Open Access Repositories:

A Global Perspective with an Emphasis on Asia

Zahid Ashraf Wani
University of Kashmir
India
zahidrais@gmail.com

Sumeer Gul
University of Kashmir
India

Javeed Ahmad Rah
University of Kashmir
India

ABSTRACT: This paper throws light on the growth and development of open repositories registered with OpenDOAR database. The study explores various facets of open repositories and tries to present a lucid picture of their overall development. The study provides a detailed description of repositories in terms of continent, country and subject wise distribution. The study further delves deep into the Asian contributions and brings to light detailed profiles of Asia.

I. Introduction

Repositories and their relationship with the Open Access (OA) movement constitute a new trend in scholarly communication worldwide. The need for even wider access to scientific data with the objective of democratizing the dissemination of research results and the cost, especially, of scientific journals, coupled with the reduced library budgets, has given rise to a strong movement that aims at free online access to digital materials for all. Universities and research institutions throughout the world are investigating, piloting, and developing systems for building collections of digital resources and learning materials in the shape of open access repositories (Bell & Rothery, 2007). An Open Access repository is more than a simple document storage, as it uses metadata to enable the users to find suitable materials. Many universities and colleges world over have initiated projects to develop repositories that will enable faculty and researchers to upload and download scholarly literature and use them to share resources with each other either within the institution or across the region, or more widely still. Sharing materials in this way may lead to an improved quality of teaching and research, the sharing of good practice, greater consistency, and an enhanced sense of community (Barker, James, & Knight, 2004).

II. Literature Review

Jacso (2006) believes that Open Access repositories (OARs) are beneficial for all the stakeholders, including publishers, editors and authors as they can substantially increase their impact and the impact factor for the source journals. Falk (2003) observes that widespread dissatisfaction among librarians with monopolistic pricing and practices in traditional journal publishing led to the development of OARs. Johnsen (2002) states that OARs are a practical, cost effective, and strategic means for institutions to build partnerships with their faculty to advance scholarly communication. Chan (2004) opines that OARs facilitate more timely and open access to research and scholarship and maximize the potential research impact of archived publications. According to Chan and Kursop (2005), OARs provide an exciting opportunity for scientific community in developing countries to make their research output public, as they can take advantage of servers anywhere in the world. Banks (2006) argues that OARs represent an exciting possibility for both the preservation and retrieval of grey literature.

Lynch & Lippincott (2005), Bravo & Diez (2007), Fang & Zhu (2006), Das, Sen, & Dutta (2005), Fernandez (2006), and Arunachalam (2008) have all discussed the deployment of OARs in various countries.

Lynch & Lippincott (2005) shared their findings of a survey in USA which reveals that 40% of respondent institutes have some type of OARs operating and that 88% of those who do no yet have a repository have planning work underway for OARs or participation in some from of a consortial repository system. In another survey, Reih, Markey, Jean, Yakel, and Kim (2007) discovered that the majority of institutions where OARs have been implemented in USA are research universities. Furthermore, the institutions which are in NP (No Planning to date) and PO (Planning only) stages are master's or bachelor's colleges and universities. While discussing the establishment of a number of OARs under the South African Research Information Services (SARIS) project, Deventer and Pienaar (2008) predicted a positive future for OARs in South Africa. Bravo and Diez (2007) found that there is little development of OARs in Spain, although a growing tendency to create them is detected. A study by Fang and Zhu (2006) discovered that Open Access is not widespread in China. The barriers are various and derived from the publishers, users, and governmental policies. Das, Sen, and Dutta (2005) found that problems of existing repositories in India are mainly due to limited availability of web servers running OARs. Fernandez (2006) noted that universities in India are at present lacking in infrastructure for establishing OARs. Arunachalam (2008) stressed the need for an OA mandate by various research organizations in India for their own research output and of projects funded by them. Chan and Costa (2005) emphasized that various stake holders in developing countries have to work in concert and promote a culture of Open Access that is not yet pervasive in most of these countries.

