (IR) using a hosted solution, Digital Commons from bepress, and to make the case that Digital Commons is still the best IR solution for smaller university libraries. Wani & Wani (2019) measured the fully open access (OA) contents cited by researchers from top three universities in the USA (MIT, CIT and Stanford) in the field of physics. This study aims to identify the core fully OA journals widely used by researchers in the field of physics and evaluate the cited fully OA resources by applying various scientometric parameters to check the credibility of these OA resources. Rasuli, et. al. (2023) in this study highlighted, only those policies are reviewed that are available, discoverable and accessible on HEIs' websites.Highlighting the detrimental effect of not managing stipulations towards embargoes clearly, the findings could be useful for national/institutional policymakers and administrators of research departments, academic libraries, institutional repositories and graduate offices. This is the first study to investigate rationales for TDs embargo practices. It creates awareness of how embargoes are managed and reflected in policy. Ultimately, it recommends further interrogation on how embargoes influence the principle of openness to scholarship. Shajitha C. (2020) found that the active participation of South Indian IRs was only seen in a few digital curation activities. However, of the 33 digital curation activities analyzed, the active participation of repositories was only seen in ten digital curation activities. The performance of preservation activities was extremely low, and disagreements were recorded by the survey participants toward several digital curation activities. The most disagreed digital curation activities were emulation and cease data curation. All the participants had assigned metadata and allowed file downloads in their repositories. Raman Research Institute had provided a good number of digital curation services in their IR. In a study it was found that the widely used metadata schema – Dublin Core (DC) is not appropriate to describe the contents of the human book. It shows that selected metadata elements from the types – person and event of schema.org can be used for describing, organizing and archiving the resources of the human library. It further highlights that existing subject entries are not sufficient to standardize the contents of these types of resources (Jana & Rout, 2022). Ukwoma, et. al.(2019) study took the form of a descriptive survey, gathering data from academics and repository managers from 15 universities that have their IR captured on OpenDOAR.

## 4. Objectives

The objective of this study is to explore the growth and usage analytics of Shodhganga and provide the insights into the various factors mentioned-below:

- To unveil the growth and development of theses submitted to Shodhganga
- To study and distinguish the usage patterns of data through DSpace and Google analytics
- To examine the Metadata Standards used in Shodhganga as compared to other Indian and Global ETDs in light of harvesting and interoperability standards.
- To make a comparison of the DSpace Metadata fields as well as the Modified Dublin Core Metadata fields adopted in Shodhganga repository.
- To highlight the status and leap in the Theses submissions by the contributing Universities into Shodhganga for National Ranking of HEIs.
- Exploring and highlighting the new features added to the repository for easy workflow and effortless submissions by the contributing Institutions to avoid duplication of work.
- To reveal the Top most viewed Theses from Shodhganga.
- To study the state wise MoU Signed and their contribution to the Shodhganga
- To discuss the challenges faced in the maintenance, workflow and policies adopted to host open access content Shodhganga.

## 5. Methodology

The present study explored various methods to analyse the data from Shodhganga in terms of Theses submissions, usage pattern, accessibility, growth and usability, metadata formats and standards, workflow management, metadata creation, data validation, text mining, Citations export tool and a file format etc. In this paper, several technical changes were also addressed which include Enhanced Submission Interfaces and Session Time Inclusion to the portal.

Usage analytics of ETDs is very important to analyse the usage pattern and accessibility of the ETD content. Monthly usage statistics of Shodhganga were generated with the help of DSpace software. Daily page view track and reports were captured with the help of Google analytics. The data is collected from Shodhganga out of 4.75+ lakhs full text Theses records contributed by 726 Universities/CFTIs/INIs in Shodhganga as on 31st July 2023.

However, in order to analyse the growth and development of the Shodhganga database, various in-house publications like Newsletters, Annual reports, Manuals, data available on Shodhganga website etc. are referred in order to analyse respective outcomes and growth rate in a quantitative way. In addition, the PostgreSQL database is utilized to extract the data submitted by a particular university during a specific period.

## 6. Data Analysis and Interpretation

### 6.1 Growth and Development of Shodhganga

As stated earlier Shodhganga has shown a progressive growth in terms of Theses submitted to the repository since 2010. The Shodhganga project,launched on January 27, 2010 by Prof. S.K. Thorat, Chairman, UGC. In the first year of inception, only 1100+ Theses were uploaded by the contributing universities and Mahatma Gandhi University, Kottyam was the first University that came forward to sign MoU and joined Shodhganga. In 2010, the total number of Theses uploaded 1100+ including 550 theses from Mahatma Gandhi University, Kottayam (Kerala), the first university to sign MoU with the INFLIBNET Centre. In 2012 Total number of theses submitted into the repository has grown to 2200+. After 10 years, as such, in 2022 was a remarkable year with 77,000 + Theses submitted to the repository. As on date (31st July 2023) total 4.75 + lakhs theses are uploaded to Shodhganga which is a remarkable figure in terms of full text PhD Theses submitted to the Indian national repository from 700+ Contributing Universities/CFTIs/INIs etc.

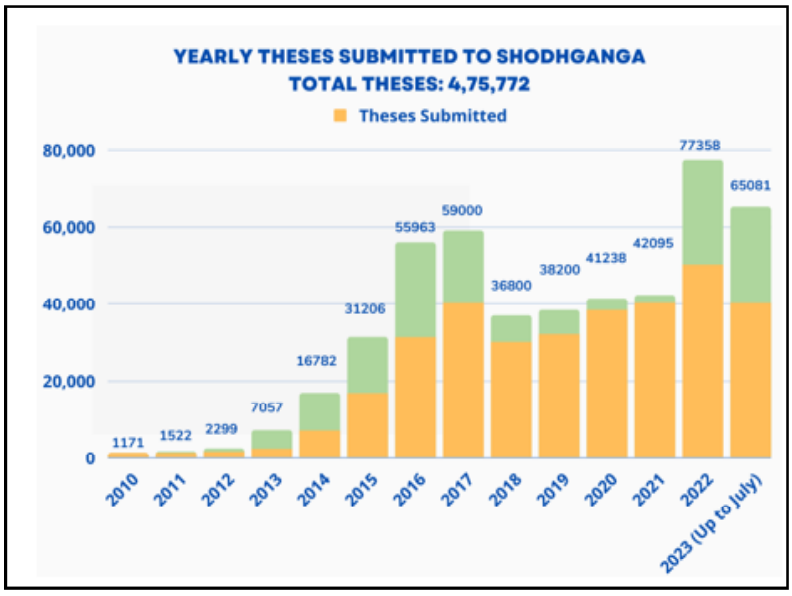



**Figure 1: Yearly Theses Submitted to Shodhganga**

### 6.2 To study and distinguish the usage patterns of data through DSpace and Google analytics

Usage statistics are crucial for repository administrators, as they provide insights into the effectiveness of the repository, user behavior, and the impact of the materials it hosts. This information can guide decisions about content management, marketing strategies, and improvements to the repository's user experience. It also justifies the expenditure on a digital repository. The popularity of a repository is reflected by its usage. If the usage is very very high then we can assume that it is a very popular repository.

