

Creative Destruction: Open Access, Institutional Repositories and the Changing Dissertation

Lisa Schiff

Technical Lead, Access & Publishing, California Digital Library, University of California

The dissertation is a cornerstone of academia, serving as evidence that the student--now the newly certified scholar--has adequately mastered the accumulated knowledge and practices of his or her field and has, in the process, made a unique contribution to an ever-growing body of scholarly research. Through this regular vetting of aspiring members of its community, academia reproduces and renews itself. Clearly the dissertation occupies a foundational space in the constant development of new researchers and in the continual fortification of the edifice of higher education. But in a curious contrast, the dissertation also resides in a liminal place, between the established and the yet-to-be accepted. In this capacity, dissertations are potential sources of constructive destabilization, since these works are always the products of minds new to a domain, who by definition bring fresh (or at least different) questions, perspectives, skills and expectations to their fields of inquiry.

Amplifying this potential for change is the reality that the familiar patterns of scholarly communication are not, in fact, etched in stone, but are instead undergoing a variety of transformations across all domains. Modes of communication are expanding to include more informal venues of exchanges, from tweets to blog posts, in addition to the traditional genres of articles, books and conference papers. Access to scholarly output has dramatically increased via the open access movement and related new entrants into the domain, such as new open access and library-based publishers, as well as individual researchers directly promoting their work across the web. Emerging areas of research are being recognized (digital humanities, for example, is both novel and “old hat”), and new classes of scholarly materials are being reconceived as valuable objects in their own right for which their creators should accrue formal, promotion-worthy credit. This changing context impacts the dissertation as well. While the dissertation of today is still most frequently a text (and a PDF document at that)¹ and is still weighted in purpose toward acting as a certifying piece of scholarship, “today” is turning into “tomorrow” before our eyes. Students are more frequently including supplementary materials with their dissertations²; a slow, but increasing number of culminating projects are created in non-textual formats, such as video, audio, or other multi-media and non-linear presentations; and

¹ Gail McMillan, Martin Halbert, and Shannon Stark, “2013 NDLTD Survey of ETD Practices.” *16th International Symposium on Electronic Theses and Dissertations, Hong Kong* (2013): p. 3, <http://hdl.handle.net/10919/50978>

² Educopia’s IMLS funded project investigating the handling of ETDs with supplemental material speaks to their increasing presence. <http://educopia.org/announcements/two-year-imls-grant-awarded-study-management-supplemental-data-etds>

expectations about the potential audiences and uses of dissertations have grown to encompass a far vaster scope of people and situations.

As the academy aims to achieve expanded access to dissertations, the institutional repository (IR) frequently surfaces as a key resource to realize that goal³. The IR is a stable access and (often) preservation platform that brings together scholarly content from across the institution, including previously published items, grey literature, electronic theses and dissertations (ETD), monographs, oral histories and data sets. Its primary roles are to bring light to this array of materials, provide a secure location for content that might have no clear public home elsewhere, and offer a sustainable and reliable alternative to the faculty website or personal computer. IRs are typically flexible and accommodating, and frequently situated in the library--that place on campus historically focused on discovery and access. At UC's [California Digital Library](#), two different repositories link together in a modular fashion in order to provide a flexible ETD service to the 10 different UC campuses, each of which has its own ETD policies, practices and expectations. CDL provides preservation services through the [Merritt preservation repository](#) (available to all UC campuses) and enables public access (for those campuses that wish to make their ETDs openly accessible) via [eScholarship](#), the University of California's open access repository and publishing platform. As of the beginning of November 2015, Merritt is managing the preservation of over 19,600 ETDs, and approximately 14,000 ETDs are available for public use in eScholarship. In the past year, these public ETDs have received on average almost seven requests per day, as compared to the just under five per day for all content in eScholarship. Clearly these items are highly sought after and their presence in the IR--eScholarship--facilitates their discovery and use.

Though now a familiar component of a university or college's scholarly communication landscape, the IR must continue to evolve to meet the changing needs of students and scholars⁴. As institutions rethink their role in stewarding and disseminating the dissertations of their students, IRs are well positioned to manage that process by leveraging existing, flexible infrastructure. IRs amplify the profile of ETDs by co-locating them with the scholarly outputs of more experienced researchers at the institution. In turn, ETDs substantially contribute to the ROI of an IR, providing a compelling justification for continued resource allocation and an excellent use-case for supporting new forms of scholarship as graduated students begin to push against the confines of existing forms.