III. Research Objectives

Due to the revolutionary development of information and communication technologies (ICT), management of various types of electronic resources has become a popular mode of information dissemination. Nowadays, creation of OARs has been initiated by institutions, research centers, libraries, and government departments for free and speedy sharing of resources. This study is an attempt to understand the distribution and magnitude of OARs in various perspectives with the following objectives:

- To explore the geographic contributions to OARs.
- To identify the subjects archived by OARs.
- To determine language diversity in OARs.
- To identify the core content types in OARs.
- To explore the operational status of OARs.
- To determine various software used for creation of OARs.
- To identify the various types of OARs.
- To determine Asian contributions along the above dimensions.

IV. Scope

The scope of the study has been confined to the "Directory of Open Access Repositories," popularly known as *Open*DOAR (http://www.opendoar.org).

V. Methodology

In order to achieve the above stated objectives, *Open*DOAR (a database that lists open access repositories from all over the globe) was selected as the source directory for tracing the development of OARs from different perspectives. A total of 1,250 repositories were identified from OpenDOAR between October 7 and 8, 2008. All the repositories identified there were thoroughly sifted and analyzed for the necessary information as per the objectives of the study.

VI. Limitations

The *Open*DOAR service provides a quality-assured listing of open access repositories around the world. *Open*DOAR maintains a comprehensive and authoritative list of institutional and subject-based repositories. *Open*DOAR is, therefore, primarily a service to enhance and support the academic and research activities of the global community. *Open*DOAR has opted to collect and provide information solely on sites that wholly embrace the concept of open access to full text resources. Thus, sites where any form of access control prevents immediate access are not included. Likewise, sites that consist of metadata records only are not included.

VII. Findings and Discussion

1. Geographical Distribution of Repositories

All the continents are maintaining OARs, but major share holders are in Europe, which contribute 599 repositories (47.92%), followed by North America with 366 repositories (29.28%). Asia emerges as the third largest contributor with 138 repositories (11.04%). While Australia maintains 73 repositories (5.84%) and South America 55 repositories (4.40%) respectively, the smallest number of repositories belongs to Africa with only 19 repositories (1.52%). The decrease in the number of repositories correlating to the Eastward move is expected as the West mostly comprises developed countries that are already way ahead in terms of technology.

Table 1: Distribution of Repositories by Continent

Continent	No. of Repositories	%
Europe	599	47.92
North America	366	29.28
Asia	138	11.04
Australia	73	5.84
South America	55	4.40
Africa	19	1.52
Total	**1,250**	**100**

2. Distribution of Repositories by Country

In the continental distribution we have seen that the major contributions are from Europe and North America. In the distribution by country, it is natural that countries from Europe and North America will be leading contributors. It was found that the USA maintains the highest number of repositories at 317 (25.36%), followed by the UK and Germany with 136 (10.88%) and 129 (10.32%) respectively. In Asia, Japan leads the continent with 69 repositories (5.52%) and ranks the fourth at the global level. The fifth position is claimed by Australia with 68 repositories (5.44%). Countries like the Netherlands, Canada, and Italy also make significant contributions. There are another 68 countries which contribute from 1 to 30 repositories each and account for 400 repositories together. For a better understanding and a clear view, see Table 2.

Table 2: Distribution of Repositories by Country

Country	No. of Repositories	%
USA	317	25.36
UK	136	10.88
Germany	129	10.32
Japan	69	5.52
Australia	68	5.44
Netherlands	45	3.60
Canada	44	3.52
Italy	42	3.36
68 other countries	400	32.00
TOTAL	**1,250**	**100**

3. Core Content Types

Traditional archives and repositories preserve various types of materials depending upon the type and policy of the repository, and the same is true about e-repositories. OARs organize contents by type such as articles, theses, multimedia software, etc. *Open*DOAR provides a comprehensive insight of content types of all repositories. The authors found that there are a large number of OARs that include more than one "content type." In order to provide a comprehensive picture, the current study has attempted to show in detail the various content types being deposited in different repositories. The majority of repositories hold "journal articles" (757), followed by "theses and dissertations" (623) and "unpublished reports and working papers" (599). The content types which were seen least are "software" and "patents" with 27 and 21 repositories preserving them respectively. Table 3 provides a complete list of "content types".

