

Electronic Theses and Dissertations Programs after Two Decades

Exploring Impact

Behrooz Rasuli¹, Joachim Schöpfel², Hélène Prost³

¹ Iranian Research Institute for Information Science and Technology (IranDoc), Tehran, Iran
rasouli@students.irandoc.ac.ir

² University of Lille, GERiiCO laboratory, Villeneuve d'Ascq, France
joachim.schopfel@univ-lille.fr

³ CNRS, GERiiCO laboratory, Nancy, France
helene.prost007@gmail.com

Abstract: Objective: Electronic Theses and Dissertations (ETDs) programs have been recognized as a relatively new and effective channel through which information resources can be made available to academics and beyond. This study aims to provide empirical evidence of the impact of ETD programs in the scientific communities.

Method: Records containing citations to theses and dissertations (TDs) in Scopus database from 1970 to 2017 are retrieved through a structured query. Trends of the citations of TDs in academic literature in different subject areas are examined. In addition, online search of a random sample of citation records on Google Scholar and BASE provided more insight of the impact of ETD programs, with a focus on open access.

Findings: According to the Scopus database, there is an obvious increasing trend in citing TDs in academic literature. In all four broad subject areas, there is an increasing trend in citing TDs, in particular from 1996 onward. The analysis of the random sample and the survey provide complementary elements for a better understanding of this trend, of the role of ETD programs and of the impact of accessibility via institutional repositories.

Keywords: Electronic Theses and Dissertations (ETDs), ETD programs, Academic Impact, Citation Analysis, Open Access.

INTRODUCTION

Electronic Theses and Dissertations (ETDs) programs are a new channel to distribute students' research. ETDs programs were launched in the early 1990s and there is an increasing interest in implementing such programs around the world. ETD programs' advantages including cost and space savings, increasing access to research, increasing visibility, increasing stability and authority, making connections, increasing authorial control, and monitoring usage statistics.

Another key promise of ETD programs is increasing impact of students' works (Bangani 2018; Ratanya 2017; Buehler 2013: 65). Through increasing the visibility of research, ETDs programs can help researchers, practitioners, business owners, and other stakeholders to search and retrieve new findings in different fields. Therefore, it can be expected that ETDs programs increase the impact of TDs. However, research impact includes academic, health, economic, societal, cultural, legal, environmental, etc. (Bornmann 2012). Among them, academic impact of research is much easier to measure, because counting citations is one of the methods to measure such impact. Then, the citations can indicate academic impact of previous research. A new method entitled 'citation analysis' has been developed to study of citations in academic works and citation indexes, such as Web of Science and Scopus, have implemented such methods in their designs.

However, citation analysis is complicated sometimes since research shows that citations in academic publications are not always correct. For example, O'Connor (2002) and Lukic et. al.

(2004) believe that citation errors are common in medical journals. Lukic et. al. (2004) study revealed that 27% of the 199 randomly selected references from articles of three major gross anatomy journals include errors. Al-Benna et. al. (2009) found that 3.3% of cited references in two major burns surgery journals are incorrect. According to Adhikari and Acharya (2010), more than 40% references of medical journals' articles published in Nepal are incorrect. But the errors in cited references are not limited to medical journals. In a study on library and information science's research articles published in academic journals, Davies (2012) believes that error rates in the journals are considerable. Her findings showed the overall error rates were: author (56 percent), page number (22 percent), article title (15 percent), volume (3 percent), publication year (2 percent) and journal title (2 percent). After the study of citations in five leading environmental science journals, Lopresti (2010) found that more than 24% of citations contain errors. The accuracy of cited references in unpublished materials, including theses and dissertations, is lower than journals. A study on five psychology theses submitted to the University of Mysore showed that seventy seven percent of the references had citation errors (Harinarayana, Chikkamanju, and Raju 2011). Rahmani, Asnafi, and Erfanmanesh (2016) reported that about 78% of cited references in Shahid Beheshti University's theses and dissertations contain errors.