Shodhganga usage analytics is extracted using various methods such as DSpace auto generated usage statistics which gives data about Page Views, Item Views, Community and Collection Views etc. as well as the Usage Analytics extracted through Google Analytics.

Each time a page of a file gets requested, this request is being logged. Here the log files are very important to generate statistics. Statistics are generated from files in [dspace]/log directory. Additionally, the Shodhganga interface is seamlessly integrated with Google Analytics, which offers even more comprehensive insights than the default DSpace statistics. It collects a broader range of data including page views, user demographics, geographic locations, referral sources, events, conversions, and user interactions with specific website elements (e.g., buttons, forms, videos).

The data from DSpace reveals a staggering 108 million total page views, with a monthly average of 9 million page views. However, the monthly page views/hits, as measured by Google Analytics, average around 147 million page views with a monthly average 12.3 million page views. These statistics were recorded during the period from April 2022 to March 2023.

**Table 1: DSpace and Google Analytics Page Views in Lacs (April 22 to March 23)**

| Month | Page Views (Lacs) | |
|---|---|---|
| | DSpace | Google Analytics |
| **April-2022** | 32.66 | 142.77 |
| **May-2022** | 19.27 | 149.73 |
| **June-2022** | 60.52 | 138.81 |
| **July-2022** | 55.76 | 131.61 |
| **Augst-2022** | 49.22 | 118.94 |
| **September-2022** | 129.28 | 121.25 |
| **October-2022** | 68.92 | 104.18 |
| **November-2022** | 178.58 | 119.75 |
| **December-2022** | 71.17 | 93.65 |
| **January-2023** | 137.95 | 112.37 |
| **February-2023** | 124.41 | 110.49 |
| **March-2023** | 155.88 | 133.02 |
| **Total** | **1083.62** | **1476.57** |

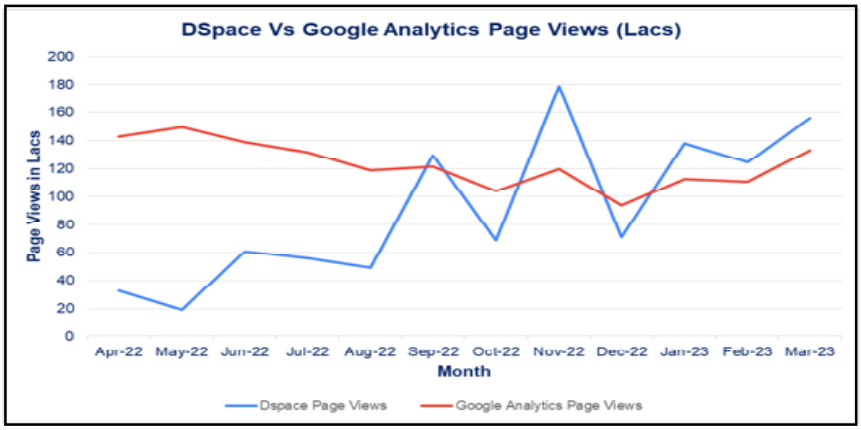



**Figure 2: Graph of DSpace and Google Analytics Page Views**

### 6.3 Metadata Standards used in Shodhganga as compared to other Indian and Global ETDs

Metadata is a set of standard Terms used to indicate or retrieve required data/ information from the digital documents. Metadata is essential to define the Institutional repositories (IRs) and ETDs digital content in the Theses and Dissertations too. Shodhganga is using DSpace, open source software that uses internationally recognized protocols and interoperability standards. Shodhganga repository also based on Dublin Core Metadata standards with few Qualifiers attached to clearly define the Metadata terms as modified Dublin Core metadata standards (DCMS). Universities based on the MoU signed are granting non-exclusive rights under INFLIBNET Centre for hosting their ETDs in Shodhganga Creative Commons Licence Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) (Attribution-NonCommercial-ShareAlike 4.0 International**).**

