

The Global Contribution of E-Theses and Dissertations (ETD): Growth and Development in the Digital Era

Sumaira Jan*, Mudasir Khazer**

Abstract

In the 21st century, world is becoming more and more conscious about information particularly the novel one that emerges through rigorous research conducted by research base of a nation. However, the ideas/information generated through the research requires a proper channel through which its communication as well as dissemination among others would be accomplished. Accordingly, to fulfil the purpose research results are usually published in the form of articles, proceedings, Theses and Dissertations etc. However, among the various forms of research communicating channels Theses and Dissertation are the most valuable mediums for exploring the nascent ideas and information emerged because of thorough research. Since as a major source of information, Theses and Dissertations are predominantly valuable to researchers. Furthermore, they represent the highest intellectual asset of the individuals in a particular field, or a faculty, or subject, thus unswervingly contributing to the growth and development of nation as a whole. Yet, despite the fact, these valuable resources remain hidden due to lack of proper communicating channels and usually rot in anonymity in university libraries and archives. However, the advents of technological transformations have helped a great deal to explore these treasures of knowledge. In this regard the emergence of Electronic theses and dissertations (ETDs) are a somewhat an innovative approach of research and scholarly communication paving way to explore the otherwise a hidden treasure across the globe. However, their visibility solely depends upon how well they have been made accessible to others and through which channel. In above perspective emergence open access E-theses repositories proved to be boon for scholarly communication.

Keywords: Repositories, Open Access (OA), Open Access (OA) Repositories, ETD, DOAR, Information Technology

Introduction

In contemporary era, information is regard as the most valuable asset which leads to the real development of a society and nation as whole. Accordingly, the generation of new and novel information has become indispensable of the information based societies. However, the information is valuable when it is managed in the most effective way i.e., it's worth can only be judged when it is disseminated and utilized to gain knowledge and ultimately the benefits of nurturing intellects on global level. In line with the same, throughout the world new and nascent ideas are being generated that are recorded and preserved for dissemination in one or other form. However, among them the most important is in the form of Thesis and Dissertations but these treasures of knowledge remain hidden due to lack proper channels of dissemination. However, with the advent of technological revolution these barriers tend to diminish to a lot of extent combine with revolutionary concept of Open Access (OA). Accordingly, the development of OA digital repositories proved to be boon for the same. Nowadays, research based institution are striving hard to develop the digital repositories of their knowledge treasures so that they could be disseminated in present as well as preserved for the future generation to come. This paved way for the development of electronic Thesis and Dissertations (ETDs).

In above context, the electronic Theses and Dissertations (ETD's) are key source of novel information embraced of novel ideas, innovations along with nascent information on the particular matter. These ETDs consist of Masters or

* Departmental Librarian, University of Kashmir, Jammu & Kashmir, India. Email: sumaira.mattoo@gmail.com

** College Librarian, Higher Education Department, Jammu & Kashmir, India.

Doctoral research work that is submitted and then archived electronically by an organization, either on an internal network or on the web. To provide a comprehensive view about the concept of ETD's, UNESCO ETD Guide website (as cited in Jasimudeen & Maghesh, 2011) states that *"An ETD is a document that explains the research or scholarship of a researcher/student. It is expressed in a form simultaneously suitable for machine archives and worldwide retrieval. The ETD is similar to its paper predecessor. It has figures, tables, footnotes, and references. It has a title page with the author's name, the official name of the university, the degree sought, and the names of the committee members. It documents the author's years of academic commitment. It describes why the work was done, how the research relates to previous work as recorded in the literature, the research methods used, the results, and the interpretation and discussion of the results, and a summary with conclusions"*. Moreover, to further, explain their importance, Hirwade (2011) states that these provide an deep insight into the research output in specific subject areas as well as also come up with suitable resolutions to the on hand predicaments. They also put forth the latest trend on a particular topic. Besides, these are practical resources for future research and accordingly must be made available as well as accessible worldwide. However, in most of the research-based organizations the access is restricted to any particular geographical area only. Besides, based on the content type such as e-journals, ETD's research publications, datasets, books, patents, software, learning objects, multimedia, etc archived by the repositories they are by and large accordingly categorized. However, despite of immense importance of Theses and Dissertations there are often very less communication channels for the same comparison to other sources posing limitations to the dissemination of productive and intellectual knowledge. But, owing to the various benefits of electronic Theses and Dissertations (ETDs) repositories, they are rapidly gaining significance over their traditional print format which has limited means for dissemination. Moreover, the repositories can be either Open Access or subscription based that depends upon the access policies of host organization that are responsible for archiving the content in the repositories.

