# ETDs Metadata Practices of Indian Institutes of Technology: A Comparative Study

**Sakshi Devi[1] and Dimple Patel[1]**

[1]Research Scholar Central University of Himachal Pradesh, Himachal Pradesh

**Abstract**

*The main objectives of this study are to investigate the metadata practises employed by the Institutional Repositories (IR) of the Indian Institutes of Technology (IITs) in relation to Electronic Theses and Dissertations (ETDs). The first step involved an analysis of the websites of the Indian Institutes of Technology (IITs) to ascertain which of these institutions had established repositories within their organisation. Additionally, the study determines the repositories that were both publicly open and accessible. The analysis focused on repositories that are publicly available and accessible, and the resulting secondary data was systematically structured in tabular format. Subsequently, in order to ascertain the metadata practises adhered to by these repositories, primary data was gathered using a questionnaire. The findings indicate that a significant proportion (over 80%) of Indian Institutes of Technology (IITs) have established their own institutional repositories. However, it is observed that only 48% of these repositories are available to the public, and they predominantly employ DSpace software for the purpose of managing digital information. The study additionally indicated that the Dublin Core™ Metadata Element Set (DCMES) is used for the purpose of describing Electronic Theses and Dissertations (ETDs). The study used criteria of completeness, correctness, and consistency to study the quality of metadata in electronic theses and dissertations (ETDs) records and identifies concerns. The study reveals that out of the 23 Indian Institutes of Technology (IITs), only four of them have Electronic Theses and Dissertations (ETDs) metadata that adhere to the FAIR (Findable, Accessible, Interoperable, and Reusable) principles. In general, the study emphasises the present condition of IITs IR in effectively managing their digital content ETDs in particular and proposes that the institute should prioritise the maintenance of metadata quality to guarantee the discoverability, accessibility, and interoperability of the digital information. The questionnaire was in general about IITs repositories, and the secondary data have been collected for metadata. Therefore, the major study has been done through secondary data observation.*

**Keywords:** Dublin Core, Electronic Theses and Dissertations (ETD), Indian Institutes of Technology (IITs), Institutional Repositories, Metadata practices, Metadata Quality

---

Corresponding Author: Sakshi Devi, Email: sakshidevi322@gmail.com and Dimple Patel, Email: dimple@hpcu.ac.in

## 1. Introduction

Theses and dissertations are the research outcomes of hard work put in by scholars, which are published and archived by university repositories or by centralized repository such as Shodhganga (Digital reservoir of Indian ETDs) for use in the foreseeable future. And for the purpose of supplying, supporting, disseminating, and preserving scholarly content such as articles, ETDs, conferences, etc., institutional repository has become a significant feature, where metadata plays a crucial role for making the digital asset known, or in other words, available and accessible to the scholarly community (Park & Richard, 2010).

Metadata is developed, gathered, and maintained in order to facilitate and enhance the usage of data, which enables that data (refer the digital resources) to be FAIR: findability, accessibility, interoperability, and re-use of digital content (Go FAIR, n.d.). The FAIR principles are designed to serve as a roadmap for making digital resources more Findable, Accessible, Interoperable, and Reusable for both computers and people (Jacobsen et al., 2019). Furthermore, metadata is required for the accessibility and preservation of digital content over the long term. Even the choices and practices that librarians and other information professionals make with regard to metadata have an effect on how easily resources may be located online (Moulaison et al., 2015). Hence, accessing digital information will be more difficult if it is not organised and preserved efficiently, which brings the concept of metadata quality. Metadata quality depends on how well it performs bibliographic functions like discovery, usage, origin, currency, integrity, and management (Park, 2009). Moreover, institutions can improve the discoverability and accessibility of their digital resources and offer enhanced support for research by providing high-quality metadata (Osman et al., 2023). However, users are having difficulty accessing digital content despite the fact that there are so many standards available to describe digital content. This is because of a variety of reasons, but the most important reason that can be considered for this problem is the poor metadata practises that have been followed (Beyene & Godwin, 2018).

Therefore, the purpose of this study is to determine the current metadata practices and study the quality of metadata through ETDs.

## 2. Review of Literature

Moulaison et al. (2015), said for the long-term access and preservation of digital content, the development of metadata for research and repository content is crucial to the scholarly communication process. And the findability of digital resources is also impacted by metadata decisions and practices. Lubasa (2009) in her study, said ETDs reside inside a hybrid environment. The people in charge of these collections must maintain a balance between consistency and adaptability, while also enhancing their understanding of user behaviour. Park (2009) studied the fundamental measurement criteria that have been most extensively used in the literature and discovered that the most common criteria used in assessing metadata quality in the literature are accuracy, completeness, and consistency. Lihitkar and Lihitkar (2014) did a study using Open DOAR to identify 41 Indian institutional repositories that had generated ETDs. The author compared Indian ETDs by software, format, search, and access. Their findings showed that most Indian IR use DSpace and

archive ETDs as PDF. And it was suggested that to boost research awareness, each institution should create an IR and host ETDs. Steele and Sumpcrethar (2016) conducted an online survey of academic IR in the United States to determine which metadata they use for description and how quality control is implemented. Their research revealed that there is no general accepted method of quality control among IRs and that the metadata IRs use today is constrained by the time and resources of their parent library. These constraints also influence the level of metadata quality control. Osman (2023) conducted a study to examine the metadata quality of institutional repositories for Electronic Theses and Dissertations (ETDs) at three research universities. The study evaluates the completeness, accuracy, and consistency of 1138 metadata records from three institutional repositories using metrics from Bruce and Hillmann. The research revealed several information completeness, correctness, and consistency issues and emphasizes the need for standardized ways to maintain metadata quality. The study also highlights the need for data model that can be used for IR. In another study Choudhary et. al (2023) explored the techniques for the automated identification, rectification, and standardization of scholarly information. They focused on the application of these methods of electronic theses and dissertations (ETDs). In this study, a framework that leverages innovative artificial intelligence techniques to enhance the metadata quality. In order to assess the effectiveness of MetaEnhance, a benchmark for evaluating the quality of metadata was constructed. This benchmark consisted of a collection of five hundred Electronic Theses and Dissertations (ETDs), which were obtained by merging subsets that were sampled based on various criteria. The results indicated that the proposed approaches had exceptional performance in detecting errors.