Since one of the most important goals of ETD programs is to increase visibility, usability, and impact of TDs (Ferrerias-Fernández, et. al. 2016; Ezema, and Igbo 2016; Ahmed, Alreyaee, and Rahman 2014), then measuring success of these programs is important. This study aims to measure the impact of ETD programs through analyzing citations to theses and dissertations (TDs). Increasing citations to TDs distributed through ETD programs can indicate the impact of such programs and make programs' manager sure that their resources are not wasted.

METHODOLOGY

Our approach combines two different related methods. (1) The first step was a citation analysis with the Scopus database. The Scopus scientometric database (one of Elsevier's product) has been used to identify TDs cited in academic literature (articles, reviews, book chapters, conference papers, etc.). Scopus citation database supports cited reference searches for different document types and indexes more literature than similar databases. To extract citations to TDs, the most popular citation styles in academic works has been surveyed.¹ (2) Based on a representative random sample of TD records, in the second step try to estimate the contribution of ETD programs, via combined searches with the Bielefeld Academic Search Engine (BASE) and Google Scholar (GS) in order to determine the part of ETD in institutional or aggregating repositories, portals or other platforms (eg NDLTD Global Search, DART Europe, EThOS, Thèses.fr, Shodhganga, and Theses Canada). Particular attention will be paid to accessibility (open access, embargo etc.) and retrodigitization programs. Studying the trend of citations and

¹ search query: (((REFSRCTITLE(("Dissertation)" OR "[dissertation]" OR "[master's thesis]" OR "[master's thesis]" OR "[thesis]" OR "dissertation" OR "Doctoral dissertation" OR "doctoral thesis" OR "honors thesis" OR "M Eng. Thesis" OR "M.S. thesis" OR "M.Sc. Thesis" OR "MA thesis" OR "Master" OR "Master's thesis" OR "Master's thesis." OR "Masters level dissertation" OR "master's theses" OR "masters thesis" OR "Master's thesis" OR "MS thesis" OR "MSc Thesis" OR "MSc. Thesis" OR "Ph.D. diss." OR "Ph.D. dissertation" OR "Ph.D. thesis" OR "PhD [dissertation]" OR "PhD diss." OR "PhD dissertation" OR "PhD thesis" OR "PhD. Dissertation" OR "Senior thesis" OR "thesis" OR "M.A. Thesis" OR "M. A. Thesis" OR "dissertation on the Internet" OR "dissertation on microfiche")) AND PUBYEAR < 2018)); search date: July 2018

random samples that whether they contain citations to TDs which are available in an ETD database or not, the real impact of ETD programs will be clarified.

FINDINGS

A preliminary search on the Scopus database retrieved 869969 documents from 1923 to 2017 (all documents = 66839703) which already provide a general trend of the impact of TDs on academic publications. Worded differently, about 1.5% of all indexed documents on Scopus database contain (at least) one citation to TDs. Figure 1 presents the percent of documents which have cited TDs from 1970 to 2017.