Table 3: Content Types

Rank	Content Type	No.
1	Journal articles	757
2	Theses and dissertations	623
3	Unpublished reports and working papers	599
4	Conference and workshop papers	451
5	Books, chapters and sections	388
6	Multimedia and audio visual materials	298
7	Other special item types	212
8	Learning objects	184
9	Bibliographic references	183
10	Datasets	65
11	Softwares	27
12	Patents	21

4. Operational Status

It is well known that ICT created various new communication channels for sharing and pooling of knowledge. At the same time, it has posed various challenges. One of them is the inconsistency of resources available on the internet. Hence, this study deemed it necessary to analyze the operational status of OARs listed in *Open*DOAR. It was found that out of the 1,250 repositories registered with *Open*DOAR, 1,135 (90.80%) are operational, 91 (7.28%) are on trial, and 16 (1.28%) have been closed (not accepting depositions) while 8 (0.64%) repositories have been declared broken, i.e., technically malfunctioning.

The findings clearly reveal that repositories are not a crazy wave of ICT. Rather, they have been developed and maintained with missionary zeal and will show an upward trend with arithmetical progression from here onward.

Table 4: Operational Status of Repositories

Status	Number	%
Operational	1,135	90.80
Trial	91	7.28
Closed	16	1.28
Broken	8	0.64
TOTAL	**1,250**	**100**

5. Software Usage

There are quite a few open source software available on the internet with a range of features and capabilities that could be used for repositories. DSpace turns out to be the most popular software among OARs, with 345 repositories (27.60%) using DSpace. Other software have not been identified by *Open*DOAR and, therefore, put under the category of "Unknown" by 286 repositories (22.88%). EPrints is being used by 239 repositories (19.12%) and stands third among the most popular software. BePress is used by 55, Opus 51, HTML 24, Wildfire 23, ETD-db 21, and 59 other software are used by 206 repositories. A complete picture can be seen in Table 5.

Table 5: Software Usage of Repositories

Software	Number	%
Dspace	345	27.60
Unknown	286	22.88
Eprints	239	19.12
Bepress	55	4.40
Opus	51	4.08
HTML	24	1.92
Wildfire	23	1.84
ETD-db	21	1.68
59 others	206	16.48
Total	**1,250**	**100**

6. Repository Type

There are different types of repositories available in open access mode. This study has categorized them by type. Out of a total of 1,250 repositories, 1,001 (80.08%) are institutional, 166 (13.28%) are disciplinary, 58 (4.64%) are aggregated (a repository aggregating data from several subsidiary repositories), and 25 (2%) are governmental in nature. See Table 6.

The reason why institutional repositories take the lion's share could be that academic and research institutions are more eager to disperse research findings and information products for publicity and the benefit of co-researchers and professionals.

Table 6: Types of Repositories

Repository type	Number	%
Institutional	1,001	80.08
Disciplinary	166	13.28
Aggregating	58	4.64
Governmental	25	2
TOTAL	**1,250**	**100**

7. Subjects Archived by Repositories

There is a range of subjects archived by repositories worldwide. It was found that multidisciplinary repositories are highest in number (698), followed by history and archeology (92), health and medicine (90), science general (83), and then computers and IT (81). A few repositories are organized by subject such as psychology (20), architecture (12) and civil engineering (11). What is worth noting here is that in numerous cases one repository hosts more than one subject. Table 7 gives details regarding other subjects dealt with by repositories registered with *Open*DOAR.

Besides Science, Medicine and Technology (SMT), history and archeology have also taken a lead. This reveals that scholars in those fields are conscious of the technological revolution and keep pace with the time. In coming years, we expect to see a greater number of archives in the disciplines of social sciences and humanities.