### 3. Objectives

❖ Identify the current metadata practices followed by Indian Institutes of Technology (IITs) to describe the collection of electronic theses and dissertations.

❖ Study the metadata quality of ETDs in IITs.

### 4. Methodology

The study was conducted in two stages: first, websites of all IITs were examined, and then an online questionnaire was administered. Initially the 23 IITs websites were analysed to determine the IITs that have developed repositories in their institution and how many of them are available and accessible and it is found that 18 IITs have developed IRs and out of which 11 IITs repositories are both available and accessible on internet. The secondary data collected from the website investigation was then analysed and organised in a tabular form for better understanding (See Tables 1-4). The primary data was collected through structured online questionnaire and was sent to librarians in charge of IITs IR through an email. Nine responses were received and all of them were considered for the analysis and are presented in pie-chart.

The questionnaire administered was in general about IITs repositories to ascertain availability and accessibility and the secondary data have been collected for metadata. Therefore, the major study has been done through secondary data observation.

## 5. Data Analysis

The primary data collected from the various IITs IR and then analysed in tabular form using a spreadsheet to make it easier to understand.

### 5.1 Availability and Accessibility of IITs IR

In total, there are 23 IITs situated in different regions of India. Eleven (11) IITs have their own institutional repositories (IR), which are available as well as accessible in public domain; 7 IITs have given links to their repository however when researchers tried to access, they were not accessible and in rest of the 5 IITs, the researcher could not find any link/URL of IR in their institutional/library website/webpage. This data is presented in Table 1 and 2.

**Table 1: Availability and Accessibility of IITs IR**

| IR Available & Accessible in public domain | IR Available but not Accessible in public domain | IR Not Available in public domain |
|---|---|---|
| IIT Kharagpur, West Bengal | IIT Madras, Tamil Nadu | IIT Tirupati, Andhra Pradesh |
| IIT Delhi | IIT Kanpur, Uttar Pradesh | IIT Bhilai, Chhattisgarh |
| IIT Roorkee, Uttarakhand | IIT Bhubaneshwar, Odisha | IIT Dharwad, Karnataka |
| IIT Ropar, Punjab | IIT Gandhinagar, Gujarat | IIT Jammu, Jammu and Kashmir |
| IIT Hyderabad, Telangana | IIT Patna, Bihar | IIT Goa, Goa |
| IIT Jodhpur, Rajasthan | IIT Dhanbad, Jharkhand | |
| IIT (BHU), Varanasi | IIT Palakkad, Kerala | |
| IIT Indore, Madhya Pradesh | | |
| IIT Bombay, Maharashtra | | |
| IIT Guwahati, Assam | | |
| IIT Mandi, Himachal Pradesh | | |

**Table 2: Status of the IITs IR**

| | Number Inst. | .% |
|---|---|---|
| Available | 18 | 78.26 |
| Accessible | 11 | 47.82 |
| Not Accessible | 7 | 30.43 |

The list of IITs IR with their URL is displayed in the Table 3, the table also comprises the list of IR software used by the IITs in managing their digital collection along with the status of the repositories that are available as well as accessible.

**Table 3: Digital IR Software used and URL of the IITs Repository**

| S.N. | Institution Name | IR Software | Repository URL available in public domain | Status of IR | ETDS collection |
|---|---|---|---|---|---|
| | | | (Accessed on 16-05-23) | | available |
| 1. | IIT Kharagpur (IITKGP) | DSpace | http://www.idr.iitkgp.ac.in/xmlui/ | Accessible | Yes |
| 2. | IIT Delhi(IITD) | DSpace | http://eprint.iitd.ac.in/ | Accessible | Yes |
| 3. | IIT Roorkee(IITR) | DSpace | http://shodhbhagirathi.iitr.ac.in: 8081/jspui/ | Accessible | Yes |
| 4. | IIT Ropar(IITRPR) | DSpace | http://dspace.iitrpr.ac.in:8080/xmlui/ | Accessible | Yes |
| 5. | IIT Bhubaneshwar(IITBBS) | | http://idr.iitbbs.ac.in/jspui/ | Not Accessible | |
| 6. | IIT Gandhinagar(IITGN) | | http://repository.iitgn.ac.in/ | Not Accessible | |
| 7. | IIT Hyderabad(IITH) | Eprints | https://raiith.iith.ac.in/ | Accessible | Yes |
| 8. | IIT Jodhpur(IITJ) | DSpace | http://theses.iitj.ac.in:8080/jspui/ | Accessible | Yes |
| 9. | IIT Patna(IITP) | | http://idr.iitp.ac.in/jspui/ | Not Accessible | |
| 10. | IIT Indore(IITI) | DSpace | http://dspace.iiti.ac.in:8080/jspui/ | Accessible | Yes |
| 11. | IIT Bombay(IITB) | DSpace | http://dspace.library.iitb.ac.in/jspui/ | Accessible | Yes |
| 12. | IIT Mandi(IITMD) | DSpace | http://odr.iitmandi.ac.in:8080/xmlui/ | Accessible | Yes |
| 13. | IIT(BHU) Varanasi | DSpace | https://idr-lib.iitbhu.ac.in/xmlui/ | Accessible | Yes |
| 14. | IIT Madras(IITM) | | http://irepose.iitm.ac.in:8080/jspui/ | Not Accessible | |
| 15. | IIT Kanpur(IITK) | | https://etd.iitk.ac.in:8443/jspui/ -Electronic theses | Not Accessible | |
| 16. | IIT Guwahati(IITG) | DSpace | http://gyan.iitg.ernet.in/ | Accessible | Yes |
| 17. | IIT Palakkad(IIT PKD) | | http://10.64.15.4:8080/jspui/ | Accessible on intranet only | |
| 18. | IIT Dhanbad (IITISM) | | http://172.16.200.91:8080/xmlui/ | Not Accessible | |
| 19. | IIT Dharwad | | Not available | | |
| 20. | IIT Goa (IITGOA) | | Not available | | |
| 21. | IIT Jammu (IITJMU) | | Not available | | have Thesis & Dissertation option under collection option in library website |
| 22. | IIT Bhilai (IITBH) | | Not available | | |
| 23. | IIT Tirupati | | Not available | | |