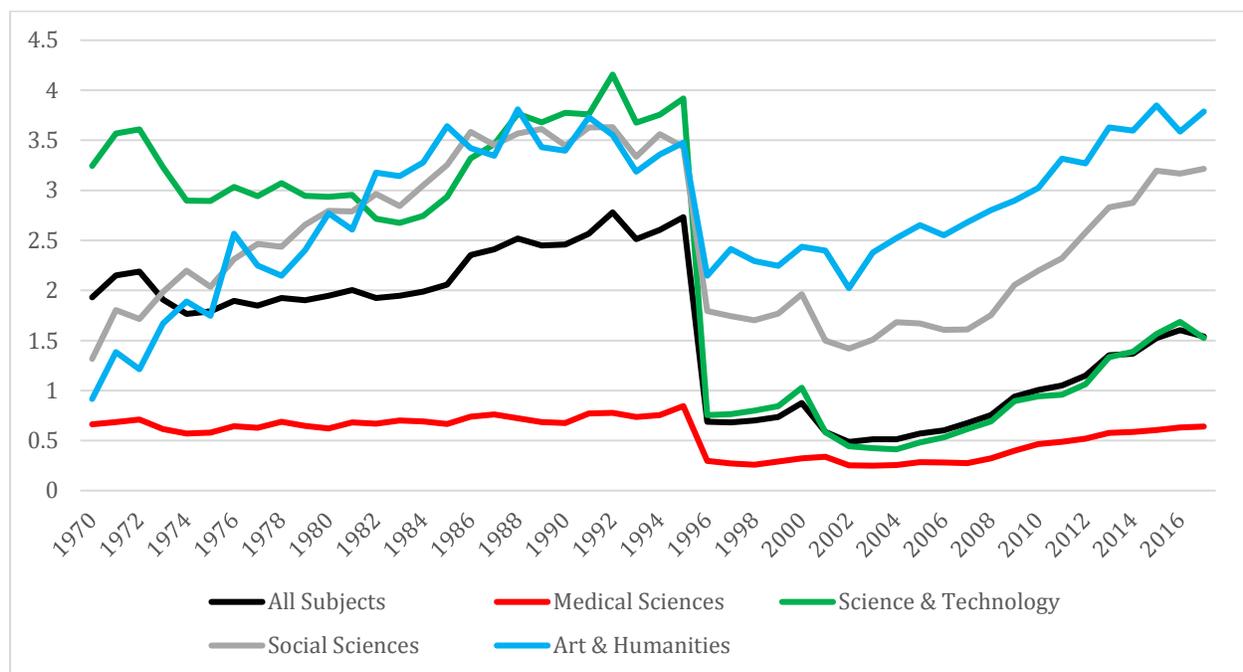


Figure 1. Percent of documents cited theses and dissertation in Scopus (1970-2017)

According to figure 1, there is an obvious increasing trend in the percent of documents containing (at least) a citation to TDs (black line). As well, this increasing trend occurred in different broad subjects areas (red, green, blue, and grey lines).² As it is showed in the figure, Art and Humanities academic literature are more willing to cite TDs in their works. In 2015, about 4% of documents in this subject area included (at least) one citations to TDs, while in the same year only 0.6% of medical sciences' documents contained citations to TDs.

According to the findings, more than 10% of indexed books in the Scopus contained (at least) on citations to TDs, followed by book chapters (2.59%), retracted (1.51%), articles (1.41%), reviews (1.08%), and conference papers (0.61%). Although, among countries with more than 100000 publications in Scopus, Nigeria is more willing to cite TDs in academic literature; more than 3.1% of all documents published by Nigerian researchers contain (at least) on citations to TDs, followed by Russian Federation (3.07), South Africa (2.94), Netherlands (2.46), New

² The sudden decrease of citations between 1995 and 1996 is probably related to Elsevier's database production and not to authors' behaviour. We asked the possible reasons of this atypical curve from Scopus team, we have not received a certain answer yet.

Zealand (2.23), Malaysia (2.22), Egypt (2.12), Sweden (2.1), Canada (1.99), Norway (1.97), Australia (1.93), Israel (1.81), United Kingdom (1.79), Finland (1.77), United States (1.7), Germany (1.68), Portugal (1.65), Iran (1.64), Pakistan (1.61), and Ukraine (1.56).

The second analysis is based on a random 384 sample³ of all publications citing TDs from 16696 to 2017. For this smaller corpus, we assess the format and the accessibility of the cited TDs. In other words, we tried to determine for each cited TD if it has digital format, and if it is a digital format (ETD), if the TD is freely available or not, for instance via an institutional repository.

