Open Access Repositories in the Field of Computers and Information Technology An Analytical Study of the OpenDOAR

Ikhlaqur Rehman[1] and Midhat Mehraj[2]

[1]Ph.D. Scholar, Department of Library and Information Science, University of Kashmir
[2]Ph.D. Scholar, Department of Library and Information Science, University of Kashmir
E-mail: [1]ak.edu05@gmail.com, [2]sheikhmidhat654@gmail.com

Abstract—*The present study aims to identify the status of Open Access Repositories (OARs) in the field of Computers and Information Technology (IT). The data was collected from the Directory of Open Access Repositories (OpenDOAR). Data collected was analysed on different parameters such as geographical distribution, software usage, content type, repository type, and language diversity. As on 15th May, 2019 OpenDOAR holds 198 repositories in the field of Computers and IT. The findings further reveal that the maximum number of repositories belong to Europe. Also, in terms of Country-wise distribution of repositories, UK holds the maximum, share.*

Keywords: *Open Access; Open Access Repositories; OpenDOAR; Digital repositories; Computers and IT repositories*

1. INTRODUCTION

Open Access (OA) publishing is a newer method of publishing that has been utilized to varying degrees by scholarly community since 1998. OA access offers free, unrestricted public access to scientific information [1]. OA makes research available to end user at no cost and opens storehouses of knowledge to public. The Budapest Open Access Initiative (BOAI) defines open access as "free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited" [2]. Two other additional definitions emerged through the October 2003 Berlin and June 2003 Bethesda statements of open access' definitions. Suber[3]is of the opinion that these three definitions "are the most central and influential to open access movement". OA has altered publishing methods and ways in which people now have access to research and costs associated with it [4]. OA accelerates research, enhances education and shares learning across rich and poor nations. The two routes to OA are OA journals and OA repositories. These two routes are also known as "Gold" and "green" routes respectively [5]. Golden OA makes the final version of an article freely and permanently accessible to everyone, immediately after publication. It involves authors publishing in a fully OA journal (where all the content is published OA) or hybrid journals (a subscription-based journal that offers an OA option which authors can chose if they wish). Also, the copyright sits with the authors of articles and most of the permission barriers are removed. Green OA, also referred to as self-archiving, involves placing author's manuscript into a repository, making it freely accessible to everyone. Copyright usually sits with publisher, of, or the society affiliated with, the title and there are restrictions as to how the work can be reused [6]. OA repositories have bought a paradigm shift in scholarly communication worldwide. Pinfield [5]defines repository as "A repository may be defined as a set of systems and services which facilitates the ingest, storage, management, retrieval, display, and reuse of digital objects. Repositories may be set up by institutions, subject communities, research funders, or other groups. They may provide access to a variety of digital objects, including peer-reviewed journal articles, book chapters, theses, datasets, learning objects, or rich media files". In 2005, Directory of Open Access Journals (OpenDOAR) was launched as a result of collaboration between University of Nottingham and Lund University. It is an authoritative global directory of open access repositories. It enables the identification, browsing and search for repositories, based on a range of features, such as location, software or type ofmaterial held (http://v2.sherpa.ac.uk/opendoar/information.html).

2. LITERATURE REVIEW

A number of studies have been carried that highlight the significance of open access repositoriesCiting the benefits of

Open Access (OA), Corrade[7] states that OA includes lower costs, greater accessibility and better prospects for long term preservation. Falk (2003) states that growing trend in digital documents. Evaluating the role of OA Institutional repositories, Chan[8]states that institutional repositories encourage more timely and open access to research and scholarship, and that they maximize the potential research impact of archived publications. According to Pinfield[9], OA repositories and Institutional repositories in particular play an important role in scholarly communication.Pinfield, et al., [10] studied the development of the global repository structure from 2005-2012 and found that the initial repository development was focused on North America, Western Europe and Australasia. Followed by East Asia, South America and Eastern Europe. However, areas such as Africa and Central Asia, and countries such as China and Russia, have experienced relatively low levels of growth throughout the period. Joint [11], states that the repositories were seen as a way of providing open access to research without supplying valuable academic intellectual property to the private sector who would sell it back to the research community at a considerable profit margin.Banks [12] argues that OARs represent an exciting possibility for both the preservation and retrieval of grey literatureAccording to Cullen and Chawner [13], institutional repositories have become popular among librarians, professionals, academics, scholars and readers for the communication and awareness of their research results. Falk [14], states that growing tide of digital documents created by faculty and students, and also because of librarians' discontent towards rising prices and practices in traditional journal publishing has led to the growing interest in the creation of digital archives in colleges and universities at national and international governmental institutions.