Among the 11 IITs repositories that are available and accessible in Internet, it has been observed that most of the IITs repositories use DSpace software (See Table 3) for organising the digital content stored in their repository except IIT Hyderabad, which uses EPrints digital library software for the management of their digital content. IITKGP, IITD, IITH, IITMD, IITRPR, IITJ, IITI, IIT-BHU, IITB, IITG, IITR, etc., have ETDs collection in their IR. And, it has been observed that IIT-BHU, IITG, IITR provide complete access to the full-text of ETDs, while some IITs such as IITRPR, IITJ, and IITI provide restricted access (RA) whereas IITKGP, IITD, IITGN and IITMD makes few pages of the ETDs available on their repositories, such as the front page, certificate page, acknowledgement page, content page, list of figures or tables. And no files are associated has been observed in some of the IITJ ETDs collection as well. Restricted access (RA) requires a login and password to open the digital content, which only allows institution members from the respective institution that has imposed the restriction. However, repositories such as IITKPR, IITRPR, IITI. IITJ, IITM, provide users an option where they can mail them asking for full-text access to any ETDs.

It has been observed Indian Institute of Technology Jammu (IITJMU), does not yet have a repository available on internet, although they have added an option of collections (includes e-databases, e-books, faculty publications, theses, and dissertations, etc.,) on their library website. These collections are now empty, but it is anticipated that they will soon be available. The researcher had emailed the library staff of IITs whose IR is available on the internet but inaccessible to inform them that their IR was inaccessible/not working. Only three responses were received, IIT Madras responded that a problem occurred while upgrading and migrating to the VMWare server. IIT Palakkad responded that their IR does not provide open access, while IIT Dhanbad stated that it is only accessible via their institution's IP range. The library staff of those IITs whose IR is not available in the public domain were also contacted via email. However, the researcher had received no responses.

### 5.1.2 ETDs

From the collected data, it has been observed that Dublin Core™ Metadata Element Set (DCMES) is the metadata schema that has been used by the IITs' repositories to describe their ETDs (See Table 4). The table represents the DC elements used to describe the ETDs collection in various IITs IR. Majority of the IITs IR are using Qualified Dublin Core™ for the description of ETDs collection whereas IITH IR uses Simple Dublin Core™. IITJ IR is the only repository to use extensive metadata elements and many unique metadata elements among other IITs IR; the numerous metadata elements used by IITJ are: dc.creator.researcher, dc.date.awarded, dc.date.registered, dc.type.degree, dc.identifier.accession, dc.description.note, dc.rights, etc,.

Common metadata elements observed among IITs IR while describing ETDs are: dc.contributor.author, dc.date.accessioned, dc.date.available, dc.date.issued, dc.identifier.uri, dc.language, dc.title, dc.type, dc.subject, dc.publisher. Whereas, some unique metadata elements are used by various IITs IR are: dc.terms.publisher used for the description of publisher by IITB; dc.guide and dc.accession number are

used for the description of guide name and accession number respectively by IITR; dc.identifier.other used by IITG that describes the roll no. of the student (See Table 4).

**Table 4: DC Elements used to Describe ETDs in various IITs IR**

| | IITKGP | IITB | IITD | IITG | IITJ | IITMD | IITI | IIT(BHU) | IITH | IITR | IITRPR |
|---|---|---|---|---|---|---|---|---|---|---|---|
| dc elements | | | | | (RA) | | (RA) | | (RA) | | (RA) |
| dc.creator | | | | | | | | | ✓ | | |
| dc.creator.researcher | | | | | ✓ | | | | | | |
| dc.contributor.advisor | | ✓ | ✓ | | ✓ | ✓ | ✓ | | | | |
| dc.contributor.author | ✓ | ✓ | ✓ | ✓ | | | ✓ | ✓ | | ✓ | ✓ |
| dc.guide | | | | | | | | | | ✓ | |
| dc.date | | | | | | | | | ✓ | | |
| dc.date.accessioned | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ |
| dc.date.available | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ |
| dc.date.issued | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ |
| dc.date.registered | | | | | ✓ | | | | | | |
| dc.date.awarded | | | | | ✓ | | | | | | |
| dc.language.iso | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ |
| dc.subject.ddc | | | | | ✓ | | | | | | |
| dc.subject | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | | ✓ | ✓ |
| dc.type | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| dc.type.degree | | | | | ✓ | | | | | | |
| dc.title | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| dc.identifier | | | | | | | | | ✓ | | |
| dc.identifier.uri. | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ |
| dc.identifier.other | | | | ✓ | | | ✓ | | | | |
| dc.identifier.citation | | | | | ✓ | | | | | | |
| dc. Identifier.gov-doc | ✓ | | | | | | | | | | |
| dc.identifier.issn | | | | | | | | | | | |
| dc.identifier.accession | | | | | ✓ | | | | | | |
| dc.description | | | | ✓ | | | | | ✓ | | |
| dc.description.abstract | | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ |

| | IITKGP | IITB | IITD | IITG | IITJ | IITMD | IITI | IIT(BHU) | IITH | IITR | IITRPR |
|---|---|---|---|---|---|---|---|---|---|---|---|
| dc elements | | | | | (RA) | | (RA) | | (RA) | | (RA) |
| dc.descriptionstatementofresponsibility | | | | | ✓ | | | | | | |
| dc.description.note | | | | | ✓ | | | | | | |
| dc.description.sponsor | | | | | | | | | | ✓ | |
| dc.language | | | | | | | | | ✓ | | |
| dc.title.alternative | | | | | | | | | | | |
| dc.publisher | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | | | ✓ |
| dc.publisher.place | | | | | ✓ | | | | | | |
| dc.publisher.department | | | | | ✓ | | | | | | |
| dcterms.publisher | | ✓ | | | | | | | | | |
| dc.relation | | | | | | | | | ✓ | | |
| dc.relation.ispartofseries | | | ✓ | ✓ | | | ✓ | | | | |
| dc.relation.uri | | | | | | | | | | | |
| dc.rights | | | | | ✓ | | | | | | |
| dc.format | | | | | | | | | ✓ | | |
| dc.format.extent | | | | | ✓ | | | | | | |
| dc.format.mimetype | | | | | | | | | | | |
| dc.format.accompanyingmaterial | | | | | ✓ | | | | | | |
| dc.accession.number | | | | | | | | | | ✓ | |

