

## **ETDs and Open Access for Research and Development: Issues and challenges**

### **Bhojaraju Gunjal**

Head & Deputy Librarian,  
Central Library,  
National Institute of Technology Rourkela,  
Odisha 769008 INDIA.  
[bhojaraju.g@gmail.com](mailto:bhojaraju.g@gmail.com)

### **Panorea Gaitanou**

Database & Information Systems Group (DBIS)  
Ionian University  
Ioannou Theotoki 72, 49100  
Corfu, Greece  
[rgaitanou@ionio.gr](mailto:rgaitanou@ionio.gr)

### **Abstract**

Electronic Theses and Dissertations (ETDs) are the most frequent document type found and consulted for use in the Digital Era. During the last decade, their growth globally has increased tremendously among universities and other organizations. Moreover, there is an increase in ETDs Initiatives and ETDs Collections. Open access and open source technologies help ETDs and other research material repositories to be developed. ETDs can play an important role in research by adopting knowledge organization system (KOS) architecture to enhance information retrieval (IR) systems and their performance. The paper will refer to the latest trends in ETDs, which have adopted various technologies, including discovery services, to meet the users' needs. It will also describe some tendencies concerning ETDs, as observed during explorations of institutional and other repositories along with associated issues and challenges. Global efforts in a broad range of ETDs initiatives will be covered as well within the scope of this paper. Finally, the paper concludes with some remarkable results, regarding the literature review study and its analysis.

### **Keywords**

Electronic Theses and Dissertations, ETDs, Open Access, Discovery Services, Information Access, Information Retrieval, Knowledge Organization Systems, KOS

### **1. Introduction**

With the enormous growth of digital information over the web, including Digital Libraries (DLs), the issue of organizing, managing and disseminating information has attracted attention and led to many efforts, including the Semantic Web. Apart from journal articles and e-prints, Electronic Theses and Dissertations (ETDs) are the most frequent document type found and consulted for use in the Digital Era. It is important to highlight the fact that making ETDs freely available to the users, this has clear benefit to student authors, since readability of their researches is enhanced and they become much more visible in the research community. What's more, open access to ETDs is of help to universities, since it increases the awareness of their research activities around the globe.

During the last decade (2005 - 2015), the growth of ETDs globally has increased tremendously among universities and other organizations. Moreover, there is an increase in both, number of ETDs Initiatives and ETDs Collections. Open access and open source technologies seems to help ETDs and other research material repositories to be developed. It is generally agreed that ETDs can play an important role in research by adopting knowledge organization system (KOS) architecture to enhance information retrieval (IR) systems and their performance. They are mostly metadata driven and facilitate search with either by their own or through third-party search engines. Nevertheless, KOSs concepts like Thesaurus, Visualization, Auto-Categorization, Text Categorization and Search Strategy need to be adopted carefully so as to achieve the desired results.

The paper aims to refer to latest trends in ETDs, which have adopted various technologies including discovery services to meet the users' needs. It will also describe some tendencies concerning ETDs. Global efforts in a broad range of ETDs initiatives will be covered as well within the scope of this paper. Finally, the paper will conclude with some remarkable results, regarding the literature review study and its analysis.

## **2. ETDs: A brief review**

As already mentioned, ETDs are the most frequent documents found in the Digital Era. They are well defined and well referenced administrative documents and their deposit and citation are generally established on a national level [24]. ETDs are generally submitted by students who support their candidature for academic degrees. Also, they can be used in order to enrich their curriculum vitae and add value to their professional qualifications [15]. ETDs can either be in print or in digital form. Some universities and institutions manage both forms. Undoubtedly, digital libraries of ETDs offer an alternative to the waste of valuable scholarship [14].

Contrary to the other grey literature, ETDs' cataloguing and preservation are covered by international standards. A very interesting research that provides guidance for the lifecycle management of ETDs can be found in [15]. This paper also describes and analyzes the different types of stakeholders that have different interests and concerns in an ETD programme. Also, in [18] a programme regarding the Lifecycle Management of ETDs is analyzed.

Another interesting study is presented in [16], which refers to issues such as: websites that ETDs are available, the persons/ organizations that are responsible for publishing and providing access to, as well as their archiving and classification, etc. The findings from this study can help stakeholders in American universities to evaluate all management options for ETDs so as to arrive at a policy that best meets the needs of all interested parties.