3.   OBJECTIVES OF THE STUDY

- To explore the geographical distribution of OARs.
- To identify the various types of OARs.
- To examine different types of software used for the creation of OARs.
- To identify the types of OARs.
- To examine the content type in OARs
- To study the language diversity in OARs

4.   METHODOLOGY AND SCOPE

To accomplish the current study, data was collected from OpenDOAR on May 15th, 2019. A total of *198* repositories were identified in the field of Computers and Information Technology (IT). These repositories were thoroughly examined to accomplish the stated objectives. The scope of the study is limited to the OARs in the field of Computers and IT available on the OpenDOAR.

5.   DATA ANALYSIS AND INTERPRETATION

5.1Geographical distribution of repositories

As evident from Table-1, Europe holds the maximum share in terms of repository distribution on OpenDOAR in the field of Computers and IT accounting for *113*(57%) repositories followed by, Asia, North America and Africa with *32*(16%),*27*(14%) and *13*(7%) repositories respectively. South America and Oceania account for the least with *11*(5%) and *2*(1%) repositories respectively.

Table-1 Continent-wise distribution of repositories		
Europe	113	57
Asia	32	16
North America	27	14
Africa	13	7
South America	11	5
Oceania	2	1
Total	**198**	**100**

5.2 Country-wise distribution of repositories

Table-2 highlights the country-wise distribution of repositories on OpenDOAR in the field of Computers and IT. The UK emerges out to be the leading contributor with *31*(15.6%) repositories followed by USA and France with *23*(11.6%) and *15*(7.5%) repositories respectively. Germany and India account for *11*(4.08%) and 8(4%) repositories respectively while Croatia and Indonesia account for 7(3.54%) repositories each. Table 2 provides further details regarding the country-wise distribution of repositories

Table 2: Country-wise distribution of repositories	
Country	**Repositories**
UK	*31*
USA	*23*
France	*15*
Germany	*11*
India	*8*
Croatia, Indonesia	*7*
Poland	*6*
Brazil, Switzerland, Ukraine	*5*
Austria, China, Greece, Ireland, Portugal	*4*
Denmark, Algeria, Spain, Peru	*3*
Others	*43*
Total	**198**

5.3 Software usage by OARs

Knowledge institutions make use of various open source as well as commercial software to create their repositories and share their knowledge stock globally. Table 3 highlights various software used by institutions for the creation of their

repositories. DSpace turns out to be the most preferred software among institutions for the creation of their repositories.*69*(34.8%) repositories use DSpace followed by Eprints and Hal accounting for the creation of *58*(29.3%) and *12*(6%) repositories respectively. *6*(3%) repositories use islandoara while 4(2%) use Invenio. Pure and Opus are used in 3(1.5%) repositories each. Some of the software names couldn't be identified and thus were put under "unspecified" category. Table 3 provides the complete picture of the software usage by OARs.

Table 3: Software usage by repositories		
Software	**Repositories**	**%**
Dspace	69	34.85
Eprints	58	29.29
Hal	12	6.06
islandora	6	3.03
Invenio	4	2.02
Pure	3	1.52
Opus	3	1.52
Greenstone	1	0.51
CONTENTdm	1	0.51
mycore	1	0.51
Weko	1	0.51
Fedora	1	0.51
Unspecified	8	4.04
Others	30	15.15
Total	**198**	**100**

5.4 Type of repository

OA repositories have been categorized into four types based on the nature of their host organization i.e. Institutional, Disciplinary, Aggregating and Governmental. As evident from table- 4 majority of the repositories are Institutional (created, hosted and maintained by an institution or department) accounting for *170*(86%) followed by Disciplinary (subject) with *22* (11%) repositories. Aggregating (a repository aggregating data from several subsidiary repositories) and Governmental repositories account for the least with 4(2%) and *2*(1%) respectively.

Table-4 Repository type		
Type	**Number**	**Percentage**
Institutional	*170*	86
Disciplinary	*22*	11
Aggregating	*4*	2
Governmental	*2*	1
Total	**198**	**100**

5.5 Content Type

Figure 3 highlights different content types incorporated by institutions into their repositories. There are 12 content types archived by these repositories. Most prominent among them are *Journal articles, Thesis and Dissertations and Conference and Workshop papers*, that are archived by *147*, *118* and *108* repositories respectively. The least content type archived are *Software anddatasets* accounting for *6 and 4* repositories respectively. Table-5 shows the range of content types archived by Computer and IT repositories on the OpenDOAR.