## 5.1.3 ETDs Metadata Quality

Metadata quality dimensions frequently reported in the literature include accuracy, completeness, consistency (Park, 2009; Osman, 2023; Romero-Pelaez. et. al, 2018; Zeng & Qin, 2016). To study the metadata quality, small sample of metadata records (50 ETDs metadata records of each IITs IR, accessible in public domain) were evaluated based on Completeness, Correctness, Consistency.

## 5.1.4 Completeness

It is concerned with the administration of metadata elements for the completed records as well as the level of descriptive details that are given to any digital item. The level of detail should reflect the primary purpose of metadata, resource discovery and use (Park, 2009; Zeng and Qin, 2016). The completeness error has been observed in the dc.date.issued metadata field among the IITs IR, there may be a reason that the exact date is unknown.

**Table 5: dc.date.issued among IITs IR.**

| IITs | dc.date.issued |
|---|---|
| IIT KGP | YYYY-MM |
| IITI | YYYY-MM-DD |
| IIT-BHU | YYYY |
| IITMD | YYYY-MM-DD |
| IITR | YYYY-MM |
| IITJ | YYYY-MM |
| IITD | YYYY |
| IITRPR | YYYY-MM-DD |
| IITB | YYYY |

## 5.15  Correctness

It is concerned with the accurate description and representation of metadata fields as it focuses on correct content, correct format, correct input, correct mapping, etc, (Park, 2009; Zeng and Qin, 2016). IITG is providing supervisor name in dc.description field rather than using dc.contributor/dc.creator.

## 5.1.6  Consistency

Consistency can be measured by examining the conceptual value of data and the structural format of data (Park, 2009). The IITR's dc.type metadata field,  has been observed to be inconsistent, which defines dc.type as Thesis/Theses/Other for various ETDs collection.

From the collected data, not much error has been observed in the metadata quality of ETDs collection as a small sample size has been chosen for the assessment. However, based on our observation for diverse assessment of metadata quality of IITs IR, whole repository content types such as articles, conference proceedings, theses, reports, images, etc., could be considered.

## 5.2  Primary Data Analysis

## 5.2.1 Importance of Institutional Repository

All the respondents agreed that the long-term preservation and management of digital content has become easier with the advent of IRs. The responses rate was (one hundred per cent) yes to the question (See Figure 1).
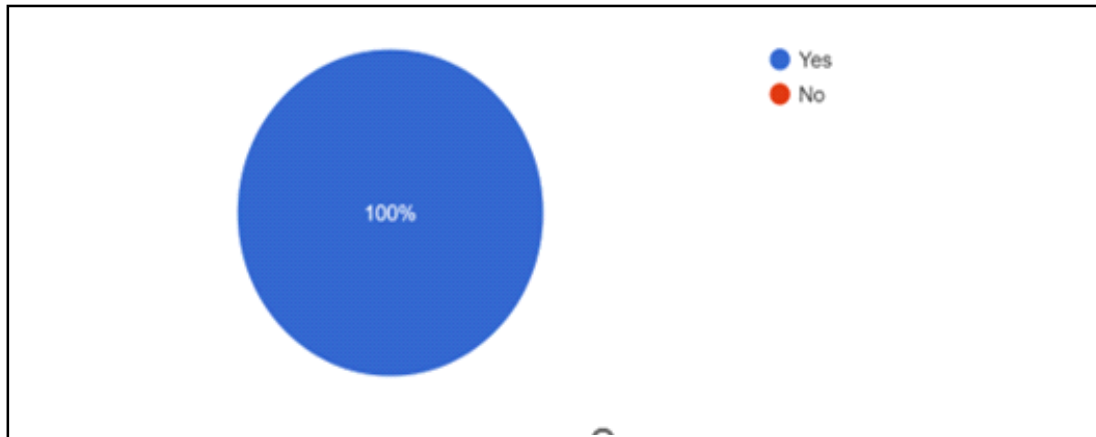
**Figure 1: Long-term Preservation Becomes Easier with IR**

### 5.2.2 Content Submission in IR

The respondents were asked about the method of content submission in IRs. From the responses received, it has been seen that majority of the contents are submitted by the repository staff with (88.9 per cent) followed by (11.1 per cent) done by author self-archiving (See Figure 2).
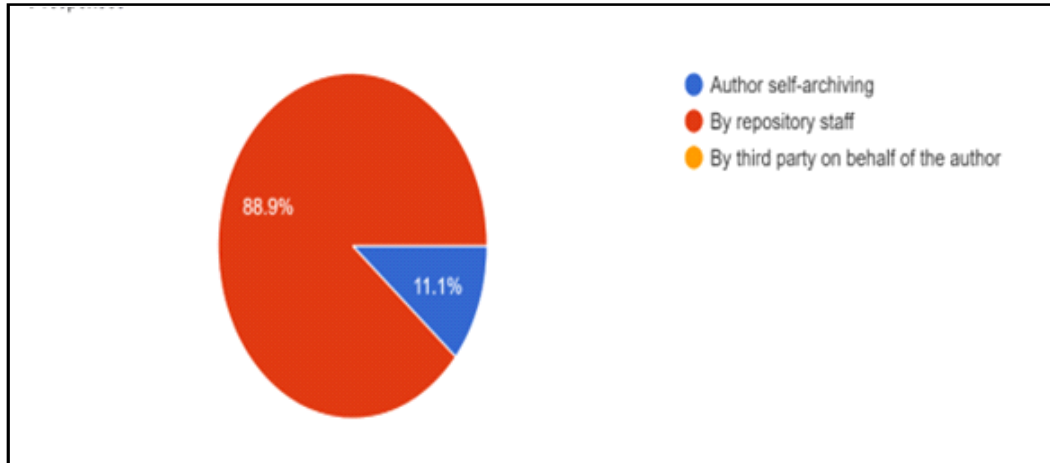