### **3. Influential factors for ETDs management and their retrieval**

Some of the influential factors for ETDs management and their retrieval are as follows:

#### ***3.1 Knowledge organisation system (KOS) architecture enhances information retrieval (IR)***

Knowledge Organization Systems (KOS) is a general term referring, among other things, to the tools ('semantic tools') that present the organized interpretation of knowledge structures [1]. KOS, such as classifications schemas, gazetteers, lexical databases, ontologies, taxonomies, topic maps, thesauri etc. models the underlying semantic structure of a domain.

For over a decade, the growth of digital information over the Web including (DLs) and ETDs is tremendous and leads to more complexity in data organization and dissemination. Due to this complexity, information organization and retrieval, as well as interfaces for displaying search results have posed several challenges to information scientists in these research areas [1]. A review of the literature reveals that KOS deployed in DLs and ETDs vary and differs in their model adopted and its services. There are different levels of KOS, which they come under by drawing the line by adopting different KOS aspects.

Hodge writes [2]: 'The term *knowledge organization systems* as used in this report was coined by the *Networked Knowledge Organization Systems Working Group* at its initial meeting at the *ACM Digital Libraries 98 Conference* in Pittsburgh, Pennsylvania.'

'The term *knowledge organization system* is intended to encompass all types of schemes for organizing information and promoting knowledge management. Knowledge organization systems include classification and categorization schemes that organize materials at a general level, subject headings that provide more detailed access and authority files that control variant versions of key information such as geographic names and personal names. Knowledge organization systems also include highly structured vocabularies, such as thesauri, and less traditional schemes, such as semantic networks and ontologies. Because knowledge organization systems are mechanisms for organizing information, they are at the heart of every library, museum and archive.'

A detailed study on KOS adaptability in ETDs [3] and an implementation of prototype hybrid KOS model for Vidyanidhi Digital Library [17] to enhance the information retrieval (IR) process [4] should be added to this literature review.

#### ***3.2 Discovery services for ETDs in libraries***

Libraries invest huge amounts in building their collections and platforms to host various resources including ETDs. Nevertheless, sometimes these resources are found under-utilized due to lack of powerful search tools that provide single interfaces. Library users may also get distracted due to multiple platforms, which have been bombarded with huge volume of information. In this scenario, library professionals play a prominent role in providing their users with tools that can make accessible such resources and enhance their usage. Hence, libraries identified web-scale discovery services as a potential tool by providing a single search interface, which could serve the information needs of diverse users.

Discovery services were first started in late 2007. Web-scale discovery services for libraries provide deep discovery to a library's local and licensed content by representing evolution (perhaps a revolution) [8] and aim to aide users to search across a wide range of local and remote library. Web-scale discovery (WSD) tools for libraries also help to (i) connect users with the relevant content from different sources, (ii) search quickly across various platforms and (iii) provide relevancy-ranked results in an intuitive way to meet the user's expectations.

Various tools on discovery services (commercial and open source tools) are available, which boost the search and support researchers in organizations. Some of the major commercial WSD tools are ProQuest, EBSCO, OCLC and ExLibris. Similarly, there are some open source tools or locally controlled tools, such as VuFind, Blacklight, Mango, Scriblio and XC. eXtensible Catalog.

Furthermore, some of the issues and challenges are associated with these discovery tools and how various initiatives around the world are working together to provide solutions to such issues [9].

#### **4. Impact of ETDs and Open Access on Research and Development**

Factors affecting Open Access and ETDs contributions in organizations should always be encouraged by the management. This will help research and development (R&D) activities. Many ETDs are still not made available openly or if made available with embargo, this will hinder R&D activities and affect the education system.

In an Indian scenario, the study by Das and others (2007) [10] indicates that the national ETD initiatives are still in developmental phase where more action plans are needed. The stakeholders are mostly aware of the implications of open access, but are awaiting some concrete policy frameworks by the national accredited and granting agencies. Open access to scholarly literature in India helps rather to produce social goods and social applications, than merely to provide research degrees to the beneficiaries [10].

Similarly, Copeland (2008) states that the situation is now changing, as efforts to promote 'open access' are leading many universities to develop institutional repositories, which contain the full text of ETDs, journal articles, book chapters, conference proceedings, reports etc. [11]. Finally, Jones and Andrew (2005) examine how the synergy of open access and open source has been used at Edinburgh University Library to design and implement an e-thesis service [22].