However, 384 sample documents contain 15340 references, which 705 references of them include TDs. After deep analysis, we found 65 documents with no citation to TDs (16.93% of 384 documents). These documents contain some terms of search query (for example, "synthesis", "synthesis", "prostheses", "hypothesis", "master plan", "Master Abstr Int", "master brewers", "masters of the big house", "master plan", "master lecture", "master course"). However, after removing these records from the sample, 319 correct documents were analyzed, which contain 476 references. More analysis showed that 76 references (~16% of 476 references) do not contain a title; many of them have just author name, year, and the level of TDs (for example, "Schulze, J., (2009) PhD Thesis"). As well, these references withdrew from the sample records.

Analysis of 400 references shows that the most of them cited a doctoral dissertation (~70%). Figure 2 shows the frequency of citations to TDs of different educational level.

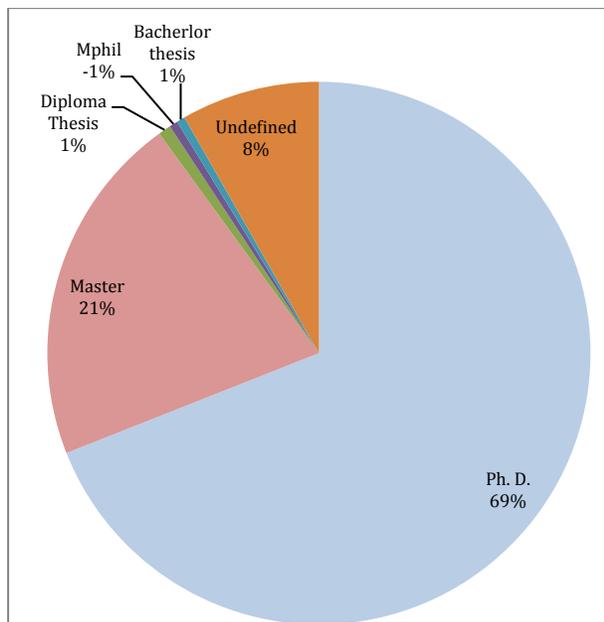


Figure 2. Citation to TDs of different educational level; Scopus (1996-2017)

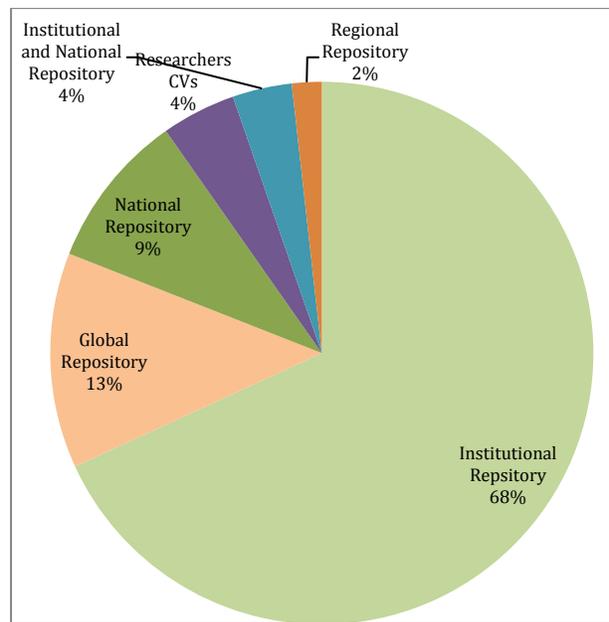


Figure 3. Citing TDs of different educational level; Scopus (1996-2017)

In additions, according to the findings, 226 records of 400 references (~57%) are accessible on the web (through Google Scholar or BASE engines), which about 79% of them (179 records) are open access. According to figure 3, the most of accessible references are findable in an institutional repository (68%). Although, global repositories (such as Eric, ResearchGate, PsycINFO, CiteSeerx, Arxiv, WorldCat, ProQuest, etc.) hosted 13% of references; followed by

³ According to 'Sample Size Calculator' at surveysystem.com/sscalc.htm the number of samples of 869969 records is 384 (Confidence Interval = 5).

national repositories (such as EThOS, Trove, HAL, Theses.fr, CiNii, INFOUNIVERSIDADES, BDTD, etc.), researcher's CVs pages, and regional repositories (such as, DART, DiVa portal, OATAD, Ohiolink, etc.).