**Figure 2: Content Submission in IR**

### 5.2.3  Areas that require more attention for the Preservation of Digital Content

88.9% of respondents said digital preservation training, policy implementation, and awareness need the most attention. However, 77.8% of respondents agreed on the wide acceptance of metadata and need of standardization of metadata practices (See Figure 3).
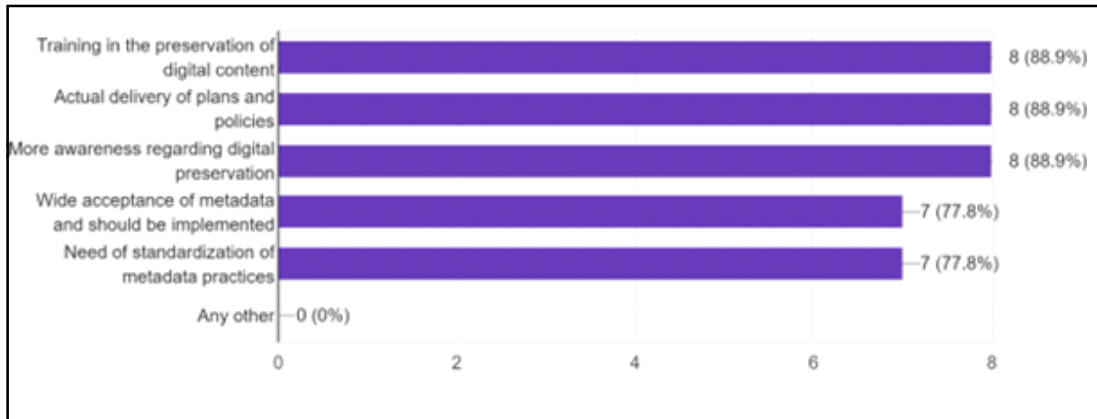
**Figure 3: Areas need more Attention for Preservation of Digital Content**

## 6. Importance of Metadata in Preservation of Digital Content

Respondents were asked if metadata is important for preservation of digital content. Out of all the responses received, majority (88.9 per cent) of the respondents accepted that metadata is crucial for preservation of digital content and (11.1 per cent) of them did not agree to the statement (See Figure 4).
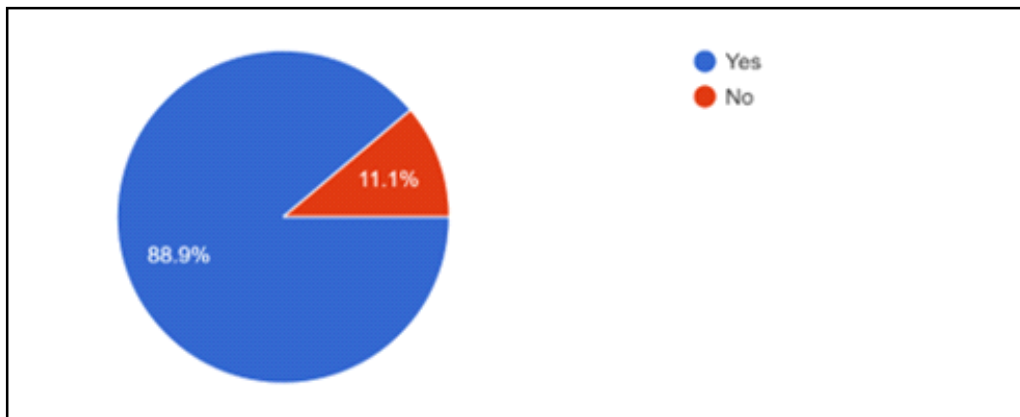


**Figure 4: Metadata is Important for Digital Preservation**

### 6.1 Metadata Schema used in IRs by IITs

Respondents were asked what metadata format their institutional repository uses to organise digital content. Dublin Core was the most popular metadata schema, used by 88.9% of respondents, followed by Qualified Dublin Core (QDC) at 22.9 percent and Metadata Encoding and Transmission Standard (METS) at one institution (See Figure 5).
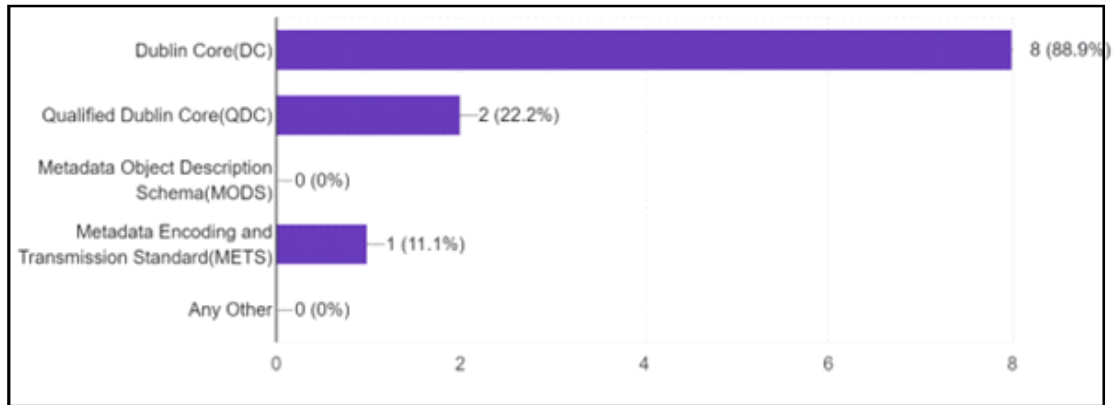
**Figure 5: Various Metadata Schema used in IITs IR**

