International Journal of Information Science and Management
Vol. 20, No. 1, 2022, 15-39
https://dorl.net/dor/20.1001.1.20088302.2022.20.1.2.5

Original Research

Collection Development and Organization in Institutional Digital Repositories: From Policy to Practice

Bijan Kumar Roy
University of Calcutta- Library and Information Science, Associate Professor – India
Corresponding Author: bkroylis@gmail.com
ORCID iD: https://orcid.org/0000-0001-9735-9586

Subal Chandra Biswas
The University of Burdwan - Library and Information Science, Professor {Retired}- India
scbiswas_56@yahoo.co.in
ORCID iD: https://orcid.org/0000-0002-6047-369X

Parthasarathi Mukhopadhyay
University of Kalyani -Library and Information Science, Professor – India
psmukhopadhyay@gmail.com ORCID iD: https://orcid.org/0000-0003-0717-9413

Received: 25 August 2020
Accepted: 2 March 2021

Abstract

This paper analyzes open access (OA) self-archiving policies of open access repositories of different organizations registered in OpenDOAR, ROAR, and ROARMAP databases. The policies relating to content policies, collection organization policies, metadata policies, submission policies, and multilingual policies, are required to be correctly formulated for the smooth functioning of any repository system on a global scale. The objective is to recommend institute-specific model policy with global recommendations and best practices. The methodology is twofold – first, to get an overview of policy issues as reflected in existing literature, and second, to analyze policy issues recorded in global registries and individual repositories. It was found that most of the organizations do not have a stated policy. Even within a specific policy, several key issues were missing. The outcomes of this research paper may help future researchers by providing a roadmap towards the successful policy implementation of open access repositories (OARs) in higher academic institutions. The paper may be helpful to the library professionals in devising institute-specific policy and may be a guiding tool to the policymakers.

Keywords: Open access repository, digital library, digital repository, digital archive, open access, self-archiving policy.

Introduction

Self-archiving open access (OA) policy is defined in the literature as "*governing how a digital library is instantiated and run*" (Innocenti et al., 2011). Self-archiving policies are important to identify repositories' perspectives for metadata, data, content, submission, and preservation. The present research study has identified several areas where proper OA self-archiving policy documentations are essential for the smooth running of a repository system. Developing a repository using any software is not so difficult, but the hard part is developing

policy documentation. It was found in a study that most of the repositories registered in global databases like OpenDOAR, ROAR, or ROARMAP do not have supporting policy documentation (Millington, 2006). The paper addresses five key policies (e.g., Contents Policy, Collection Organization Policy, Metadata Policy, Submission Policy, and Multilingual Policy) along with several key issues in their policy documentation. The purpose of this paper is two-fold: first, to provide an overview of the policy documentation in the five areas as mentioned above based on existing published literature; and second, to suggest best practices in the line of global recommendations. Therefore, the objective of this paper is to study the pre-defined policies (i.e., Content Policies, Collection Organization Policies, Metadata Policies, Submission Policies, and Multilingual Policies) based on the researchers' points of view and their implementations. The specific details of the policy studies are as follows:

- ➔ Content Polices — Type of materials covered
- ➔ Collection Organization Policies — Organization and management; categories and sub-categories; community rights
- ➔ Metadata Policies — Access to metadata; re-use of metadata; eligible depositors/creators; harvesting framework, authentication; schema used
- ➔ Submission Policies — Eligible contributors; deposition rules; moderation, workflows, waive deposit
- ➔ Multilingual Policies — Incorporation of Indic-script based documents; Browsing & searching of multilingual resources; multilingual subject access support system

Methodology

This research aims to explore implementations of policies related to collection development and organization at a global scale. The specific parameters for this analytical study of practical implementations of OARs policies are content, collection organization, metadata, submission, and multilingual policies.

Content Policies

Content guidelines or policies govern what sorts of material are eligible for inclusion in an institutional digital repository (IDR). This policy ensures the relevance and quality of material deposited to IDR. The content policies generally define how scholarly content is selected for inclusion in an IDR. It is difficult to identify what kind of publications are preserved and by whom, i.e., content contributors. Another critical question is whether repositories contain only full text with metadata or include metadata-only records without archiving the related digital object in full-text form.

Collection Organization Policies

These policies define how digital collections are organized and managed inside the repository and whether the communities have the right to design their policies and guidelines regarding the specific content or deposited items in submission processes. Therefore, the digital resources need to be organized under suitable categories and sub-categories to reflect and cover the community members' thrust areas or information demand areas. These policies cover two important issues: i) *how collections are organized and managed* and ii) *whether the communities have the right to design their policies*.

Metadata Policies

In metadata policies, questions arise about who can enter or edit metadata or which metadata standards must be followed. Will IDRs follow different metadata schema for different types of documents? Will IDRs follow any indexing system? Will metadata of withdrawn items be searchable? Will IDRs allow metadata harvesting of dataset descriptions by other institutions following the OAI-PMH guidelines or other harvesting protocols? Most of the IDRs registered in global databases have covered five key issues: i) *existence of stated metadata policy,* ii) *metadata creation,* iii) *metadata standards used*, iv) *indexing systems followed*, and v) *metadata harvesting framework*.

Submission Policies

It is essential to define who will be able to submit content to the IDR. A significant decision in the submission policy is whether your institution would assist in the submission process. And, if you will mediate the submissions, how far would you go? Or, what would be the workflow for submission? Or, can the deposit in IDRs be waived? These policies cover three key issues: i) *who would be the eligible submitter and who will provide assistance at the time of submission, ii) publishers' and funders' embargo,* and iii) *workflow design and management*.

Multilingual Policies

These policies deal with developing a Unicode-compliant script-based interface for processing and managing objects written in languages other than English. The most important question in this area is 'availability of Unicode-compliant script-based interface' for IDRs.

The broad groups of works associated with this study include a) investigation of experts' opinions as recorded in literature against the major issues/areas mentioned under each of the OA policies related to collection development and organization; the opinions of experts are then grouped and analyzed to crystallize a set of recommendations and/or best practices; b) identification of global repositories that adopted the policy issues/areas as identified as important considerations for the said OA policies. The first part of work (a) is based on extensive literature reviews sourced from various scholarly resources, spanning from 1972 to 2019 and covering 82 papers, including 39 journal articles. And the second part (b) includes the study to measure implementation of recommendations/best practice guidelines as provided by experts for the policy issues adopted by the listed OARs and as reflected in the OA policy databases. The policy databases at the global scale are presently OpenDOAR (University of Nottingham), ROAR (University of Southampton.), and ROARMAP (University of Southampton). There are overlaps in coverage of OA repositories in these three global datasets. Therefore, an exclusive list has been prepared after due duplication-checking of OA repositories

from respective databases. The unified list of OA repositories has then been analyzed against a set of pre-defined parameters based on experts' opinions.

The following steps have been taken to accomplish this task:

1. Consultation of three global databases on open access repositories (OARs), namely OpenDOAR, ROAR & ROARMAP databases;
2. Preparation of unified list of OARs as listed in all three above-stated databases after due duplication-checking;
3. Collecting policy datasets of each listed OARs from OpenDOAR, if available (this policy tool is now archived here presently (https://web.archive.org/web/20070831193103/http://opendoar.nottingham.ac.uk/tools/en/policies.php), available up to December 2019);
4. Consulting individual OAR to collect policy data if not available comprehensively on OpenDOAR policy tool; and
5. Developing a dataset related to pre-defined policy issues as derived in a tabular format derived from steps 3 and 4 tabulated for analysis.

The above methodology may be illustrated in Table 1A.

Table 1A
Thematic representation of the methodology

Policies and issues	Review of Literature	Practical Implementation
Content Policies *(types of content)*	Articles (13) Books/Chapters (2) Databases (2) Online Resources (1) Ph.D. Theses (1)	OpenDOAR Policy Tool (https://v2.sherpa.ac.uk/opendoar/policytool/) ROAR (http://roar.eprints.org/) ROARMAP (http://roarmap.eprints.org/)
Collections Organization Policies *(organization and management of resources)* *(setting permission policies for collections)* *(community rights)*	Articles (7) Books/Chapters (1) Databases (2) Ph.D. Theses (1)	
Metadata Policies *(eligible submitter of metadata)* *(metadata standards to be followed)* *(indexing systems to be followed)* *(metadata harvesting protocol)*	Articles (9) Books/Chapters (3) Conference/Seminar papers (1) Databases (2) Online Resources (7) Ph.D. Theses (1)	
Submission Policies *(eligible submitter)* *(assistance in submission)* *(workflow for submission)*	Articles (15) Books/Chapters (1) Conference/Seminar papers (1) Databases (2) Online Resources (1) Ph.D. Theses (1)	

Policies and issues	Review of Literature	Practical Implementation
Multilingual Policies (*Unicode-compliant script based interface*)	Articles (10) Books/Chapters (2) Conference/Seminar papers (9) Databases (2) Online Resources (7) Ph.D. Theses (1) Presentations/Talk (2)	

Discussion

This section deals with the analysis and interpretation of different issues associated with the above-stated policies under two broad headings: r*esearchers' point of view* and *Practice point of view*. The first one, as stated earlier, is concerned with a review of existing scholarly literature, and the second one is related to the study of IDRs registered in the global databases mentioned above. An overview of the policies under study has already been given briefly in the methodology section.