In above milieu despite of the fact that the ETD's are well thought-out to be an imperative part of information resources in any institution, yet these valuable unpublished

resources are often limited to parent institution only so far as access to the full content is concerned. So in order to overcome the predicament, a vast amount of Literature clearly reveals that the research Institutions are by and large addressing this problem by making their ETD's available online or by achieving membership of various online repositories like Shodhganga, OpenDOAR, etc and thus not only increasing its visibility but access also (Sengupta, 2012). Accordingly, keeping in view the immense importance of OA ETD repositories for the development of societies and nation as whole, current study is an effort to highlight the global contribution of OA ETD repositories while evaluating their features as well as content management policies to know about the status of OA ETD repositories at individual country level.

Literature Review

In view of gaining importance of ETD repositories more and more researchers globally are conducting studies related to the various parameters of OA ETD repositories. This section, present an overview of studies conducted by various researchers globally in orders to gain an insight about OA ETD repositories.

Hirwade (2011) intricate that the Electronic Theses and Dissertations (ETD's) are key source of novel ideas, innovations and latest information that has been documented by the researchers after tireless effort on the particular topic of interest that could invaluablely benefit the society as a whole. As these provide a thorough insight into the research output in particular areas that ultimately come up with solutions to the existing predicaments of the society. Schopfel (2013) while attempting to highlighting the importance of ETD's states that these not only highlight the intensive research but also represents a mounting part of accessible open content archived in institutional repositories where they potentially add to the general impact as well as ranking of their particular institutions. Moreover, Sengupta (2012) observed that ETD's are considered to be one of the most significant sources of scholarly communication across the globe. Authors and researchers are trying hard to highlight research findings readily available to the people by means of uploading their ETD's on Open Access Platforms. In view of the fact that people have been more and more gaining awareness about the importance of ETD'S owing to its various benefits of over the print format, various institutions are developing their individual

E-thesis repositories. Besides, Hirwade (2011) pointed out that the utmost advantages of ETDs is that these help in shunning of duplication in research work, thereby making certain swift retrieval of information, endorsing resource sharing, as well as providing a everlasting solution to the space problem. Ghosh (n.d.) divulges that throughout the world an increasing number of higher education institutions have made it mandatory for graduate students to submit electronic Theses and Dissertations (ETDs) subsequently they provide OA to these works in online repositories. To mention a few Countries such as Korea, China, Brazil, Australia, have executed, or are in the phase of implementing, national policies for the development of national ETD repositories. Moreover, Ahmed, Alreyaee and Rahman (2014) highlights the status of OA repositories with particular reference to Asia reveals that although some of the Asian countries are lagging behind in terms of development of OA ETD repositories however, at the same time, OA ETD system is budding fast in a number of Asian countries like India, Japan, China and Taiwan indicating that regions have become sensitized about the importance of OA ETD repositories.

Objective (s)

To assess and highlight the status of open access (OA) E-Theses repositories globally based on select parameters viz; geographical distribution, software usage, language diversity, operational status, repository type, subject coverage and content management policies.

Methodology

Current study is based on the data congregated from openDOAR. Data congregated were systematically analyzed based on selected parameters viz: geographical distribution, software usage, language diversity, operational status, repository type, subject coverage and content management policies.

Scope

The scope of the study is limited to E-Theses repositories registered in the open-DOAR (Directory of open access repositories).

Purpose

The present study attempts to reveal the prominence of Open Access (OA) E-Theses repositories globally to

highlight the status as well as growth and development of said repositories in different parts of the world.

Data Analysis and Interpretation

Continental Distribution of Repositories

Europe emerges to be the leading contributor of ETD repositories followed by Asia and North America (Table 1).

Table 1: Continental Distribution of Repositories

Continent	No. of Repositories
Europe	700 (43.4)
Asia	333 (20.6)
North America	270 (16.7)
South America	166 (10.3)
Africa	88 (5.5)
Australasia	41 (2.5)
Central America	9 (0.6)
Other	6 (0.4)
Total	1613

Figures in parentheses indicate percentage

Country - Wise Distribution of Repositories

Among the top three global contributors USA holds no. 1 rank through the contribution of major share of 215 (13.3) repositories followed by Germany ranking no. 2 and Japan ranking no. 3 with 122 (7.6%) and 90 (5.6%) respectively. Moreover, the analysis reveals that there are another 105 countries, which contribute from 40 to 1 repository in aggregate account for 722 (44.8%) repositories respectively (Table 2).