Table 7: Subjects Archived by Repositories

Rank	Subject	Number
1	Multidisciplinary	698
2	History and archeology	92
3	Health and medicine	90
4	Science General	83
5	Computers and IT	81
6	Technology general	77
7	Social sciences general	72
8	Business and economics	65
9	Geography and regional studies	61
10	Law and politics	59
11	Biology and biochemistry	58
12	Library and information science	52
13	Education	48
14	Physics and astronomy	45
15	Ecology and environment	42
16	Mathematics and statistics	41
17	Fine and performing arts	39
18	Arts and humanities general	32
19	Chemistry and clinical technology	31
20	Language and literature	30
21	Mechanical engineering and materials	29
22	Philosophy and religion	28
23	Earth and Planetary sciences	26
24	Agriculture, Food Veterinary	24
24	Management and planning	24
25	Electrical and electronic engineering	20
25	Psychology	20
26	Architecture	12
27	Civil engineering	11
	Total	**1,250**

8. Language Diversity

It is observed that a large number of repositories use more than one language as an interface. English has been used by 1,060 repositories, followed by German, French, and Spanish with 151, 82, and 81 repositories respectively. There are 21 languages which are represented by 1 to 3 repositories. Even Urdu has been used by one of the repositories. Table 8 displays the various languages used by repositories.

Popularity of English as an international language is obvious and therefore used by repositories world over as the main language of the interface. But other regional, national, and sub-national languages are coming up strongly as the main or one of the interface languages, which is good sign to overcome the language barrier at regional, national or sub-national level.

Table 8: Language-Diversity of Repositories

Rank	Language	Number	Rank	Language	Number
1	English	1060	**18**	Arabic	3
2	German	151	**18**	Croatian	3
3	French	82	**19**	Afrikaans	2
4	Spanish	81	**19**	Czech	2
5	Japanese	70	**19**	Estonian	2
6	Dutch	45	**19**	Hebrew	2
7	Italian	40	**19**	Icelandic	2
8	Swedish	33	**19**	Lithuanian	2
9	Portuguese	29	**19**	Maori	2
10	Chinese	16	**19**	Pashto	2
11	Finnish	14	**19**	Persian	2
12	Polish	12	**19**	Welsh	2
13	Norwegian	10	**20**	Azerbaijani	1
13	Russian	10	**20**	Bulgarian	1
14	Greek	9	**20**	Georgian	1
15	Catalan	8	**20**	Kannada	1
15	Ukrainian	8	**20**	Malay	1
16	Danish	6	**20**	Romanian	1
16	Hungarian	6	**20**	Serbian	1
16	Korean	6	**20**	Slovenian	1
17	Hindi	4	**20**	Urdu	1
17	Latin	4	**20**	Yiddish	1
17	Turkish	4			

9. Asian Scenario

Asia is not only the biggest land mass on earth but also the largest populated continent with most diversified cultures and ethnic groups.

10. Type of Repositories and Use of Software in Asia

As already noted, Asia contributes 138 open access repositories to the *Open*DOAR directory. In Asia, the biggest contributor is Japan, with 69 repositories, followed by India with 30, while the rest of the Asian countries maintain between 1 and 6 repositories. Details can be seen in Table 9.

Out of a total of 138 repositories, 121 of them are institutional repositories, 9 are disciplinary, 6 are aggregators, and 2 are governmental repositories.

Moreover, DSpace (95) again turns out to be the leading software used by repositories in Asia, followed by EPrints (15), while XooNIps (a Japanese e-repository software) is used by 5 repositories. Table 9 provides a detailed picture of country, type of repository, and software used.

Table 9: Type of Repository and Software Used in Asian Countries

Country	No. of Repositories	Type of Repository				Software Used							
		Institutional Repositories	Disciplinary	Aggregators	Governmental	Dspace	Eprints	XooNIps	BEPRESS	green stone	Mikrobeta	InfoLib	Unknown
Afghanistan	1		1							1			
Azerbaijan	1	1				1							
Bangladesh	1		1			1							
China	6	5	1			5							1
Georgia	1			1									1
India	30	26	4			21	9						
Indonesia	2	2											2
Israel	2	1		1			1						1
Japan	69	65	1	3		53	2	5	1			1	7
Kazakhstan	1			1						1			
Korea	3	2	1			2							1
Kyrgyzstan	1	1				1							
Malaysia	4	4				3				1			
Pakistan	2				2		1			1			
Philippines	1	1				1							
Saudi Arabia	2	2					1						1
Singapore	1	1				1							
Taiwan	6	6				6							
Turkey	4	4					1				1		2
Total	**138**	**121**	**9**	**6**	**2**	**95**	**15**	**5**	**1**	**4**	**1**	**1**	**16**

