EDITED BY VIRGINIA KUHN AND ANKE FINGER

SHAPING THE DIGITAL DISSERTATION

KNOWLEDGE PRODUCTION IN THE ARTS AND HUMANITIES

https://www.openbookpublishers.com

Cover image: Erda Estremera on Unsplash, https://unsplash.com/photos/eMX1aIAp9Nw. Cover design by Anna Gatti.

11. Highly Available Dissertations

Open Sourcing Humanities Scholarship

Lisa Tagliaferri

The moment an early career scholar decides how to deposit a dissertation often comes during a fraught period of transition. For those who have options, they must weigh the often unquantifiable benefits of open access against the fear of not being able to turn their dissertation into a viable book for publication to earn tenure. Much of this anxiety stems from concerns about university press policies and copyright with respect to digital publication. Currently, there is not much data around open-access dissertations being published by presses. Additionally, scholars without a publication track record may not be mentally prepared for an additional level of scrutiny from the public after rounds of reviews with their committee. Though dissertations are vetted by important stakeholders in the field, they are often not structurally overhauled by professional editors and polished by copyeditors beyond the writer's own departmental resources. This lack of a final quality assurance process during an already stressful time can cause the dissertator to err on the side of caution and embargo their dissertation for several years if they have the option.[1]

Though I had similar concerns during my dissertation process, I deposited my dissertation in 2017 in the open access repository set up by the libraries of my alma mater (CUNY Academic Works) and licensed my work with a Creative Commons license. Because my dissertation was an interdisciplinary digital humanities project that included source

---

1 Each institutional repository will have its own policies; some repositories may allow embargo while others do not.

code for data visualizations, text mining and a static website, this choice may seem a little more natural, stemming from the fact that the open-source software movement has championed open repositories as a common practice in the tech community for decades. By conceiving of my project as open source, I was also able to think about version control and opportunities for collaboration on a structural level. As I have long been a member of the public university writ large, contributing back to broader publics was an important legacy I wanted to leave following my doctoral training. The licensing was an aspect that I had to think about more significantly. Due to having done a great deal of work in the tech sphere under a Creative Commons license, I was less intimidated than I may have been had I not had that experience. However, I still relied on mentors who had been in the university and publishing space longer than I had, and was fortunate to be able to discuss my options with very knowledgeable people.

That said, open access, open source and Creative Commons licensing is not for everyone. My choices have been grounded in my own particular opportunities and environment, so I will explain the benefits that can be achieved through doing research this way, but believe that every dissertator should come to their own conclusions based on their own experiences, level of comfort and professional goals.

This chapter explores open access dissertations, beginning with contextualizing openness in the humanities and discussing the histories of open source, open access and Creative Commons. The next section will offer guidance for humanities PhD candidates, their committees and institutions. Stakeholders should consider the multiple levels of open access that can be beneficial for dissertations, and they may wish to adopt additional methodologies from open-source software projects. Finally, I will offer my own dissertation as a case study in the interest of sharing data, beginning with an examination of the open access repository where my dissertation was deposited (CUNY Academic Works), and delving into analytics around my dissertation in the interest of data sharing. This chapter engages the ways that increased openness in the humanities can facilitate innovation in the field, and what the benefits and challenges may be from starting a public-facing research career at that vital moment as a scholar: the dissertation defense and deposit.

The Open Humanities

Within the university, there has been a push towards the public humanities recently, including increased engagement with broader communities through digital humanities, and a turn towards openly sharing articles through scholarly repositories in the field. The Modern Language Association (MLA) received several National Endowment for the Humanities (NEH) grants beginning in 2014 to support its open access initiative, Commons Open Repository Exchange (CORE), which is currently in its beta release and open to all humanities fields. The Open Library of Humanities, a non-profit open access publisher for the humanities and social sciences funded by the Mellon Foundation, was launched in 2015 and supports journals including its own multidisciplinary journal. In terms of the dissertation, open access has become integrated into universities' electronic theses and dissertation (ETD) management systems, with an increasing number of institutions, like Duke University and the City University of New York (CUNY), requiring deposit into an open access repository.[2]

Despite the elevated enthusiasm for the public turn of the humanities, it is not a new interest. Humanistic tradition has long participated in open dialogues with the public: dialectic and public disputation during Antiquity transitioned into public scholarly disputation as a form of pedagogy in the Middle Ages.[3] With the movement from humanism into the Renaissance, increased use of the vernacular in both speech and text opened opportunities for marginalized and uneducated communities to participate in broader intellectual participation within a Latinate world. As Jill Cirasella and Polly Thistlethwaite note, the Italian humanist Colluccio Salutati considered that *disputatio* was an essential form of education as it would be 'absurd to talk with oneself between walls and in solitude'.[4] The printing press allowed increased

---

2   Duke University, The Graduate School, 'ETD Availability', *Duke Graduate School*, https://gradschool.duke.edu/academics/theses-and-dissertations/etd-availability; The Graduate Center Library, 'Dissertations and Theses: Deposit Procedure', *Mina Rees Library Research Guides*, https://libguides.gc.cuny.edu/dissertations/deposit-procedure