#### **5. Growth of ETDs, Open Access, and ETD Initiatives**

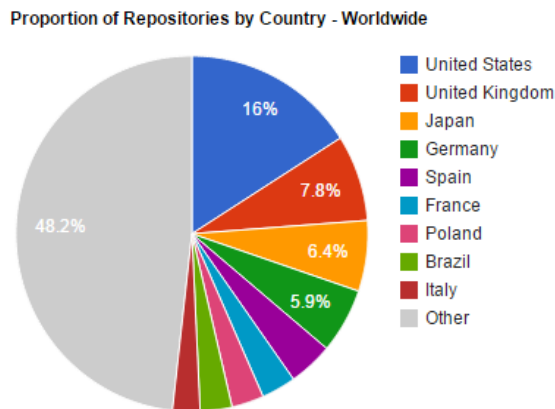
It is observed that the growth of ETDs in the world during 2005 - 2010 has increased enormously among universities and other organizations and that there is an increase in *ETDs Initiatives* by 46 and in *ETDs Collections* by 619.073 [3]. Whereas a drastic growth of ETDs and items is noticed during 2010 - 2015 i.e. open access repositories by 2944 and items in the repositories by 6,229,261. A growth statistics is shown in Table 1.

	August 5, 2015	March 4, 2010	November 24, 2005
No. of repositories	2944	100	54
No. of items in the repositories	62,29,261	8,36,636	2,17,563

**Table 1. Growth of ETDs during 2005 - 2015** (2944 repositories till August 5, 2015)  
Source: Adapted from Gunjal, B. and Urs, S.R. (2010)

The Networked Digital Library of Electronic Theses and Dissertations (NDLTD) is an initiative that has gained the support of several international and American colleges and universities since 1996. It plays now an international leadership role in ETD initiatives by promoting the adoption, creation, use, dissemination, and preservation of ETDs. It supports electronic publishing and open access to scholarship in order to enhance the sharing of knowledge worldwide; it helps to preserve this important material and finally empowers graduate students to convey a richer message through the use of multimedia and hypermedia technologies [14]. The NDLTD Union Catalogue project focussed on theses metadata via the Protocol for Metadata Harvesting (OAI-PMH) [12]. One of its main goals was to maintain an OAI harvestable public repository of ETDs for all institutions that register their ETD OAI site with the Networked Digital Library of Theses and Dissertations (NDLTD). Recently, NDLTD recently announced Global Electronic Thesis and Dissertation Search.

Finally, OpenDOAR [13] shows the growth of ETDs up to 2015 by presenting the proportion of repositories by country worldwide.



**Figure 1: Repositories worldwide**

Source: <http://www.openoar.org/onechart.php?groupby=c.cCountry>

## 6. Issues and challenges

Cayabyab (2015) explained common issues and challenges associated with ETDs analytically [7]. Also, trends and challenges are discussed in [19], [21] and [20]. Some of the major issues and challenges to provide probable solutions are tabulated in Table 2.

Attributes	Issues and Challenges
Policy	No standard policies adopted across the globe
Management Support	Lack of management support
Metadata	Adopting ETD Metadata Standard (ETD-MS) and Dublin Core (DC) Element Set
Interoperability	Protocols such as Open Archives Initiative for Metadata Harvesting Protocol (OAIPMH), MARC21 and OAI-ORE) and controlled vocabularies
Information Retrieval	Semantic Web, Linked Data
Embargo	This will hinder the open access motive
Open Access	Lack of interest in making content open by most of the organizations will hinder the motto of open access
Training	Staff and user training
Scalability	System should be scalable and easy to upgrade with latest features
Copyright/IPR	Policies should be framed to comply IPR
Discovery Tools	Implementation of discovery tools to explore the content from various sources of content
Duplication	Avoid working on duplication of similar projects
Outreach	Proper indexing and harvesting, social plug-ins to reach out large numbers of prospective users
Preservation	Long-term preservation should be addressed in order to achieve the perpetuity of ETDs

**Table 2. Issues and Challenges of ETDs**

## 7. Solutions proposed to overcome ETDs issues

Taking into consideration all the above, we could propose several solutions to overcome the aforementioned issues.

