

## Lessons of a thousand: two years of required electronic dissertation submission at the University of Texas at Austin

Tim Brace

Georgia Harper

University of Texas at Austin

[timbrace@mail.utexas.edu](mailto:timbrace@mail.utexas.edu)

[GHARPER@utSYSTEM.edu](mailto:GHARPER@utSYSTEM.edu)

Main Building 101, University of Texas, Austin, TX 78712

[www.utexas.edu/ogs/](http://www.utexas.edu/ogs/)

Keywords: Brace, Harper, copyright

### Abstract

*The University of Texas at Austin has required electronic dissertations from its graduating doctoral students since the summer of 2001. During that time, over 1,000 etds have been submitted, with topics spanning the gamut of almost 100 disciplines, and formats ranging from traditional PDF versions of word processing documents to non-linear etds consisting of text and multimedia files embedded within html pages. Along the way, numerous problems have arisen (as one can imagine). This presentation will focus on technical and legal issues that may be present with traditional paper dissertations, but have become larger issues with etds. These include copyright and other issues associated with the university's becoming a digital publisher, working with traditional dissertation publishing companies, fonts as proprietary software, formatting issues, implications for future publication of all or part of an etd, and ownership and proper use of data and information previously published. Issues themselves as well as principles used to approach solutions will be discussed. This presentation is appropriate for those considering etd implementation, as well as those whose programs are already implemented.*

### Preface

The University of Texas at Austin has required electronic dissertations (etds) from its graduating doctoral students since the summer of 2001. Since then, over 1,000 etds have been submitted, with topics running the gamut of almost 100 disciplines, and formats ranging from traditional PDF versions of word processing documents to non-linear etds consisting of text and multimedia files embedded within html pages. Along the way, numerous problems have arisen (as one can imagine). This presentation will focus on technical and legal issues that may be present with traditional paper dissertations, but have become larger issues with etds. These include copyright and other issues associated with the university's becoming a digital publisher, working with traditional dissertation publishing companies, fonts as proprietary software, formatting issues, implications for future publication of all or part of an etd, and ownership and proper use of data

and information previously published. Issues themselves as well as principles used to approach solutions will be discussed. This presentation is appropriate for those considering etd implementation, as well as those whose programs are already implemented.

My name is Tim Brace, and I am Senior Systems Analyst in the Office of the Dean of Graduate Studies at the University of Texas at Austin. I am not an administrator, nor a librarian. I make my living in the technical field of computers, but my work with the etd initiative at Texas is essentially that of liaison among all the people who do the real work: the administrators, student service staff, students, technical staff, and faculty. I understand each of their roles, and provide the communication necessary across these various groups to get the initiative going and have it succeed. I present my role modestly, but I think such a person or persons can be of great help in getting an etd initiative off the ground and in keeping it flying once it takes off.

The University of Texas is one of the largest graduate schools in the United States, with a graduate population of around 11,000. We are a "full service" university in that we graduate Ph.D.s in a broad range of fields each year, from Classics to Chemistry, from Music to Mechanical Engineering and Mathematics, and from Geology to Germanic Studies. Because of this broad range, we have faced most of the problems - administrative, technical, and legal - that an etd initiative presents. We use ProQuest (formerly UMI) as our primary distributor or publisher of Texas dissertations.

For more information on our administrative processes, how we developed them, and how we maintain them, I refer you to a paper delivered by Terry Kahn at the etd 2002 Conference ([http://www.utexas.edu/ogs/etd/project/etd2002\\_Kahn.pdf](http://www.utexas.edu/ogs/etd/project/etd2002_Kahn.pdf), or go to <http://www.utexas.edu/ogs/rc/OpEd.html> and click on "One Year with Digital Dissertations"). There you will find description of the history of our initiative, our requirements, the principles behind them, and some of the hurdles we made it over in order to make our initiative successful

## Principle-based Administrative Decision-making

We have found it best to have principles drive our technical and administrative decisions. There are as many different problems of a technical nature as there are dissertations, it seems. If you can elevate your dialogue to a discussion of goals and principles, it can help you deal with the minutiae in a more effective way.

For example, one of our principles is that the dissertation be "freely cross-platform readable." In other words, the client machine that accesses the dissertation should be able to do so on any commonly used platform using commonly used SYSTEM software. Applications used to present the dissertation should be freely available. For this reason, we prefer PDF or HTML documents. Any multimedia files should also conform to this principle.

When we developed our initiative, we also had the following goals:

It should involve the least possible change in business processes for our staff; and

It should involve the least possible hassle for our students.

One decision on the level of principles that, once made, cleared up several areas of contention and/or confusion was whether the binary file or the paper printout is the "real" dissertation – the standard, the goal. This decision led to answers to whole sets of questions regarding many formatting regulations. Our doctoral evaluator, early in the process of moving to digital dissertations, often objected to a particular "page" in a dissertation (viewed online) because "it won't look right when it prints out." It took several discussions to realize that we were starting with a different goal in mind. For her, the printout was still the "final" state of the dissertation, to which all questions of formatting referred; for me, the digital file is itself the end product. If the digital file is the final state of a dissertation, it doesn't matter that the viewer might have to "zoom" in or out to see various sections of text or graphic - that's what computers do, and they do it well.

