User Interface Features and Knowledge Organisation Systems in Institutional Repositories: A Case Study of Centrally Funded Technical Institutions in India

Dibya Kishor Pradhan and Bulu Maharana

ABSTRACT

Institutional Repositories (IRs) are effective systems for managing and disseminating institutions’ scholarly communication. More specifically, an IR enhances the visibility and discoverability of the content and validates the repository’s importance. Knowledge Organisation System (KOS) strengthens the digital content organisation, connects users with collections, and improves information retrieval functionalities. This paper investigates the present status of user interface features and incorporates KOS in the institutional repository of technical institutions, restricted to Centrally Funded Technical Institutions in India. A group of twenty-four web-accessible institutional repositories was identified, and their KOS and user interface features were evaluated. It was found that user interfaces of all IRs under study comply with essential search and navigation functionalities, such as simple and advanced search, browsing, faceted or filtering approaches, and integration with multiple KOS. Only a few of them include complex KOS, such as control vocabulary. All repositories show their search results in both normal-text and metadata views. Some have specific display features, such as highlighting the query or displaying a thumbnail. Google is one of the most popular search engines that indexes IR content for visibility and discoverability, and approximately 90 per cent of repositories are linked with NDLI. Global visibility and impact participation are moderate, and they require attention.

Keywords: Institutional repository; Knowledge organisation; Knowledge organisation system; Information retrieval; User interface; Digital library

1. INTRODUCTION

Academic libraries set up Institutional Repositories (IRs) to preserve, manage, and disseminate the generated knowledge as scholarly communication. The goal of creating IRs is to archive, organise, and disseminate the intellectual works of faculty and students in the form of research and teaching materials and document all institutional activities. Clifford A. Lynch defined the standard definition of institutional repositories, commonly accepted everywhere, as “A set of services that a university offers to its community members to manage and disseminate digital materials created by the institution and its community members. It is essentially an organisational commitment to stewardship these digital materials, including long-term preservation where appropriate organisation and access or distribution”. The open-access institutional repositories benefitted the academic institution by establishing a broader profile and scholarly communication to a wide range of communities which helps enhance the visibility of its research and education. The libraries can establish this best opportunity by shaping and organizing scholarly communication for broader dissemination. The SPARC position paper stated that “the potential impact of institutional repositories on academic libraries occurs on both strategic as well as tactical levels. Establishing an institutional repository indicates that the library’s role moves beyond a custodial to a contributory for the evolution of scholarly communication. Institutional repositories offer an adequate response for libraries with an organisational imperative to invest in the future”. The IRs are more than just organisational tools; they can help the institution’s visibility and status. The potential of the IRs is determined based on their usage. The user’s information needs and the content in the repositories can be linked through the Knowledge Organisation System (KOS) schematic framework that organises the digital content for dissemination and retrieval. The KOS has been integrated and optimised to make information retrieval easier for the user. The IRs implement the KOS for organizing and managing precise and accurate information retrieval through a systematic search and browsing interface for greater user experience. This study attempts to assess the current status of the user interface features and analyses the KOS incorporated with the institutional repositories under the purview of this study.

2. LITERATURE REVIEW

There are four types of repositories: Subject, Research, National, and Institutional. The institutional repositories evolve as online storehouses where research outputs are archived according to their institutional requirements or settings. The
scholarly communication supports knowledge advancement and is parallel to the existing journal’s infrastructure; hence, the repositories are essential in research communication. Mainly, the IRs are not only offering dedicated service to a specific community; instead, they play a significant role in the institution’s assessment. Furthermore, the library’s part is essential to compile and measure the impact\(^9\). The organisation of the digital documents and control authority is necessary for the institutional repository for consistency and knowledge sharing\(^7\). The scholarly communication support through institutional repositories emphasises the collaboration of different professions and depicts that the library is familiar with organizing knowledge access\(^8\). Discoverability of the content is one of the essential aspects of the repository in an institution similar to publisher databases. The development of the metadata is vital to improve discoverability and usage\(^9\). The libraries followed interoperability, harvesting, and standards for the information retrieval function, and the Dublin Core is the most popular, well-recognised, and widely used standard\(^10\). However, various challenges are found in the metadata dimension, such as insufficient resources to create metadata, the quality, interoperability among schemes, the lack of controlled vocabulary, etc. A single metadata scheme, e.g., Dublin Core, is insufficient, and few other metadata schemes are available to represent only a specific domain\(^11\).