Contents Policies

Many OA advocates have emphasized devising content policies for IDRs. For example, Jones (2007) said that building technological infrastructure is not hard, but "deciding on content is the matter of concern". Generally, there is no consensus on contents type, and different institutions have different content structure (OpenDOAR, 2019; ROAR, 2019). Prost and Schöpfel (2014) study reported that more than half of the documents were journal articles. In several studies, Roy, Biswas and Mukhopadhyay (2011, 2012 and 2013) also reported that most of the IDRs covered mainly textual documents, but full-text documents were not available in some cases. Several experts (Genoni, 2004; Gibbons, 2004b; Lynch, 2003; Crow, 2002) described potential digital contents for the IDRs. Shearer (2002) reported that MIT's DSpace allows only publisher-accepted materials, and the University of California's e-Scholarship accepts only materials that authors have not submitted to publishers. Smith (2008) details a "wide variety of materials in digital form, such as research journal articles, pre-prints and post prints, digital versions of theses and dissertations, and administrative documents, course notes or learning objects." As stated in the Research Information Network's report, there are "hugely varied kinds of digital research data – from texts and numbers to audio and video streams" in IDR (RIN, 2008). Yeates (2003) reported that IDRs generally cover the following content: pre-prints, other work-in-process, peer-reviewed articles, monographs, enduring research material, datasets, other ancillary research material, conference papers, electronic theses, and dissertations, and grey literature. Another group of experts *(*Loan & Rather, 2007; Lone, Rather & Shah, 2008*)* supported this view and revealed that repositories archive both published and unpublished documents like articles, research papers, conference/seminar papers, theses, books, etc. Roy (2014), in his research study, opined that there is no hard and fast rule regarding content types because the structure of the archive depends on the software, technical support, vision, and resources of the IDR. He further stated that 'content contributors' may include outsiders if they are co-authoring with the university authors or are affiliated closely with the university. In another work (Roy, 2015), he showed that IDRs host varieties of materials depending on the institutional preferences. The types of contents can range from dissertations and articles to raw research data and datasets, post-prints (peer-reviewed research articles), book chapters, working

papers, theses, etc. Based on the opinions of OA experts, this section now checks the ground realities as reflected in ROARMAP database. After analyzing ROARMAP database, it was found that all IDRs (a total of 439 IDRs as of June 2013 as reported by Roy in 2014) contain different types of textual and non-textual documents. Though collections were mainly textual, only a few IDRs hold non-textual objects such as software, specials, patents, multimedia, etc. The present scenario is almost the same as it was five years before. A review of a total number of 5395 IDRs (as of August 2020) registered in ROARMAP database shows that institutional policies regarding contents vary substantially, and documents are mainly text-based. About 3897 IDRs contain peer-reviewed articles; 3321 IDRs cover ETDs (electronic theses and dissertations), and 2135 IDRs hold books. Though in the case of a few IDRs, pieces of evidence of holding complex digital materials like datasets, software, patents, etc are available. As per OpenDOAR database, about 73.9% (out of 5395 IDRs as of August 2020) of repositories did not define content policy (Fig. 1) and mainly contained textual materials (Fig. 2).

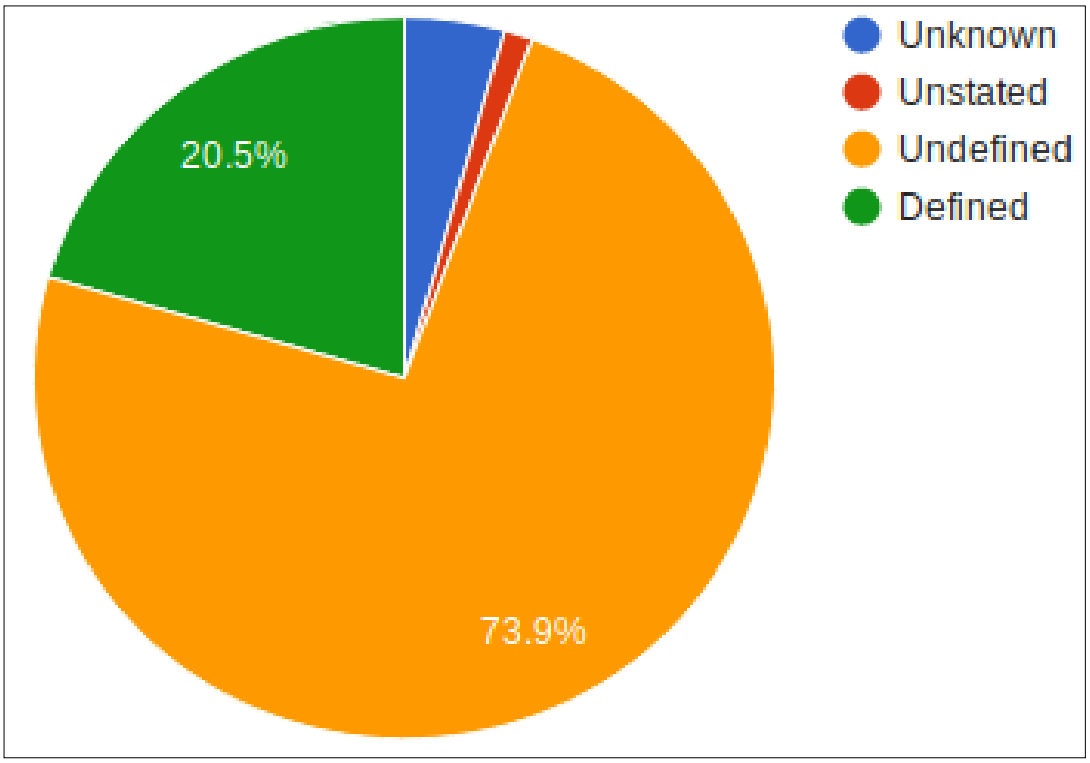

*Figure 1:*Content Policies – Worldwide (Source: OpenDOAR, 2019)

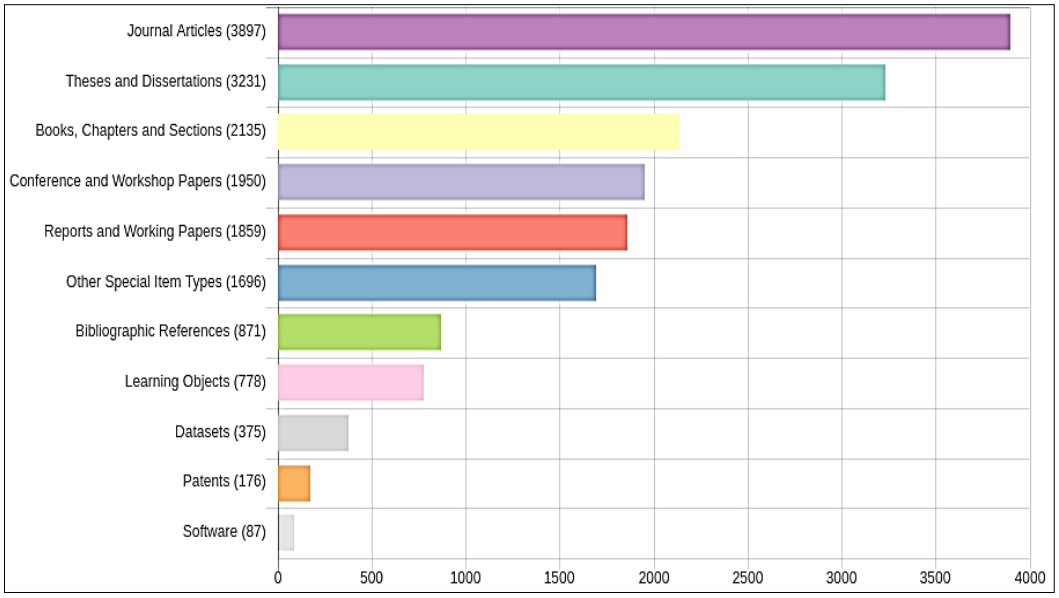

Figure 2: Content type of Repositories (Source: OpenDOAR, 2020)