3   See, for example, Alex J. Novikoff, *The Medieval Culture of Disputation: Pedagogy, Practice, and Performance* (Philadelphia: University of Pennsylvania Press, 2013), https://doi.org/10.9783/9780812208634

4   Jill Cirasella and Polly Thistlethwaite, 'Open Access and the Graduate Author: A Dissertation Anxiety Manual', in *Open Access and the Future of Scholarly*

dissemination of ideas, which would eventually change the course of scholarship through enabling the journal article and monograph as vehicles for knowledge production. In the nineteenth century, Humboldt University led the modern conception of the PhD that would later be adopted at institutions in the United States. From these beginnings, Kyle Courtney and Emily Kilcer explain, 'at the heart of the conferral of the PhD is an affirmation of an individual's substantial, original, and *public* contribution to one's field'.[5]

Open Source, Open Access and Creative Commons

While the humanities have historically tended to keep public engagement and a commitment to openness at its core, more recently software developers and computer scientists have led the way towards open collaboration that has been facilitated through the spread of the internet. Through sharing source code (the lines of code that implement a program), open source promotes public access to programming files, blueprints, and documentation. Open-source software is freely available to use and redistribute, welcomes collaboration by others, and has been integral to the history of computation.

From the 1950s through the early 1970s, most software was produced in collaboration between academia and industry. This software was typically shipped as public-domain software with source code included, which was necessary in order to implement the software on specific machines.[6] However, computer programs were declared to be protected by copyright within the United States in 1974, causing a decline in publicly available source code.[7] With the rise of proprietary software,

---

Communication: Implementation, ed. by Kevin L. Smith and Katherine A. Dickson (Lanham: Rowman and Littlefield, 2017), pp. 203–24 (at 206), https://www.grad.miami.edu/_assets/pdf/open-access-and-the-graduate-author_-a-dissertation-anxiety-manua.pdf

5   Kyle K. Courtney and Emily Kilcer, 'From Apprehension to Comprehension: Addressing Anxieties about Open Access to ETDs', in *Open Access and the Future of Scholarly Communication: Implementation*, ed. by Kevin L. Smith and Katherine A. Dickson (Lanham: Rowman and Littlefield, 2017), pp. 225–44 (at 226, emphasis mine).

6   Erik von Hippel and Georg von Krogh, 'Open Source Software and the "Private-Collective" Innovation Model: Issues for Organization Science', *Organization Science*, 14.2 (2003), 209–23, https://doi.org/10.2139/ssrn.1410789

7   Jan L. Nussbaum, 'Apple Computer, Inc. v. Franklin Computer Corporation Puts the Byte Back into Copyright Protection for Computer Programs', *Golden Gate*

developers began to create free/libre and open-source alternatives beginning in the 1980s. Richard Stallman and Linus Torvalds became central figures among developers working to make non-proprietary alternatives, developing the GNU Project and Linux, respectively.[8] Serving as the foundation of most web servers and the basis of Android smartphones, Linux has tens of millions of users worldwide.

As the internet promoted collaboration and transparency in the computer science field through open source, the sciences at large began to evaluate their mechanisms of exchange in response to networked computers offering the cheap and easy distribution of digital content. Especially with regard to research funded by the public, debate arose around paid models that hinder the dissemination of scientific ideas.[9] The open access movement began in the 1990s, and continues to proliferate throughout scholarship in STEM, the humanities and social sciences. Open access (OA) refers to the free and unrestricted online access to any published research output, with key statements about open access coming from several open access meetings in the early 2000s.[10]

Along with open access comes the question of licensing. For some, having a less restrictive copyright is more consistent with open values. The popular Creative Commons Attribution license took inspiration from computing's GNU General Public License of the Free Software Foundation (GNU GPL). Founded in 2001, Creative Commons was designed for works that are not software, allowing creators to either make their work available in the public domain, or retain their copyright but make the work free for certain uses and conditions.[11] Among the more restrictive Creative Commons licenses are the Non-Commercial and Share-Alike provisions, allowing creators to keep more rights reserved if that is their preference.

University Law Review, 14.2 (1984), 281–308, https://digitalcommons.law.ggu.edu/cgi/viewcontent.cgi?article=1344&context=ggulrev

8   Kathleen Juell, 'A Brief History of Linux' (October 27, 2017), *DigitalOcean*, https://www.digitalocean.com/community/tutorials/brief-history-of-linux

9   Mikael Laakso et al., 'The Development of Open Access Journal Publishing from 1993 to 2009', *PLoS ONE*, 6.6 (2011), e20961, https://doi.org/10.1371/journal.pone.0020961

10 Cold Spring Harbor Laboratory, Library & Archives LibGuides, 'Guide to Open Access: What Is Open Access (OA)?', *CSH*, https://cshl.libguides.com/c.php?g=474046&p=3243847

11 Creative Commons Wiki, 'History' (2011), *Creative Commons*, https://wiki.creativecommons.org/wiki/History

Open source, open access and Creative Commons seek to make work more discoverable and implementable by broader publics. These each can intersect with the choices that are made during a dissertation process (especially a digital dissertation process). Learning more about the histories of each of these frameworks can support the decision making that needs to be done when depositing and disseminating a dissertation.