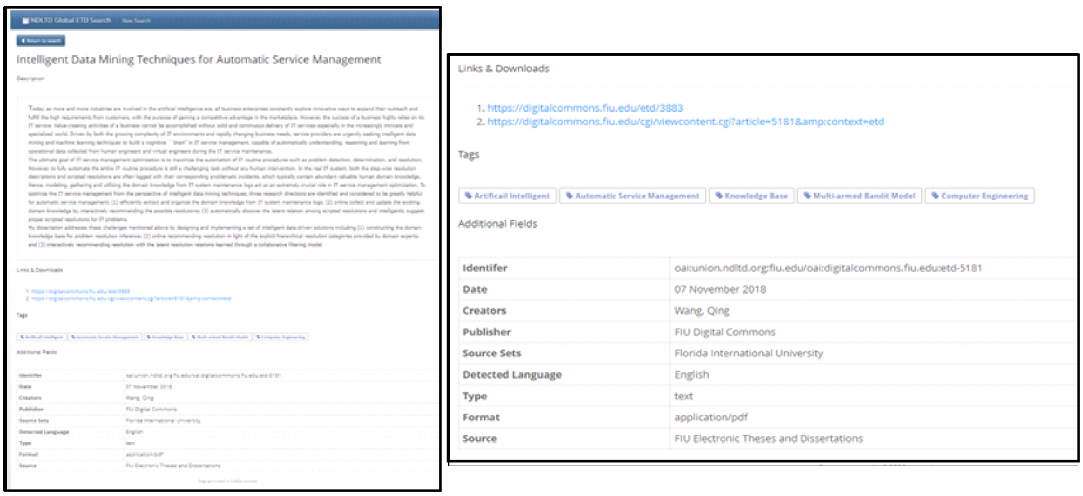



**Figure 3&4: Metadata fields from NDLTD ETD Search**

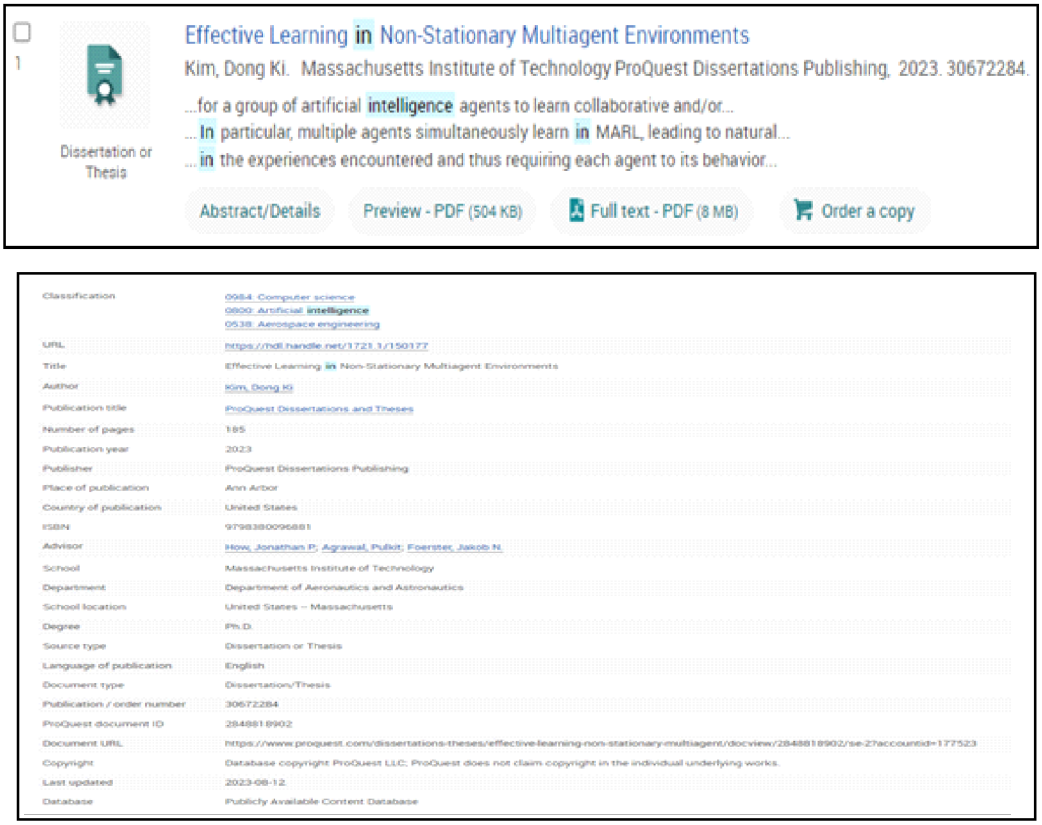



**Figure 5&6: ProQuest Theses Database**

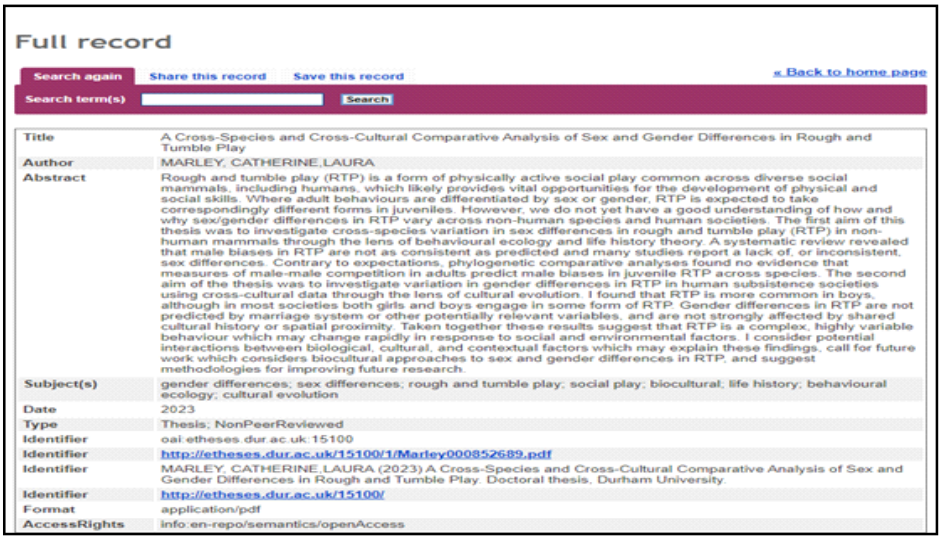
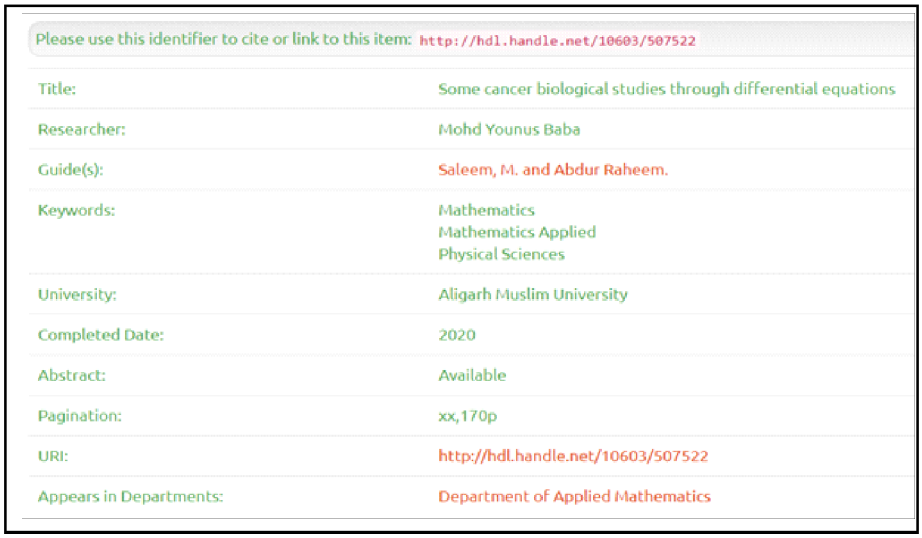



**Figure 7: DART Repository**



Please use this identifier to cite or link to this item: http://hdl.handle.net/10603/507522

| | |
|---|---|
| Title: | Some cancer biological studies through differential equations |
| Researcher: | Mohd Younus Baba |
| Guide(s): | Saleem, M. and Abdur Raheem. |
| Keywords: | Mathematics<br>Mathematics Applied<br>Physical Sciences |
| University: | Aligarh Muslim University |
| Completed Date: | 2020 |
| Abstract: | Available |
| Pagination: | xx,170p |
| URI: | http://hdl.handle.net/10603/507522 |
| Appears in Departments: | Department of Applied Mathematics |

**Figure 8: Shodhganga Metadata Standard**

### 6.4 DSpace Metadata fields as well as the Modified Dublin Core Metadata fields

Shodhganga repository using Dublin Core Metadata standards with few Qualifiers attached to clearly define the Metadata terms as modified Dublin Core Metadata Standards (DCMS). Total 31 modified Dublin Core Metadata Standards including 15 basic DCMS Standards are applied to define the Theses uploaded to the Shodhganga repository.

