

Introduction

Open access publishing appears to be an important value for the worldwide ETD community. The term 'Open Access' has been a prominent theme of most every international ETD conference since 2004, and appears in the titles of numerous presentations and papers shared at these conferences. The importance of open access ETD's has been discussed in numerous threads on the international ETD-L list, and touted on the web pages of the Networked Digital Library of Theses and Dissertations (NDLTD). The opening lines of the NDLTD website (2012) state "We support electronic publishing and open access to scholarship in order to enhance the sharing of knowledge worldwide." Moreover, the *ETD Guide* produced by NDLTD leaders states in its "Why ETD's" section: "The main goals of the ETD initiative are "...for universities and graduate students to more effectively engage in open access electronic scholarly communications" (NDLTD, 2011). It remains to be seen, however, whether the widely-held community value for open access ETD's has actually translated into practice. Has the period governed by the ETD movement (1998-current) seen an increasing trend toward OA-published ETD's? Little research has yet been conducted to answer this question.

To address that gap in knowledge, the author is assessing the state of open access publishing for ETD's. The current paper reports on a preliminary study to measure the extent to which North American theses and dissertations are being published via open access as defined by the Budapest, Bethesda, and Berlin Open Access declarations (Suber, 2006a). The findings of this early, small-scale study begin to shed light on the larger question of Open Access ETD publishing, with clear data reflecting very low uptake of BBB-compliant OA publishing in North American ETD's. The reasons for this trend, and some strategies for addressing it, are provided at the end of this paper.

Background

The Budapest, Bethesda, and Berlin declarations, known colloquially as BBB, represent three international public statements on open access publishing that form the foundation of the worldwide Open Access movement. According to professor and scholar Peter Suber, editor of the *SPARC Open Access Newsletter*, BBB represent community consensus on the definition of

the term open access and have "unparalleled stature and influence within the OA movement" (Suber, 2004).

According to the BBB definition, Open Access publishing fulfills two essential requirements by providing:

- Access to the full content of the work at no cost to users by removing monetary barriers
- Access to the full content in order to reproduce, adapt, distribute, display or perform that content, by removing permission barriers

The first element of BBB Open Access – cost-free access to scholarly literature on the Internet – is the most familiar and commonly implemented version of OA in scholarly publishing today. There are myriad monographs, serials and other scholarly works published without charge to the user, while remaining fully governed by copyright restrictions with "All rights reserved" by the owner. Such a circumstance equally holds true within the North American ETD community. For example, the *Theses Canada* portal, designed to serve as "the source to find Canadian theses," states as one its goals "To provide free access to as many Canadian electronic theses and dissertations as possible" (Library and Archives Canada, 2012). The Texas Digital Library's *Federated ETD collection*, one of several regional ETD portals in the US, advertises that the service enables Open Access to scholarship (TDL, 2012). Similarly, the OhioLINK *ETD Center* describes itself as "a free, online database of Ohio's masters and doctoral theses and dissertations from participating OhioLINK member schools." (OhioLINK, 2012). On the commercial side, the Open Access ETD database offered by dissertation distributor ProQuest/UMI -- *PQDT Open* -- "provides the full text of open access dissertations and theses free of charge." The company's website elucidates its vision that "Open Access is a term used to describe content that a reader can access free of charge." (ProQuest, 2012). In all of the cases described above, it is evident that North American ETD distributors consider the definition of Open Access to comprise the cost-free element of the BBB OA definition.

The second element of BBB OA refers to copyright and other restrictions that prevent legitimate scholarly re-uses of an OA work, including activities such as text mining and translation. Suber points to the Budapest statement as offering a complete vision for OA (2011):

By "open access" to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited."

Suber acknowledges that the second element of OA described above is harder to explain and to understand, resulting in considerably lower uptake than the first BBB element. In the journal literature, for example, only about 10 percent of research papers are published in open access journals according to a recent study by investigators at CERN, the European Organization for Nuclear Research (Dallmeier-Tiessen, 2011). Many OA journals are BBB-compliant because they not only remove price barriers (user access fees) but also permission barriers. Specifically, they require the authors to distribute their OA articles using the most liberal of Creative Commons licenses, CC-BY, permitting almost any form of reuse for scholarly and even commercial purposes as long as attribution to the original and author are provided.