How To Make a Dissertation Highly Available

With an understanding of open source, open access and Creative Commons licensing, a dissertator can make an informed decision about how best to share their work. In addition to learning about each of the frameworks, they would be well served to have a sense of how each of these is being implemented by scholars, and how open-access works are being received in their given field. Unfortunately, there is not a vast body of quantitative research in this area, but important future work should come through increased interest.

Assuming that a doctoral candidate wants to make their dissertation highly and widely available, they will want to deposit it in an open-access repository if they are able to do so. Many universities in the United States currently have this deposit option available, and a soon-to-be PhD graduate should check in with their institutional library about what their options are in terms of open access. In some cases, open-access repositories may enable a time-bound embargo, preventing public readership of the dissertation until a certain date. In a study of University of Salamanca theses deposits from 2006–2011, researchers found that, across knowledge areas, only the humanities fields had fewer open-access deposits than non-open access.[12]

Opting for an embargo period is especially common in humanities fields because the academic monograph is exceptionally tied to obtaining a long-term job (traditionally in the form of a tenure-track position) or receiving tenure. The issue of the embargo is made more urgent due to the early career precarity for scholars in the current

---

12   Tránsito Ferreras-Fernández et al., 'Providing Open Access to PhD Theses: Visibility and Citation Benefits', *Program: Electronic Library and Information Systems*, 50.4 (2016), 399–416 (at 408). https://doi.org/10.1108/prog-04-2016-0039

climate—there are exceptionally more PhDs produced in the United States than there are full-time academic positions (including tenure-track and non-tenure track roles, as well as term-limited fellowships). For example, in the 2014–15 academic year there were 1,145 new history doctorates awarded, and these recipients could immediately compete for the 572 job advertised in the 2015–16 academic year.[13] With tenure-track positions in research institutions still considered to be the metric of success on the academic job market, candidates find themselves in a desperate situation to err on the side of caution in order to not jeopardize their job prospects in any way.

However, the available research demonstrates that allowing the dissertation to be open access upon deposit does not preclude future publishing opportunities. A recent survey found that when considering an open-access dissertation for a manuscript, university presses were generally receptive: 9.8% indicated that these manuscripts are 'always welcome', 43.9% would consider them on a 'case-by-case basis' and 26.8% would like to see dissertations edited to be 'substantially different' prior to consideration.[14] Cirasella and Thistlethwaite expand on these figures: 'Graduate students might initially be alarmed […] but it is important to remember that publishers consider *all* manuscripts on a case-by-case basis. Similarly, just about all publishers expect dissertation-based manuscripts to differ significantly from the original dissertation...'[15] Because open-access scholarship is often rendered more visible, an open-access dissertation may prove to be more attractive to presses for a variety of reasons, some of which can be quantified via an open access platform.

---

13   Robert B. Townsend and Emily Swafford, 'Conflicting Signals in the Academic Job Market for History' (January 9, 2017), *Perspectives on History*, https://www.historians.org/publications-and-directories/perspectives-on-history/january-2017/conflicting-signals-in-the-academic-job-market-for-history; Scott Jaschik, 'The Shrinking Humanities Job Market: New Analysis Finds the Number of Doctorates Awarded Keeps Rising, even as Number of Job Openings Drops' (August 28, 2017), *Inside Higher Ed*, https://www.insidehighered.com/news/2017/08/28/more-humanities-phds-are-awarded-job-openings-are-disappearing

14   Marisa L. Ramirez et al., 'Do Open Access Electronic Theses and Dissertations Diminish Publishing Opportunities in the Social Sciences and Humanities? Findings from a 2011 Survey of Academic Publishers', *College & Research Libraries*, 73.4 (2013), 368–80 (374), https://doi.org/10.5860/crl-356

15   Cirasella and Thistlethwaite, 'Open Access and the Graduate Author', p. 206 (their emphasis).

One way to encourage an academic publisher to consider a monograph is through providing them with a common metric of academic success: citations. Allowing a work to be open access 'increases citation rates by 50 percent or more'.[16] Open-access work that is more discoverable allows it to be more readily cited by future researchers, and helps aggregators like Google Scholar count those citations, showing the impact factor of scholarly work. Other data that can support a manuscript proposal include page views, downloads, and other metrics that open-access repositories often make available for authors. As noted in Tránsito Ferreras-Fernández et al.'s study, this is not the case for works that are not open access: 'OA repositories can obtain information on the use (visibility), and on the citation (impact) of doctoral theses, this information cannot be obtained in the case of theses that are not on OA'.[17] While it is necessary to consider that not all academic publishers may be receptive to monograph adaptations of dissertation work, the increased visibility and potential for impact can assist in launching one's career after graduate school.

Though humanities fields still privilege single-author publications as the primary measure for success, an open-access dissertation with no embargo period can help to facilitate collaboration. While some areas of humanities specialization, like the digital humanities, are more receptive to collaborative work, a movement towards more research partnerships, and interdisciplinary ones, can work to advance innovation across the humanities and scholarship at large. Though this open-access outcome is more challenging to measure, the increased visibility can enable opportunities for not only collaborative research projects, but also for traditional humanities alliances: conference panel submissions, invited talks, invitations to edited volumes and more. Allowing the dissertation to be visible upon deposit will provide the early career scholar with the time that is often needed in academia to establish a reputation and foster professional relationships that can support their development.