## 6.2 Metadata Entries

Figure 6 shows that 77.8% of metadata entries in IIT repositories are manually entered and 22.2% are imported from external databases. In addition, respondents that import metadata items were asked which database(s) they use, and only one cited Scopus, Web of Science (WOS).
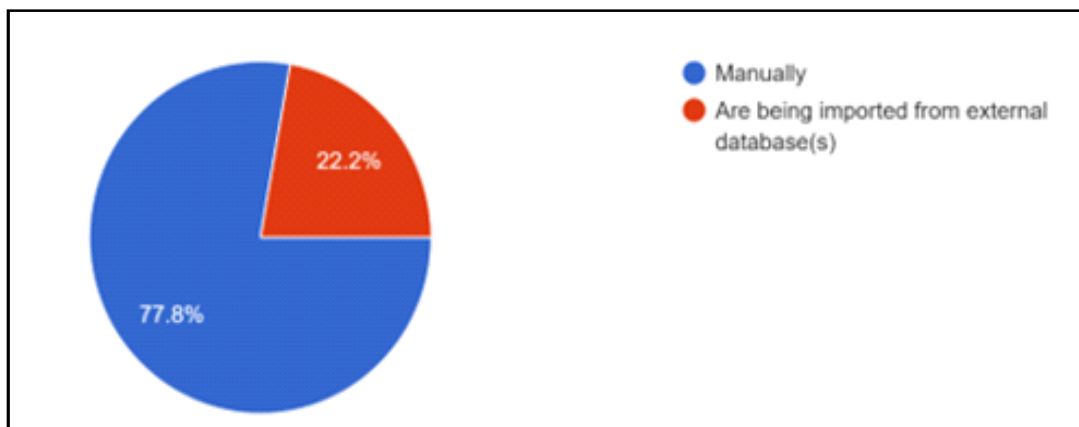


**Figure 6: Metadata Entries**

## 6.3 Metadata Standardization Policy

33.3 per cent of the IITs repositories follow metadata standardization policy for their content while 66.7 per cent did not have any standardization policy for metadata (See Figure 7). Following this, respondents were asked to mention the specific policy/policies followed by their institution for metadata standardization. To which, one of the respondents said they accept their own library practice, another said they follow DCMES, another two responses were not clear as respondents had given repository URL.
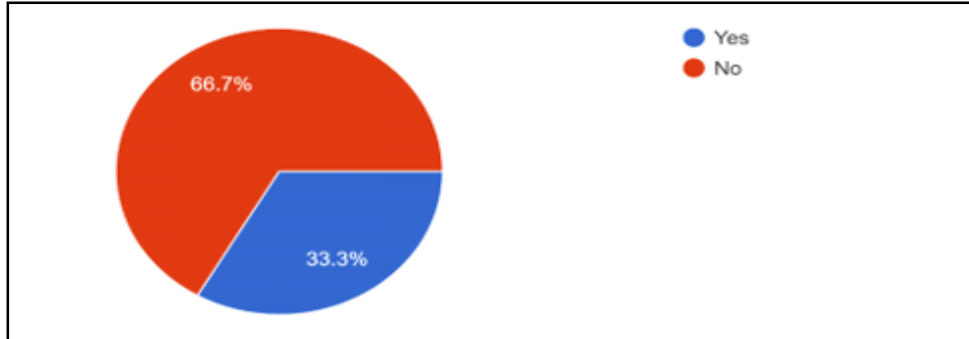
**Figure 7: Standardization Policy for Metadata**

### 6.4 OAI-PMH Compliance

The results revealed that 90 per cent of the respondents' responses were yes to the question while 10 per cent said no (See Figure 8).
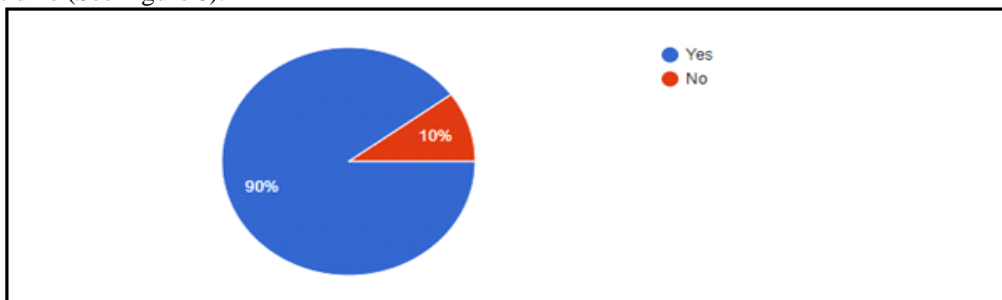


**Figure 8: OAI-PMH compliant IR**

### 6.5 OAI-ORE Compliance

80 per cent of the respondents said yes to the question while 20 per cent said no (See Figure 9).
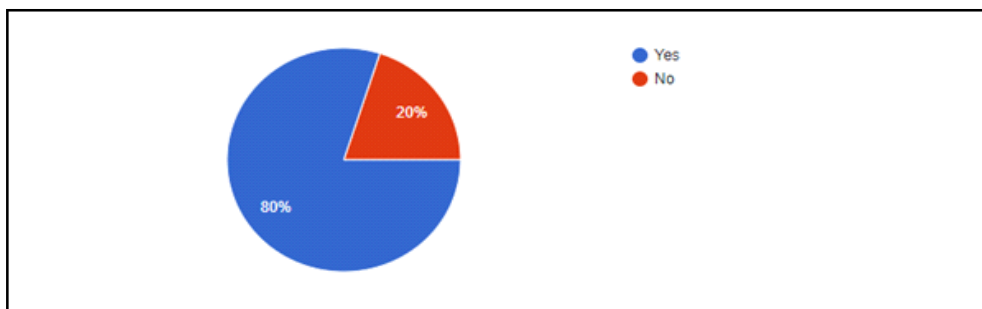


**Figure 9: OAI-ORE Compliant IR**

**6.6 Role of metadata in long-term digital preservation**

Descriptive question was asked which was how do you perceive the role of metadata in long-term preservation of digital contents of your IR? And the received responses are:

❖ Very crucial

❖ It plays an important role to preserve content in IR and it helps to search, find documents etc.