While it is not surprising that low uptake of BBB-compliant OA in the scholarly publishing community would spill over to ETD publishing at this point in time, Suber has argued that there is no rationale for this circumstance. Speaking at the 9th International Symposium on ETDs in 2006 in Quebec City, Suber (2006b) pointed out that theses and dissertations are a perfect fit for OA-publishing because they are sufficiently rigorous to make them valuable and worth disseminating and using; and sufficiently unconventional to avoid a publisher's investment and barriers to OA. In his presentation to the ETD community, he underscored the importance of distributing these valuable and unique works of scholarly in a manner compliant with BBB OA.

Apparently, Suber's argument found resonance with the NDLTD, which now defines open access in terms well aligned with the BBB definition. As represented in its document *ETD Terms and Definitions*, NDLTD (2010, p. 8) defines Open Access within the context of ETD publishing as:

“Information readily available on the Internet at no cost and without access restrictions... Open access (OA) literature is digital, online, free of charge, and free of most copyright and licensing restrictions.”

One can infer from the definition published in *ETD Terms and Definitions* that NDLTD shares the vision of the worldwide OA movement. In the NDLTD vision, access to scholarly literature, including ETDs, must be free not only of cost barriers but of permission barriers as well. Whether ETD authors, their universities and ETD distributors have put into practice the principles of BBB/NDLTD OA is the driving question underlying this study.

Research questions and methodology

Are North American ETD's fulfilling the principles of the BBB and NDLTD vision for OA publishing? To address this question, a methodology has been developed to measure the number of North American ETDs published since 1998 without cost or permission barriers. The first element means that the ETD must be fully free for users to access without payment or authorization. The second element means that the copyright owner of the ETD has granted permission for users to reuse the work in some capacity as long as proper attribution is given to the ETD author and his/her work.

A search strategy was designed to be run in one or more databases that index and distribute at least some portion of North American ETD's via Open Access. Three databases were identified that met the criteria: *Theses Canada*; *NDLTD Catalog*; and *ProQuest PQDT Open*. However, limitations in search functionality precluded use of the first two databases in this preliminary phase of the investigation. In *Theses Canada* and the *NDLTD Catalog*, there is no way at present to limit the search to the necessary parameters detailed in Table 1. For that reason, the preliminary searches were conducted in *PQDT Open* which, by definition, meets the first element of the BBB/NDLTD definition of Open Access because it charges nothing to access the citations, abstracts and full text of the ETD's held there.

Table 1. Search strategy devised to identify North American OA ETDs

	Include	Exclude	Notes/Assumptions
Document Type	Born digital	Digitized legacy	Look for evidence that thesis or

	theses; Born digital dissertations	theses or dissertations	dissertation was scanned to digital from an analog copy. Anything predating 1998 automatically excluded
Publication date	1998 or later	1997 or earlier	1998 represents beginning of ETD movement
Geography	United States Canada	English language but not produced in US or Canada	
Cost-free	Any work to which an anonymous (public) user can gain automatic access to view citation and abstract or the full text	Any work for which an anonymous (public) user must authenticate or remit payment in order to gain access	Items under embargo, where only a citation and abstract were viewable, not excluded because it was assumed that when the embargo period ended, the item would be provided without charge
Permission-free	Any work to which an open license (Creative Commons or GNU) was attached, or any work for which the copyright statement included the phrase "Some rights reserved."	Any work for which the copyright statement included the phrase 'All rights reserved' and/or to which no open license (Creative Commons or GNU) was attached. copyright	Items under embargo, where only a citation and abstract were viewable, not excluded if the search detected the presence of a Creative Commons or other open license. ----- Items held within a database that posted system-wide copyright notices with "All Rights Reserved" or which indicated usage was subject to licensing conditions that prohibited reuse rights were still included in the study with the assumption the author's OA statement

		statement included the phrase "Some rights reserved."	overrode the publisher's copyright claims.
--	--	---	--

The specific search string applied in *PQDT Open* was designed to limit ETD's to those born digital from 1998 forward from either Canada or the United States:

```
sch(SCHLOC(united states) or SCHLOC(canada)) and PDN(>1998) and PDN(<2012) and ("creative commons" or "some rights reserved")
```

The phrases "Creative Commons" or "Some rights reserved" appeared in the full-text file operating behind the page images of the document. This feature was apparent for both ETD's that were available in full text, and those for which the full-text was restricted by an embargo.