Full metadata record

| DC Field | Value |
|---|---|
| dc.coverage.spatial | xx,170p |
| dc.date.accessioned | 2023-08-16T10:59:11Z |
| dc.date.available | 2023-08-16T10:59:11Z |
| dc.identifier.uri | http://hdl.handle.net/10603/507522 |
| dc.description.abstract | Available |
| dc.format.extent | xx,170p |
| dc.language | English |
| dc.rights | university |
| dc.title | Some cancer biological studies through differential equations |
| dc.creator.researcher | Mohd Younus Baba |
| dc.subject.keyword | Mathematics |
| dc.subject.keyword | Mathematics Applied |
| dc.subject.keyword | Physical Sciences |
| dc.contributor.guide | Saleem, M. and Abdur Raheem. |
| dc.publisher.place | Aligarh |
| dc.publisher.university | Aligarh Muslim University |
| dc.publisher.institution | Department of Applied Mathematics |
| dc.date.completed | 2020 |
| dc.date.awarded | 2020 |
| dc.format.accompanyingmaterial | DVD |
| dc.source.university | University |
| dc.type.degree | Ph.D. |
| Appears in Departments: | Department of Applied Mathematics |

**Figure 9: Shodhganga Modified Dublin Core Metadata Standards with Qualifiers**

**Table 2: Shodhganga Modified Dublin Core Metadata Standards with Qualifiers**

| DC Field | Value |
|---|---|
| **dc.coverage.spatial** | xx,170p |
| **dc.date.accessioned** | 2023-08-16T10:59:11Z |
| **dc.date.available** | 2023-08-16T10:59:11Z |
| **dc.identifier.uri** | http://hdl.handle.net/10603/507522 |
| **dc.description.abstract** | Available |
| **dc.format.extent** | xx,170p |
| **dc.language** | English |
| **dc.rights** | university |
| **dc.title** | Some cancer biological studies through differential equations |
| **dc.creator.researcher** | Mohd Younus Baba |
| **dc.subject.keyword** | Mathematics |
| **dc.subject.keyword** | Mathematics Applied |
| **dc.subject.keyword** | Physical Sciences |
| **dc.contributor.guide** | Saleem, M. and Abdur Raheem. |
| **dc.publisher.place** | Aligarh |
| **dc.publisher.university** | Aligarh Muslim University |
| **dc.publisher.institution** | Department of Applied Mathematics |
| **dc.date.completed** | 2020 |
| **dc.date.awarded** | 2020 |
| **dc.format.accompanyingmaterial** | DVD |
| **dc.source.university** | University |
| **dc.type.degree** | Ph.D. |
| **Appears in Departments** | Department of Applied Mathematics |

## Table 3: Metadata fields DSpace and Shodhganga (Comparative) modified Schema

**Metadata Fields in DSpace and Shodhganga modified Schema**

| | **DC Fields (15 Main)** | | **DC Fields in Shodh-ganga** | **Qualifiers (Shodh -ganga)** | **Other Qualifiersin Shodhganga** | | | **D Space Default DC fields** | | **D Space Default DC fields** | **Qualifiers (D Space)** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Contributor | 1 | Contributor | Guide | | | 1 | Contributor | 1 | Contributor | |
| 2 | Coverage | 2 | Coverage | Spatial | | | | | 2 | Contributor | Advisor |
| 3 | Creator | 3 | Creator | Researcher | | | | | 3 | Contributor | Author |
| 4 | Date | 4 | Date | Registered | Accessioned | system generated | | | 4 | Contributor | Editor |
| | | 5 | Date | Completed | Available | system generated | | | 5 | Contributor | Illustrator |
| | | 6 | Date | Awarded | Issued | system generated | | | 6 | Contributor | Other |
| 5 | Description | 7 | Description | Abstract | Provenance | | 2 | Coverage | 7 | Coverage | Spatial |
| | | 8 | Description | Note | Sponsor (Adv Search) | | | | 8 | Coverage | Temporal |
| | | 9 | Description | Release | | | 3 | Creator | 9 | Creator | |
| 6 | Format | 10 | Format | Medium | | | 4 | Date | 10 | Date | |
| | | 11 | Format | Extent | | | | | 11 | Date | Available |
| | | 12 | Format | Dimensions (in cm) | Size (page numbers) | | | | 12 | Date | Accessioned |
| | | 13 | Format | Accompanying Material | | | | | 13 | Date | Copyright |
| 7 | Identifier | 14 | Identifier | URI | Series (Adv Search) | | | | 14 | Date | Created |
| | | 15 | Identifier | Thesis number | | | | | 15 | Date | Issued |
| | | 16 | Identifier | Handle | | | | | 16 | Date | Submitted |
| 8 | Language | 17 | Language | | | | 5 | Description | 17 | Description | |
| 9 | Publisher | 18 | Publisher | Place | | | | | 18 | Description | Abstract |
| | | 19 | Publisher | University | | | | | 19 | Description | Provenance |
| | | 20 | Publisher | Institution | | | | | 20 | Description | Sponsorship |
| 10 | Relation | 21 | Relation | | | | | | 21 | Description | Statement of Responsibility |
| 11 | Rights | 22 | Rights | | | | | | 22 | Description | table of Contents |
| 12 | Source | 23 | Source | URI | | | | | 23 | Description | URI |
| | | 24 | Source | University | | | 6 | Format | 24 | Format | |
| | | 25 | Source | Selfsubmission | | | | | 25 | Format | Extent |
| | | 26 | Source | Guide/supervisor | | | | | 26 | Format | Medium |
| | | 27 | Source | INFLIBNET | | | | | 27 | Format | Mimetype |
| 13 | Subject | 28 | Subject | Keyword | | | 7 | Identifier | 28 | Identifier | |
| 14 | Title | 29 | Title | | | | | | 29 | Identifier | Citation |
| | | 30 | Title | Alternative | | | | | 30 | Identifier | govdoc |

| 15 | Type | 31 | Type | Dcmitype | | | | | 31 | Identifier | isbn |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Citation Reference | | | | 32 | Identifier | ismn |
| | | | | | Appears in Department | | | | 33 | Identifier | issn |
| | | | | | | | | | 34 | Identifier | other |
| | Highlighted cells don't match | | | | | | | | 35 | Identifier | sici |
| | | | | | | | | | 36 | Identifier | uri |
| | | | | | | | 8 | language | 37 | language | |
| | | | | | | | | | 38 | language | ISO |
| | | | | | | | 9 | Publisher | 39 | Publisher | |
| | | | | | | | 10 | Relation | 40 | Relation | |
| | | | | | | | | | 41 | Relation | Has part |
| | | | | | | | | | 42 | Relation | has version |
| | | | | | | | | | 43 | Relation | is based on |
| | | | | | | | | | 44 | Relation | is format of |
| | | | | | | | | | 45 | Relation | is part of |
| | | | | | | | | | 46 | Relation | is part of series |
| | | | | | | | | | 47 | Relation | is referenced by |
| | | | | | | | | | 48 | Relation | is replaced by |
| | | | | | | | | | 49 | Relation | is version of |
| | | | | | | | | | 50 | Relation | replaces |
| | | | | | | | | | 51 | Relation | requires |
| | | | | | | | | | 52 | Relation | URI |
| | | | | | | | 11 | Rights | 53 | Rights | |
| | | | | | | | | | 54 | Rights | uri |
| | | | | | | | 12 | Source | 55 | Source | |
| | | | | | | | | | 56 | Source | uri |
| | | | | | | | 13 | Subject | 57 | Subject | |
| | | | | | | | | | 58 | Subject | Classification |
| | | | | | | | | | 59 | Subject | ddc |
| | | | | | | | | | 60 | Subject | lcc |
| | | | | | | | | | 61 | Subject | LCSH |
| | | | | | | | | | 62 | Subject | Mesh |
| | | | | | | | | | 63 | Subject | Other |
| | | | | | | | 14 | Title | 64 | Title | |
| | | | | | | | | | 65 | Title | alternative |
| | | | | | | | 15 | Type | 66 | Type | |