In addition to deciding whether or not to make a dissertation open access, the author should determine how to license it. Some universities,

---

16 Hillary Corbett, 'Out of the Archives and into the World: ETDs and the Consequences of Openness', in *Open Access and the Future of Scholarly Communication: Implementation*, ed. by Kevin L. Smith and Katherine A. Dickson (Lanham: Rowman and Littlefield, 2017), pp. 187–202 (at 198).

17 Ferreras-Fernández et al., 'Providing Open Access to PhD Theses', 403.

like Duke, require a Creative Commons license,[18] so candidates will need to understand the terms of their university's submission policies. An all rights reserved copyright is a more cautious choice, and should certainly be considered when thinking through future publishing opportunities. However, reserving all rights is generally seen as being against the ethos of open access and will likely prevent healthy dissemination of the text. Using a more flexible Creative Commons license can enable others to share the text to various degrees, encouraging collaboration, idea building, and future research in the specific area of the dissertation. There are several different options for Creative Commons licenses (see https://creativecommons.org/licenses/), and an institutional dissertation librarian can often provide insight on license options. In the end, the candidate should be comfortable with the license they are choosing based on their needs and their expectations for the future lives of their dissertation.

Once an open-access dissertation is available in one repository, one of the best things that a scholar can do to make the work more highly available is through making that publication redundant. Digital media continues to present challenges to archivists, and it is important for submitters to do their own due diligence to make their work redundant and highly available through multiple venues and archival channels to ensure continued access. Disseminating the work widely in different spaces can ensure not only the discoverability of the dissertation (increasing opportunities for downloads, citations, collaborations), but will also ensure that the work persists through multiple digital archive channels. Ferreras-Fernández et al.'s study found that submission across platforms did enable increased access and citations. They write, 'PhD theses disseminated through repositories are benefited through interoperability, which allows their dissemination through multiple portals, sites and search engines, thereby increasing their visibility and making them likely to be cited'.[19] When thinking about increased distribution, the author should consider several spaces: their own personal website and servers, field-specific repositories such as MLA's

---

18   Duke University, The Graduate School, 'ETD Copyright Information', *Duke Graduate School*, https://gradschool.duke.edu/academics/theses-and-dissertations/etd-copyright-information

19   Ferreras-Fernández et al., 'Providing Open Access to PhD Theses', p. 413.

CORE, traditional open-source repositories if applicable, institutional or public OER repositories as appropriate, and submission to sites like archive.org and other free eBook databases. In order to maintain as much control over the work as possible, a PhD candidate should seek out open access scholarly repositories that were developed within universities and built with open-source software, this way those submitting scholarship can even improve the platform in which that scholarship resides. That said, when dissertations become highly available in a redundant way, it becomes more difficult to keep track of the number of engagements that are taking place, whether those are pageviews, downloads or citations, but ultimately redundant copies are the most effective way to ensure the work is openly accessible.

For dissertations that incorporate the work of others, there are some more factors to consider with regard to open access. Long-form translations of writing may need to request permissions from the copyright holder or original publisher prior to making the work open access. Requesting permissions can demonstrate good will even in cases when work is licensed under Creative Commons. In addition to university resources, a candidate may wish to consult with organizations like the American Translators Association for guidance. Art historical dissertations or those that include images and other media files will similarly need to follow guidelines for fair use or fair dealing, and potentially consult the rights holder of any media used. Depending on the period and nature of the work, the proper authority may vary: for manuscripts and early printed books one can consult with the holding archive, for the visual arts including film, a holding museum could be a first step but for living artists or artists with foundations one may decide to make contact via official websites. That said, there are an increasing number of open access and Creative Commons-licensed repositories of media that originate through museums, libraries and organizations like Wikimedia Commons. A candidate can read terms of service and licensing information through websites to ensure that their use case in the dissertation falls within given bounds, and, if so, should cite the origin of incorporated media. In all cases, good judgment should be exercised and citation practices appropriate to scholarly endeavors should be followed.

Dissertations that include data, whether collected or accessed by the doctoral candidate, will also need special consideration. In some cases, the author can cite data that is collected by other sources, but they will have to look into terms of service and sharing policies. When candidates have data collected from human subjects, they will have to ensure that they are in compliance with the Institutional Review Board (IRB) and any other relevant bodies as set out by their university. Doctoral students who find and clean data themselves, through web scraping or other methods, should comply to the data store's terms of service. Candidates can decide whether or not to release the data upon deposit of the dissertation. Making the data publicly available immediately could support others' research and will facilitate the peer review of the dissertation findings. However, if the candidate is planning to do more with that data, there may be good reasons for withholding the full data set until a later date. Data can be released in a software repository, be printed in an appendix, or made available as downloads.