Results

The total number of open access US theses and dissertations published since 1998 in *PQDT Open* equals 15,382 titles. The total number of open access Canadian theses and dissertations published since 1998 in *PQDT Open* equals 2 titles. It can therefore be concluded that at least 15,384 North American ETD's have been published without cost barriers.

How many of these also conform to the second element of BBB/NDLTD OA – removal of permission barriers? Based on the occurrence of either the term "Creative Commons" or "Some rights reserved," a total of 115 BBB/NDLTD ETD's were found in *PQDT Open*. However, many of these were actually governed by 'All rights reserved,' and contained the string 'Creative Commons' as a passing reference in the body of the text. Only 44 ETD's in the results set actually were distributed with a Creative Commons or other open license allowing some reuse rights. That number represents less than half of one percent (.0005) of the ETD's in the database.

Further examination of the open licensing terms attached to the 44 BBB/NDLTD-compliant OA ETD's in *PQDT Open* reveals a variety of re-use rights granted to users, represented by the following types of Creative Commons licenses: (Example ETD's with each type of license are illustrated below in Figure 1a-d):

- CC BY-ND, meaning that users may reproduce, distribute, display or perform the work with attribution, but no derivative works may be produced from the ETD
- CC BY-NC-ND, meaning that users may reproduce, distribute, display or perform the work with attribution, but not for commercial purposes, and no derivative works may be made
- CC BY-NC-SA, meaning that the above-listed reuse rights are permitted, as well as the right to make derivative works, for all but commercial purposes. And, in the case of derivative works, they must be distributed via OA in the same manner as the original ETD.
- CC-BY, meaning that all types of reuse of the ETD are permitted – even commercial— with attribution.