Collection Organization Policies

Many researchers have advocated the need for collection organization and management policies of IDR. Another set of researchers has raised two views regarding these policies, such as organization and managing collections and setting different permission policies for different collections, including community rights. There is no hard and fast rule, and resources could be organized according to institutional preferences (Roy, Biswas & Mukhopadhyay, 2011, 2012, and 2013). Resources are generally organized by departments (e.g., *product units*) and by subject categories (e.g., *contents*) (Roy, 2014, 2015). For example, OA resources may be arranged by 'subject' (Vidyanidhi), by 'collection type' (IIT, Bombay), or by 'departments and school' (Cochin University of Science and Technology) (OpenDOAR, 2019; ROAR, 2019). Probets and Jenkins (2006) advocated that contents could be organized under 'communities' because organizing content in this way (under 'communities') has the advantage of providing departmental flexibility. Another study (Pappalardo, Fitzgerald, Fitzgerald, Kiel-Chisholm, O'Brien & Austin, 2007) suggested that content be organized in many different ways (e.g., subject, departments, authors or years of publications) so that the repository is browsable through as many different ways as possible. Controlled vocabulary searching and authority control needs serious reworking in IDR environment. Roy (2014), in his research work, developed a mechanism by which content could be organized as per standardized vocabulary control device such as DDC (Dewey Decimal Classification). In another work (Roy, 2015), he stated that it is much more logical to arrange content according to open standard-based subject access tools to satisfy the subject approaches of the users. In a recent study, Roy, Biswas and Mukhopadhyay (2017) reported integrating the subject access system, viz., DDC in the IDR system for organizing and retrieving (browsing and searching) scholarly objects in both English and Bengali languages. Now it is time to link the views of experts with the ground realities as reflected in ROARMAP database. There is no consistency in organizing IDR resources, and it varies from a repository to repository. It was found in ROARMAP database that only 9 (2%) IDRs (out of 439 as of June 2013) as reported by Roy in 2014, explicitly stated how their collections were organized and managed (e.g., under departments, units, or subject categories)

(Table 1). Presently the database lists 5395 IDRs (as of August 2020), but data regarding this issue is not available. Setting up permission policies for different collections is essential, but no expressive study has been conducted in this area. There is a lack of published literature in this area. Probets and Jenkins (2006) recommended that administrators set different permission policies for different collections (read, remove, add, write, etc.) for different types of users. Roy (2014), after analyzing OpenDOAR, ROAR, and ROARMAP databases, reported that community members or content contributors could set content-specific policies. In another work, he suggested that administrators set permission policies for different users (Roy, 2015). There is no such information regarding permission policies in ROARMAP database, and the actual scenario differs from IDR to IDR. In an earlier study, after analyzing ROARMAP (an international registry to record open access mandates and policies adopted by OARs) database, it was found that only in 3 (0.6%) IDRs (out of 439 as of June 2013) the communities can make their own decisions on issues such as the type of materials to be deposited (Roy, 2014)**.** The database currently shows 5395 IDRs (as of August 2020), but no information related to this issue is available.

Table 1
Collection Management Policies

Name of the Repository	Managed and Organized by	A policy designed and owned by
Anglia Ruskin Research Online	Communities and Collections	NA
University of California (CDL)	Departments or Units	NA
Goddard Library Repository	Projects, Organizational code, Subject, or Author	NA
HKUST Institutional Repository	Communities and Collections based on Departments or Units	NA
Massachusetts Institute of Technology (MIT)	Departments or Units	A community may set its policies and guidelines regarding the specific content/deposit of content and submission processes
Queensland University of Technology	Subject categories	NA
The University of Hawaii at Manoa	Departments or Units	Each collection can have its submission process & Authorization (Contributors) set out in Community policies
University of Kansas	Communities and Collections	A community may set its policies and guidelines regarding the specific content/deposit of content and submission processes
University of Utah	The structure depends on the software and technical support	NA

Metadata Policies

An IDR needs to have a stated metadata policy. Devising metadata policies is quite essential in developing any IDR system. Several key issues have to be taken into consideration before devising such policies. Here, five key issues (for example, access, use, and re-use of metadata; eligible depositors/creators of metadata; designing harvesting framework for extraction of metadata, metadata schema used to display an object) associated with these policies have been analyzed against five different viewpoints. Many authors reported that most of the IDRs have not mentioned these policies. Roy (2014) reported that most of the repositories registered in globally recognized databases have no metadata policy. For example, a total of 16 (31%) repositories have a stated metadata policy of different OARs affiliated to COAR (Confederation of Open Access Repositories) (Roy, Biswas & Mukhopadhyay, 2018). Roy (2014) further suggested providing metadata based on the recognized global standards at the point of deposition of items in IDR. In another work, he recommended following basic cataloging standards such as AACR or RDA (Roy, 2015). However, the real-life situation, as evident in ROARMAP database, is quite different from the views of OA advocates. After analyzing ROARMAP database, it was found that only 33 (7.5%) IDRs (out of 439 as of June 2013) as reported by Roy in 2014, have this policy. As per OpenDOAR (OpenDOAR, 2019) database, more than 78% (4208 IDRs out of 5395 IDRs) repositories did not define metadata policy (Fig. 3). Almost all the IDRs clearly stated that anyone might access metadata free of charge. Nevertheless, some metadata may be restricted based on the requirements of the collections. The same database of 5395 IDRs (as of August 2020) does not provide any data on this section. Another key issue has already been discussed in literature by several experts that standard metadata schemas (generic and domain-specific) need to be used by IDR system. Only a few studies (Graaf & Eijndhoven, 2008; Barton & Walker, 2002) have suggested that academicians (generally authors) enter simple metadata, such as author name, title, and keywords. Mediated services may be provided where the administrators add metadata. Other studies (DINI, 2003; Pinfield, Gardner & MacColl, 2002; UGC, 2005) advocated that standardized metadata should be created and provided for exchange and harvesting services. Roy (2014) recommended that authors provide basic metadata at the time of deposition, and if required, additional metadata may be provided by library staff. In another work, he reported that content creators or authors are responsible for creating metadata of a particular object deposited to IDR (Roy, 2015). However, experts' opinions are quite different from the existing operational frameworks of IDRs. In an earlier study, after analyzing ROARMAP database, Roy (2014) reported that only 21 (4.7%) IDRs (against 439 IDRs as of June 2013) suggested metadata be created and provided by authors or eligible contributors. Library staff, if necessary, may create or edit additional metadata (Table 2). The same database lists 5395 IDRs (as of August, 2020), but information is not available regarding this issue. Many experts suggest applying domain-specific metadata schemas as desirable for describing different types of resources that demand special treatments like maps, cultural objects, learning objects, etc. It is logical to use different metadata schemas for the different types of content existing in the repository (Gonzalez & Porcel, 2007). Crow (2002) advocated for standardized metadata formats to provide access to users outside the university. Green, Macdonald and Rice (2009) reported that IDR systems differ widely in handling metadata schema and repositories may need to put in place additional metadata schema to support ingest, management, and use of data in their collections. Graaf and Eijndhoven (2008) concluded that most repositories follow the Qualified Dublin Core standards