As the digital humanities increase in popularity, more dissertations will include source code. If software is created through collaboration with others, it is necessary to agree on whether or not the software project will be open source and under which license it will be released. The Open Source Initiative includes guidance on different types of licenses and best practices. For those coding alone, an understanding of the open-source ecosystem will help guide the project and its release upon deposit of the dissertation. The software code can live in a Git repository[20] separate from the text of the dissertation, and can also be added to the appendix of the dissertation (within reason), or as zip files as part of the institutional deposit or other disseminations. Where to house a Git repository has become more fraught for humanities researchers as large corporations are increasingly controlling large open-source stores. In addition to housing code on a platform like GitHub, software developers may consider using their own servers to run Git. Wherever the code is housed, once it is released as an open-source project on the internet, there are a number of things for the author or other software

20 Git refers to an open-source version-control system that allows for collaborating on computer files. Git repositories are popular for housing code and hosting open-source projects. GitHub, GitLab and SourceForge are popular choices to store software projects and source code.

maintainer to consider. While this is outside of the scope of the current discussion, I provide an overview in my article on 'How to Maintain Open-Source Software Projects'.[21]

An open-access dissertation that is part of a well-maintained repository is well set up to be discovered. However, it is possible for the author to further increase the visibility of the dissertation through discussing it in public digital channels. Traditional social media platforms like Twitter can help to spread the open-access dissertation, as can blogs and even video. A landing website for the project can house all of the various elements, especially if there is source code and other media involved. Institutions may have their own blogging platforms, like CUNY Academic Commons, that can enable either a landing page, or offer the opportunity for multiple posts about the research. Wider, non-institutional networks like HASTAC and Humanities Commons can also serve to share research publicly while establishing connections with other scholars and potential collaborators. The dissertation author should take care to evaluate opportunities against their own comfort levels with wider engagement.

CUNY Academic Works

As a doctoral student of the City University of New York's Graduate Center, I was required to deposit in CUNY Academic Works, an open access repository that holds over 20,000 papers written by the CUNY community (current CUNY faculty, students, and staff may submit their scholarly and creative works to this repository). A service of the CUNY Libraries, CUNY Academic Works collects scholarly papers and provides free public access to these as part of the Library's efforts to advance the mission of CUNY as a public university. To date (February 2021), CUNY Academic Commons' papers have been downloaded approximately 5.5 million times collectively—an average of about 119 downloads per paper.[22] This number is a bit higher than the average number of copies purchased of academic monograph titles in the United

---

21   Lisa Tagliaferri, 'How to Maintain Open-Source Software Projects' (October 6, 2016), *DigitalOcean*, https://www.digitalocean.com/community/tutorials/how-to-maintain-open-source-software-projects

22   CUNY Academic Works, *CUNY Academic Works*, https://academicworks.cuny.edu/

States, which was noted as 83 by Michael Zeoli in 2015 (interestingly, the average number of units sold is the same for both monographs of revised dissertations and those that were new works).[23]

Among the benefits listed on its 'About' page,[24] CUNY Academic Works states that papers that are submitted to the repository will become more discoverable by search engines, be securely hosted on a server, have a dedicated URL for long-term access, and be freely accessible to the public including those who may have limited access to scholarship. This multi-pronged approach to open access taps into the affordances of search engine optimization through metadata and cataloguing efforts, digital archiving for a persistent information store of the research, and a commitment to maintain a perpetually free and public website.

While CUNY Academic Works requires that all Graduate Center dissertations, theses and capstone projects be submitted to the repository, not all are immediately available to read and download, as authors are able to set an embargo period. As noted above, the embargo is intended to allow students to keep their research private prior to publishing the work through an academic press or through journal articles.

In addition to submitting to CUNY Academic Works, I was also required to submit to ProQuest as a doctoral recipient at my institution. Many university ETD management systems in the United States still require some form of submission to ProQuest, an information-content and technology company that was founded in 1938. However, as Gail Clement and Fred Rascoe note, there is a growing number of ProQuest optional or 'NoQuest' institutions (requiring no ProQuest submission), including the University of Michigan, Brown University and Stanford University.[25] While CUNY Academic Works is an open-access initiative unlike ProQuest, the repository uses commercial, proprietary software called Digital Commons produced by Bepress, a commercial software firm that was founded in 1999 and is now owned by the RELX Group. Digital Commons is one of the three software platforms recommended

---

23 Cirasella and Thistlethwaite, 'Open Access and the Graduate Author', p. 208.
24 CUNY Academic Works, 'About', *CUNY Academic Works*, https://academicworks. cuny.edu/about.html
25 Gail P. Clement and Fred Rascoe, 'ETD Management & Publishing in the ProQuest System and the University Repository: A Comparative Analysis', *Journal of Librarianship and Scholarly Communication*, 1.4 (2013), 1–28 (at 5), https://doi. org/10.7710/2162-3309.1074

by Google Scholar for academic repositories to be aggregated by the Google service, and the only one that is not a non-profit (see https://scholar.google.com/intl/en/scholar/inclusion.html).