## 6.5 Status and Leap in the Theses submissions by the contributing HEIs into Shodhganga for National Ranking of HEIs

The Ministry of Education (MoE) has announced through its Notification dated 30th November, 2022 to capture the Theses data on number of PhD students graduated for Academic Year 202l-2022 from Shodhganga, for India Rankings 2023. Based on the Notification, all the CFTIs/INIs including IITs, IIMs, NITs, IIITs, IISc,IISERs, etc. have come forward for signing the MoU and joining Shodhganga for depositing the soft copies of their PhD Theses to the single national repository. As such, different CFTIs/INIs like IITs, IIMs already maintain their respective IRs and ETD repositories. They requested for a mechanism and portal for fetching the Metadata of the Theses records directly from their repositories to Shodhganaga for interoperability and avoidance of duplication of work and hence to solve this issue a new Bulk Metadata upload portal is designed to facilitate the Institutes to upload the already created metadata from their IRs and submitting into Shodhganga with full text record. A total 6236 metadata of Theses were uploaded using bulk metadata interface from 12 Institutions (IISc, IITs, IIMs and IISER). Due to this, there was a huge contribution by all the CFTIs/INIs. In Dec 2022 (12,450) and in January 2023, the highest number of theses (i.e. 20,369), since its inception was uploaded by all Institutions.

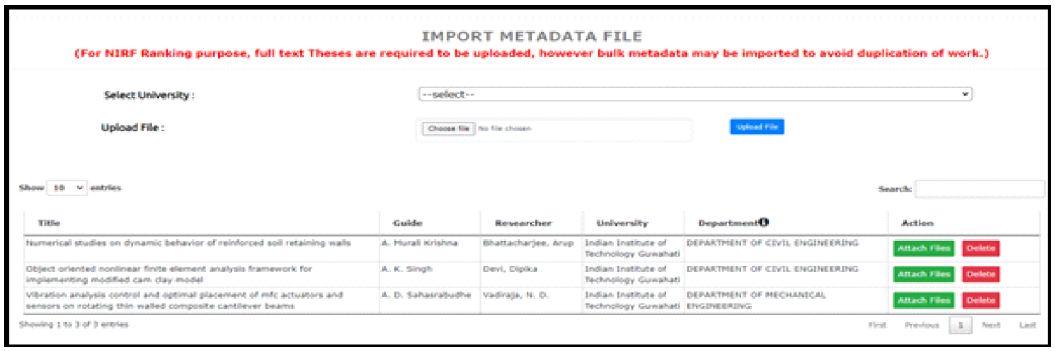
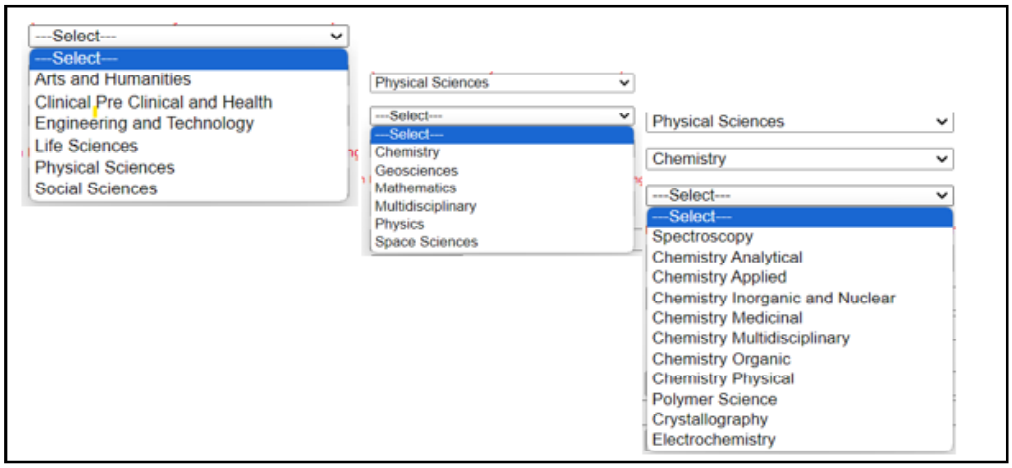



**Figure 10: Bulk Metadata upload Interface**

## 6.6 Exploring and highlighting the new features added to the repository for easy workflow and effortless submissions by the contributing Institutions to avoid duplication of work

Shodhganga as a national repository has been increasing with an exponential rate at an average rate of 53,000 Theses annually (based on last 8 years submitted data). The submissions and maintenance of Theses into Shodhganga reflects the academic structure of the Universities/Institutions in terms of Departments/Schools. Thus in order to maintain and to handle the ever increasing repository certain features have been incorporated to the repository from time to time to host, display, maintain and to provide easy searching, browsing and retrieving the Theses records by the users as well as for easy workflow, effortless submissions by the contributing Institutions to avoid duplication of work. Theses new features incorporated are as discussed below:

- The initiative got momentum when eligible universities were provided access to Plagiarism Detection Software (PDS) to all the Universities/CFTIs/INIs through the ShodhShuddhi project.
- Shodhgangotri:Initiated in 2013, Shodhgangotri is a repository of research in progress that hosts 1249 Synopses/MRPs/PDFs/Emeritus Fellowship.
- To organise the Theses records Subject Indexing is embedded into Shodhganga based on the major six subject categories are Arts and Humanities, Clinical, Pre-Clinical and Health, Engineering and Technology, Life Sciences, Physical Sciences and Social Sciences. Shodhganga consists of 06 major subject categories at the first level and 28 subject categories at the second level. The subjects are further divided into 260 micro level categories based on the Web of Science scheme.




- For assigning the Keywords The subjects are then taken from the Library of Congress Subject Heading List (LCSH) in the form of Keywords for standardizing vocabulary control and avoiding improper indexing.
- Provision of Subject Search platform for easy searching, browsing and to demonstrate the Theses records through State wise, Subject wise, Language wise etc. for easy searching and browsing (https://sgsubjects.inflibnet.ac.in/).
- Incorporating the BiBTex Citation format for importing and downloading the Citations for easy referencing and citing the Theses records in Mendeley and Zotaro. It facilitates to download the citation in a common exchangeable standard format such as BibTeX
- Uploading and maintaining of 80_Recommendation files/Chapter in the record. It is done to extract and capture the 80_Recommendaton the files containing Recommendation, Suggestions, Observations, future scope of work etc. so that the a new interface/platform to be designed to equip with the provision of hosting and accessing recommendations and conclusions of the Theses suggested by the researcher in a particular discipline.

- ❖ Enhanced Submission Interfaces and Session Time Inclusion to the portal so that, this ensures that metadata and files are reliably uploaded to the server without any interruptions, and it helps maintain a seamless connection with the university coordinator email.
- ❖ Incorporating the Social Media Plugin Tags like Facebook, Whatsapp, Twitter etc. for sharing the Theses records among the users and aspirant researchers.