and Unqualified Dublin Core standards. UGC, India (2005) supported this view and suggested using other metadata schema to meet users' needs and provide better access and management of different digital objects. Teli (2015) reported the development of a metadata harvesting model for Indian IDRs where he used Dublin Core Metadata standard. Roy (2014), in his research work, suggested using a qualified version of the Dublin Core schema as a descriptive metadata standard for its popularity. In another work, he recommended following community/domain-specific metadata elements where no suitable element or element refinement exists (Roy, 2015). Gibbons (2004a) reported that a few IDRs had created their metadata schemas, and others have used or adopted existing schemas, such as Dublin Core or MARC. Another study (DINI, 2003) reported that some IDRs started supporting other metadata standards (such as Onix). A study of ROARMAP database has shown that only 17 (3.8%) IDRs (out of 439 as of June 2013 as reported by Roy in 2014) stated using metadata standard adopted by the system (Table 2). Most of the IDRs follow Dublin Core standards as a generic metadata schema. But a few repositories implemented additional or extended metadata schema for domain-specific datasets. The database lists 5395 IDRs (as of August 2020), and no information regarding this issue is available in its present structure. Some experts advocated for adopting and use of standard vocabularies by IDR system. But, the published information about using controlled vocabularies is limited. A few studies (DINI, 2003; Nolan & Costanza, 2006; UGC, India 2005) have recommended that standard vocabularies be followed. Roy (2014) developed an OA information retrieval (IR) system where DDC (Dewey Decimal Classification) was incorporated. In another work, the author advocated using a domain-specific vocabulary control device in the IDR system (Roy, 2015). A comparison of ROARMAP data with the views of OA researchers has demonstrated that only two 2 (0.4%) out of 439 (as of June 2013) IDRs (such as Kwame Nkrumah University of Science and Technology Institutional Repository and University of Abertay Dundee) as reported by Roy in 2014 suggested using controlled vocabularies (Table 2). The database covers 5395 IDRs (as of August 2020), but no information is available regarding this issue. Supporting interoperability in order to exchange and share data from system to system has been suggested by many researchers. Crow (2002) stated that the system should support interoperability, and a proper metadata harvesting protocol should be in place to provide access via search engines and other searching tools. Kellogg (2004) surveyed several open-source OAI harvesting tools and concluded that PKP was an excellent metadata harvesting and presentation tool. Sarkar and Mukhopadhyay (2010) proposed a harvesting model (for e-theses) using PKP harvester and explained how metadata is extracted from different OAI-PMH compliant repositories. In another paper, they reported designing a prototype union catalogue of ETDs on Health and Medicine through the application of metadata harvesting from OAI-PMH compliant repositories (Sarkar & Mukhopadhyay, 2012). Jayakanth and Minj (2012) shared their practical experiences in setting up a prototype metadata harvesting service using the PKP harvesting software for the OAI-compliant repositories in India. Teli (2015) designed a metadata harvesting model using the OAI-PMH standard version 2.0. Roy (2014) reported the development of a metadata harvesting framework using PKP harvester by which metadata of OAI-PMH compliant IDRs could be harvested. In another work, he advocated for a federated search option for multiple IDRs from a single-window search interface (Roy, 2015). ROARMAP database versions (2014 and 2020) differ a lot in this context. For example, Roy (2014), after analyzing 439 (as of June 2013) IDRs registered in ROARMAP database, concluded that not a single repository has any published documents on

OAI compliance issues. However, repositories, in principle, are expected to support metadata harvesting from OAI/PMH-compliant repositories. Now, the database lists 5395 IDRs (as of August 2020), and no information regarding this issue is available in its present shape (Table 2).

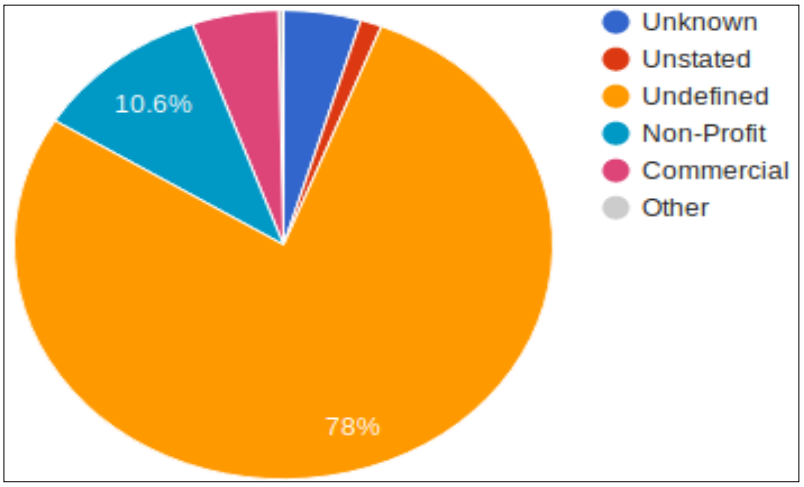

Figure 3: Metadata Policies (Source: OpenDOAR, 2019)

Table 2
Metadata Policies

Name of the Repository	Policy related to Metadata		
	Metadata Schema Used	Bibliographic Metadata	
		Provided by	Created or Edited by
Anglia Ruskin Research Online	Simple Dublin Core		Library staff
Brandeis Institutional Repository	NA	Eligible contributor	NA
Brigham Young University Library	Simple Dublin Core	NA	NA
Centre for Environmental Data Archival Repository	NA	√	NA
Cornell University (eCommons)	NA	NA	√
Edith Cowan University	Unqualified Dublin Core	NA	NA
Goddard Library Repository	GEMS (own)	NA	NA
Griffith University	NA	√	NA
Harvard University Library	NA	authorized submitter	NA
Katholieke Universiteit Leuven	NA	√	NA
Kwame Nkrumah University of Science and Technology Institutional Repository (KNUSTSpace)	Qualified Dublin Core	NA	√
Loughborough University	NA	√	NA
Massachusetts Institute of Technology (MIT)	NA	Eligible contributor/ depositors	NA
Northeastern University Libraries Institutional Repository	METS schema & Qualified & Unqualified Dublin Core	NA	NA

Name of the Repository	Policy related to Metadata		
	Metadata Schema Used	Bibliographic Metadata	
		Provided by	Created or Edited by
St John University	NA	√	NA
Teesside University's Institutional Repository (TeesRep)	Dublin Core	NA	NA
Trento University	NA	√	NA
University of Abertay Dundee	Dublin Core	Authors/ or delegated agents	√
University of Calgary: Library and Cultural Resources	NA	√	NA
University of Cambridge	Qualified Dublin Core	NA	NA
University of East Anglia	NA	√	NA
University of Kansas	Dublin Core Library Application Profile (DC-Lib)	NA	NA
University of Melbourne Eprint Repository	Simple Dublin Core	NA	NA
University of Queensland	NA	√	NA
University of Reading	NA	√	NA
University of Rochester's	Dublin Core & locally defined DTDs	NA	NA
University of Salford	Dublin Core	NA	√
University of South Australia	MARCXML & Dublin Core	NA	NA
University of Starling (STORRE)	Dublin Core	NA	NA
University of Sydney	Qualified Dublin Core	NA	NA
University of Utah's institutional repository	Dublin Core	NA	NA
University of Westminster	NA	√	NA
York St John University	NA	√	NA

Submission Policies

Several researchers have expressed concern about five key issues related to submission policies – eligible contributors to submit items in IDR; deposition rules; moderation, workflow steps in the submission process, and waive-off deposit feature. It was also found in OpenDOAR database that 73.1% (out of 5395 IDRs) repositories did not mention the submission policies (Fig. 4), and only authorized users are eligible to submit documents in IDRs (OpenDOAR, 2019; ROAR, 2019). Only a few researchers (Horwood, Sullivan,Young & Garner, 2004; Ashworth, Mackie & Nixon, 2004) suggested using a 'mediated deposit' system to assist the contributors in the submission process. Armbruster (2011) insisted that librarians and OA advocates archive on behalf of the authors. A survey of IDRs managers by Hanlon and Ramirez (2011) indicated that most IDRs follow a mediated deposit process. Another study (Carr &

Brody, 2007) stated that authors should upload documents directly rather than in massive submissions by technical staff. SHERPA (2007) emphasizes that the validity and authenticity of the content of submission (all materials submitted by the depositor, including full data and metadata) is the sole responsibility of the depositors, and the repository would not check it. Roy (2014) reported that only authorized users would submit objects to a particular collection where they are authorized to submit. In another work, author suggested that library staff assist the content contributors in the submission process (Roy, 2015). It was found after tallying ROARMAP data and experts' views that only 76 (17.3%) IDRs out of 439 (as of June 2013) stated that community members or their delegated agents may submit documents to the IDRs (Roy, 2014). Library staff, if required, may assist them in the submission process (Table 3). Presently the same database lists 5395 IDRs (as of August 2020), but the information is not available regarding this particular issue. Managing the retention period for an object submitted in IDR is essential to avoid any legal issue. Generally, many publishers – but certainly not all – stipulate an embargo period before an article can be made OA. It is found that 66% of publishers allow some form of self-archiving of the final version of an article but require an embargo period during which OA is not allowed (SHERPA/RoMEO, 2011). Roy (2014), in his research study, mentioned three types of institutional self-archiving policies (such as *Immediate Deposit with immediate Open Access; Later deposit after the embargo period* and, *immediate deposit with optional access*), and a repository can follow any one of them as per institutional preferences. In another study, the author concluded that institutions are therefore adopting type 3 policy (i.e., *immediate deposit with optional later access*) due to its popularity in the academic world (Roy, 2015). As stated above, the authors' views have been compared with ROARMAP database to show the ground realities. After analyzing ROARMAP database, it was found that only 68 (15.4%) IDRs out of 439 (as of June 2013), as reported by Roy in 2014, mentioned the retention period (Table 3). Almost 65 IDRs out of a total of 68 stated that items could be deposited at any time but would be made publicly visible after the embargo period has expired. Only 3 IDRs mentioned that items would be deposited only after publishers' or funders' embargo. The same database lists 5395 IDRs (as of August 2020), but the information is not available regarding this issue. Designing workflow steps for different types of objects in the submission process is a technical issue as it is associated with the software used by the IDR. The scholarly literature on this area considers the submission system one of the main barriers to greater repository adoption by researchers because it is too complex and time-consuming and seems to require some technical knowledge (Kim, 2010, 2011; Covey, 2011). Another group of experts (Foster & Gibbons, 2005; Lam & Chan, 2007) emphasized the availability of user-friendly pages for submission. Because there are several types of workflow in a typical repository depending upon the types of documents and software used (Hulse, Cheverie & Dygert, 2007; Campbell, Blinco & Mason, 2004; Nolan & Costanza, 2006). Generally, these include workflows to manage user registration and administration, authorization and permissions within the repository, and various administrative workflows to allow for maintenance and software updates. Madsen and Oleen (2013) described the staffing and workflow of a maturing IDR. Roy (2014) recommended following different workflow patterns at different submission processes/stages to avoid irrelevant items in the IDRs. In another study, the author reported that administrators, if required, could define or avoid workflow stages, but IDRs follow the same workflow as specified by the software (Roy, 2015). In another study, Roy, Biswas, and Mukhopadhyay (2016) emphasize E-persons' different roles at different