If the final state of a dissertation is paper, for example: Margins are of great significance, as the printout might lead to a bound copy.

The size of each page is significant, as oversize pages might not print appropriately

The issue of landscape view versus portrait could cause trouble, depending on

Your printer's configuration

If the final state is digital, however:

What the file looks like printed out is irrelevant

The standard is based on viewability on a computer monitor; the viewer can take advantage of the scalability of the viewing software.

The current policy at Texas is that the digital file is the final state of the dissertation. We Advise the student to keep some appropriate margins, just in case printing the file is Desired; but adherence to strict rules for margins is less important than with a printed copy. As long as the file is viewable on a computer monitor via means reasonable for the greatest number of potential viewers, we will accept the file.

Though we have had less experience with audio-visual files, we have successfully used the above principles in dealing with the examples we have come across. There are no universally adopted quality standards regarding multimedia files, or for images, for that matter. The digital library at the University of Texas is working on a proposal for standards regarding image quality in dissertations, and again, the Office of Graduate Studies, if it adopts these, will make sure they do not add an unfair burden to the student or our staff. At the same time, because electronic documents are fundamentally different from paper documents in that they can present themselves differently on different "reading machines," we are sympathetic to the need for standards. In this area, the electronic dissertation world is truly still in its infancy.

One issue that involves both technical and legal fields is the use of proprietary fonts in a PDF file. Specialty fonts, such as those used in Classics or in Slavic Studies, are proprietary software developed by companies many of which are, in my opinion and from my experience, ill-prepared for the digital world. Fonts cannot be "stolen" from a paper document, but they can be lifted from PDF files in which they have been embedded. Embedding proprietary software such as this raises copyright issues, and I have found some of these font providers to be very nervous about allowing the embedding of their products in PDF files that might be distributed worldwide. There are technical ways around this problem, depending on the individual situation, and I would be happy to share them with you at a future time via telephone or email.

Other issues we have dealt with take us into the realm of intellectual property. Who owns the data? What are the future publication implications of my having put my electronic dissertation on my web page for free download worldwide? How much of an image, excerpt from a poem or musical score can I include in my dissertation without asking for permission from the owner, and is this changing due to the easy distribution and availability (at least potentially) of electronic dissertations?

To address some of these questions, please allow me to introduce Georgia Harper, Manager of the Intellectual Property Section of the Office of General Counsel of the University of Texas System (which oversees 9 academic campuses and 6 health institutions throughout the state of Texas). Georgia will address electronic dissertation issues specific to intellectual property.

## Copyright for University Publishers

Most institutions require that students "publish" their dissertations, and in some cases, their theses. In the analog world, students published through commercial enterprises that handled all the details. Now, the digital environment offers students and their institutions a chance to accomplish publication's goals without an outside vendor, so it was natural that we would begin to examine how we might do this and how it might improve access to the documents.

As colleges and universities began to consider their options for publishing theses and dissertations electronically (hereafter, "etd" or "etds"), we recognized that there were two models for analyzing the copyright issues associated with the use of others' materials within the documents: either we would be publishing etds in the sense that books, journals, magazines and newspapers were published, that is, we would be participating in commercial publishing, or we would be posting them like other academic papers and project results, on the Web, with limited distribution and usually for a limited time. Since we had a long-standing model for etd publishing in the paper environment, it was natural that we would look to the company that was then known as UMI, in its historical role as the principal "publisher" of etds, to see what they required.

UMI's policy on permissions has probably evolved over time, but at the time we were considering these issues, and currently, UMI requires permission to include others' works in an etd beyond a narrow definition of fair use. See *New Media, New Rights, and Your New Dissertation* by Kenneth D. Crews at [www.umi.com/hp/Support/DServices/copyright/](http://www.umi.com/hp/Support/DServices/copyright/). This seemed to make sense. Publishing is publishing and liability is pretty straightforward as well. As a publisher, an institution could expect to stand in the same shoes as any other publisher, and would have to be willing to take full financial responsibility if an etd were found to infringe someone's rights. Thus, we classified the preparation and distribution of etds as commercial publishing. This was not profitable publishing to be sure, but similar to scholarly nonprofit publishing. In fact, our scholarly presses served as another example for a copyright permissions protocol for this type of publishing, making it easy to conclude that our scope of fair use would contract dramatically for etds.