Table 2: Country-Wise Distribution of Repositories

Country	No. of Repositories
United States	215 (13.3)
Germany	122 (7.6)
Japan	90 (5.6)
U.K	80 (5)
France	60 (3.7)
Spain	60 (3.7)
Brazil	49 (3)
Turkey	47 (2.9)
Italy	45 (2.8)

Country	No. of Repositories
Canada	41 (2.5)
Norway	41 (2.5)
Sweden	41 (2.5)
(105) Others	722 (44.8)
Total	1613

Figures in parentheses indicate percentage

Repositories According to Host Organization

Depending upon the nature of host organization of OA repositories, these are grouped into the four categories which include Institutional, Disciplinary, Aggregating and Governmental. Majority of 1498 (92.9%) repositories are Institutional (an institutional or departmental archive) followed by 56 (3.5%) Disciplinary (across institutional subject repository) repositories. On the other hand, least no. of, 45 (2.8%) and 14 (0.9%) of repositories are contributed each by Aggregating (archive aggregating data from several subsidiary repositories) and Governmental (repository for governmental data) organization (Table 3).

Table 3: Repositories According to Host Organization

Repository Type	No. of Repositories
Institutional	1498 (92.9)
Disciplinary	56 (3.5)
Aggregating	45 (2.8)
Governmental	14 (0.9)
Total	1613

Figures in parentheses indicate percentage

Software Usage by Repositories

Amid the varied range of softwares majority of 779 (48.3) repositories prefer DSpace software followed by EPrints, Digital Commons, OPUS preferred by 222 (13.8%), 99 (6.1%) and 61 (3.8%) repositories respectively. While as 172 other softwares are used by 370 (22.9%) repositories. However, 82 (5.1%) repositories that have not specified the type of software used by them have been put under the category "Unknown" (Table 4).

Table 4: Software Usage by Repositories

Software Type	Number of Repositories
DSpace	779 (48.3)
EPrints	222 (13.8)
Digital Commons	99 (6.1)
OPUS	61 (3.8)
Other (172)	370 (22.9)
Unknown	82 (5.1)
Total	1613

Figures in parentheses indicate percentage

Operational Status of Repositories

Repositories have been categorized based on their functionality into four categories viz: Operational that signifies fully operational repositories, Trail that signifies repository on trail basis, Broken that signifies technically malfunctioning repositories and Closed that signifies repositories not accepting depositions/statements of the organization. The highest no. of 1523 (94.4%) repositories are operationally functional, followed by 52 (3.2%) trail & 33 (2%) broken repositories. On the other hand, 5 (0.31%) repositories are undetermined that represent those whose status is not described by the host organization (Table 5).

Table 5: Operational Status of Repositories

Type	Number of Repositories
Operational	1523 (94.4)
Trail	52 (3.2)
Broken	33 (2)
Closed	0
Other	5 (0.31)
Total	1613

Figures in parentheses indicate percentage

Content Type Archived by Repositories

An array of content types are archived by the repositories, which include Journal articles, Conference and workshop papers, Theses and Dissertation etc. A broad classified listing of content types archived by ETD repositories apart from them is shown in Table 6. Of the total 11 content types identified, the findings reveal that majority

of repositories hold content in the form of journal articles 1179 followed by conference, workshop papers 674. On the other extreme least archived content type are patents and Softwares with 70 and 2 repositories archiving them respectively (Table 6).

Table 6: Content Type Archived by Repositories

Rank	Content Type	Number of Repositories
1	These and dissertations	1613
2	Journal article	1179
3	Conference and workshop papers	674
4	Unpublished reports and working papers	622
5	Books, chapters and sections	642
6	Multimedia and audio-visual materials	322
7	Bibliographic references	231
8	Learning objects	291
9	Other special item types	225
10	Datasets	72
11	Patents	70
12	Softwares	24
	Total	5965

Note: Since, majority of repositories hold several Content Types, so the no. of repositories for content type exceeds the actual no. of repositories.

Top Ten Language Interface of Repositories

Language interface of OA repositories has been developed in diverse languages in order to support the users of their respective language. However, among all of them English emerges to be the most preferred language of repositories interface. Since, of the total 1613 repositories, 1174 repositories were in English language followed by 211 repositories in Spanish and 144 in German language interfaces respectively. Table 6 shows the list of top ten languages particularly used by OA repositories (Table 7).