workflow stages. This issue can now be linked with the ROARMAP database to show the actual scenario covered in the literature. After reviewing 439 (as of June 2013) IDRs registered in the ROARMAP database, Roy (2014) reported that most repositories follow the workflow as included in the software used. Nevertheless, it is not clear whether repositories follow all steps or avoid any one of them. The database lists 5395 IDRs (as of August 2020), and no information regarding this issue is available in its present structure. Below are the explanations based on the following table (3) prepared.

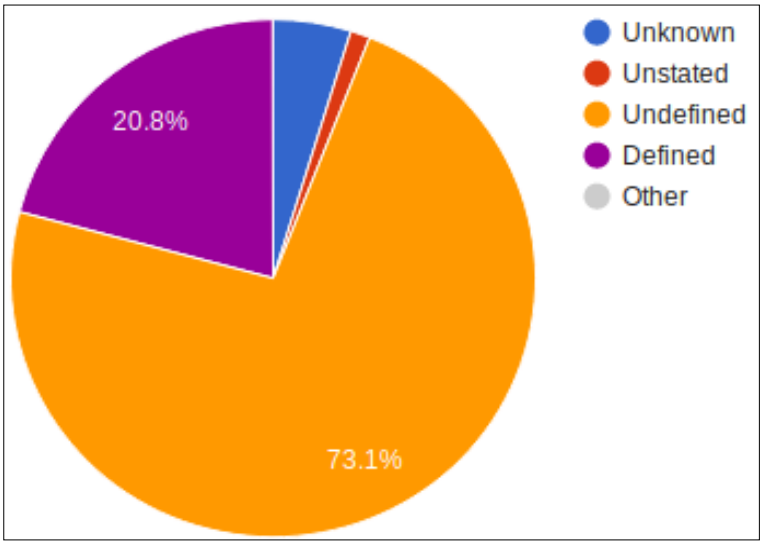

Figure 4: Submission Policy (Source: OpenDOAR, 2019)

Table 3
Submission Policies

Name of the Repository	Policy related to Submission		
	Eligible Depositors	Deposition Rules	Publishers' and Funders' Embargo
Anglia Ruskin Research Online	Accredited members of the organization or their delegated agents	NA	NA
Archive ouverte UNIGE	A	Authors may only submit their work for archiving	B
Arts and Humanities Research Council	A	√	B
ANU Research repository	A	NA	NA
Arizona State University Digital Repository	A	NA	NA
Aston University Research Archive	A	√	B
Bond University	A	NA	NA
Brigham Young University Library	A	NA	NA
Canadian Health Services Research Foundation	A	√	B

Name of the Repository	Policy related to Submission		
	Eligible Depositors	Deposition Rules	Publishers' and Funders' Embargo
Canadian Institutes of Health Research	A	√	B
California Digital Library (CDL)	A	NA	NA
California Institute of Technology	either personally or with the help of library staff	NA	NA
Canadian Cancer Society	A	√	B
Concordia University	A	√	B
Centre for Environmental Data Archival Repository	A	NA	B
Charles Stuart University Institutional Repository	A	NA	B
Covenant University Repository	A	√	B
Council of Scientific and Industrial Research	A	√	B
Curtin University	A	NA	B
Edith Cowan University	A	NA	B
European Heads of Research Councils	A	√	B
European Research Advisory Board	A	√	B
European Research Council	A	√	B
European University Association	A	√	B
Fonds de la recherche en sante Quebec	A	√	B
Fonds zur Foerderung der wissenschaftlichen Forschung	A	√	B
Genome Canada	A	√	B
Goddard Library Repository	A	NA	B
Griffith University	A	√	B
Heart and Stroke Foundation of Canada	A	√	B
HKUST Institutional Repository	Faculty, Staff & PhD students	√	C
Joint Information Systems Committee	A	√	B
Katholieke Universiteit Leuven	A	√	B
Khazar University	A	√	B
Kwame Nkrumah University of Science and Technology Institutional Repository (KNUSTSpace)	A		B
Leeds Metropolitan University institutional repository	A	NA	B

Name of the Repository	Policy related to Submission		
	Eligible Depositors	Deposition Rules	Publishers' and Funders' Embargo
Leicester Research Archive	A	√	B
Loughborough University	A	√	B
Michael Smith Foundation for Health Research	A	√	B
Monash University	A		B
National Research Council	A	√	B
Natural Environmental Research Council	A	NA	B
Natural Sciences and Engineering Research Council of Canada	A	√	B
Northern Melbourne Institute of TAFE	A	√	B
University of Nottingham	A	√	B
Ontario Institute for Cancer Research	A	√	B
Queensland University of Technology	A	NA	NA
Stanford University: School of Education	A	√	B
University of Strathclyde Institutional Repository (Strathprints)	A	√	B
Teesside University's Institutional Repository (TeesRep)	A	√	B
Trento University	A	√	B
Universidad Nacional de Colombia	A	√	B
University of Abertay Dundee	A	NA	B
University of Birmingham Research Archive	A	NA	B
University of Bath	A	NA	B
University of Calgary: Library and Cultural Resources	A	√	B
University of East Anglia	A	√	B
University of Edinburgh	A	NA	B
University of Florida Institutional Repository		NA	C
University of Helsinki		NA	B
University of Huddersfield	Staff, researcher or postgraduate student	NA	NA
University of Kansas	Faculty, staff along with researchers	NA	NA
University of Lincoln	A	√	B
University of Melbourne	A	NA	B

Name of the Repository	Policy related to Submission		
	Eligible Depositors	Deposition Rules	Publishers' and Funders' Embargo
Eprint Repository			
University of New South Wales	A	NA	B
University of Nottingham	A	NA	B
University of Queensland	A	NA	B
University of Pittsburgh	A	√	B
University of Reading	A	√	B
University of Salford	A	√	B
University of Starling (STORRE)	A	√	B
University of Surrey	A	√	C
University of Sydney	A	NA	NA
University of Virginia	A	√	B
University of Westminster	A	NA	B
Victoria University Institutional Repository	A	√	B
Wake Forest University Institutional Repository	A	NA	B
Warwick Research Archive Portal	A	NA	B
York St John University	A	NA	B

***[Legend:** A = Accredited members of the organization or their delegated agents; B = Items can be deposited at any time but will not be made publicly visible until publishers' or funders' embargo period has expired; C = Items may not be deposited until any publishers' or funders' embargo period has expired.*]

Multilingual Policies

Many OA advocates have recommended multilingual user interface of an IDR system to support access in vernacular languages. So, designing such a user interface for IDR is an important parameter covered by these policies. Language issues in digital libraries are multifarious (Borgman, 1997). One of the major responsibilities of multilingual digital libraries is to support multilingualism. So far, in the digital library sector, most research and development activities have concentrated on monolingual environments, and, in the large majority of the cases, the language employed has been English. For example, this issue is too important in India, where several dialects are used (Chatterjee, 1972; Prasad, 2003). India has 418 languages, of which 407 are living, and 11 are extinct (Maitra, 2002), and less than 5% of people can read and write English (Technology Development for Indian Languages Group, 2003). As languages diversify our country, only 5% to 10% of the population is aware of English and can read or write English. Over 90% to 95% population is normally deprived of the benefits of English-based information technology (Vikas, 2005). Loan and Rather (2007) found that digital repositories contain different types of content, and most are in the English language. Several other experts (Roy, Biswas & Mukhopadhyay, 2011, 2012, 2013) support this view. In the case of India, out of 79 IDRs, only 24 IDRs cover documents other than the English language, such as Sanskrit, Hindi, Arabic, Urdu, Gujarati, Malayalam, and Kannada.