Figure 1a. ETD with CC BY-ND License in *PQDT Open*

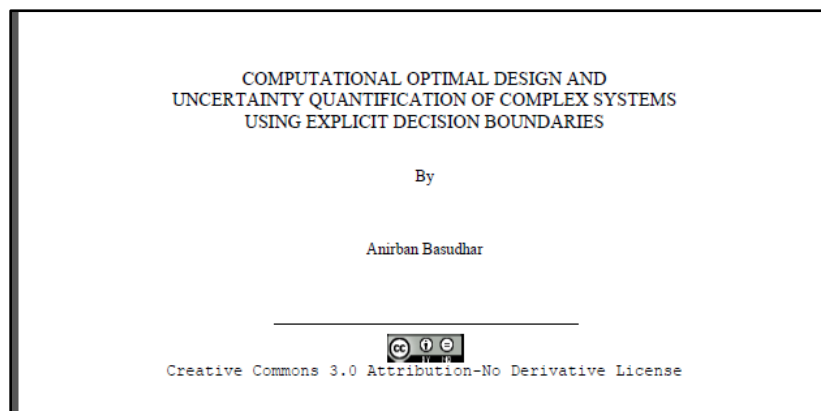


Figure 1b. ETD with CC BY-NC-ND License in *PQDT Open*

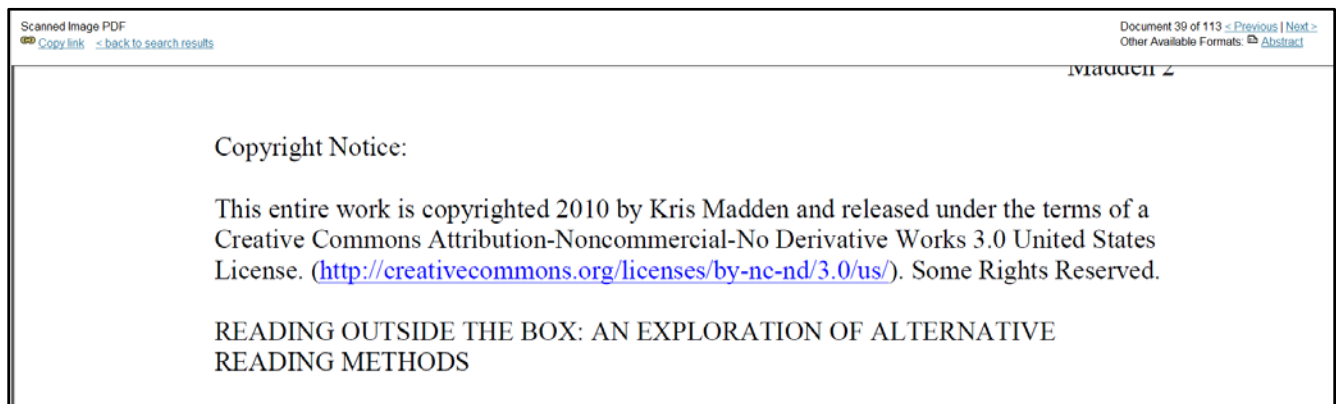
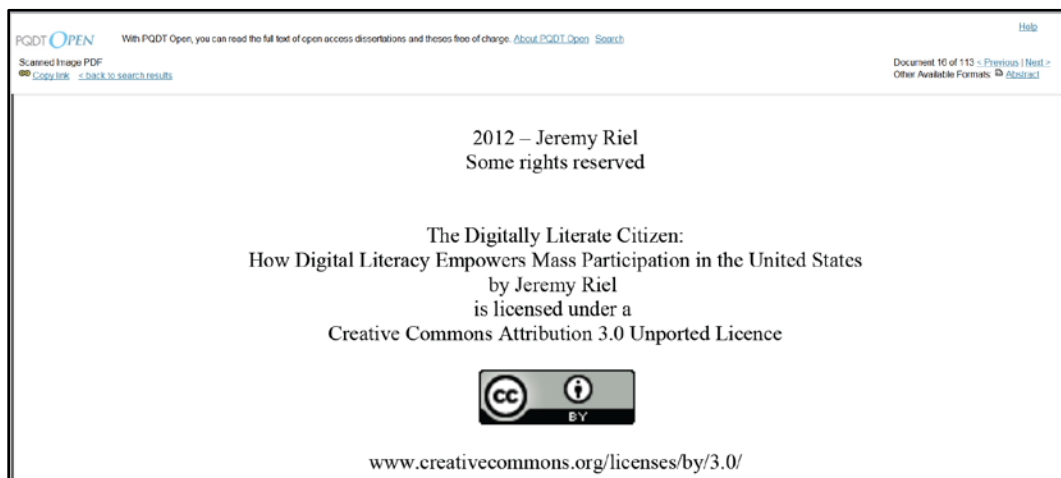


Figure 1c. OA ETD with CC BY-NC-SA License in *PQDT Open*



Figure 1d. OA ETD Licensed with CC BY License in *PQDT Open*



Discussion

The results of this study show that over 15,000 North American graduate students have chosen to publish their graduate works via free Internet access through the *PQDT Open* database over the last decade and a half. They have each willingly paid an author's publishing fee of \$95 to provide this form of Open Access for their users. This practice of "author side fees" is becoming increasingly popular in journal publishing. What's more, a tiny fraction of these ETD authors

have also chosen to remove permission barriers, providing for liberal reuse of their graduate scholarship by users around the world. This is encouraging data.

What is less encouraging, however, is that the BBB/NDLTD Open Access choices made by the 44 ETD authors are undermined by the publisher itself. For each of these 44 ETDs, Proquest has added a pop-up warning that the work is copyrighted by the author, with "All Rights Reserved." This warning is in addition to the company's own copyright notice stamped into every ETD included in the database.

Of equal confusion and concern for the 44 BBB/NDLTD-compliant ETDs is that the vendor license associated with *PQDT Open* prohibits a range of re-uses that are explicitly permitted by the ETD author through Creative Commons licenses. For example, the *PQDT Open* license explicitly limits legitimate scholarly reuses such as printing and redistribution of the ETD, in direct contradiction to the re-use rights explicitly permitted by the ETD authors.