When I submitted my dissertation, I chose to make the dissertation available without an embargo through both CUNY Academic Works and ProQuest to increase its discoverability and to provide academics more familiar with ProQuest the ability to find it there. Dissertation authors can choose to embargo with neither, either, or both the CUNY Academic Works and ProQuest databases to a greater or lesser amount of time.[26]

An institution's repository and the infrastructure platform that it hooks into are discrete entities, just as a library is a community of people that hooks into an infrastructure of books and other resources. As some institutions have more resources than others, the platform that repositories are located in will vary, and it is the onus of the institution to make choices that will best serve their community based on the resources they have available. In my opinion, an ideal open-access scholarly repository would be one that originated within universities and built with open-source software that could be iteratively improved upon by wider communities. However, I understand that this is not always feasible due to many different challenges and limitations. The fact that CUNY Academic Works leverages the platform of a for-profit company was not clear to me at the time of my submission, and I did not know that the owner of the platform could change. The RELX Group, formerly Reed Elsevier, acquired Bepress in August 2017, shortly after my submission.[27] Although Bepress has leaders with roots in academia, it was not built as an open-source and non-profit organization, leaving it susceptible to acquisition by a large conglomerate. However, through their contract, CUNY ensured that this platform provider did not own the content and the metadata associated with the repository.[28] Ann Hawkins,

---

26   The Graduate Center Library, 'Dissertations and Theses: A Note about Databases and Embargoes', *Mina Rees Library Research Guides*, https://libguides.gc.cuny.edu/dissertations/embargoes

27   Robert Cookson, 'Reed Elsevier to Rename Itself RELX Group' (February 26, 2015), *Financial Times*, https://www.ft.com/content/4be90dbe-bd97-11e4-9d09-00144feab7de; David Bond, 'Relx buys Bepress to Boost Academic Publishing' (August 2, 2017), *Financial Times*, https://www.ft.com/content/c6f6c594-7787-11e7-a3e8-60495fe6ca71

28   Megan Wacha, Scholarly Communications Librarian, Office of Library Services (CUNY), personal communication, August 9, 2018.

Miles Kimball and Maura Ives recommend that universities 'make their ETD policies and information about ETDs available prominently and conveniently on their web sites and in their practices', which I believe will only help PhD students navigate their options.[29] Increased transparency and documentation around dissertation depositing can go a long way towards further empowering graduate students, who are often a marginalized population in the university landscape.

A Dissertation and Its Afterlife

At the moment, my dissertation has not been refactored for publication as a monograph or a series of articles, so I cannot yet detail what that process may look like. I can, however, speak to the process of making this digital humanities project open access with open-source code, and the discoverability aspect of it as a digital dissertation, and extrapolate how this may inform the decision making of others.

Treating the subject of the Italian mystic writer Catherine of Siena and seeking to exert her status as a literary author in her own right through traditional literary analysis paired with digital humanities techniques, 'Lyrical Mysticism: The Writing and Reception of Catherine of Siena' is a comparative literature dissertation that I began working on in 2015 and defended on April 19, 2017. My dissertation work was deposited shortly thereafter, and the full text was made available in the CUNY Academic Works repository on May 17, 2017. I licensed the work under the Creative Commons Attribution-Noncommercial 4.0 License, allowing it to be freely shared and adapted for noncommercial purposes. I spoke to librarians at my institution to decide on a license, and those who are looking to use a Creative Commons license should consider the various options against their hopes for discoverability and opportunities for collaboration.

My dissertation consists of 211 textual pages, as well as a digital component that is housed under the website caterina.io and a GitHub code repository (available at https://github.com/ltagliaferri/dissertation).

---

29   Ann R. Hawkins, Miles A. Kimball and Maura Ives, 'Mandatory Open Access Publishing for Electronic Theses and Dissertations: Ethics and Enthusiasm', *The Journal of Academic Librarianship*, 39.1 (2013), 32–60 (at 38), https://doi.org/10.1016/j.acalib.2012.12.003

The code repository, which is a result of the dissertation being a digital humanities project, consists of several different elements. First, there is the full final text of the dissertation as well as the version-controlled text of the dissertation—for example, if you wanted to see all the revision commits of chapter 1, you can see them via this link: https://github.com/ltagliaferri/dissertation/commits/master/Chapter1.txt. I decided to make revisions publicly visible (as repository commits) in order to show the development of the work over time, as well as provide a record of how feedback was being incorporated. Experimenting with how Git can reveal the evolution of a text could be a fruitful endeavor to show the progression of a text over time. (I am still a bit terrified of providing this level of access to my dissertation.) Next, there are the textual files of the writing of Catherine, Dante and Petrarch, which I used to complete comparative analysis through programming. These files were acquired through web scraping digital versions of primary sources (NB: always read the terms of service prior to web scraping), and I performed some programmatic work to add consistency across the files and bundle files together as needed. Next, there are the actual programming files of the digital humanities work I completed in Python, R and JavaScript. Finally, there is the source code of the website that includes interactive visualizations (these were web recorded by the Graduate Center Library for it to be archived in the repository).

Completing a solo programming project—leveraging data analysis, system administration and web development skills that I developed outside of my PhD program—while also conducting medieval and early modern archival research and completing a traditional dissertation to advance the knowledge of the field was challenging. In retrospect, the care around maintenance and archiving my writing and programmatic work along the way were what suffered the most during this endeavor. Ideally, I would have hosted my dissertation Git repository on my own server (rather than on GitHub's company servers) that could have been archived separately by CUNY librarians, but this would have added an additional level of complexity during an already stressful time. Because of my status as a student and GitHub's generous Student Developer Pack offering, I was able to keep my dissertation work private prior to my defense without having to pay for the privilege, which assuaged my fear of being 'scooped'. Even as someone with significant technical

knowledge, I would have benefited immensely from a greater university ecosystem that supports digital humanities projects with teams to help with version control, Git and other digital best practices. For others completing software development projects as part of humanities dissertations, I would encourage seeking out best practices from existing open-source projects. Institutions that are committed to digital humanities projects should consider ways to connect humanities scholars with others in technical fields for support in code reviews, repository stores, testing, and continuous integration and delivery. Increasingly, it is important for universities to hire in-house programmers and network administrators in order to best support the technical research of the university that in industry is usually carried out by full teams.