### 6.7 The Top most viewed Theses from Shodhganga

Data is extracted from Shodhganga DSpace backend dataset for Top most viewed 15 Theses from Shodhganga during 1 Jan 2022-30 April 2023. The table shows that the most viewed Theses is with handle ID (10603/408944) viewed 21,726 times whereas the Thesis record with Handle ID (10603/421252) is viewed 21,642 times.

**Table 5: Top most viewed Theses from Shodhganga**

| SNo | Item/Handle | Number of views |
|---|---|---|
| 1 | Chemical and phytochemical studies of Lepidium sativum with reference to antioxidant and antimicrobial property (10603/408944) | 21,726 |
| 2 | Microbial degradation of lignocellulosic biomass by co_digestion with organic fraction of municipal solid waste into value added products (10603/421252) | 21,642 |
| 3 | Trace based comparison of mobility models for routing in Manet (10603/30248) | 19,032 |
| 4 | A Contribution of Information and Communication Technology ( ICT) in the Development of Grant in aid University Libraries of Gujarat: A Study (10603/63521) | 9,210 |
| 5 | Experimental studies on concrete using fly ash rice husk ash and egg shell powder (10603/141242) | 8,496 |
| 6 | Synopsis of my proposed work on the fictional art of Rabindranath Tagore (10603/268279) | 8,436 |
| 7 | Ideological Underpinnings in Select Malayalam Commercial Films of the Post Liberalization Era (10603/343491) | 7,766 |
| 8 | Dissertation (10603/157251) | 7,028 |
| 9 | 0 distributive almost lattices (10603/365907) | 6,952 |
| 10 | Multiple query processing using soft computing tools (10603/75389) | 6,714 |
| 11 | Cyber crime in India a critical study in modern perspective(10603/189479) | 5,312 |
| 12 | Impact of Social Media on Everyday Life of the Youth (10603/370014) | 4,986 |
| 13 | Impact of Digital Marketing Communication on Consumer Buying Decision Process A Study of Indian Passenger Car Market(10603/206479) | 4,594 |
| 14 | [1,5]-Halo Migratory Aptitudes in 1,3-Pentadiene Derivatives,Conformational Surface of Hexachlorophene Derivatives andReactivity of Cation Bound Capillin Analogues – Computational Study (10603/351580) | 4,518 |
| 15 | Attitude of undergraduate students towards E Learning in COVID Era (10603/358512) | 4,506 |

### 6.8 To identify the State wise MoU Signed and Shodhganga Contribution

A Memorandum of Understanding (MoU) is a prerequisite for identifying organizational ties and formalizing agreements. By specifying roles and duties, it makes collaboration, the creation of partnerships, and project implementation easier. The categories and State-by-State MoU signed along with Theses contribution into Shodhganga are shown in Figure 11 & 12 and Table 6, respectively. As of July 31, 2023, a total of 828 Memorandums of Understanding (MoUs) had been signed with the Shodhganga by institutions and universities belonging to different categories, including State Universities (35.14%), Private Universities (34.06), Deemed Universities (13.29), Institutes of National Importance (10.87), Central Universities (05.68), and Others (Standalone Universities) (0.97).

The undersigned MoUs of all states are further outlined in Table 6 and Figure 12, along with the status of each State's MoUs signed and their contribution to its theses as of July 31, 2023. Its statistics reveal that in terms of MoUs, Uttar Pradesh has signed (72, 8.70%) MoUs and secured first position over all states, followed by Rajasthan (69, 8.33%), Gujarat (66, 7.97%), Maharashtra (60, 7.25%), Tamil Nadu (57, 6.88%), Madhya Pradesh (55, 6.64%), Karnataka (53, 6.40%), Haryana (39, 4.71%), West Bengal (37, 4.47%), Odisha (29, 3.50%) and so on. Furthermore, only one MoU (1, 0.12%) was signed by Chandigarh and Ladakh. According to the State Contribution to Theses report, Ladakh served as zero (0) theses, whereas Tamil Nadu contributed the highest number of theses to Shodhganga 82143 (17.27%) and took top overall. In second place was Uttar Pradesh (73867, 15.23%), followed by Maharashtra (47106, 9.90%), West Bengal (29816, 6.27%), Gujarat (25125, 5.28%), and so forth.