At the Australian National University institutional repositories, the user interfaces of DSpace and Eprints are based upon three primary functions such as simple search, advanced search, and browsing\(^12\). The study on the user interface of the Canadian post-secondary institutional repositories incorporated KOS, whereas few IRs have complex KOS, such as controlled vocabularies. Still, all essential information retrieval features, such as searching and browsing options, are available in all IRs. The information retrieval, organisation, and representation should focus on controlled vocabularies and user drive metadata\(^9\). The study on 32 Digital Libraries (DLs) illuminated that each DL based on their collection, uses different KOS, and their search options are either own or third-party search engines. Categorisation is deployed in some DLs, whereas the KOS plays a vital role in DL architecture to enhance the information retrieval system and performance\(^13\).

The study on 33 Canadian digital libraries identified the use of KOS in their digital library collection, where thesauri, subject heading lists, and classification schemes are widely used\(^14\). The 269 North American digital library collections survey found that the most popular subject representation was the Library of Congress Subject Headings (LCSH), followed by domain-specific thesauri. Almost half of the digital library collections utilised the locally developed taxonomies, whereas only a few DLs used the Dewey Decimal Classification (DDC) and alphabetical indexes\(^15\). The LCSH is the most widely acknowledged controlled vocabulary; it features rich cross-referencing links of its interrelated structure, indicating associations between terms with BT for broader terms, NT for narrower terms, RT for related terms, and UF use-for. The most common criticism about LCSH is that it is difficult for cataloguers and users to use. The Dewey decimal system is the most popular for organizing digital collections in the DL. There are several classification schemes, and many people use more than one. The thesauri are one of the vital KOS, which offers a complete alphabetical index of subject access terms used to represent resources in the database. The locally produced subject descriptions faced interoperability issues\(^16\).

The visibility and discoverability of institutional repositories and their contents are vital to scholarly communication and validate the repository’s importance. Search engines like Google and others and specialised search engines like Google Scholar feature and establish them as a paradigm for making world information accessible\(^17\). Scholarly aggregators such as CORE and BASE are supported by leading repository software such as Eprint, DSpace, Digital Commons, OJS, and others for improved web visibility and usability\(^18\). However, apart from the popular search engine, a library catalog is essential for locating library materials. Adding item lists to OCLC WorldCat makes the items more discoverable\(^19\). To increase the visibility and impact of any repository and to encourage open access, the Ranking Web of World Repositories (https://repositories.webometrics.info/en) initiative launched by Cybermetrics Lab is an indicator of webometrics ranking of institutional repositories. It has established a benchmark for measuring impact\(^19\).

3. OBJECTIVES OF THE STUDY

The primary objective of this study is to look into the current state of user interfaces of accessible IRs for searching and retrieving information. Furthermore, from the perspective of the search interface, the incorporation of KOS to organise the documents or document representations is being investigated. However, research has been conducted to learn about the KOS aspect in IRs. This present study focuses specifically on institutional repositories developed by academic libraries of technical institutions in India, with a focus on Centrally Funded Technical Institutions (CFTIs) only by attempting to make this study unique in the Indian context with the following main objectives:

- To identify the IR initiative and its status;
- To evaluate and analyse the user interface features and to determine how the KOS concept is integrated into IRs;
- To explore the visibility and discoverability of the repository content.

4. SCOPE AND LIMITATIONS OF THE STUDY

This study considered the CFTIs as mentioned in the Ministry of Education, Government of India website (https://www.education.gov.in/en/technical-education-1)\(^23\). The collection of data was carried out during the month of August and September 2021. The IRs, which were not accessible on the Web and had only campus-wide access, were not covered under its purview. The major focus of the study is to consider different features of user interface and KOS of IRs, and therefore, there is scope for further research in terms of various standards used in these IRs.

5. METHODOLOGY

The study used the following methodology to identify the
IRs and their user interface features, as well as to evaluate the KOS framework:

- A checklist was developed following a systematic review of literature in the digital library related to IRs and KOS applications.
- Identification of IRs by consulting institute websites of CFTIs and also directories such as OpenDOAR and ROAR.
- Each potential IR was visited and observed to identify the user interface features that support searching and browsing, as well as the display of results. In addition, other observations, such as the software or platform used, institution type and content type, were noted.
- Examined how KOS was implemented in the IR’s user interface and evaluated the primary attributes of KOS.
- Each repository under study was searched using different index and aggregator search interfaces in order to investigate the content discoverability and visibility.
- The observed data were tabulated and analysed using MS Excel spreadsheet for interpretation.