Because of data that is available to me as the author of my dissertation, I do have a sense of how often it is downloaded and where. CUNY Academic Works provides readership reports through Bepress's Author Dashboard that authors can access through the repository software. Having ready access to data as an author is one of the benefits that is often packaged into open access repositories, allowing you to track downloads and other data points around your dissertation. What is important to note is that each repository where your dissertation is included will have its own data metrics, and you may not have access to all of them. ProQuest is much less transparent in terms of download tracking data than most open access repositories, so you may be unable to keep tabs on every download of your dissertation. It is also challenging to have a sense of how frequently dissertations are downloaded across fields or through various services, so comparative analyses through anecdota are what tend to persist.

To add to the available data, I will share mine here. Based on what I have available from CUNY Academic Works beginning on the date of its deposit on May 17, 2017 until the day of this chapter's revision for publication (February 2021), my dissertation was downloaded over 1,500 times. With ranges between 11 to 62 monthly downloads, there has been no consistent decline. To have an understanding of who is accessing my dissertation more broadly, the text has been downloaded in 90 different countries, and by 180 different organizations (including universities and industry). Dissertation authors may notice that both academic and corporate institutions that are interested in hiring the author for an open role are among those downloading their dissertation.

Although unlike open source, we often think of single-authored dissertations as less collaborative in nature, I think we overlook all the ways that collaboration exists in different degrees and in different directions. As I was working on my dissertation within a traditional humanistic field, the text that I produced was heavily scrutinized by advanced scholars, and I am glad that it was. Librarians at the Graduate Center Library provided considerable support in the deposit of both the text and the digital component of the work, and offered important guidance. The larger Graduate Center community and the broader academic networks of my fields also fostered the advancement of my scholarship in many different ways, and I do not wish to take all of this community nurturing for granted. Still, the development side of my dissertation project did not benefit from code reviews, and I did not have any collaborators to help with the digital manifestation of the work in a hands-on manner. It is possible that I did not fully seek out additional assistance, but a ready framework for this did not exist that I could find in the same manner that I found the considerable documentation and guidance for the more traditional parts of the dissertation. Humanities departments and institutions that encourage digital humanities research should work to support it in meaningful ways, providing resources and direction akin to what they provide for traditional research.

If we return to open-source code development as a framework for open access, there are takeaways from a technology approach that can be applied to an open-access digital dissertation. Transparent and clear structures for writing, review, revision, depositing and caring for the afterlife of the dissertation (whether through making the text redundant, marketing the text, or offering recommendations for filing for infringement) can help to support the dissertation writer as they navigate this process. Open-source code development is often done completely in public over time, while dissertations are often completed in a relatively closed-off manner (apart from conference papers, etc.). The differing practices are a result of the fields, as explicated by the Graduate Center's dissertation research librarian Roxanne Shirazi, who writes, 'In recent years, it would seem that humanities and social science scholars are worried about getting publications out of a dissertation, while STEM folks are increasingly concerned with getting publications

into a dissertation'.[30] Having the opportunity to open dialogues with public readership can allow authors to incorporate feedback and iterate on the text for a more vigorous, living book that can be developed over time. Open source has a saying, 'commit early and commit often', speaking to the iterative nature of development projects. What could we, as humanists, gain by engaging early and often?

Bibliography

Anthes, Gary, 'Open Source Software No Longer Optional', *Communications of the ACM* 59.8 (2016), 15–17, https://doi.org/10.1145/2949684

Bond, David, 'Relx buys Bepress to Boost Academic Publishing' (August 2, 2017), *Financial Times*, https://www.ft.com/content/c6f6c594-7787-11e7-a3e8-60495fe6ca71

Budapest Open Access Initiative, 'Read the Budapest Open Access Initiative' (February 14, 2002), http://www.budapestopenaccessinitiative.org/read

Cirasella, Jill, and Polly Thistlethwaite, 'Open Access and the Graduate Author: A Dissertation Anxiety Manual', in *Open Access and the Future of Scholarly Communication: Implementation*, ed. by Kevin L. Smith and Katherine A. Dickson (Lanham: Rowman and Littlefield, 2017), pp. 203–24, https://www.grad.miami.edu/_assets/pdf/open-access-and-the-graduate-author_-a-dissertation-anxiety-manua.pdf

Clement, Gail P., and Fred Rascoe, 'ETD Management & Publishing in the ProQuest System and the University Repository: A Comparative Analysis', *Journal of Librarianship and Scholarly Communication*, 1.4 (2013), 1–28, https://doi.org/10.7710/2162-3309.1074

Cold Spring Harbor Laboratory, Library & Archives LibGuides, 'Guide to Open Access: What Is Open Access (OA)?', *CSH*, https://cshl.libguides.com/c.php?g=474046&p=3243847

Cookson, Robert 'Reed Elsevier to Rename Itself RELX Group' (February 26, 2015), *Financial Times*, https://www.ft.com/content/4be90dbe-bd97-11e4-9d09-00144feab7de

Corbett, Hillary, 'Out of the Archives and into the World: ETDs and the Consequences of Openness', in *Open Access and the Future of Scholarly Communication: Implementation*, ed. by Kevin L. Smith and Katherine A. Dickson (Lanham: Rowman and Littlefield, 2017), pp. 187–202.