11. Subjects Archived (Asia)

There are various subjects archived by repositories in Asia. Subjects like "mathematics and statistics" and "health and medicine" have been organized by 6 repositories each, while "technology general" is organized by 5 of them. Subjects like "science general" and "physics and astronomy" are maintained by 4 repositories. There are 103 repositories that have shown a multidisciplinary approach. An analysis is provided in Table 10, showing subjects archived by repositories in Asia.

Table 10: Subjects Archived by Repositories in Asia

SUBJECTS	Number
Multidisciplinary	103
Mathematics and Statistics	6
Health and Medicine	6
Technology General	5
Science General	4
Physics and Astronomy	4
Chemistry and Chemical Technology	3
Computers And IT	3
Library and Information Science	3
History and Archaeology	3
Business and Economics	2
Biology and Biochemistry	2
Ecology and Environment	2
Mechanical Engineering and Materials	1
Law & Politics	1
Agriculture, Food and Veterinary	1

12. Content Types (Asia)

There are 12 content types organized by e-repositories in Asia. Most prominent among them are articles, theses, and unpublished manuscripts, which are maintained by 108, 64 and 59 repositories respectively. Content types that are being organized least are datasets (3) and software (2). Table 11 shows the range of content types in Asian repositories.

Table 11: Content Types organized by Repositories in Asia

Content Type	Number
Articles	108
Theses	64
Unpublished	59
Conferences	48
Special	24
Books	23
Conferences	22
Learning Objects	19
Patents	7
References	5
Datasets	3
Softwares	2

13. Language Diversity (Asia)

Asia does not provide a different view as far as popular language in e-repositories is concerned. English is widely used, with 113 repositories using English as the only language or one of the languages, followed by Japanese, which is being used by 65 e-repositories, and Chinese, which is the third largest language in Asia, used by 11 repositories. In addition, Georgian, Arabic, Kazakh, Spanish, Malay, Pashto, Persian, and Azerbaijani is represented by 1 repository each. Table 12 gives further details regarding other languages used by repositories in Asia.

Table 12: Language Diversity in Asia

Language	Number
English	113
Japanese	65
Chinese	11
Turkish	4
Hindi	3
Kannada	3
Korean	3
Russian	3
Arabic	1
Azerbaijani	1
Georgian	1
Kazakh	1
Malay	1
Pashto	1
Persian	1
Spanish	1

VIII. Conclusion

The OA is gaining momentum. Given the above statistics on the growth and diversity of open access repositories, one can easily conclude that like the invention of paper and Gutenberg printing press which revolutionized the dissemination and preservation of knowledge in the past, the OA movement, especially the Open Access repositories, is going to set new standards for information sharing and management. The trend to set up OA repositories worldwide is inevitable and needs to be encouraged as well, particularly in the emerging world like Asia, which has every reason to excel, given the strong ICT background and mushrooming of quality academic and research institutions with high research output.

References

Andrew, Rothery., & Bell, Viv. (2006). *E-sharing: developing use of e-repositories and e-libraries for learning and teaching*. Retrieved on October 14, 2008 from http://eprints.worc.ac.uk/48/

Arunachalam, Subbiah. (2008). Open access to scientific knowledge. *DESIDOC Journal of Library and Information Technology*, *28*(1), 7-14.

