

To Restrict or Not to Restrict Access: The PhD Candidate's Intellectual Property Dilemma

Kimberly Douglas

California Institute of Technology

kdouglas@caltech.edu

MC 1-43, Pasadena, CA 91125

<http://library.caltech.edu>

Keywords: copyright withhold access theses patent publishers scholarly communication

Abstract

The NDLTD software allows PhD candidates to restrict access to their thesis or portions thereof on different levels. At the same time there is an increasing demand that research results be openly available. In fact, such has been re-affirmed as a core value of the research community (See AAAS 2002 Report on Authorship). At Caltech in the U.S., a study was undertaken to examine the motivations and impact that the restriction decision had for the earliest submitters, the group of PhD candidates who, in 2001-02, voluntarily added their theses to the newly created Caltech ETD repository.

Problem Statement

The last five years have been characterized by continuing escalation of concern and alarm regarding the cost of scholarly journals and resulting barriers to scholarly communication.¹ Graduate students now completing their PhDs have been exposed to the discussion of alternatives for scholarly communication and they are of a generation that has enjoyed the benefits of the World-Wide Web throughout their post-secondary education.² The behavior of this class and future graduates will undoubtedly be instrumental in the creation and subsequent success or failure of any initiatives in the exchange and sharing of scholarly information.

To gather more factual information about the evolving behavior and underlying motivations of the next class of researchers, this paper looks at the decisions of recent PhD candidates regarding restriction of all or a portion of their thesis from world-wide access.

Methodology

The Caltech candidates in 2001 and 2002 who voluntarily submitted an electronic thesis, presumably indicating a

desire for world-wide distribution but who restricted all or partial access, were asked the following questions in January 2003:

1a. What was the specific concern that you addressed by restricting access to your ETD to within the Caltech network only?

Or

1b. What was the specific concern that you addressed by withholding access to your ETD.

2. Was the action to restrict (or withhold) a joint decision with your thesis advisor?

3. Do you have a date in mind as to when you will release access to your thesis? If so, what date? If not, any reason?

4. Please feel free to make any other comments regarding the treatment of your electronic thesis.

Data

Of the 63 theses that were voluntarily submitted to the Caltech ETD from the graduating classes of 20001 and 2002, two (2) were withheld and fifteen (15) restricted portions or all of their thesis to just the Caltech community.

The ETD-db application allows a candidate, when submitting a thesis, to submit but withhold presentation and access to the thesis via the public interface. In this instance the existence of the thesis is not listed in the browse screens nor is the metadata included in the searching algorithm. The thesis is totally hidden. A candidate may alternatively choose to limit access to the thesis to a specific IP range established by the ETD System administrator. The range is usually that of the school maintaining the thesis repository and grantor of the PhD degree. In this case, the existence of the thesis is presented publicly. A researcher may discover the thesis by any number of searching tools available and will be able to view the service or front-matter page. However, when

¹ The high cost of scholarly journals is well-documented in a number of studies over the last 5 years. Especially see . Recent studies such as that of the AAAS, see, refer to the high cost of scholarly journals while also exhorting the research community to take back control in order to advance the sharing models that work for them.

² Current graduating PhD's are about 27 years old. In 1994, the advent of the World-Wide Web they were 19-20 yrs. old and were most likely juniors in college.

the icon for the pdf is clicked the researcher will not be able to access the file if he or she is working at a computer not registered within the specified IP address range. In a variation on that model, candidates can package their thesis into different files which may each have a different access configuration. In that case, it's possible to restrict

access to only a portion of a thesis rather than to the whole work.

The subject areas for this group of restricted or withheld theses at Caltech breaks down as follows:

	Total Theses Submitted	Withheld	Restricted Whole Thesis	Restricted Portion of Thesis	Total Withheld or Restricted Theses	Survey Respondents
Biology	9	0	3	0	3	0
Chemistry	12	2	2	1	5	3
Engineering	22	0	4	1	5	4
Geology	8	0	1	1	2	1
Mathematics	1	0	1	0	1	1
Physics	11	0	1	0	1	0
Total	63	2	12	3	17	9

All 17 students were individually contacted by email. Nine (9) responded. There was no follow-up with the 8 students who did not respond. One Engineering thesis was removed from consideration in this study when it was determined that only the source file, the .tex file had been restricted. This had been an error in the submission process. The .pdf file of the thesis was fully accessible. There had been no response from the biology or physics candidates. That left 8 responses for analysis.