❖ Very important

❖ Metadata is a map that shows how to make IR. In IR, metadata tells us what the IR objects are. It is a directory that helps people to find information in an IR.

❖ Very much needed and crucial.

**2. Findings and Observations**

❖ Majority of IITs IR does not provide open access.

❖ No one single record has included all the elements of the Dublin Core for the description of ETDs. As DCMES use fifteen core elements to define the digital resources such as audio-video, text, images, webpages, etc. And these fifteen elements are used for the creation of metadata as well, the elements are contributor, date, type, title, subject, publisher, source, identifier, language, format, relation, coverage, rights, description, and creator (Dublin Core™ Metadata Initiative, n.d.).

❖ Repositories such as IITJ, IITGN, IITR, IITD have enriched the ETDs metadata in their IRs by extending the DCMES by using qualifiers. However, IITJ provides the rich metadata description for the ETDs, IITJ have used the maximum qualifiers for defining theses in their collections (See Table 4).

❖ Common metadata elements observed among IITs IR while describing ETDs are: dc.contributor.author, dc.date.accessioned, dc.date.available, dc.date.issued, dc.identifier.uri, dc.language, dc.title, dc.type, dc.subject, dc.publisher.

❖ Some unique metadata elements observed among IITs IR for the description of ETDs are: dc.guide, dcterms.publisher, dcterms.publisher, dc.accessionnumber, dc.type.degree, dc.creator.researcher, dc.identifier.accession, dc.date.registered, dc.date.awarded.

❖ Ninety percent of the respondents agreed that their IR is OAI-PMH compliant and 80% agreed that their IR is OAI-ORE compliant. However, none of the IITs have made their base URL (OAI-PMH URL) available on their website or IR, and on self-testing, it was found that only the Indian Institute of Technology Hyderabad (IITH), Indian Institute of Technology Roorkee (IITR), Indian Institute of Technology Delhi (IITD), Indian Institute of Technology Gandhinagar (IITGN) are OAI-PMH compliant (See Table 6). However, some issues were encountered in these repositories when the researchers

attempted to harvest metadata from these repositories. For instance, in case of IITD IR the researchers were only able to harvest 1705 records out of 6845 records from their theses collection; but after that, harvesting stopped and is not continuing. In IITR, the researchers were able to harvest, but no files were associated with the item in their repository, whereas in IITH, the researchers are harvesting records, but it is not bringing any results correctly.

**Table 6: OAI-PMH/OAI-ORE Compliant IITs on Self-testing**

| S.N. | IITs | OAI-PMH Compliant | OAI-ORE Compliant | Base URL |
|------|------|-------------------|-------------------|----------|
| 1 | IIT Delhi | Yes | Yes | http://eprint.iitd.ac.in/oai/request |
| 2 | IIT Gandhinagar | Yes | Yes | https://repository.iitgn.ac.in/oai/request |
| 3 | IIT Roorkee | Yes | Yes | http://repository.iitr.ac.in/oai/request |
| 4 | IIT Hyderabad | Yes | Yes | https://raiith.iith.ac.in/cgi/oai2 |
| 5 | IIT Guwahati | Yes | Yes | http://gyan.iitg.ernet.in/oai/request |
| 6 | IIT Mandi | No | No | Not Available |
| 7 | IIT-BHU | No | No | |
| 8 | IIT Ropar | No | No | |
| 9 | IIT Jodhpur | No | No | |
| 10 | IIT Indore | No | No | |
| 11 | IIT Bombay | No | No | |

❖ From the observations, it has been noted that only four IITs such as IITD, IITH, IITR, IITG ETDs metadata are FAIR principles compliant.

## 3. Discussion

The study reveals that management of ETDs has become easier with the advent of IR where metadata plays important role to preserve content and helps user to search, find and retrieve information in an IR. However, it has been observed, some IITs repository are not available as well as accessible in public domain while some provide restricted access to their ETDs. But people should be able to access the scholarly content without any restrictions. The distribution of scholarly content should not be limited to those who have access to Institutional Repositories. It should be accessible to anyone interested in the academic content. As already rising prices of databases such as SCOPUS, WOS, etc., has become a barrier in the way of individuals accessing the scholarly content.

The study also revealed that not a single record has used all of the DCMES for the description. However, IITs IR have enriched their metadata by using qualifiers, where IITJ IR has used maximum no. of metadata elements for the description of ETDs. IR should use as many as metadata for the description because rich metadata of data (digital resource) allows your computer to automatically accomplish tedious sorting and enables a user to find data even if the identifier is not given/does not work. So, describe each record with rich metadata (Go FAIR, n.d.).

From the findings of the study, it has been noted that except IITH, all other IITs repository are using QDC for the description of ETDs collection, as they have enriched the metadata in their IR by extending the DC elements by using qualifiers such as dc.contributor.author, dc.date.accessioned, dc.identifier.uri, etc., (See Table 4).

Variations have been observed in the metadata elements of ETDs among IITs for defining the same identity such as for the description of author name, metadata filed, dc.contributor.author is used by the IR of IITD, IITKGP, IITMD, IITB, IITI whereas dc.creator.researcher is used by IITJ. Similarly for the guide name, metadata field, dc.contributor.advisor is used by the IR of IITD, IITJ, IITB, IITMD, IITI whereas dc.guide is used by IITR IR for the same. And for accession number metadata field, dc.identifier.accession is used by IITJ IR and dc.accession.number is used by IITR IR (See Table 4).