With this framework established, providing education about copyright for graduate school staff, faculty advisors and graduate students became an important undertaking to limit institutional exposure to infringement liability. We have many educational tools online, including the Copyright Crash Course at [www.utSYSTEM.edu/ogc/intellectual-property/copypol2.htm](http://www.utSYSTEM.edu/ogc/intellectual-property/copypol2.htm) and the Crash Course Tutorial at <http://www.lib.utSYSTEM.edu/copyright/>. We required that those seeking advanced degrees, among others, take the

Tutorial and pass a short test at the end. We also developed a set of fair use guidelines for those who review dissertations and theses for conformance to standards, including conformance to requirements to get permission for the use of others' materials. In short, we adopted commercial standards for fair use; we required permission for everything but short quotations and single isolated images.

Needless to say, this created problems in many cases. Images were the biggest problem: art history dissertations commonly include 100 to 200 images and permissions are impossible to obtain in many cases. Students of architecture and many other social sciences and humanities disciplines produce etds rich in images. Music was also a problem because the music industry is not prepared to provide permissions to include clips in etds; thus, again, permissions are difficult, sometimes impossible, to obtain. The same is true for audiovisuals. Fonts can be very problematic as well, so we advise students to avoid the use of proprietary fonts. In many cases, however, their use is unavoidable. We sometimes find that the cost to license the use of a proprietary font can be prohibitively expensive or the copyright owner may require technological protections that are overly restrictive and burdensome for the student and institution.

The difficulty associated with getting permission to include others' works in certain etds created a circumstance very familiar to us as educators: there are times when a requirement to get permission would make it impossible for us to use certain materials for educational purposes, either because permission is too difficult to obtain, too expensive, or even impossible. But that is when our reliance on fair use is on its firmest footing. In fact, that is the essence of fair use: to make it possible for educators to use others' works without their permission when our use furthers the goals of copyright without undermining the incentive the law provides to authors and their assignees. This understanding led us to a second look at our conclusion that posting etds is the equivalent of commercial or scholarly publishing.

Upon re-examination, we concluded that making etds available even on publicly available web sites is not very much like commercial publication. On the other hand, it is very similar to other typical educational uses. This suggested to us that a strong fair use argument could be made in support of the use of larger amounts and kinds of others' works in etds without permission. Many facts about etds support this conclusion. For example, the images used are generally small and low resolution, sound and audiovisual elements are usually brief and of low quality, all sources are painstakingly documented and in many cases links to sources are provided. These are all practices for which we in academe rely upon fair use. In fact, a US court of appeal recently concluded that even the commercial use of small low resolution images is a fair use. See *Kelly v. Arriba Soft Corp.*, 280 F.3d 934 (9th Cir. 2002).

On the other hand showing how little these distributions resemble commercial publishing, sales statistics from UMI suggest that the average dissertation is purchased 1.7 times. Current usage statistics from our own servers where UT dissertations are available free to the UT community show that, while some students' dissertations are accessed frequently, the vast majority are accessed 1 or fewer times. This results in part from the fact that etds are created to fulfill institutional requirements for degrees, as are other papers and class projects, they are unedited, and they are rarely of a quality that attracts public attention or has commercial value. The UMI statistics quoted above certainly attest to the lack of commercial value.

Further supporting the notion that electronic posting is different from commercial publishing, many academic and scholarly publishers permit their authors to include articles already published, or in some cases, yet to be published, in etds published by institutions. They distinguish this from "commercial publication." (See *authors.eelsevier.com/getting\_published.html?dc=Ci#internet*.)

Nevertheless, some publishers still consider posting to a web site to be prior publication, and this raises another important issue. Institutions should have a procedure to identify students who may have an opportunity to publish commercially, to determine the publisher's policy regarding prior publication in etds, and to protect that opportunity if necessary. The Academic Journal Policy Database (<http://www.etsd.uc.edu/journal/>) provides welcomed assistance to institutions making these determinations and may even encourage policy change at those journals that still consider electronic publication of etds to be prior publication.

Finally, one issue seems to continue to arise from time to time, regarding ownership of the copyrights in etds. All

copyright laws provide that as an initial matter, creators are the authors and owners of their works. In the US, there are only a limited number of ways for an institution to own a copyright in a work created by an individual: it must either employ that individual and the work must be within the scope of his or her employment, it may contract with an individual to create a work and satisfy the rigorous statutory work made for hire requirements associated with contractual commissions, or it must obtain the creator's copyright by assignment. As might be readily determined, in nearly all circumstances, none of these would apply to etds. Thus, any US institution whose policies state that it owns dissertations or theses would not find support for such an assertion in the US Copyright Act. There continues to be confusion on this point, however, both among students themselves, and sometimes among administrators. It seems we must continue our educational efforts, focusing on ownership of etds as well as the application of fair use in this context.

In summary, as a nonprofit educational institution, we have concluded that we may claim a reasonable scope of fair use in etds as student works with a limited distribution and no commercial value, even when they are posted to publicly accessible web sites. We augment this claim with policies and procedures that provide opportunities for university staff, faculty and students to learn about the scope of fair use in etds, establish protection for students who wish to publish all or part of their dissertation commercially and explain the application of the law's ownership principles. Thus, while we will assert our rights to claim fair use, we are cognizant of the responsibility we have to help our community understand the limits of that claim.