Table-5 Content type[1]	
Content	**Repositories**
Journal Articles	*147*
Thesis and dissertations	*118*
Conference and workshop papers	*108*
Unpublished reports and working papers	*90*
Books, chapters & sections	*77*
Multimedia & audio video	*39*
Learning objects	*38*
Bibliographic references	*37*
Other special items	*23*
Patents	*16*
Software	*6*
Datasets	*4*

5.6 Language diversity of repositories

As far as the language diversity of repositories in the field of Computers and IT on the OpenDOAR is concerned, English is the most widely used language. *148* repositories use English as the only language or one of the languages followed by French and German used in *20* and *12* repositories respectively. In addition, Spanish and Portuguese are used by *8* repositories each while Croatian is used in *7* repositories. Table 6provides further details regarding the language interface of these repositories

Table 6: Language diversity[2]	
Language	**Number**
English	148
French	20
German	12
Portuguese, Spanish	8
Croatian	7
Russian, Indonesian, Chinese	6
Ukrainian	5

[1]Since, the majority of repositories hold multiple content types so, the no. of repositories for content type exceeds the actual no. of repositories.
[2]Since, repositories develop interface in multiple languages so, the no. of repositories with multiple language interfaces exceeds the actual no. of repositories

Arabic, Danish, Greek(modern), Hungarian, Korean, Polish	2
Finish, Irish, Bulgarian, Czech, Malay, Japanese, Slovene, Turkish, Latin	1

6. CONCLUSION

Open Access is a new shift in scholarly communication that aims at providing free access to scholarly literature over the internet and has attracted a lot of attention in recent years. The study concludes that majority of the repositories are contributed by Europe. Also, in terms of country-wise contribution, developed nations hold the lion's share. This may be attributed to the fact that developed nations have realized the need and importance of knowledge dissemination that helps in development of real knowledge-based nations. So, developing countries need to be sensitized about the importance of OA repositories especially in present knowledge-based society. In terms of repository type, institutional repositories account for the maximum share that can be attributed to the fact that academic and research institutions are enthusiastic to disseminate research findings and information products for public interest and the benefit of co-researchers and professionals. As far as language diversity of repositories is concerned, English is the widely used language. Since, English language has the highest number of speakers in the world, the popularity is obvious and therefore used by repositories world over as their main language of interface.

REFERENCES

[1] T. DeAngelis. (2004). *Debating access to scientific data.* [online] Available at: https://www.apa.org/monitor/feb04/debating

[2] *Budapest Open Access Initiative* (2002). [online]. Availableat:

http://www.soros.org/openaccess/read.shtml

[3] P. Suber. (2010). *Open Access Overview.* [online]. Available at:https://legacy.earlham.edu/~peters/fos/overview.htm

[4] E.E. Kirk (2010, March 7). *Open-access publishing: 'An attractive alternative.* [online]. Available at : http://chronicle.com.ezproxy1.lib.asu.edu/article/Open-Access-PublishingAn/64517/

[5] S. Pinfield. "Journals and repositories: an evolving relationship?" *Learned Publishing*, *22*(3), 165- 175. 2009

[6] *What is open access.* (2019). [online]. Available at: https://www.springer.com/gp/authorseditors/authorandreviewertutorials/open-access/what-is-open-access/10286522

[7] E. M. Corrado. (2005). *The importance of open access, open source, and open standards for libraries.* [online]. Available at: http://www.istl.org/05-spring/article2.html

[8] L. Chan. "Supporting and enhancing scholarship in the digital age: The role of open access institutional repositories." *Canadian Journal of Communication*, *29*, 277-300. 2004

[9] S. Pinfield. (2005). *A mandate to self-archive? The role of open access institutional repositories.* [online]. Available at: http://eprints.nottingham.ac.uk/152/

[10] S. Pinfield, J. Salter, P.A. Bath, B. Hubbard, P. Millington, J. H. Anders, et.al., "Open☐access repositories worldwide, 2005–2012: Past growth, current characteristics, and future possibilities." *Journal of the association for information science and technology*, *65*(12), 2404-2421. 2014

[11] N. Joint. "Current research information systems, open access repositories andlibraries: ANTAEUS." *Library review*, *57*(8), 570-575. 2008

[12] M. A. Banks. "Towards a continuum of scholarship: the eventual collapse of the distinction between grey and non-grey literature." *Publishing research quarterly*, *22*(1), 4-11. 2004

[13] R. Cullen, B. Chawner. "Institutional repositories, open access, and scholarly communication: a study of conflicting paradigms." *The Journal of Academic Librarianship*, *37*(6), 460-470. 2011

[14] H. Falk. "Digital archive developments." *The Electronic Library*, *21*(4), 375-379.2003