Only a few researchers (Das, Das, Kar & Chatterjee, 2005; Mukhopadhyay & Azim, 2006; Mukhopadhyay, 2007) developed Bengali script-based digital library systems through the application of open standards and open-source software (OSS). In other papers, Mukhopadhyay (2006a, 2006b) has demonstrated the designing FLOSS-based software framework for public libraries and Web-enabled multilingual community information service. In another paper, the same author has demonstrated the methodology of designing a Bengali script-based interface of two popular software, e.g., DSpace and Greenstone (Mukhopadhyay, 2006c). In a recent study, he reported the development of a Unicode-compliant FLOSS-based framework for Indian IDRs (Mukhopadhyay, 2010). During 2011 and 2012, Mukhopadhyay (2011a, 2011b, 2012) reported a series of development related to Bengali-script-based digital community information systems (CIS). Chandrakar (2002) described 'Unicode' standard UTF-32 that uses four bytes (32 bits) to encode all possible (millions) characters. In another paper, he explained Unicode as a multilingual standard and the related technology available for localizing the Indian language materials (Chandrakar, 2004). Tripathi (2012) described problems and prospects of Hindi language search and text processing. Idicula and Peter (2005) developed an Open Agent Architecture (OAA) for multilingual information retrieval where a query for information retrieval could be given in plain Hindi or Malayalam. This system architecture is designed in a structured way to be adapted to other regional languages of India. Jasimudeen, Maghesh Rajan and Suresh Kumar (2012) examined the multilingual search capabilities of Mahatma Gandhi University's online theses archive using the first digital library solution with capabilities for information retrieval from full texts in Indian languages. The study suggested that the multilingual search features of the Mahatma Gandhi University archive would help the user with more relevant hits than in any similar systems used in India. Chavan and Mukhedkar (2012) tried to develop a model library automation multilingual package, which may be helpful to save on costs incurred, allow to enter details in local language, and source code will be made available to make modifications (if required) in it. Another study (Swapna, Hareen Kumar & Padmaja Rani, 2012) reported the development of a new information retrieval system in a multilingual environment and the functionality of processing and retrieval of Indian languages such as Hindi and Marathi, Telugu, Bengali, Malayalam, Kannada etc. Roy (2014), in his research study, developed an Indic-script-based multilingual information representation and retrieval (MIRR) system where a user may search and browse documents in the Bengali language. In another work (Roy, 2015), he opined that a Unicode-compliant multilingual information system is essential in India and concluded that the system could be developed in any other national or local language. Qin, Zhou, Chau and Chen (2006) investigated the feasibility of employing various cross-lingual information retrieval techniques for developing and evaluating a bilingual Web portal. Chen, Lin, Yang and Lin (2006) developed a framework for handling multilingual-named entities. In a recent study, Roy, Biswas and Mukhopadhyay (2016) reported the development of a Unicode-compliant Bengali script-based IDR system to support the integrated searching and browsing of Bengali language-based resources. Now, the opinion of experts is connected with the ROARMAP database to show the existent scenario. After reviewing 439 (as of June 2013) IDRs registered in the ROARMAP database, Roy (2014) reported that only a few IDRs have provision to incorporate multilingual documents into their system to support searching and retrieving multilingual objects. As per OpenDOAR database, about 70% IDRs cover English documents (OpenDOAR, 2019), as depicted in Figure 5. The database lists 5395 IDRs (as of August 2020) and no information regarding this issue is

available in its present structure.

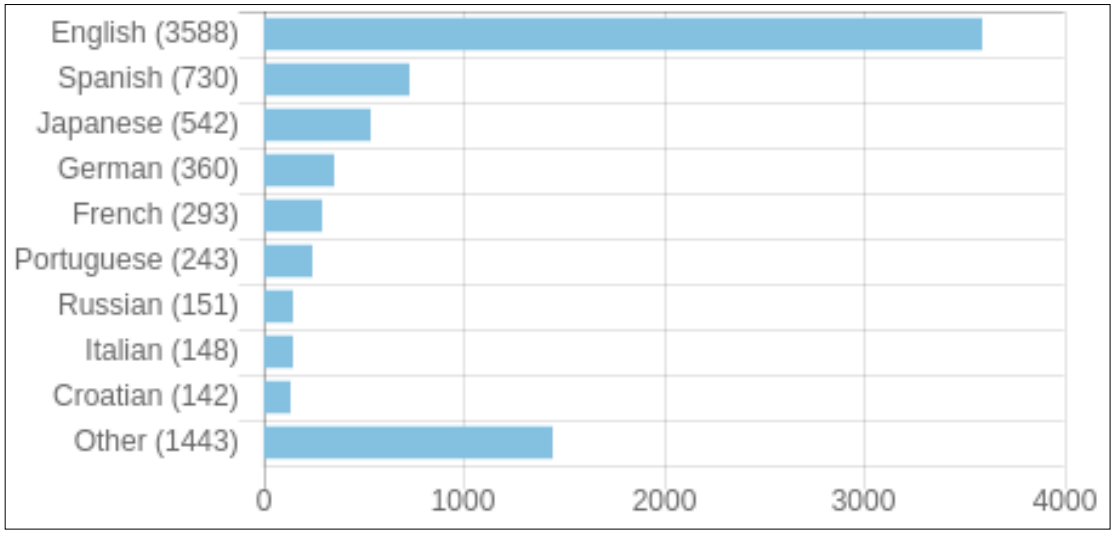

Figure 5: Repositories with most frequent languages (Source: OpenDOAR, 2020**)**

Conclusion

Based on the above discussion, which is based on the experts' view and assessment of the practical situation of ROARMAP database, the conclusion may be drawn to single out policy-wise respective guidelines against each policy discussed above. There are variations among the researcher's opinions regarding content type, which is also reflected in the ROARMAP database. However, most researchers agree that an IDR should hold journal articles, theses and dissertations, unpublished reports and working papers, conference and workshop papers, books, and chapters, as depicted in Figure 2. There is a lack of literature on collection organization policies focusing on using standard subject schemes or vocabulary control tools in organizing collections. No specific guideline regarding permission policies for different collections is available. So, it may be suggested that collections be organized department-wise along with proper permission with community rights to the administrators. In addition, adopting a standard controlled-vocabulary tool may increase efficiencies and ease of searching. Metadata policy needs extensive refinement as most issues are not covered in scholarly literature and the real situation as shown in the ROARMAP database is not satisfactory. The policy-related best practice guidelines are – metadata should be based on the recognized global standard; basic cataloging standards such as AACR or RDA are to be followed; Qualified version of the Dublin Core schema as well as domain-specific metadata elements may be used for describing particular objects such as for ETD (ETD-MS), Learning Objects (IEEE-LOM), Journal articles (Qualified DCMES); mediated services may be provided; IDR should support metadata extraction from OAI-PMH compatible repositories, and controlled vocabularies may be used to maintain consistency and to enhance the quality of records exposed to search and browse services.

Regarding submission policies, many issues like workflows in the submission process, embargo periods or retention period, suitable format for submission, deposition of standard metadata require re-engineering. A set of recommendations may be drawn based on existing literature and keeping in mind the ground realities. The suggestions are – authors may submit

their work in digital format, preferably in PDF format, standard metadata may be provided by library staff, submission workflow may be at least three steps, the repository should respect embargo or retention period for each object. In multilingual policies, it is expected to cover all major languages. The IDR system should be Unicode-compliant to incorporate documents in any local/regional language. In addition, it should support retrieval and processing of Web-based IDR resources in different languages and should have a mechanism to switch user interfaces from English to any national/regional languages and maybe vice versa.

It is well established that devising IDR using open source software (OSS) is not difficult, but formulating policy documentations is challenging for all repository managers. Some institutions have well-designed, structured, and instructive policy documentation, but most of the IDRs are in the pipeline or pilot stages of policy formulation. All these issues need further reworking against existing best practice guidelines because these may affect the repository services and impact policy formulation.