However, to investigating the impact of ETD programs, the trend of accessible TDs in past years can provides a better insights. In figure 4, the percent of accessible TDs in different years from 1996 to 2017 is presented.

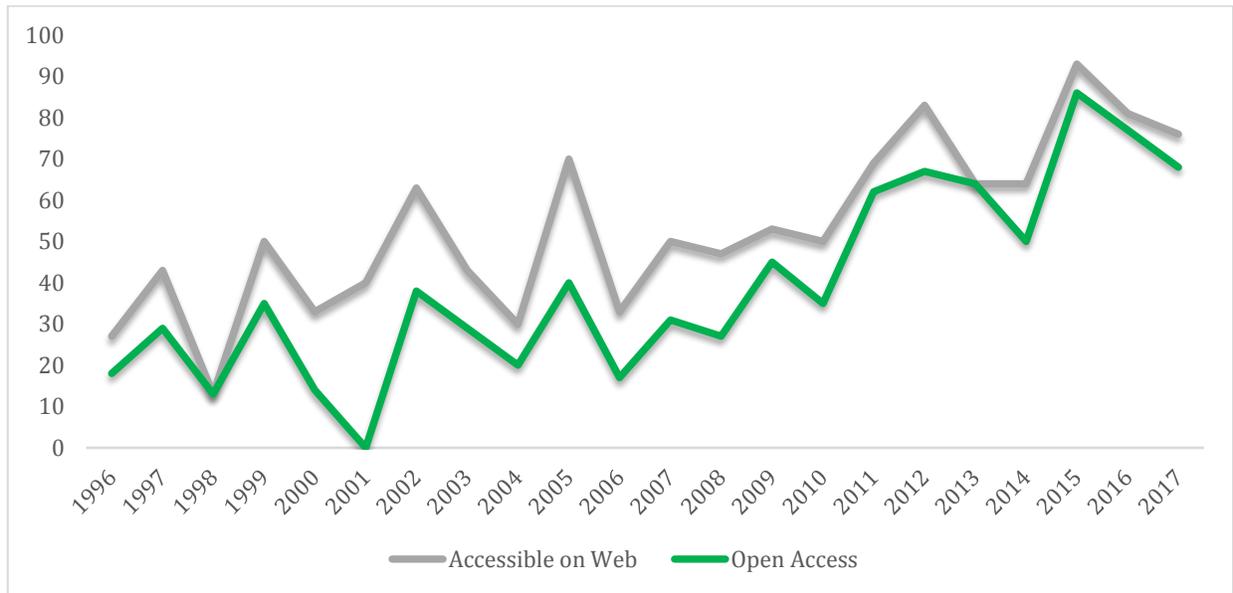


Figure 4. Percent of accessible and open access reference (TDs) on the web; Scopus (1996-2017)

According to figure 4, only 27% of references in 1996 are accessible on the web through library catalogues, repositories, social networks, etc., which 18% of them are open access. While, in 2015 about 93% of the references are accessible through the web and 86% of them are open access.

CONCLUSIONS

This study is a first attempt to assess ETD programs' impact after more than two decades from establishment of these programs. The originality of our study is the application of scientometric measures to assess the academic impact of ETD programs. Trends of citations to TDs in academic literature can be categorized in three time periods:

1) From 1970 to 1992: In this period we observe increasing trends in citations to TDs. Probably, the main reason of increasing trends is the increasing universities and academic institutions with more and more students, and accordingly increasing theses and dissertations. 2) From 1993 to 2002: In this period we observe decreasing trends in citations to TDs (ignoring the sudden decrease of citations in 1996). Probably, the main reason is empowering access to journals through electronic, as well as online channels. But TDs in this period were on the libraries' shelves. Researchers and authors of research articles had to go to the libraries to read and cite TDs. Therefore, because of restricting access to TDs the number of readers and citing authors have been decreased. 3) From 2003 onward: In this period we observe increasing trends in citations to TDs again. It can be argued that the main reason of the increasing trends is ETD

programs, which were developed in the beginnings of 90s and increased gradually from 2000s onwards.