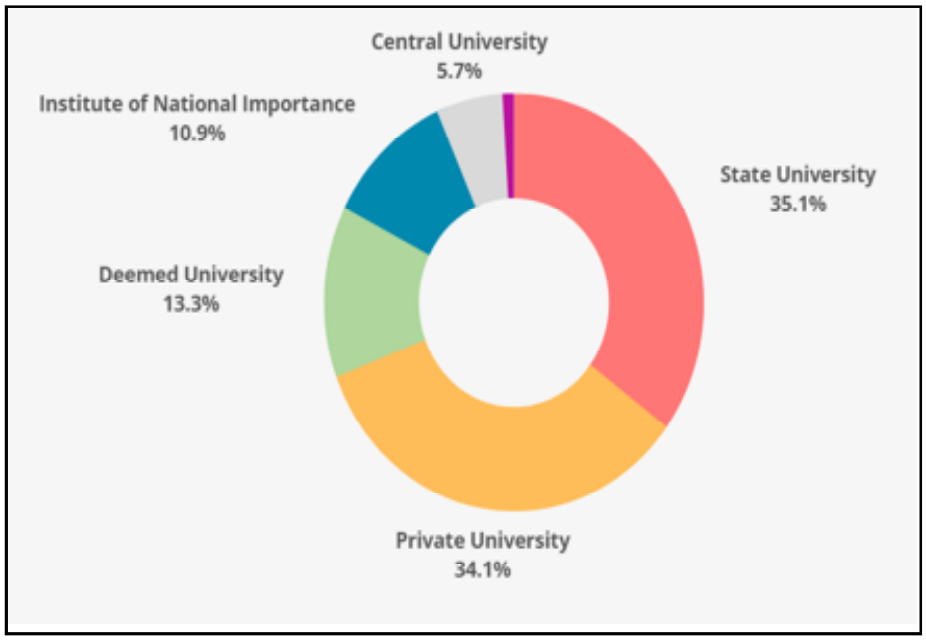



**Figure 11: Category wise MoUs signed**

**Table 6: Category wise MoUs signed and Theses Contribution**

| State | MoUs | Theses Contribution (31st July 2023) | % Contribution |
|---|---|---|---|
| **Tamil Nadu** | 57 | 82143 | 17.27 |
| **Uttar Pradesh** | 72 | 73867 | 15.53 |
| **Maharashtra** | 60 | 47106 | 9.90 |
| **West Bengal** | 37 | 29816 | 6.27 |
| **Gujarat** | 66 | 25125 | 5.28 |
| **Andhra Pradesh** | 24 | 23541 | 4.95 |
| **Karnataka** | 53 | 22717 | 4.77 |
| **Rajasthan** | 69 | 19081 | 4.01 |
| **Delhi** | 28 | 18183 | 3.82 |
| **Bihar** | 26 | 18081 | 3.80 |
| **Kerala** | 16 | 17149 | 3.60 |
| **Assam** | 23 | 12084 | 2.54 |
| **Madhya Pradesh** | 55 | 10106 | 2.12 |
| **Odisha** | 29 | 10442 | 2.19 |
| **Punjab** | 29 | 9886 | 2.08 |
| **Haryana** | 39 | 9208 | 1.94 |
| **Chandigarh** | 1 | 8912 | 1.87 |
| **Telangana** | 20 | 7895 | 1.66 |
| **Uttarakhand** | 25 | 5082 | 1.07 |
| **Himachal pradesh** | 23 | 4810 | 1.01 |
| **Chhattisgarh** | 20 | 4335 | 0.91 |
| **Jammu and Kashmir** | 9 | 3847 | 0.81 |
| **Meghalaya** | 5 | 2842 | 0.60 |
| **Puducherry** | 3 | 2385 | 0.50 |
| **Manipur** | 4 | 1564 | 0.33 |
| **Jharkhand** | 16 | 1526 | 0.32 |
| **Goa** | 2 | 1188 | 0.25 |
| **Mizoram** | 2 | 793 | 0.17 |
| **Arunachal Pradesh** | 5 | 702 | 0.15 |
| **Tripura** | 3 | 591 | 0.12 |
| **Nagaland** | 3 | 482 | 0.10 |
| **Sikkim** | 3 | 283 | 0.06 |
| **Ladakh** | 1 | 0 | 0.00 |
| **Total** | **828** | **475772** | |

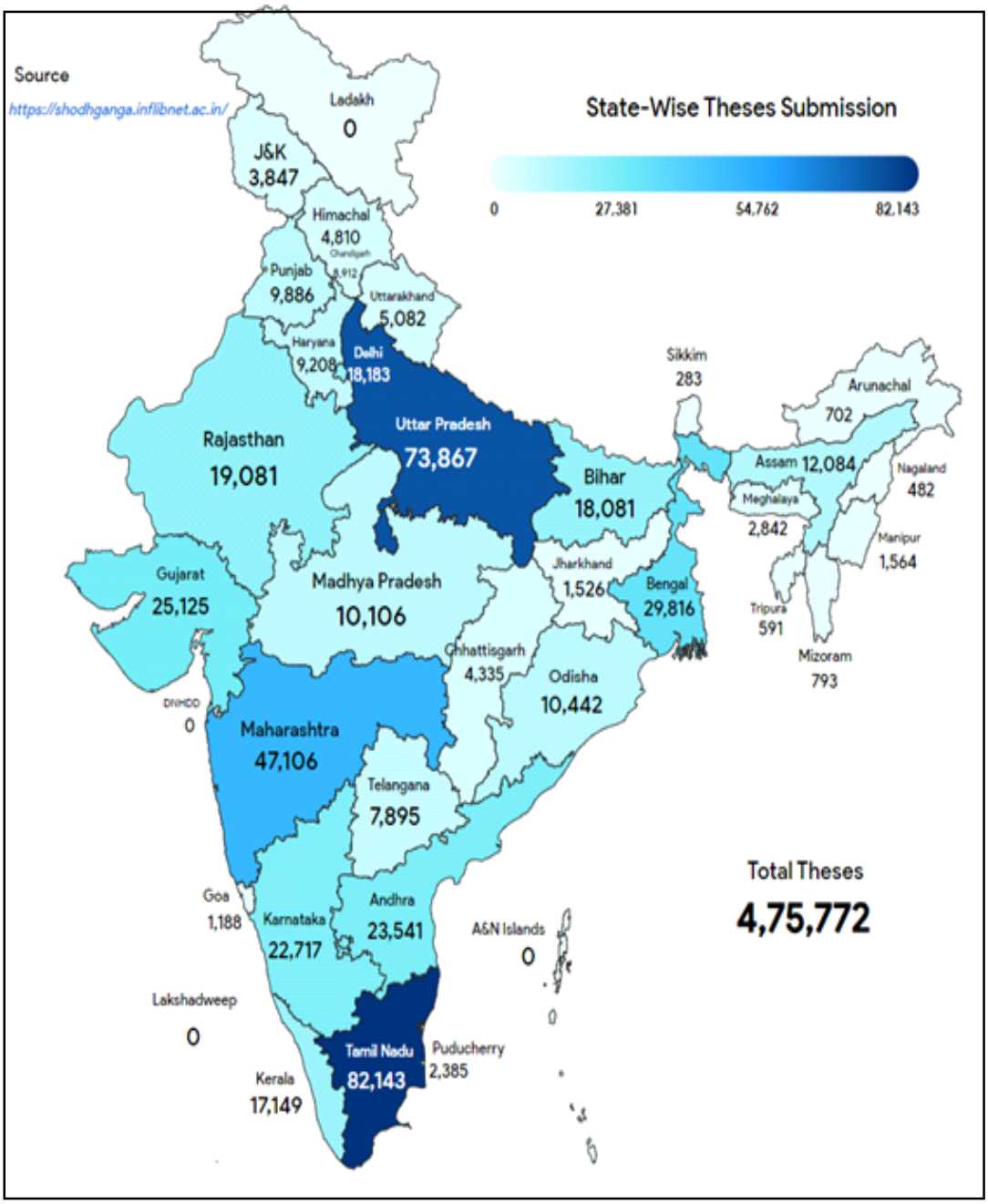



**Figure 12: State wise Theses Contribution**

### 6.9 To discuss the Challenges faced in the Maintenance, Workflow and Policies adopted to host open-access content into Shodhganga.

Open access content although provides an extensive and enormous coverage to the content of the research output in terms of online PhD theses and Di11ssertations maintained and uploaded to the open access repositories, however, it has various challenges and restrictions in terms of the legal issues and obligations on the copyright of the content bearing uploaded into the open access format. Since Shodhganga is an open access repository maintaining the soft copies of the PhD Theses, there are also various challenges and issues faced in hosting the soft copies of the Theses contributed from 700+ Universities/Institutions/ CFTIs/INIs. It has been reported by some of the research scholars and guides that the Theses uploaded into Shodhganga have been unethically and illegally published by the publishers in the form of eBooks and sold commercially on their e-commerce websites. In another case, the Theses content available in open access mode can be easily downloaded and copy paste and hence plagiarised by other researchers. Some cases have been noted and informed by the scholars that their PhD Thesis content is unethically reproduced by other scholars without citations and acknowledgements and also research papers have been published out of the content available through ETDs. This may be the serious concern and debatable issue and should be addressed in the open Symposiums and Conferences where the experts and managers of ETDs are joint together to formulate some guidelines and standards for protecting and safeguarding the hard work of the researchers and academicians in the form of PhD research work and Theses. As far as Metadata creations and workflow is concerned, Shodhganga is also facing challenges and issues since various Universities are not strictly following the standards and formats specified for creation of Metadata and uploading of Chapters in PDF file formats. Hence, it is recommended in this study that all Universities should follow the standardised Metadata formats for easy uploading, harvesting and exchange of data among the academic communities.