Advisor involvement in the decision to withhold and restrict the theses were:

Chemistry yes (3)
 Engineering no (3)
 Geology yes (1)
 Math no (1)

The reasons for withholding or restricting access to the theses were:

- 2 - Fear of theft of unpublished work involved in a patent process
- 4 - Waiting for publication of peer-reviewed paper
- 1 - Couldn't remember why.
- 1 - Did not have permission to publish data in electronic format

When will thesis be fully released:

- 3 Chemistry - 1 year / after patent is awarded / when paper is published
- 3 Engineering-5 years "to be safe" / can be done now / NA
- 1 Geology - in a few months
- 1 Math- in March or April 2003

Discussion

Clearly the sample and data in this study are so small. Observations and conclusions must naturally be modest and circumspect. Yet there are some observations can be made that provide direction for student and faculty education regarding electronic theses and how such programs should differ depending on the subject area of the thesis.

During the period in question only two chemistry theses were completely withheld from visibility on the network out of the seventeen voluntarily submitted but with access restrictions. It is interesting to note that both theses came from the same research group: A team that seeks new polymeric structures which have useful mechanical, electronic or optical properties. Clearly this is an area of patent potential and advisors are directly involved in the handling of the information in such theses.

PhD candidates may, in fact, have conflicting objectives. On the one hand, as budding researchers, the need to share results is an instinctive objective. On the other, there may be intellectual property that must be protected in order for the research group members to fully benefit from their work. One student described this dilemma in his response. Theses, at least a properly written theses, synthesize known information from separate and distinct published and known sources with the new observations generated by the students' research to tell a complete story. Traditionally, the availability of a thesis could lead to patent difficulties and competitors unfairly benefiting by selectively citing previous work. The grad student's response is to create electronic theses to accomplish two goals: One - to make the thesis work more available to the general public and Two - to make it more available to coworkers in the field.³ Thus, some students do see uni-

3 Quoted from survey response.

versal access and the high visibility it generates as a protection against unfair use.

In other subject areas, it appears that students took it upon themselves to decide whether and how much to restrict access to a thesis. Most of the decisions to limit the access to a thesis were done with little or no conscious understanding of author's rights. Usually the student made the decision based on a perception of the safest conservative option with the intent to avoid any possible later difficulties. Since the thesis could and would eventually be released the decision was not a binding restriction and therefore posed low if any downstream risk.

Of the 63 theses voluntarily submitted only 27% were restricted in some way. This low number might strengthen the thought that graduate students are biased to making their work openly available. However since the sample of the theses were those voluntarily submitted it is a group biased in that direction. It will be more instructive to calculate this number for the 1st class in which the etd is required at Caltech.

Conclusion

In Chemistry there is a clear need to include information about the patenting process in electronic thesis education programs. Working with the campus Technology Transfer organization to develop the content of such a program would allow the library to insure that students and faculty were receiving the same message. Faculty would gain a understanding of the etd environment.

Candidates, as authors have certain intellectual property rights. It does not appear to be generally known or

at all understood that the factual results of research are not copyrightable. Only the expressions of the facts in different publications may be protected by copyright. Therefore, a candidates expression of their research results in the thesis can and should be different from that of the formally published peer-reviewed paper. In the United States, a more in-depth understanding of this at an early stage of the thesis preparation would appear to allow a prompter release of full theses

Another approach developed at the University of Cincinnati is to provide a portal to an Academic Journal Policy Database.⁴ When more publishers contribute their information it will certainly provide a welcome resource. However, such an approach puts the understanding of copyright completely in the hands of the publisher. With theses, librarians have an excellent opportunity to educate the next generation of researchers and authors as to their rights and the consequences of their decisions.

Overall the ETD effort would be enhanced by increased proactive management of the option to withhold or restrict access of theses whether by advising and counseling candidates or even by active monitoring of the length of time that a thesis remains hidden from global access.

Bibliography

- BRANIN, JOSEPH J., MARY CASER : Reforming Scholarly Publishing in the Sciences: A Librarian's Perspective
In: Notices of the AMS 45/ 1998 4 P. 475-486
- FRANKEL, MARK S. : Seizing the Moment: Scientists' Authorship Rights in the Digital Age Washington, DC: American Society for the Advancement of Science Directorate for Science & Policy July 2002

4 See Doug Hott's work at <http://www.etd.uc.edu/journal/>