Banks, Marcus A. (2006). Towards a continuum of scholarship: The eventual collapse of the distinction between grey and non-grey literature. *Publishing Research Quarterly, 22*(1). Retrieved November 28, 2007 from http://eprints.rclis.org/5803/1/GL7Paper_Final.pdf

Barker, E., James, H., & Knight, G. (2004). *Long-Term Retention and Reuse of E-Learning Objects and Materials*, report commissioned by the JISC. Retrieved October 14, 2008 from www.jisc.ac.uk/uploaded_documents/LTR_study_v1-4.doc

Bravo, B. R., & Diez, M. L. A. (2007). E-science and open access repositories in Spain. *OCLC Systems & Services, 23*(4). Retrieved April 17, 2008 from http://www.emeraldinsight.com/10.1108/10650750710831501

Chan, Leslie. (2004). Supporting and enhancing Scholarship in the digital age: the role of open-access institutional repositories. *Canadian Journal of Communication, 29*, 277-300. Retrieved November 28, 2007 from http://eprints.rclis.org/2590/1/Chan_CJC_IR.pdf

Chan, Leslie., & Costa, Sely. (2005). Participation in the global knowledge commons: challenges and opportunities for research dissemination in developing countries. *New Library World, 106*(3/4). Retrieved May 27, 2008 from http://www.emeraldinsight.com/10.1108/03074800510587354

Das, A. K., Sen, B. K., & Dutta, C. (2007). *ETD policies, strategies and initiatives in India: A critical appraisal*. 10th International Symposium on Electonic Theses and Dissertations, Uppsala, Sweden. Retrieved November 29, 2007 from http://eprints.rclis.org/archive/00010657/01/Das_Dutta_Sen_India__ETD_2007_Paper.pdf

Deventer, Martie van., & Pienaar, Heila. (2008). South African repositories: Bridging knowledge divides. *Ariadne, 55*. Retrieved May 14, 2008 from http://www.ariadne.ac.uk/issue55/vandeventer-pienaar

Falk, Howard. (2003). Digital archive developments. *The Electronic Library, 21*(4).Retrieved April 17, 2008 from http://www.emeraldinsight.com/10.1108/02640470310491603

Fang, Conghui., & Zhu, Xiaochun.(2006). The open access movement in China. *Interlending & Document Supply, 34*(4). Retrieved April 21, 2008 from http://www.emeraldinsight.com/10.1108/02641610610714777

Fernandez, Leila. (2006). Open access initiatives in India: An evaluation. *The Canadian Journal of Library and Information Practice and Research, 1*(1). Retrieved November 22, 2007 from http://journal.lib.uoguelph.ca/index.php/perj/article/view/110/172

Jacso, Peter. (2006). Open access to scholarly full-text documents. *Online Information Review, 30*(5). Retrieved April 17, 2008 from http://www.emeraldinsight.com/10.1108/14684520610706442

Johnson, Richard K. (2002). Institutional repositories: Partnering with faculty to enhance scholarly communication. *D-lib Magazine, 8*(11). Retrieved May 23, 2007 from http://www.dlib.org/dlib/november02/johnson/11johnson.html

Lynch, Clifford A., & Lippincott, Jaon K. (2005). Institutional repository deployment in the United States as of early 2005. *D-lib Magazine, 11*(9). Retrieved December 1, 2007 from http://www.dlib.org/dlib/september05/lynch/09lynch.html

Rieh, S. Y., Markey, K., Jean, B. S., Yakel, E., & Kim, J. (2007). Census of institutional repositories in the U.S.: A comparison across institutions at different stages of IR development. *D-Lib Magazine, 13*(11/12). Retrieved May 14, 2008 from http://www.dlib.org/dlib/november07/rieh/11rieh.html

Authors:
Dr. Zahid Ashraf Wani, Faculty Department of Library & Information Science, University of Kashmir. Email: zahidrais@gmail.com

Dr. Sumeer Gul, Assistant Professor, Department of Library & Information Science, University of Kashmir.

Javeed Ahmad Rah, Faculty Department of Library & Information Science, University of Kashmir.

Submitted to CLIEJ on 18 May 2009.

Wani, Zahid Ashraf; Gul, Sumeer; & Rah, Javeed Ahmad. (2009). Open access repositories: A global perspective with an emphasis on Asia. *Chinese Librarianship: an International Electronic Journal*, 27. URL: http://www.iclc.us/cliej/cl27WGR.pdf