## 7. Results and Findings

Results of the study indicate that the Shodhganga usage has seen a significant increase since its inception, due to the leap in submission of Theses by the Contributing Universities after the Ministry Notification for submitting the digital copy of the Theses to Shodhganga for National Ranking purposes and also there is an increase and change in usage pattern after the pandemic. As such, most of the research scholars, academicians are now consulting open access resources and learning materials for their research work which leads to increased research visibility, enhanced accessibility, research collaborations etc. The study suggests that academicians and Institutions should adopt policies to support open access publishing of scholarly resources and should come forward voluntarily to join Shodhganga to make it the largest global repository. The ETD repository initiated with only 1,171 Theses records. In 2016, (after 5 years) the number of Theses grew to 1,16,000, with more than 100,000 Theses in 5 years. However, due to COVID-19 and lockdown of many Universities and research activity being slowed down due to the closure of Institutions, the status of Theses additions in Shodhganga shows a sum of 42,000+ Theses added in the year 2021. In the

last 2 years, the ETD project had seen tremendous growth by addition of 1,57,120+ Theses during (2021-2023) while 77,358 Theses were uploaded only in the year 2022, while in FY 2022-23 total 98,129 theses were added.

## 8. Conclusion

The findings of the study highlight the substantial impact of Shodhganga as a national repository on the academic and research community, shedding light on its usage patterns and trends. The study also pointed out various challenges faced by the Shodhganga in terms of hosting and adopting the open access model, copyright issues, plagiarism issues and also legal challenges raised by the research scholars and academicians on their content. The study seeks the appropriate solutions by unfolding the procedures followed by the global ETDs to overcome these legal challenges. It also highlights the issues faced in submission of authenticated data/content by the contributing Institutions hosted on Shodhganga to bring transparency into the research environment by avoiding the duplication of work. Another important issue of copying the content and publishing the thesis leading to severe plagiarism and there was no mechanism to trace and find it out. In order to solve this issue ShodhShuddhi was introduced in 2019 which provides plagiarism tools to 1100+ HEIs. The challenges and issues faced by the project Shodhganga in terms of the copyright and intellectual property rights are also discussed in this paper.

## Reference

Arlitsch, K., Wheeler, J., Pham, M. T. N., & Parulian, N. N. (2021). An analysis of use and performance data aggregated from 35 institutional repositories. Online Information Review, 45(2), 316-335. https://doi.org/10.1108/OIR-08-2020-0328.

Bhat, M. I., Mudhol, M. V., & Mahesh, V. (2014). Importance of Electronic Theses and Dissertations (ETD's) on Internet: A survey of Indian ETD repository Shodhganga. International Journal of Library and Information Studies, 4(2), 53-61.

Brush, D. A., & Jiras, J. (2019). Developing an institutional repository using Digital Commons.

Digital Library Perspectives, 35(1), 31-40.. https://doi.org/10.1108/DLP-08-2017-0028.

Chapepa, G. G., Ngwira, F., & Mapulanga, P. (2023). Metadata creation practices at the Lilongwe University of Agriculture and Natural Resources library's institutional repository. Digital Library Perspectives, 39(2), 205-219. https://doi.org/10.1108/DLP-09-2022-0074.

Dublin Core metadata element set, version 1.1: DCMI. https://dublincore.org/specifications/dublin-core/dces/

Fox, E. A., Eaton, J. L., McMillan, G., Kipp, N. A., Weiss, L., Arce, E., & Guyer, S. (1996).

National digital library of theses and dissertations. D-Lib Magazine, September.

Gul, S., Bashir, S., & Ganaie, S. A. (2020). Evaluation of institutional repositories of South Asia.

Online Information Review, 44(1), 192-212. https://doi.org/10.1108/OIR-03-2019-0087.

INFLIBNET Centre Publications: Annual Report, Retrieved from https://www.inflibnet.ac.in/publication/annualreport/AR_2009-10.pdf

https://www.inflibnet.ac.in/publication/annualreport/AR_2020-21.pdf & https://www.inflibnet.ac.in/publication/annualreport/AR_2021-22.pdf

INFLIBNET Centre Publications: Newsletter, Retrieved from https://www.inflibnet.ac.in/publication/newsletter/jan_mar_2023/jan_mar_2023.pdf

https://www.inflibnet.ac.in/publication/newsletter/apr_jun_2023/apr_jun_2023.pdf

Jana, A., & Rout, R. (2022). Designing metadata schema for a human library: a prototype. Digital

Library Perspectives, 38(3), 346-361. https://doi.org/10.1108/DLP-07-2021-0052.

Paskali, L., Ivanovic, L., & Ivanoviæ, D. (2021). Use analysis of the digital library of PhD

dissertations defended at the University of Novi Sad. The Electronic Library, 39(1), 97-117. https://doi.org/10.1108/EL-11-2019-0268.

Rasuli, B., Schöpfel, J., Boock, M., & Van Wyk, B. (2023). Access and impact barriers to academic publications: a global study of thesis and dissertation embargo policies. Online Information Review. https://doi.org/10.1108/OIR-09-2022-0497.

Shajitha, C., & KC, A. M. (2021). Institutional repositories in South India: An exploratory study. Global Knowledge, Memory and Communication, 70(8/9), 817-841.. https://doi.org/10.1108/GKMC-07-2020-0099.

Shajitha, C. (2020). Digital curation practices in institutional repositories in South India: a study. Global Knowledge, Memory and Communication, 69(8/9), 557-578. https://doi.org/10.1108/GKMC-10-2019-0125

Shodhganga: a reservoir of Indian Theses, Retrieved from https://shodhganga.inflibnet.ac.in/.

Sivakumaren, K. S. (2015). Electronic thesis and dissertations (ETDs) by Indian universities in Shodhganga project: a study. Journal of advances in library and information science, 4(1), 62-66. http://jalis.in/pdf/4-1/Sivamit.pdf.

Ukwoma, S. C., Osadebe, N. E., & Dim, C. L. (2019). Management of institutional repositories (IR) in higher education perspective. Library Management, 40(8/9), 543-557. https://doi.org/10.1108/LM-12-2018-0094.

Wani, Z. A., & Wani, S. A. (2019). Mapping the use of open access resources by doctoral students in the USA by employing citation analysis. Collection and Curation, 38(4), 112-118. https://doi.org/10.1108/CC-05-2018-0014.

Yiotis, K. (2008). Electronic theses and dissertation (ETD) repositories: What are they?

Where do they come from? How do they work?. OCLC Systems & Services: International digital library perspectives, 24(2), 101-115.

Zhang, Y., Lee, K., & You, B. J. (2001). Usage patterns of an electronic theses and dissertations system. Online Information Review, 25(6), 370-378. https://doi.org/10.1108/EUM0000000006536