- There are various policies around the world. The ideal would be that a single uniform policy would exist, possibly following the NDLTLD so as to maintain uniformity irrespective of region or type of libraries. Also, the Registry of Open Access Repository Mandates and Policies (ROARMAP) can be referred to, for policies adopted by various institutes. ROARMAP is a searchable international registry charting the growth of open access mandates and policies adopted by universities, research institutions and research funders that require or request their researchers to provide open access to their peer-reviewed research article output by depositing it in an open access repository [5]. This would lead to an ‘One Policy Concept’ linked with all countries universally and accepted by all.
- Uniform policy itself would provide solution to most of the common aforesaid issues and challenges.
- Open Access should be encouraged
  - through in-house publication cell of all scholarly assets by Universities/ Institutions to avoid commercialization of scholarly content, which may hinder the R&D activities in future.
  - to enrich the usage of resources, enhance reading capability and research and strengthen the education system.
- Embargo needs to be adopted with no/minimal period (if required). A very interesting research regarding this issue can be found in [23].

- Skill development and training is finally essential to all stakeholders for ETDs creation and usage.

Generally, the decision whether an institution will provide electronic copies of theses, or it will make these openly accessible to the users depends on its policies. Nevertheless, taking into consideration now-a-days the increased movement and support towards open access around the world, we see that Japan has recently introduced a country-wide mandate [6].

## 8. Conclusion

From all the above, it is obvious that access to ETDs should be enhanced in all possible ways (not only from the point of view of metadata, but the full text as well). It is of great importance that ETDs can be found and explored both by the university community and all researchers. Within this context, libraries should play a leading and evolving role. Moreover, ETDs present particularly several other technical and legal issues related to digital publication. Furthermore, there are several other issues as well, such as the open access of ETDs, the use of multimedia files etc. that can create implications regarding the intellectual property rights [15]. All these issues need to be thoroughly addressed and managed by all stakeholders, so as to achieve the best results for the dissemination of this really important research material.

## 9. References

1. Hjørland, B. (2008): Knowledge organization systems (KOS). In: *Hjørland, B. Lifeboat of knowledge organization.* 2008. Available at [http://www.iva.dk/bh/lifeboat\\_ko/CONCEPTS/knowledge\\_organization\\_systems.htm](http://www.iva.dk/bh/lifeboat_ko/CONCEPTS/knowledge_organization_systems.htm) (Accessed on July 30, 2015)
2. Hodge, G. (2000): *Systems of Knowledge Organization for Digital libraries. Beyond traditional authority files.* Washington, DC: the Council on Library and Information Resources. Available at <http://www.clir.org/pubs/reports/pub91/contents.html> (Accessed on July 25, 2015)
3. Gunjal, B. and Urs, S. R. (2010): Knowledge Organisation Systems in ETDs: A Comparative Study. In *Proceedings of 13th International Symposium on Electronic Theses and Dissertations (ETD 2010); June 16 -18, 2010, Austin, TX, USA.* Available at <https://conferences.tdl.org/utlibraries/etd2010/paper/view/28> (Accessed on July 25, 2015)
4. Gunjal, B. and Urs, S. R. (2012): A Hybrid Approach Incorporating KOS Techniques for Focused Information Retrieval in the ETDs / DLs. *Journal of Indian Library Association*, 48(3) pp. 24-31. (Accessed on July 27, 2015)
5. ROARMAP. Available at <http://roarmap.eprints.org/> (Accessed on July 25, 2015)
6. Kingsley, D. (2013). Open Access ETDs: Providing Readership and Opportunities for Graduate Students. *USETDA* (Blog-July 16, 2013). Available at <http://blog.digitalcommons.bepress.com/2013/07/16/open-access-etds-providing-readership-and-opportunities-for-graduate-students/> (Accessed on July 25, 2015)
7. Cayabyab, T. A. C. (2015): A Review of Emerging ETD Initiatives, Challenges and Future Developments. *International Journal of Information and Education Technology*, 5(10). Available at <http://www.ijiet.org/papers/609-D031.pdf> (Accessed on July 28, 2015)

