ETDs in the 21st Century

Twenty years ago, the first electronic thesis/dissertation (ETD) was required to earn a master’s or doctoral degree, so this seems an appropriate time to look at the ETD’s past and future. We have come a long way from considering ETDs to be a new form of online scholarship and from the early days of “Oh my gosh, the files will be too large to download” or “File formats become obsolete so quickly.” Many of the early goals of ETDs have been realized, including graduate students saving money and libraries saving shelf space. We have seen institutional repositories use ETDs as exemplars of public access to information, and we’re no longer surprised about the number of “visitors” and “sessions” (in Google Analytics terms) received by an ETD. However, we still haven’t seen the majority of ETDs transformed into new media with sophisticated interactive graphics and/or embedded sound and video demonstrating research findings.

Innovative ETDs
On the other hand, some standouts have been recognized through the annual (since 2004) NDLTD (Networked Digital Library of Theses and Dissertations) Awards. “Innovative” ETDs have largely been PDF files accompanied by multimedia (and not just in STEM fields). In the recordings of the NDLTD Awards, the authors point out that voices and visuals often say more than text and that they want to entertain as well as educate. They explain how they want to reach an audience beyond academia and, for instance, show the way a garment moves and the fabric flows. They add that multimodal ETDs can “change learning.” Award-winning ETDs have been accompanied by full-length documentary videos demonstrating health care and human rights and showing parents voicing concerns about their sons’ and daughters’ educational opportunities. Some ETDs have even evolved into enhanced “books.”

Developing Technologies
These outliers had no best practices to follow. The Educopia Institute has helped to remedy that with funding from the Institute of Museum and Library Services (IMLS). For example, the ETDplus Project, a multi-university endeavor, developed guidelines and built tools for graduate students who are creating and submitting works beyond the PDF file, including data sets, videos, websites, and other results that cannot be captured in writing alone. Various activities involve ETD systems development. Since ETD-db, Vireo, and ProQuest ETD Administrator, we haven’t seen ETD-specific systems developed. However, librarians at Texas Tech designed SAVE (Streaming Audio and Video Experience), a tool specifically for handling music ETDs. The IMLS-funded Educopia project ETD Lifecycle Management Tools developed openly licensed micro-services software that can be used independently or incorporated into existing repository workflows.

Digital library service developments led by computer scientists and researchers see ETDs as a vast resource for experimentation in artificial intelligence, information extraction, information retrieval, machine learning, natural language processing, and more. An exemplar is ETDseer, based on CiteSeerX.

Access
From the earliest days, ETDs have also been recognized as a way to advance other agendas, particularly public access to information. Today the many interests of the ETD community are shared through conference programs, publications, and the ETD Forum. Periodically the community hosts discussions and presentations about whether ETDs should have immediate public access or limited embargoes. Restricting access during patent applications is a no-brainer, but many faculty advisors continue to be concerned that allowing online, publicly accessible dissertations will make it more difficult for authors to get published and join the academic ranks of traditionally published faculty. Naysayers deny the studies that show, as well as the editors who have stated in public venues, that an ETD requires an extensive transformation and rigorous peer review before it will be published as a journal article or a book.

In the ETD Forum archives, no topic comes close (in terms of both posts and views) to the discussion of access to ETDs, with almost 100 posts and nearly 900 views in the last four years. The debate continues about how long the embargo should be for an ETD that contains creative writing. There was much buzz when the American History Association recommended six-year embargoes. However, a study by Midge Coates showed that this may have increased the publication of history books based on ETDs by only 4 percent.

Articles
An increasing number of graduate students are publishing articles prior to completing their dissertations and are
including those articles in their ETDs. Some colleges and universities have what is often referred to as “manuscript style” ETDs. The ETD chapters are articles that have been published, are under consideration by a journal, or have been prepared for submission. These require that graduate students have a greater understanding of copyright law, as well as a willingness to read, and not just sign, publishers’ author contracts.

This format raises concerns when an article has more than one author. Graduate school policies at universities such as Southern Illinois and Virginia Tech require that the ETD author’s contribution be clearly described. Of course, policies vary. At the University of British Columbia, multi-authored ETDs are not allowed, but the near-same chapter can appear in more than one thesis if this is clearly explained and each student’s contribution is described. Universities such as Loyola (Chicago) and Central Florida allow multi-authored ETDs from their Schools of Education with EdD programs, which have group doctoral research projects.

Data
Data is trending—not only data accompanying or within ETDs but also their metadata. The peer-reviewed literature about ETDs focuses heavily on ETDs in institutional repositories and on ETD metadata (including cataloging). While good metadata improves the discoverability of ETDs, it is also a significant body of data itself, as Lucas Mak and his colleagues at Michigan State University Library have pointed out. ETDS metadata can be used to measure academic activity within and outside the administrative structure (e.g., departments), and it can reveal related communities beyond the college/university. ETDs can expose disciplinary trends and networks of academic activity, as well as potential research advisor genealogy.

Text and data mining is beginning to appear as a topic in many venues, including those outside the ETD community. For example, at the February 2018 International Digital Curation Conference in Barcelona, my colleague Zhiwu Xie presented “Advancing Policy and Changes for Graduate Data Management” and recommended that graduate students prepare data-management plans so that from the early stage of their research, there will be periodic review-and-revise with their advisors and submission of the data along with the ETD for review before the defense, all as integral components of the students’ plan of study and final exam procedure.

Another change for graduate students is the move of many colleges and universities away from requiring ETD authors to give nonexclusive copyright licenses to ProQuest. There are now many online and searchable ETD databases, including Global ETD Search from the NDLTD, OADT (Open Access Theses and Dissertations), and EBSCO Open Dissertations.

Preservation
Another scholarly communication issue that ETDs have exemplified is long-term preservation. Many who are concerned about digital libraries (including Clifford Lynch and Henry M. Gladney) have pointed out that the difficulty with preservation issues are largely organizational, economic, and institutional. The mechanics are relatively easy, as evidenced by, among others, the MetaArchive Cooperative. Since 2004, when the Cooperative received initial funding from the Library of Congress’s National Digital Information Infrastructure and Preservation Program (NDIIPP), there has been a practical and affordable strategy for ETD preservation available. The NDLTD lists this and many other preservation strategies.

While I celebrate the anniversary of the ETD requirement at Virginia Tech, I know that other colleges and universities are still wondering when, not if, they will require ETDs. Unfortunately, the issues surrounding ETDs continue to be more about an institution’s political climate than about its technical expertise. I’m still longing for the “innovative” ETDs to evolve into the “typical” ETDs, though I do enjoy referring to Timur Oral’s 1997 ETD, in which he incorporated digital video (accessible, though primitive) and lovely color illustrations. Although progress is slow, I’m anticipating a time when the norm for a graduate student’s legacy will be a demonstration of the research, not just words describing it.

Notes

Gail McMillan (gailmac@vt.edu) is Director of Scholarly Communication at Virginia Tech Libraries.

© 2018 Gail McMillan. The text of this article is licensed under the Creative Commons Attribution 4.0 International License.