Table 7: Language Interfaces of Repositories

Rank	Language	Number of Repositories
1	English	1174
2	Spanish	211
3	German	149
4	French	119

Rank	Language	Number of Repositories
5	Japanese	92
6	Portuguese	90
7	Chinese	67
8	Italian	50
9	Turkish	47
10	Swedish	43

Note: Since, repositories develop Interface in more than one language, so the no. of repositories with Different language interfaces exceeds the actual no. of repositories.

Top Subjects Covered by ETD Repositories

Current study observes top ten subjects Categories under which most of the records are preferred to be archived by the repositories. As most of the large institutions cover all subjects areas in their repositories, they are therefore, categorized as multidisciplinary. Majority of 1123 repositories archive records under multidisciplinary heading which is obvious due to the fact that it is combination of more than one subject. Apart from multidisciplinary, Business and economics emerges out to be predominant subject archived by (148) repositories followed by Health and Medicines by (144) repositories (Table 8).

Table 8: Top Ten Subjects Covered by the Repositories

Rank	Open DOAR Subjects	Number of Repositories
1	Multidisciplinary	1123
2	Business and economics	148
3	Health & medicine	144
4	Technology general	137
5	Science –general	114
6	Law & politics	104
7	Social sciences general	99
8	Computers & IT	94
9	Education	91
10	Arts and Humanities General	88

Note: Number of subjects covered exceeds no. of repositories for the reason that one-repository archives content on more than one subject area.

Content Management Policies of Repositories

The most significant part of OA repositories are the policies associated with organization and management of the contents (records) that cater the information needs of users in a most convenient way. In this milieu, following content management policies of OA repositories have been considered for evaluation.

Meta-Data Re-use Policy

Metadata re-use policies refer to those policies that are concerned with re-use of information related to the items in the repository. Majority of 1097 (77.8%) repositories have 'undefined' policies, followed by 138 (9.8%) repositories that authorize metadata reuse for 'non-profit' whereas only 88 (6.2%) repository authorizes 'commercial' reuse of metadata. Further, least number of 15 (1.1%) and 67 (4.8%) repositories have 'Unstated' and 'Unknown' policies respectively (Table 9).

Table 9: Metadata Re-use Policy of Repositories

<i>Metadata Re-use Policy</i>	<i>Number</i>
Undefined	1097 (77.8)
Non-Profit	138 (9.8)
Commercial	88 (6.2)
Unstated	15 (1.1)
Unknown	67 (4.8)
Total	1409

Figures in parentheses indicate percentage

Note: Not all entries in OpenDOAR have policy information; therefore, the breakdown repositories with metadata re-use policies is less than the actual number of repositories.

Content Policies of Repositories

Content policies are those Policies that are proposed for type of document & data set held by the repositories. Majority of 1039 (73.7%) repositories have an undefined policy while 289 (20.5%) repositories have 'defined' policy (Table 10).

Table 10: Content Policies of Repositories

<i>Content Policy</i>	<i>Number</i>
Undefined	1039 (73.7)
Defined	289 (20.5)
Unstated	17 (1.2)
Unknown	65 (4.6)
Total	1409

Figures in parentheses indicate percentage

Note: Not all entries in OpenDOAR have policy information; therefore, the breakdown of repositories with content policies is less than the actual number of repositories.

Submission Policies of Repositories

Submission policies are associated with depositors, quality & copyright. Since, it is highly imperative for OA repositories to make it evident that who is authorized to submit material to repositories and what are the terms and conditions for submission of an item. Maximum no. of 1012 (71.8%) repositories has "undefined policies" while as merely 311 (22.1%) repositories have "defined policies" (Table 11).

Table 11: Submission Policies of Repositories

<i>Type of Policy</i>	<i>Number</i>
Undefined	1012 (71.8)
Defined	311 (22.1)
Unstated	15 (1.1)
Unknown	72 (5.1)
Total	1409

Figures in parentheses indicate percentage

Note: Not all entries in OpenDOAR have policy information; therefore, the breakdown of repositories with submission policies is less than the actual number of repositories.

Preservation Policies of Repositories

Preservation Policies are associated with the management of costs, techniques etc involved in ensuring long-term access to contents archived by the repositories. It is

observed that just 101 (7.2%) repositories have “defined” policies for the preservation of information contents. On the other hand maximum no. of 1060 (75.2%) repositories has “undefined” preservation policies. While as just 67 (4.8%) repository have “unknown” preservation policy table (Table 12).