8. Vaughan, J. (2012): Investigations into library web scale discovery services. *Information Technology and Libraries*, 31(1), pp.32-82 Available at: [http://digitalscholarship.unlv.edu/lib\\_articles/44](http://digitalscholarship.unlv.edu/lib_articles/44) (Accessed on July 25, 2015)
9. Thompson, J. L., Obrick, K. S. and Abate, L. E. (2013). Web-scale discovery in an academic health sciences library: Development and Implementation of the EBSCO discovery service. *Medical Reference Services Quarterly*, 32(1), 26-41.
10. Das, A., Sen, B. and Dutta, C. (2007): ETD Policies, Strategies and Initiatives in India: a critical appraisal. In *ETD 2007: the 10th International Symposium on Electronic Theses and Dissertations, Uppsala, 13-16 June*. Available at <http://epc.uu.se/etd2007/files/papers/paper-54.pdf> (Accessed on July 25, 2015)
11. Copeland, S. (2008): Electronic Theses and Dissertations: promoting 'hidden' research *Policy Futures in Education*, 6(1), pp.87-96. Available at <http://pfe.sagepub.com/content/6/1/87.full.pdf+html> (Accessed on July 25, 2015)
12. *The NDLTD Union Catalog*. Available at <http://www.oclc.org/research/themes/data-science/ndltd.html> (Accessed on July 25, 2015)
13. *OpenDOAR Statistics*. Available at <http://www.opendoar.org/new1statistics.php> (Accessed on August 5, 2015)
14. Edminster, J. R. (2002): *The diffusion of new media scholarship: Power, innovation, and resistance in academe*. Graduate Theses and Dissertations. Available at <http://etd.fcla.edu/SF/SFE0000035/etd.pdf> (Accessed on July 25, 2015)
15. Alemneh, D. et al. (2014): *Guidance Documents for Lifecycle Management of ETDs*. Atlanta, Educopia Institute. Available at [http://educopia.org/sites/educopia.org/files/publications/Guidance Documents for Lifecycle Management of ETDs\\_0.pdf](http://educopia.org/sites/educopia.org/files/publications/Guidance Documents for Lifecycle Management of ETDs_0.pdf) (Accessed on July 25, 2015)
16. Clement, G. P., Rascoe, F. (2013). ETD Management and Publishing in the ProQuest System and the University Repository: A Comparative Analysis. *Journal of Librarianship and Scholarly Communication*, 1(4). Available at <http://jlsccpub.org/cgi/viewcontent.cgi?article=1074&context=jlscc> (Accessed on July 25, 2015)
17. Urs, S. R. and Raghavan, K. S. (2001). Vidyaniidhi: Indian digital library of electronic theses. *Communications of the ACM*, 44(5). pp. 88-89
18. Stark, S., Alemneh, D., and Eisenhauer, S. (2013). The lifecycle management of ETDs Project: Multi stakeholders' national partnership. *iConference 2013 Proceedings* (pp. 869-873). Available at <https://www.ideals.illinois.edu/bitstream/handle/2142/42053/444.pdf?sequence=2> (Accessed on July 25, 2015)
19. Anubi, C. V. and Onyebinama, C. O. (2011): ETDs Initiatives in Federal University of Technology, Owerri (FUTO): Successes, Challenges, Prospects. In *Proceedings of the 14<sup>th</sup> International Symposium on Electronic Theses and Dissertations, Cape Town, 13-17 September*, Available at [http://dl.cs.uct.ac.za/conferences/etd2011/papers/etd2011\\_anunobi.pdf](http://dl.cs.uct.ac.za/conferences/etd2011/papers/etd2011_anunobi.pdf) (Accessed on July 22, 2015)
20. Rodriguez, K. (2006): Electronic theses & dissertations (etds): a literature review. *Revista Puertorriquena de Bibliotecologia y Documentacion*, 8, pp.73-85
21. Satyanarayana, K.V., Babu, B. R. (2007): Trends in the development of E-Theses in India: issues, constraints, and solutions. In *10th International Symposium on Electronic Theses and*



- Dissertations, Uppsala, Sweden, 2007, Available at <http://epc.ub.uu.se/etd2007/files/papers/paper-17.pdf> (Accessed on July 24, 2015)*
22. Jones, R. and Andrew, T. (2005): Open access, open source and e-theses: the development of the Edinburgh Research Archive. *Program*, 39(3), pp.198-212
  23. Owen, T. M., Hackman, T. and Harrod, T. (2009): ETDs in Lock-Down: Trends, Analyses and Faculty Perspectives on ETD Embargoes. In *Proceedings of the 12<sup>th</sup> International Symposium on Electronic Theses and Dissertations*. Available at <http://hdl.handle.net/1903/9087> (Accessed on July 24, 2015)
  24. Stock, C. *Open access to full text and ETDs in Europe: improving accessibility through the choice of language?* Available at [http://www.greynet.org/images/GL9\\_page\\_39.pdf](http://www.greynet.org/images/GL9_page_39.pdf) (Accessed on July 24, 2015)