Metadata is accessible even when the digital resource is not available, for instance if the original data is no longer available (this has been observed in some ETDs collection of IITJ IR, where no files are associated to some of their collection), locating people, organisations, or publications involved with the original digital resource can be incredibly beneficial for an individual. Therefore, even when the digital resources are unavailable the metadata should be accessible (Go FAIR, n.d.). This displays how important a metadata is for the digital content because digital objects are fragile and require special care to remain usable; due to the rapid pace of technological change, they may become inaccessible within a few years. Consequently, IR should assure the Findability, Accessibility, Interoperability, and Reusability of both the digital content and its metadata. But, only four IITs (IITH, IITD, IITG, IITR) ETDs metadata are compliant with FAIR Principles though ETDs metadata of all the accessible IITs are findable as well as accessible

## 4. Conclusion

The study revealed that IITs have developed their own repository to manage and preserve digital content created by their students, research scholars, and faculty to ensure long-term access, as the amount of digital content produced has grown exponentially and is expected to continue. IIT repositories describe ETDs using the Dublin Core™ Metadata Element Set for active content management. 11 (47.87%) of 23 IITs' IR are accessible on Internet. The content such as ETDs are submitted by the repository staff in maximum of the IITs IR while author self-archiving is also available in IITK. 77.8% of IITs IR manually enter metadata whereas IITGN and IITM import metadata from SCOPUS and WOS. 66.7% of IIT IRs have no metadata standardization policy.

Variations have been observed in the metadata practices of IITs IR for the description of ETDs such as IITs IR are using enriched metadata elements using qualifiers, while a few using extensive metadata elements. Common metadata elements (dc.title, dc.type, dc.subject, etc.,) and unique metadata elements (dc.guide, dc.creator.researcher, dcterems.publisher, etc.,) have been observed among IITs IR while describing ETDs. Different metadata elements are used for describing the same entity among IITs.

The study also provides insights into the metadata quality of ETDs in IITs. However, a small sample size was used for this purpose. Based on the observations a suggestion is given that all content types (such as articles, reports, images, theses, etc.,) contained in the IITs IR should be considered for a comprehensive evaluation of resource metadata quality because IITs IR contains a wide range of scholarly content, various variations in metadata quality could be discovered.

On the basis of the conducted study, need of trained library staff with sufficient skills will be required to deal with things like metadata and other technical matters, thereby creating an opportunity for tech-savvy library professionals. Hence, it is essential that the curriculum of Library and Information Science (LIS) includes all facets pertaining to the management and preservation of digital information, ensuring that students acquire prior knowledge and understanding of the associated terminology. There exists a necessity to provide education to administrators as well regarding the benefits of Open Access, open repositories, and the implementation of OAI-PMH, under who the library is working. Therefore, administrators, faculties and the general public has to be made aware of the benefits of the open access so that it can be accepted in society.

IIT's repository are not compliant with FAIR Principles; some IIT IRs were inaccessible when their links were clicked while others provide restricted access to ETDs collection while others make only few pages of ETDs accessible. However, the metadata of ETDs collection among all the accessible IITs IR is findable as well as accessible. Though, the ETDs metadata of only four IITs such as IITD, IITR, IITH and IITGN are FAIR compliant. An IR is a repository where all scholarly content produced by the institution is stored and made freely accessible to users, thereby ensuring content accessibility. Therefore, IITs should frequently examine their repositories to determine if they are operational and if any digital items are absent from their collections and should be OAI-PMH compliant since it shows that the repositories are actually following the concept of Open Access by allowing others to gather their metadata.

**References**

Jacobsen, A., de Miranda Azevedo, R., Juty, N., Batista, D., Coles, S., Cornet, R., ... & Schultes, E. (2020). FAIR principles: interpretations and implementation considerations. Data intelligence, 2(1-2), 10-29. doi: https://doi.org/10.1162/dint_r_00024

Beyene, W. M., & Godwin, T. (2018). Accessible search and the role of metadata. Library Hi Tech, 36(1), 2–17. https://doi.org/10.1108/LHT-08-2017-0170

Choudhury, M. H., Salsabil, L., Jayanetti, H. R., Wu, J., Ingram, W. A., & Fox, E. A. (2023). MetaEnhance: Metadata Quality Improvement for Electronic Tbeyheses and Dissertations of University Libraries. arXiv:2303.17661. https://doi.org/10.48550/arXiv.2303.17661

GO FAIR. (n.d.). Fair Principles. Retrieved March 2, 2023, from https://www.go-fair.org/fair-principles/

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. http://jlisnet.com/journals/jlis/Vol_1_No_1_December_2013/10.pdf

Lihitkar, S. R. & Lihitkar, R. S. (2014). Electronic theses and dissertations (ETDs) in            India:          a comparative study. Library Hi-tech News, 9-14. 10.1108/LHTN-10-2013- 0061

Lubasa, R. L. (2009). Defining Best Practices in Electronic Thesis and Dissertation Metadata. Journal of Library Metadata, 9, 252-263. 10.1080/19386380903405165

Moulaison, H. L., Dykas, F., & Gallant, K. (2015). Open DOAR repositories and metadata practices. D-Lib Magazine, 21(3–4). https://doi.org/10.1045/march2015-moulaison

Osman et al., (2023). Metadata matters: evaluating the quality of Electronic Theses and Dissertations (ETDs) descriptions in Malaysian institutional repositories. Malaysian Journal of Library & Information Science, 28(1), 109-125. https://doi.org/10.22452/mjlis.vol28no1.7

Park, E. G., & Richard, M. (2011). Metadata assessment in e theses and dissertations of canadian institutional repositories. The Electronic Library, 29(6). https://doi.org/10.1108/02640471111141124

Romero-Pelaez, A., Segarra-Faggioni, V., & Alarcon, P. P. (2018). Exploring the provenance and accuracy as metadata quality metrics in assessment resources of OCW repositories. In Proceedings of the 10th International Conference on Education Technology and Computers (pp. 292-296). https://doi.org/10.1145/3290511.3290540

Steele, T. & Sump-cretha, N. (2016). Metadata for Electronic Theses and Dissertations: A Survey of Institutional Repositories. Journal of Library Metadata, 16(1), 53-68. http://dx.doi.org/10.1080/19386389.2016.1161462

Zeng, M. L., & Qin, J. (2016). Metadata. Facet Publishing.