References

Armbruster, C. (2011). Open access policy implementation: first results compared. *Learned Publishing*, 24(4), 311–324. https://orcid.org/10.1087/20110409

Ashworth, S., Mackie, M. & Nixon, W. J. (2004). The DAEDALUS project, developing institutional repositories at Glasgow University: The story so far. *Library Review*, 53(5), 259-264. https://doi.org/10.1108/00242530410538391

Barton, M. R. & Walker, J. H. (2002). *MIT libraries' DSpace business plan project: Final report to the Andrew W. Mellon Foundation*. Retrieved from http://libraries.mit.edu/dspace-fed-test/implement/mellon.pdf

Borgman, C. L. (1997). Multi-media, multi-cultural, and multi-lingual digital libraries, or how do we exchange data in 400 languages? *D-Lib Magazine*, 3(6). Retrieved from http://www.dlib.org/dlib/june97/06borgman.html

Campbell, L. M., Blinco, K. & Mason, J. (2004). *Repository management and implementation: A white paper for alt-i-lab 2004* [White paper]. Retrieved from http://www.jisc.ac.uk/uploaded_documents/Altilab04-repositories.pdf

Carr, L. & Brody, T. (2007). Size isn't everything: sustainable repositories as evidenced by sustainable deposit profiles. *D-Lib Magazine*, 13(7/8). Retrieved from http://www.dlib.org/dlib/july07/carr/07carr.html

Chandrakar, R. (2002). Multi-script bibliographic database: an Indian perspective. *Online Information Review*, 26(4), 246-251. https://doi.org/10.1108/14684520210438705

Chandrakar, R. (2004). Unicode as a multilingual standard with reference to Indian languages. *The Electronic Library*, 22(5), 422- 424. https://doi.org/10.1108/02640470410561947

Chatterjee, S.K. (1972). Origin and development of the Bengali language (Reprint of 1926 ed.). London: Allen and Unwin.

Chavan, S. P. & Mukhedkar, M. V. (2012). Developing a model library automation multilingual software (LAMP – Library Automation Multilingual Package). *8th* Convention of PLANNER (pp. 32-38). Ahmedabad: INFLIBNET.

Chen, H.-H., Lin, W.-C., Yang, C., Lin, W.-H. (2006). Translating transliterating named entities for multilingual information access. *Journal of the American Society for Information Science and Technology*, 57(5), 645–659. https://doi.org/10.1002/asi.20327

Covey, D.T. (2011). Recruiting content for the institutional repository: the barriers exceed the

benefits. *Journal of Digital Information*, 12(3), 1-18. Retrieved from file:///C:/Users/drghane/AppData/Local/Temp/2068-Article%20Text-9257-1-10-20110418-1.pdf

Crow, R. (2002). *The case for institutional repositories: A SPARC position paper*. Washington, D. C: The Scholarly Publishing & Academic Resources Coalition.

Das, R., Das, B., Kar, S. & Chatterjee, S. (2005). Designing a digital library with Bengali language support using unicode. *3rd Convention of PLANNER* (pp. 83 – 88). Ahmedabad: INFLIBNET.

DINI. (2003). *DINI-Certificate for document and publication services*. Retrieved from http://nbn-resolving.de/urn:nbn:de:kobv:11-10075687

Foster, N. F. & Gibbons, S. (2005). Understanding faculty to improve content recruitment for institutional repositories. *D-Lib Magazine* 11(1). Retrieved from http://www.dlib.org/dlib/january05/foster/01foster.html

Genoni, P. (2004). Content in institutional repositories: a collection management issue. *Library Management*, 25(6/7), 300–306. https://doi.org/10.1108/01435120410547968

Gibbons, S. (2004b). Benefits of an institutional repository. *Library Technology Reports*, 40(4), 11-16. Retrieved from file:///C:/Users/drghane/AppData/Local/Temp/4376-5859-1-PB.pdf

Gibbons, S. (2004a). Establishing an institutional repository. *Library Technology Reports*, 40(4), 1-68.

Graaf, M. V. & Eijndhoven, K. V. (2008). *The European Repository Landscape*. Amsterdam: Amsterdam University Press.

Green, A., Macdonald, S. & Rice, R. (2009). *Policy making for research data in repositories: A guide*. Retrieved from http://www.disc-uk.org/docs/guide.pdf

Gonzalez, A. B. & Porcel, A. F. (2007). *Guidelines for the Creation of Institutional Repositories at Universities and Higher Education Organisations*. Valparaiso: ALFA Network Babel Library.

Hanlon, A. & Ramirez, M. (2011). Asking for permission: A survey of copyright workflows for institutional repositories. *Portal: Libraries and the Academy*, 11(2), 683-702. Retrieved from https://www.learntechlib.org/p/51573/

Horwood, L., Sullivan, S., Young, E. & Garner (2004). OAI compliant institutional repositories and the role of library staff. *Library Management*, 25(4/5), 170-176. https://doi.org/10.1108/01435120410533756

Hulse, B., Cheverie, J. F. & Dygert, C. T. (2007). ALADIN Research Commons: a consortial institutional repository. *OCLC System and Services: International digital library perspectives*, 23(2), 158-169. https://doi.org/10.1108/10650750710748469

Idicula, S. M. & Peter S, D. (2005).Intelligent Agent-based Multilingual Information Retrieval System. *Proceedings of 3rd International CALIBER* (February 2-4, 2005) (pp. 8-21). Ahmedabad: INFLIBNET Centre.

Innocenti, P., Smith, M., Ashley, K., Ross, S., De Robbio, A., Pfeiffenberger, H. & Faundeen, J. (2011). Towards a holistic approach to policy interoperability in digital libraries and digital repositories. *The International Journal of Digital Curation*, 1(6), 111-124. https://doi.org/10.2218/ijdc.v6i1.176

Jasimudeen, S., Maghesh Rajan, M. & Suresh Kumar, T V. (2012, January). Multilingual Searching in Digital Archives: A Case Study of Mahatma Gandhi University On-line

Theses Archive. Talk presented at the *Conference on Computing Paradigms & Bio-informatics*. School of Computer Sciences, Mahatma Gandhi University. Retrieved from http://eprints.rclis.org/19679/1/Search%20using%20Indic%20Script.pdf

Jayakanth, F. & Minj, F. (2012). Federated Search Service for OAI-compliant Open-Access Repositories in India. *Proceedings of the International Conference on Trends in Knowledge and Information Dynamics* (pp. 1-8). Retrieved from http://eprints.iisc.ernet.in/43176/1/fj-minj_drtc_ictk_final.pdf

Jones, C. (2007). *Institutional repositories: Content and culture in an open access environment*. Oxford: Chandos Publishing.

Kellogg, D. (2004). *Open Source OAI Metadata Harvesting Tools*. Retrieved from http://old.diglib.org/aquifer/oct2504/harvesting.pdf

Kim, J. (2010). Faculty self-archiving: motivations and barriers. *Journal of the American Society for Information Science and Technology*, 61(9), 1909-1922. https://doi.org/10.1002/asi.21336

Kim, J. (2011). Motivations of faculty self-archiving in institutional repositories. *Journal of Academic Librarianship*, 37(3), 246-254. https://doi.org/10.1016/J.ACALIB.2011.02.017

Lam, K.T., & Chan, D.L.H. (2007). Building an institutional repository: sharing experiences at the HKUST library. *OCLC Systems & Services*, 23(3), 310-323. https://doi.org/10.1108/10650750710776440

Loan, F. A. & Rather, R. (2007). Open access digital repositories: An Indian scenario. Know Genesis International Journal for Technical Communication, 2(1), 16-20. Retrieved from file:///C:/Users/drghane/AppData/Local/Temp/OpenAccess-India.pdf

Lone, F., Rather, R., & Shah, G. J. (2008). Indian Contribution to Open Access Literature: a case study of DOAJ & OpenDOAR. *Chinese Librarianship: an International Electronic journal*, 29. Retrieved from http://www.whiteclouds.com/iclc/cliej/cl26fayaz.pdf

Lynch, C. A. (2003). *Institutional repositories: Essential infrastructure for scholarship in the Digital Age*. Retrieved from http://www.arl.org/resources/pubs/br/br226/br226ir.shtml

Madsen, D.L., & Oleen, J.K. (2013). Staffing and Workflow of a Maturing Institutional Repository. *Journal of Librarianship and Scholarly Communication*, 1(3). eP1063. Retrieved from http://dx.doi.org/10.7710/2162-3309.1063

Maitra, D. (2002). *Languages and scripts of India*. Retrieved from http://www.cs.colostate.edu/~maitra/scripts.html

Millington, P. (2006, May). Moving Forward with the OpenDOAR Directory. Talk presented at the *8th International Conference on Current Research Information Systems.* Retrieved from http://www.opendoar.org/documents/BergenPresentation20060512Handouts.ppt

Mukhopadhyay, P. (2006a). Designing Web-enabled multilingual community information services: A Floss based framework for public libraries in West Bengal. *Community information service – challenges and opportunities for libraries*: Proceedings of the *National Seminar of Department of Library and Information Science, Banaras Hindu University* (pp.124-134). Varanasi: BHU.

Mukhopadhyay, P. (2006b). Public Library based Web-enabled Community Information System for Rural Development in India: Designing A Floss based Multilingual Prototype. *Open Source Movement – Asian Perspective:* Proceedings of the *22nd IASLIC National Seminar* (pp. 251-258). Kolkata: IASLIC.

Mukhopadhyay, P. (2006c). DSpace and GSDL in Bengali: A FLOSS based Methodology. Proceedings of the *National workshop on DSpace* (pp. 146–163).