What is the legal effect of the ProQuest license and their conflicting copyright notices superimposed on the author's own statement of user rights? This question is beyond the ken of the author. But it is important to point out, strictly from a customer service perspective, that the assertion by the database provider of rights they do not have is careless and misleading. It may also chase away prospective OA authors interested in publishing their ETD's through *PQDT Open*.

The requirement of an author-side fee and the imposition of the publisher's threatening, if not legally actionable, rights statements may be deterrents to OA ETD publishing in *PQDT Open*. To address this concern, it would be advisable for the company to change their practices to align with those of journal publishers implementing author-side fees, whereby the works are distributed under the CC-BY license and the publisher asserts no rights of its own.

CONCLUSION

The focus of this preliminary study brought to light factors working against BBB/NDLTD OA publishing of ETD's within the *PQDT Open* database. It is not known if similar issues also exist in the other OA ETD databases containing North American ETD's such as *Theses Canada* and the *NDLTD catalog*. Continued investigations within this research program will attempt to

measure the BBB OA uptake in those systems, in comparison with the one investigated in this paper.

Should similarly low uptake of BBB/NDLTD uptake be found in the other databases, it will be important to look at additional factors contributing to this circumstance. Is the uptake higher at certain institutions, or in certain disciplines, suggesting a higher awareness level or greater support for OA publishing in those sectors of academe? Or is the degree of uptake increasing over time, suggesting that community acceptance of reuse rights as the norm for scholarly publishing is widening? It is hoped that additional data from a more representative slice of ETD publishing systems will provide sufficient evidence to fully assess to what extent BBB OA is finding application in North American ETD's.

REFERENCES

- Dallmeier-Tiessen, Suenje *et al.* (2011) *Highlights from the SOAP project survey. What Scientists Think about Open Access Publishing*, doi: arXiv:1101.5260v2.
- Library and Archives Canada. (2012). "About Theses Canada," *Theses Canada* website. Retrieved from <http://www.collectionscanada.gc.ca/thesescanada/027007-5000-e.html>.
- NDLTD, 2012. "Networked Digital Library of Theses and Dissertations (NDLTD)." Retrieved from <http://www.ndltd.org/>
- NDLTD, 2011. "ETD Introduction –Why ETDs" In *ETD Guide*. Retrieved from http://curric.dlib.vt.edu/wiki/index.php?title=ETD_Guide:Introduction_-_Why_ETDs%3F.
- NDLTD. (2010). ETD Terms & Definitions List. Retrieved from http://www.ndltd.org/resources/Definition_of_ETD_Terms_6_10_2010_NDLTD.pdf
- OhioLINK. (2012). "Electronic Theses and Dissertations Center (OhioLINK)", *OhioLINK* web site. Retrieved from http://www.ohiolink.edu/resources/show_details.php?db=x_etdc.
- ProQuest. (2012). "ProQuest Open Access Publishing PLUS," *ProQuest* website. Retrieved from <http://www.proquest.com/en-US/products/dissertations/epoa.shtml>.
- Suber, Peter. (2011). *Budapest Open Access Initiative: Frequently Asked Questions*. Retrieved from <http://www.earlham.edu/~peters/fos/boaifaq.html#literature>.
- Suber, Peter, (2006a). "Open access to electronic theses and dissertations (ETDs)", *SPARC Open Access Newsletter*, July 2 2006. Retrieved from <http://www.earlham.edu/~peters/fos/newsletter/07-02-06.htm>.
- Suber, Peter. (2006b). "Open Access for ETDs: Unlocking Scholarly Access: ETDs, Institutional Repositories and Creators", *9th International Symposium on ETDs*, Quebec City, June 7-10, 2006 . Retrieved from <http://www6.bibl.ulaval.ca:8080/etd2006/pages/prog.jsf>.
- Suber, Peter. (2004). *SPARC Open Access Newsletter*, issue #77, September 2, 2004. Retrieved from <http://www.earlham.edu/~peters/fos/newsletter/09-02-04.htm>
- TDL. (2012). *Texas Digital Library* website. Retrieved from <http://repositories.tdl.org/tdl-ir/handle/2249.1/9387>.