6. FINDING AND DISCUSSION

6.1 Institutional Repositories of Centrally Funded Technical Institutions in India

With the advancement of ICT and information infrastructure, many universities and institutions in India are developing IRs to archive and disseminate valuable knowledge resources and documents produced by these institutions. These IRs help to increase the visibility and accessibility of these resources22. The technical institutions in India are divided into centrally funded, state-funded, and self-financed institutions; eighty-two technical and scientific-educational institutions, known as Centrally Funded Technical Institutions (CFTIs), are centrally supported21. Appendix 1 contains a list of 24 accessible IRs, along with their names and URLs, and Fig. 1 depicts the source of identification. This list excludes IRs with campus-wide access restrictions.

![Figure 1. Source of identification.](image)

Figure 1 demonstrates that, in comparison to Eprints, DSpace seems to be the preferred software for creating IRs. IITs have the most accessible IRs of any institute category examined for this study. It is also noteworthy that institutions like IISc and NIT Rourkela built up repositories on two platforms to archive their respective two types of content. Each of the study’s repositories is multidisciplinary. Consider Electronic Theses and Dissertations (ETD) as one type of content and other digital objects (journal articles, conference papers, pre-print, reports, books, book chapters, courseware, records, etc.) as another. Then, as shown in Fig. 2, 12 IRs (50%) of the participants in this study archived various digital objects, including ETD. Relatively, half of them only preserve ETDs or digital items.

6.2 Evaluation of User Interface Features & Knowledge Organisation System

6.2.1 User Interface Features

All 24 repositories’ user interface features have been observed and tabulated. To retrieve the desired document or information, all institutional repositories in this study have essential functions such as browsing, simple and advanced search, filtering options, syndication links, and relevancy ranking, as shown in Fig. 3. The advanced or expansion search incorporates filters or facets such as title, author, subject, date, Boolean logic, and other search fields to improve the precision of the information retrieved. Only six IRs support filename and file description searches. Interestingly, IIMB has integrated funder, investigator, and journal searching fields options based on their needs in advanced search.

On the DSpace platform, auto-suggestion is available in ten repositories, 19 repositories provide the latest or recently added features, 12 IRs have a help page, and 17 repositories allow personal account creation. It was discovered that none of the repositories provide an alternate term suggestion option. For syndication, all repositories have links to RSS or ATOM. ‘Has file’ can be found in 11 repositories. Browsing by author, title, subject, and date is available in 19 IRs where the repository platform is DSpace. In contrast, browsing by title is not available in 5 repositories where the platform is Eprint. Interestingly, the IR of IISER Kolkata, IISER Pune, IITGN, IITH, and NITR (ETD) have the browsing option by the department. Specifically, the IR of IITG allows users to search by accession number, whereas the repository of IITGN enables users to browse by degree.
6.2.2 Knowledge Organisation System (KOS)

All repositories use multiple KOS and categorise their digital content in one or more ways, as shown in Fig. 4. The KOS alphabetical listing, author listing, and subject alphabetical listing are all available in all 24 repositories. Twenty-one repositories have features for chronological listing, and 20 have subject features displayed hierarchically. Twenty repositories provide a title listing. Only nine of the 19 repositories where theses and dissertations are archived provide advisor information. In the author field, IITI lists the advisor as a supervisor and IISER Kolkata as additional information.

Two IRs establish subject headings or classifications based on discipline or departments, whereas the other seven IRs base their work on the research field. One IR uses DDC, and only one uses LCSH as a subject heading. Uncontrolled subject keywords are used for indexing in 22 IRs. 15 IRs developed a local taxonomy for categorizing content.

6.3 Visibility and Discoverability of the Repository Content

6.3.1 Display

The content’s display or visualisation is required to identify the desired information correctly. The display features are used to assist the user in better understanding the item after retrieving it using the repository’s various retrieval features. Only four types of display or visualisation are identified after examining different methods, as shown in Fig. 5. Each of the 24 repositories has standard text-based as well as metadata record
views. DSpace only supports one metadata record view, Dublin Core. In contrast, the Eprint platform supports a wide range of metadata record views, including Dublin Core, EP3 XML, METS, HTML, ASCII, Refer, BibTex, and others. Only ten IRs highlighted their search term in the display text, and only eight have access to the archived documents as thumbnails.