Table 12: Preservation Policies of Repositories

Policy Type	Number
Undefined	1060 (75.2)
Defined	101 (7.2)
Unstated	181 (12.8)
Unknown	67 (4.8)
Total	1409

Figures in parentheses indicate percentage

Note: Not all entries in OpenDOAR have policy information; therefore, the breakdown of repositories with Preservation policies is less than the actual number of repositories.

Findings and Conclusion

The global status of OA repositories archiving ETDs is more prominent and pervasive in developed countries like that of Europe, America etc as they contribute more than 70% of repositories as revealed above. However, on the other hand, developing countries are still wadding behind in terms of the contribution of OA repositories, which is evident from the fact that the Japan is the only developing country contributing 90 repositories holding 3rd rank at global level and first rank at Asian level. Although, at continental level Asia emerges out to at 2nd rank, but despite of that at country level much progress is required for keeping pace with developed continents. This can be attributed to fact that developed countries have become well aware of the fact the managing and preserving the primary and nascent information generated in the form of scholarly communication will be a boon for nations overall development. Consequently, these regions have become well versed with fact that their proper benefits can only be reaped when they will be disseminated through proper channels of scholarly communication for which OA repositories proved to be a most valuable asset. However, on the other side of the globe, developing countries still need to be sensitised about the significance of developing and maintaining OA ETD repositories particularly in present knowledge-based society. Among all types

of repositories, institutional repositories hold highest share as the mainstream of institutions support research based activities for which OA is the most appropriate platform for dissemination of scholarly information. While as the governmental repositories contribute low share indicating that these may not be supporting the research activities as such resulting least significance for archiving ETD's. Further, the statistics reveal that the majority of repositories develop interface in English language, that must be reconsidered and efforts must be on developing language interface in native languages other than the English in order to increase use and visibility of scholarly communication globally.

Content management policies are biggest issues faced by the repositories, which is evident from the fact that just 20-25% of repositories have well defined policies for content management. Most of the repositories are still in quandary so far as content management policies are concerned due to rapid and dynamic changes brought due to advancements in the field of ICT. However, undermining the importance of well defining the content management policies pose big threat in proper utilization of scholarly communication at global level as it hinders the users in utilization of information due to inconsistencies in content management policies. As results repositories must gave immense importance in solving the issues related to content management policies in order to increase the visibility of contents archived by the repositories. Besides, the trend of having more and more repositories from developed world only must be taken into due consideration by the developing as well as by under developed nations. They must develop proper strategies for more and more development of repositories archiving the most valuable information sources in the form of ETDS. Consequently, in order to go ahead in repository soul varied issues associated with the development and advancement of OA repositories archiving ETDs need to be dealt prudently.

References

- Ahmed, A., Alreyaee, S., & Rahman, A. (2014). Theses and dissertations in institutional repositories: An Asian perspective. *New Library World*, 115(9/10), 438-451. doi:10.1108/NLW-04-2014-0035
- Ghosh, M. (n.d.). ETDs in India: Towards a national repository with value added e-theses service (pp.

- 1-22). Retrieved from <http://epc.ub.uu.se/etd2007/files/papers/paper-20.pdf>
- Hirwade, M. A. (2011). Digitization of theses and dissertations: Status quo India digitization of theses and dissertations: Status quo India. *Proceedings of the 14th International Symposium on Electronic Theses and Dissertations, Cape Town, South Africa* (pp. 13-17), September. Retrieved from http://dl.cs.uct.ac.za/conferences/etd2011/papers/etd2011_hirwade.pdf
- Jasimudeen, S., & Maghesh, R. M. (2011). Open access archives of doctoral dissertations in the Mahatma Gandhi University: A study of the feasible access by the research scholars. *International Journal of Digital Library Services*, 1(2), 153-164. Retrieved from http://www.ijodls.in/uploads/3/6/0/3/3603729/jasmin12_144-152.pdf
- Schopfel, J. (2013). Adding value to electronic theses and dissertations in institutional repositories. *D-Lib Magazine*, 19(3/4). doi:10.1045/march2013-schopfel
- Sengupta, S. (2012). E-thesis repositories: The Asian scenario. *Chinese Librarianship: An International Electronic Journal*, 33, 1-8. Retrieved from <http://www.iclc.us/cliej/cl33shantashree.pdf>