Mukhopadhyay, P. (2007). Designing Bengali Script based Digital Library through the Application of Open Standards and Open Source Software Publication. In A.R.D Prasad & Devika P. Madalli (Eds.), *International Conference on Semantic Web and Digital Libraries* (pp. 381-396). Bangalore: DRTC.

Mukhopadhyay, P. (2010). Indic Scripts based Institutional Repositories: Designing Unicode-compliant FLOSS based Framework. *Proceedings of NACCS 2010 – National Conference on Computer Systems* (pp. 64–73).

Mukhopadhyay, P. (2011a). *Digital Community Information System: A Framework for India*. Germany: Lambert Academic Publishing.

Mukhopadhyay, P. (2011b). Multilingual Digital Community Information Services: Developing Course Curriculum and Technologies for a neglected area in LIS Education. *LIS Education Research and Training- Vision 2020.* Proceedings of *XXVIII IATLIS National Conference* (pp. 145–159). Punjab: IATLIS.

Mukhopadhyay, P. (2012). Interactive and multilingual digital community information services: towards the next level. I*nformation and Knowledge Dissemination: Present Status and Future Direction* (IKD, 2011) (pp. 57-68).

Mukhopadhyay, P., & Azim, A. (2006). Multiscript information retrieval system: A FLOSS based prototype for Indic-scripts with special reference to Bengali script. *Information management in digital libraries:* Proceedings of the *National Conference of Indian Institute of Technology, Kharagpur* (pp. 305- 316). Kolkata: Radical Impression.

Nolan, C. W., & Costanza, J. (2006). Promoting and Archiving Student Work through an Institutional Repository: Trinity University, LASR, and the Digital Commons. *Serial Review*, 32(2), 92-98. https://doi.org/10.1080/00987913.2006.10765038

OpenDOAR. (2019). *Directory of Open Access Repositories.* Retrieved from https://v2.sherpa.ac.uk/opendoar/

Pappalardo, K. M., Fitzgerald, A. M., Fitzgerald, B. F., Kiel-Chisholm, S. D., O'Brien, D. & Austin, A. (2007). *A Guide to Developing Open Access through Your Digital Repository.* Retrieved from https://eprints.qut.edu.au/9671/

Pinfield, S., Gardner, M. & MacColl, J. (2002). Setting up an institutional e-print archive. *Ariadne*, 31. Retrieved from http://www.ariadne.ac.uk/issue31/eprint-archives/

Prasad, A.R.D. (2003). Creation of digital libraries in Unicode using Indian languages. In A R D Prasad (Ed.), *Digital libraries: Theory and practice.* Bangalore: Documentation Research and Training Centre.

Probets, S. & Jenkins, C. (2006). Documentation for Institutional Repositories. *Learned Publishing*, 19(1), 57–71. https://doi.org/10.1087/095315106775122556

Prost, H., & Schöpfel, J. (2014). Degrees of Openness: Access Restrictions in Institutional Repositories. *D-Lib Magazine*, 20(7/8). Retrieved from http://www.dlib.org/dlib/july14/prost/07prost.html

Qin, J., Zhou, Y., Chau, M. & Chen, H. (2006). Multilingual Web retrieval: An experiment in English–Chinese business intelligence. *Journal of the American Society for Information Science and Technology*, 57(5), 671–683. https://doi.org/10.1002/asi.20329

RIN [Research Information Network]. (2008). *Stewardship of digital research data – a framework of principles and guidelines*. Retrieved from http://www.rin.ac.uk/data-

principles

ROAR. (2019). *Registry of Open Access Repositories.* Retrieved from http://roar.eprints.org/

ROARMAP. (2019). *Registry of Open Access Repositories Mandatory Archiving Policies.* Retrieved from http://roarmap.eprints.org/

Roy, B. K., Biswas, S. C. & Mukhopadhyay, P. (2011). An Analytical Study of Institutional Digital Repositories in India. *Library Philosophy and Practice.* Retrieved from http://digitalcommons.unl.edu/libphilprac/692

Roy, B. K., Biswas, S. C., & Mukhopadhyay, P. (2012). Open Access Repositories in Asia: From SAARC to Asian Tigers. *Library Philosophy and Practice.* Retrieved from http://digitalcommons.unl.edu/libphilprac/808

Roy, B. K., Biswas, S. C. & Mukhopadhyay, P. (2013). Global visibility of Indian Open Access Institutional Digital Repositories. *International Research: Journal of Library & Information Science*, 3(1), 182-194. Retrieved from http://irjlis.com/wp-content/uploads/2013/07/13_IR096.pdf

Roy, B. K. (2014). *Designing Institutional Digital Repository for the University of Burdwan: A FLOSS Based Prototype*. A PhD Thesis, Library and Information Science Department, The University of Burdwan, Burdwan.

Roy, B. K. (2015). *Institutional Digital Repository: From Policy to Practice*. Germany: LAP.

Roy, B. K., Biswas, S. C. & Mukhopadhyay, P. (2017). DDC in DSpace: Integration of Multi-lingual Subject Access System in Institutional Digital Repositories. *International Journal of Knowledge Content Development & Technology*, *7*(4), 71-84. http://ijkcdt.net/xml/12713/12713.pdf

Roy, B. K., Biswas, S. C. & Mukhopadhyay, P. (2016). The COAPI cats: the current state of open access repository movement and policy documentations. *International Journal of Knowledge Content Development & Technology*, *6*(1), 69-84. file:///C:/Users/drghane/AppData/Local/Temp/The_COAPI_Cats_The_Current_State_of_Open.pdf

Roy, B. K., Biswas, S. C. & Mukhopadhyay, P. (2016). Open access repositories for Indian universities: towards a multilingual framework. *IASLIC Bulletin*, 61(4), 150-161.

Roy, B. K., Biswas, S. C., & Mukhopadhyay, P. (2018). Towards An Open Access Policy Framework: A Case Study of COAR. *Liber Quarterly*, 28(1), 1-33. https://doi.org/10.18352/lq.10227

Sarkar, P., & Mukhopadhyay, P. (2010). Designing single-window search service for electronic theses and dissertations through metadata harvesting. *Annals of Library and Information Studies*, 57(4), 356-364. Retrieved from http://nopr.niscair.res.in/bitstream/123456789/11053/4/ALIS%2057%284%29%20356-364.pdf

Sarkar, P., & Mukhopadhyay, P. (2012). Developing virtual union catalogue of ETDs on health and medicine: a practical approach. Proceeding of the *National Seminar on Challenges in Library Management System* (pp. 334-340). Kolkata: Indian association for the cultivation of science.

Shearer, K. (2002). *A Guide to Setting-up an Institutional Repository.* Retrieved from http://www.carl-abrc.ca/projects/institutional_repositories/setup_guide-e.html

SHERPA. (2007). *OpenDOAR Policies Tool.* Retrieved from http://www.opendoar.org/tools/en/policies.php

SHERPA/RoMEO. (2011). *Home page of SHERPA/RoMEO.* Retrieved from http://www.sherpa.ac.uk/

Smith, K. (2008). Institutional Repositories and E-Journal Archiving: What Are We Learning? *Journal of Electronic Publishing,* 11(1). http://dx.doi.org/10.3998/3336451.0011.107

Swapna, N., Hareen Kumar, N., & Padmaja Rani, B. (2012). Information Retrieval in Indian Languages: A Case Study on Cross-Lingual and Multi-Lingual. *International Journal of Research in Computer and Communication technology*, 1(4), 139-142.

Technology Development for Indian Languages Group. (2003). *About Indian languages*. Retrieved from http://tdil.mit.gov.in/home.asp

Teli, S. (2015). Metadata Harvesting From Selected Institutional Digital Repositories in India: A Model to Build a Central Repository. I*nternational Journal of Innovative Research in Science, Engineering and Technology,* 4(4), 1935-1942. https://doi.org/10.15680/IJIRSET.2015.0404018

Tripathi, A. (2012). Problems and prospects of Hindi language search and text processing. *Annals of Library and Information Studies*, 59(4), 219-222. Retrieved from http://nopr.niscair.res.in/bitstream/123456789/15703/4/ALIS%2059(4)%20219-222.pdf

UGC [University Grants Commission], India. (2005). U*GC (Submission of Metadata and Full-text of Doctoral Theses in Electronic Format) Regulations.* Retrieved from www.ugc.ac.in/new_initiatives/etd_hb.pdf

Vikas, O. (2005, May). Multilingualism for cultural diversity and universal access in cyberspace: an Asian perspective. Talk presented at the *UNESCO WSIS Thematic Meeting*. UNESCO, Paris.

Yeates, R. (2003). Over the Horizon: Institutional Repositories. *VINE: The Journal of Information and Knowledge Management Systems,* 33(2), 96-99.