6.3.2 Discoverability and Visibility

The visibility of an open-access institutional repository among various users is essential, and repository visibility and discoverability are required for content retrieval. By indexing the metadata content with various indexing or aggregating services, the internet can be used to discover content. When evaluating the 24 repositories in this study, the schemes mentioned in Fig. 6 are taken into account from the initial literature review. The National Digital Library of India (NDLI) is a digital library that stores metadata from various educational resources and offers a single-window search interface, which is included in this study.

Google is the most popular search engine, with 23 repositories using it to index metadata, as shown in Fig. 6, while Microsoft Bing indexed 18 IRs. Similarly, twenty IRs prefer Google Scholar, while ten prefer Microsoft Academic. Eight repositories are registered as content providers, with each scholarly aggregator, CORE, and BASE. The OCLC WorldCat only collects metadata from the four repositories considered in this study. The NDLI collects metadata from 21 member repositories to provide a decentralised search service.

The web discoverability of the repository’s content for increased visibility impacts the use of IR digital content and the institute profile. Ranking web repositories attempts to rank institutional repositories, and Fig. 7 depicts the rank of those available repositories from “TRANSPARENT Ranking: Institutional Repositories by Google SCHOLAR (May 2021) 11th Edition”.

7. CONCLUSION

The study attempts to investigate the current state of user interface features incorporating KOS in academic libraries at centrally funded technical institutions in India. This study lists twenty-four publicly available institutional repositories, most from IITs (Indian Institute of Technology). The majority of the IRs in this study are hosted on the DSpace platform, all of which are multidisciplinary. This study shows that functions such as browsing, simple and advanced search, filtering, Boolean logic, and relevancy ranking are available in all IRs. However, features such as auto-suggestion, help page, the latest addition, and personal account options vary from institute to institute, depending on the platform. One of the study’s main goals is to evaluate the KOS used in these IRs. All repositories have an authority file, and most use uncontrolled subject-keyword as an index term for searching or browsing. In contrast, one-third of the repositories use subject headings or subject classifications developed locally based on the research.
area or subject discipline they provide. Out of the 24 IRs, only one, IISER Kolkata, uses the controlled subject heading LCSH. The use of the DDC by IISc (ETD) and the LCC by IISER, Kolkata, to derive the subject heading demonstrates that very few IRs use controlled vocabulary for KOS. As KOS bridges the user’s information needs and the material in the collections, the library will focus on using control vocabulary to organise the information for retrieval.

Further, it was discovered that all IRs displayed their digital content in normal text and metadata view. In contrast, query highlight and thumbnail view were platform-dependent. Their visibility and discoverability validate open access IRs on the internet’s diverse web resources. In comparison to “Microsoft Bing” or “Microsoft Academic,” “Google” is the most popular search engine for discovering additional content or metadata from the repositories investigated. “Google Scholar” is the preferred specific search engine. The majority of IRs are linked to the National Digital Library of India (NDLI) to increase the visibility and discoverability of their content. This study, however, shows that IRs have moderate global visibility and impact, a matter to be further investigated and focused on the necessary policy framework.

REFERENCES


CONTRIBUTORS

Mr Dibya Kishor Pradhan is the Head and Assistant Librarian of Biju Patnaik Central Library, National Institute of Technology Rourkela, Odisha, India, and a PhD. scholar at the Department of Library and Information Science at Sambalpur University, Odisha. His research interests include digital libraries, knowledge organisation, information retrieval, e-resource management, open-access systems, and services. His contributions to the current study are literature review, data collection, analysis, and completion of the paper, which is part of his doctoral research.

Prof Bulu Maharana is the Head of the Department of Library and Information Science at Sambalpur University, Odisha, India. His areas of expertise include information technology, library automation, and information storage and retrieval. His contributions to the current study are the conceptualisation, supervision, and review of this paper.
## Appendix 1