To sum up, it can be argued that ETDs programs at all levels (i.e. institutional, national, regional, and global) are reviving TDs' impact (at least in academic literature). The trend of the percent of open access references from 1996 to 2017 showed that there is a significant increase in the number of open access ETDs through ETD programs. However, surveying authors citing TDs in their works is needed for more clarification of their access to TDs.

The results of our study will provide a better understanding of the real impact of ETDs in terms of citations. This may be helpful for institutions and operators of ETD programs and repositories, in order to assess the academic "return on investment" of their programs and projects and to improve, if necessary, their way to produce, disseminate and exploit ETDs in order to strengthen their contribution to open science.

REFERENCES

- Adhikari, P., and R. Acharya. "Accuracy of references in different speciality medical journals of Nepal." *Nepalese Journal of ENT Head and Neck Surgery*, vol. 1, no. 2 (2010): 23-25.
- Ahmed, Aquil, Sulaiman Alreyaee, and Azizur Rahman. "Theses and dissertations in institutional repositories: an Asian perspective." *New Library World* 115, no. 9/10 (2014): 438-451.
- Al-Benna, Sammy, Prachi Rajgarhia, Safraz Ahmed, and Zeeshan Sheikh. "Accuracy of references in burns journals." *Burns*, vol. 35, no. 5 (2009): 677-680.
- Bangani, S. "The impact of electronic theses and dissertations: a study of the institutional repository of a university in South Africa." *Scientometrics*, vol. 115, no. 1 (2018): 131-151.
- Bornmann, L. "What is societal impact of research and how can it be assessed? a literature survey." *Journal of the American Society for Information Science and Technology*, vol. 64, no. 2 (2012): 217-233.
- Buehler, M. A. *Demystifying the institutional repository for success*. Oxford: Chandos Publishing. Oxford: Chandos Publishing (2013).
- Davies, Karen. "Reference accuracy in library and information science journals." *Aslib proceedings*, vol. 64, no. 4 (2012): 373-387.
- Ezema, Ifeanyi J., and Harriet U. Igbo. "Electronic Theses and Dissertations and the Promotion of Scholarship in African Universities: Prospects and Challenges." *International Information & Library Review* 48, no. 1 (2016): 21-30.
- Ferreras-Fernández, Tránsito, Francisco García-Peñalvo, José A. Merlo-Vega, and Helena Martín-Rodero. "Providing open access to PhD theses: visibility and citation benefits." *Program* 50, no. 4 (2016): 399-416.
- Harinarayana, N. S., and Raju N. Vasantha. "A study of citation accuracy in psychology theses submitted to the University of Mysore." *Annals of Library and Information Studies*, vol. 58, no. 4 (2011): 326-334.
- Lopresti, Robert. "Citation accuracy in environmental science journals." *Scientometrics*, vol. 85, no. 3 (2010): 647-655.
- Lukić, Ivan Krešimir, Anita Lukić, Vicko Glunčić, Vedran Katavić, Vladimira Vučenik, and Ana Marušić. "Citation and quotation accuracy in three anatomy journals." *Clinical Anatomy: The Official Journal of the American Association of Clinical Anatomists and the British Association of Clinical Anatomists*, vol. 17, no. 7 (2004): 534-539.
- Rahmani, M., Asnafi, A. Erfanmanesh, M. "A Survey on citation accuracy of thesis in Faculty of Psychology and Educational sciences in Shahid Beheshti University." *Iranian Journal of Information Management*, vol. 2, no. 1 (2016): 97-115.
- Ratanya, F. C.. Institutional repository: Access and use by academic staff at Egerton University, Kenya. " *Library Management*, vol. 38, no.4/5 (2017): 276-284.