---

30 Roxanne Shirazi, 'The Doctoral Dissertation and Scholarly Communication', *College & Research Libraries*, 79.1 (2018), https://crln.acrl.org/index.php/crlnews/article/view/16864/18489

Courtney, Kyle K., and Emily Kilcer, 'From Apprehension to Comprehension: Addressing Anxieties about Open Access to ETDs', in *Open Access and the Future of Scholarly Communication: Implementation*, ed. by Kevin L. Smith and Katherine A. Dickson (Lanham: Rowman and Littlefield, 2017), pp. 225–44.

Creative Commons Wiki, 'History' (2011), *Creative Commons*, https://wiki. creativecommons.org/wiki/History

CUNY Academic Works, *CUNY Academic Works*, https://academicworks.cuny. edu/

CUNY Academic Works, 'About', *CUNY Academic Works*, https://academicworks. cuny.edu/about.html

Duke University, The Graduate School, 'ETD Availability', *Duke Graduate School*, https://gradschool.duke.edu/academics/theses-and-dissertations/ etd-availability

Duke University, The Graduate School, 'ETD Copyright Information', *Duke Graduate School*, https://gradschool.duke.edu/academics/theses-and-dissertations/etd-copyright-information

The Graduate Center Library, 'Dissertations and Theses: Deposit Procedure', *Mina Rees Library Research Guides*, https://libguides.gc.cuny.edu/ dissertations/deposit-procedure

The Graduate Center Library, 'Dissertations and Theses: A Note about Databases and Embargoes', *Mina Rees Library Research Guides*, https://libguides.gc.cuny. edu/dissertations/embargoes

Hawkins, Ann R., Miles A. Kimball and Maura Ives, 'Mandatory Open Access Publishing for Electronic Theses and Dissertations: Ethics and Enthusiasm', *The Journal of Academic Librarianship*, 39.1 (2013), 32–60, https://doi. org/10.1016/j.acalib.2012.12.003

Jaschik, Scott, 'The Shrinking Humanities Job Market: New Analysis Finds the Number of Doctorates Awarded Keeps Rising, even as Number of Job Openings Drops' (August 28, 2017), *Inside Higher Ed*, https:// www.insidehighered.com/news/2017/08/28/more-humanities-phds-are-awarded-job-openings-are-disappearing

Juell, Kathleen, 'A Brief History of Linux' (October 27, 2017), *DigitalOcean*, https://www.digitalocean.com/community/tutorials/brief-history-of-linux

Ferreras-Fernández, Tránsito, et al., 'Providing Open Access to PhD Theses: Visibility and Citation Benefits', *Program: Electronic Library and Information Systems*, 50.4 (2016), 399–416. https://doi.org/10.1108/prog-04-2016-0039

Laakso, Mikael, et al., 'The Development of Open Access Journal Publishing from 1993 to 2009', *PLoS ONE*, 6.6 (2011), e20961, https://doi.org/10.1371/ journal.pone.0020961

Novikoff, Alex J., *The Medieval Culture of Disputation: Pedagogy, Practice, and Performance* (Philadelphia: University of Pennsylvania Press, 2013), https://doi.org/10.9783/9780812208634

Nussbaum, Jan L., 'Apple Computer, Inc. v. Franklin Computer Corporation Puts the Byte Back into Copyright Protection for Computer Programs', *Golden Gate University Law Review*, 14.2 (1984), 281–308, https://digitalcommons.law.ggu.edu/cgi/viewcontent.cgi?article=1344&context=ggulrev

Ramirez, Marisa L., et al., 'Do Open Access Electronic Theses and Dissertations Diminish Publishing Opportunities in the Social Sciences and Humanities? Findings from a 2011 Survey of Academic Publishers', *College & Research Libraries*, 73.4 (2013), 368–80, https://doi.org/10.5860/crl-356

Shirazi, Roxanne, 'The Doctoral Dissertation and Scholarly Communication', *College & Research Libraries*, 79.1 (2018), https://crln.acrl.org/index.php/crlnews/article/view/16864/18489

Tagliaferri, Lisa, 'How to Maintain Open-Source Software Projects' (October 6, 2016), *DigitalOcean*, https://www.digitalocean.com/community/tutorials/how-to-maintain-open-source-software-projects

Townsend, Robert B., and Emily Swafford, 'Conflicting Signals in the Academic Job Market for History' (January 9, 2017), *Perspectives on History*, https://www.historians.org/publications-and-directories/perspectives-on-history/january-2017/conflicting-signals-in-the-academic-job-market-for-history

Von Hippel, Erik, and Georg von Krogh, 'Open Source Software and the "Private-Collective" Innovation Model: Issues for Organization Science', *Organization Science*, 14.2 (2003), 209–23, https://doi.org/10.2139/ssrn.1410789