### List of institutional repositories of CFTIs

<table>
<thead>
<tr>
<th>Name of the Institute</th>
<th>Name of the Repository</th>
<th>URL</th>
<th>Software</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIMA</td>
<td>Institutional Repository@VSL(IIMA Institutional Repository)</td>
<td><a href="http://vslir.iima.ac.in:8080/xmlui/">http://vslir.iima.ac.in:8080/xmlui/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>IIMB</td>
<td>Research @IIMB</td>
<td><a href="https://repository.iimb.ac.in/">https://repository.iimb.ac.in/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>IIMK</td>
<td>DSpace at Indian Institute of Management Kozhikode (DSpace@IIMK)</td>
<td><a href="http://dspace.iimk.ac.in/">http://dspace.iimk.ac.in/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>IISC (IR)</td>
<td>Open Access Repository of IISc Research Publications (ePrints@IISc)</td>
<td><a href="http://eprints.iisc.ac.in">http://eprints.iisc.ac.in</a></td>
<td>Eprint</td>
</tr>
<tr>
<td>IISC (ETD)</td>
<td>Electronic Theses and Dissertations at Indian Institute of Science (eTheses@IISc)</td>
<td><a href="https://etd.iisc.ac.in/">https://etd.iisc.ac.in/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>IISER Kolkata</td>
<td>IISER Kolkata ePrints Repository</td>
<td><a href="http://eprints.iiserkol.ac.in/">http://eprints.iiserkol.ac.in/</a></td>
<td>Eprint</td>
</tr>
<tr>
<td>IISER Pune</td>
<td>Digital Repository at Indian Institute of Science Education and Research Pune</td>
<td><a href="http://dr.iiserpune.ac.in:8080/xmlui/">http://dr.iiserpune.ac.in:8080/xmlui/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>IITBBS</td>
<td>IDR @ IIT Bhubaneswar</td>
<td><a href="http://idr.iitbbs.ac.in/jspui/">http://idr.iitbbs.ac.in/jspui/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>IITB</td>
<td>DSpace at IIT Bombay</td>
<td><a href="http://dspace.library.iitb.ac.in/jspui/">http://dspace.library.iitb.ac.in/jspui/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>IITD</td>
<td>Eprints@IIT Delhi</td>
<td><a href="http://eprint.iitd.ac.in">http://eprint.iitd.ac.in</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>IITGN</td>
<td>IIT Gandhinagar Digital Repository</td>
<td><a href="http://repository.iitgn.ac.in/">http://repository.iitgn.ac.in/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>IITG</td>
<td>Lakshminath Bezbaroa Central Library Digital Repository</td>
<td><a href="http://gyan.iitg.ernet.in/">http://gyan.iitg.ernet.in/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>IITH</td>
<td>Research Archive of Indian Institute of Technology Hyderabad (RAIITH)</td>
<td><a href="http://raith.iith.ac.in/">http://raith.iith.ac.in/</a></td>
<td>Eprint</td>
</tr>
<tr>
<td>IITI</td>
<td>IITI Institutional Digital Repository</td>
<td><a href="http://dspace.iiti.ac.in:8080/jspui/">http://dspace.iiti.ac.in:8080/jspui/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>IITJ</td>
<td>IIT Jodhpur Theses Repository</td>
<td><a href="http://theses.iitj.ac.in:8080/jspui/">http://theses.iitj.ac.in:8080/jspui/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>IITKGP</td>
<td>Institutional Digital Repository (IDR)-IIT Kharagpur</td>
<td><a href="http://www.idr.iitkgp.ac.in/xmlui/">http://www.idr.iitkgp.ac.in/xmlui/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>IITM</td>
<td>IRepose IIT Madras</td>
<td><a href="http://irepose.iitm.ac.in:8080/jspui/">http://irepose.iitm.ac.in:8080/jspui/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>IITP</td>
<td>Institutional Repository of IIT Patna</td>
<td><a href="http://idr.iitp.ac.in/jspui/">http://idr.iitp.ac.in/jspui/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>IITR</td>
<td>Sodhabhagirathi</td>
<td><a href="http://shodhabhagirathi.iitr.ac.in:8081/jspui/">http://shodhabhagirathi.iitr.ac.in:8081/jspui/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>IITRPR</td>
<td>Institutional Digita Repository (DSpace Repository) IIT Ropar</td>
<td><a href="http://dspace.iitrpr.ac.in:8080/xmlui/">http://dspace.iitrpr.ac.in:8080/xmlui/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>NITK</td>
<td>IDR@NITK</td>
<td><a href="https://idr.nitk.ac.in/jspui/">https://idr.nitk.ac.in/jspui/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>NITR (IR)</td>
<td>DSF@NITR (Institutional Repository)</td>
<td><a href="http://dspace.nitrkl.ac.in/dspace/">http://dspace.nitrkl.ac.in/dspace/</a></td>
<td>DSpace</td>
</tr>
<tr>
<td>NITR (ETD)</td>
<td>ethesis@nitr</td>
<td><a href="http://ethesis.nitrkl.ac.in/">http://ethesis.nitrkl.ac.in/</a></td>
<td>Eprint</td>
</tr>
</tbody>
</table>