³Julia Lovett and Andrée Rathemacher, "Open Access and the Institutional Repository," In *Proceedings of the Querying the Library: Digitization and its Impact Conference*, ed. Mark J. Caprio (Providence, R.I.: Rhode Island College, James P. Adams Library, 2014), 22-29. http://digitalcommons.ric.edu/ebook_gallery/29/

⁴Joachim Schöpfel, Adding Value to Electronic Theses and Dissertations in Institutional Repositories. *D-Lib Magazine* 19 no 3&4 (2013). <http://dx.doi.org/10.1045/march2013-schopfel>

The benefits ETDs receive from IRs demonstrate the fulfillment of the IR's essential mission--to capture and increase exposure of the variety of scholarship generated by the scholarly community within which it is situated. Given that content contributors are strongly motivated by the desire for increased use of their publications, they will generally only deposit their work in IRs that provide an effective mechanism for generating significant use. IRs, then, have a survival imperative to deeply engage in a variety of dissemination activities core to the flow of scholarly communication, beginning with the responsibility to follow basic domain practices such as the use of unique identifiers for primary content (e.g. DOIs, ARKs, or Handles), supporting programmatic discovery interfaces such as OAI-PMH, and providing human and machine readable licensing information, such as Creative Commons license marks and metadata. Of more interest to authors are the activities most obvious to end users, for instance ensuring consistent and solid indexing in Google, Google Scholar and other search engines, and inclusion of content in third party free and commercial discovery platforms, such as Research Papers in Economics (RePEc), OCLC's Worldcat Local, and EBSCO.

Because of its focus on discovery and use and because of its inclusive tendency to accept local content of all sorts, the IR is a strategic service for realizing the scholarly communication goals of the institution. We have seen this play out over the last several years as faculty at an increasing number of universities and colleges across North America have adopted Open Access (OA) policies. Where they already exist, the IRs at those institutions can naturally serve as the locus for the archiving of scholarly works that fall under those policies^{5 6}. Not only do these IRs provide a pre-existing destination for the author's version of a published item, but they already have the processes in place to expose that content to the world. If one of the transforming goals associated with dissertations is to share that body of work with the public, then the IR is an efficient and logical place to make that happen.

Moving access out from the library to the web is of course not the only change taking place with ETDs. Formats of scholarly outputs are evolving, including those of dissertations, meaning that the infrastructure underpinning that content will also have to evolve. At the CDL, we find ourselves at two contrasting positions in the spectrum of potential infrastructure development. On the one hand, our preservation strategy will remain effective and will require no change, because the UC3 Merritt repository is agnostic regarding format. The conceptual structure and

⁵Ellen Finnie Duranceau and Sue Kriegsman, "Implementing Open Access Policies Using Institutional Repositories," in *The institutional repository: Benefits and Challenges*, eds. Pamela Bluh and Cindy Hepfer (Chicago: American Library Association, 2013), 75-97. <http://nrs.harvard.edu/urn-3:HUL.InstRepos:10202474>

⁶Ellen Finnie Duranceau and Sue Kriegsman, "Campus Open-Access Policy Implementation Models and Implications for IR Services," in *Making Institutional Repositories Work*, eds. Burton B. Callicott, David Scherer, and Andrew Wesolek. (West Lafayette, IN: Purdue University Press, 2016). <http://hdl.handle.net/1721.1/99738>

physical nature of the dissertation can wildly transform, but those transformations would in no way inhibit Merritt's ability to accept complex objects; to version those objects; to perform bit-level auditing and more. New formats present a more stubborn access challenge for IRs because of their somewhat limited capacity to support extensive customization of display. While eScholarship, like all IRs, supports a variety of genres, the presentation of those items is templated, varying only in the display of metadata fields considered essential for one type of publication versus another. The benefits of this simple approach are a reduction in startup and maintenance costs; a low burden of effort for content contributors; and a relatively minimal marginal cost to bring in new sets of material, assuming those materials don't differ in extreme ways from existing content types. The weaknesses of this approach are the lost opportunities to display varied content in the unique ways most suited to that content and the related challenge of staying in step with new forms of scholarship that are producing new forms of content. eScholarship, no doubt like other IRs, has had the unfortunate experience of having to say "No" to unique, valuable artifacts from its local scholars because the work involved in providing an acceptable display for those materials could not be extended to other items and, therefore, justified. Evolving formats for dissertations currently pose the same risk of being idiosyncratic "one-offs," but no crystal ball is required to realize that these new formats will, soon enough, be common for dissertations and scholarly works in general. The challenge for eScholarship and IRs overall is to develop new, malleable infrastructure that is therefore more long-lived. The first step is to build in solid support for HTML encoded publications while at the same time continuing to effectively serve PDF documents, which are likely to remain the bulk of our content for quite some time. Without investing resources to explore and devise solutions to this complex environment today, IRs risk becoming irrelevant and failing in their goal to support the many scholars who seek a robust OA platform for the dissemination of their work.

Though the potential display challenges posed by new formats are understandably of interest and concern for many involved with ETDs, the most frequent source of complications arising from CDL's dissertation service is the graduate student herself, who accidentally discovers that her dissertation is now widely accessible, but does not remember agreeing to make it so. Despite the variety of efforts on the part of graduate divisions and campus libraries to inform students about the preservation and eventual public display of their institutions' dissertations, the message is not effectively reaching everyone, no doubt because it often comes near the final stages in a student's education, at the moment when they are most tired and anxious and ready to be done. In reality, even if students take the time to conduct a close reading of their dissertation submission forms, at that point it is already too late. Decisions about research topics, methodologies and perhaps even future publishing goals have already been made, and made in the absence of understanding that dissertations are now a more vocal part of the scholarly conversation, even if the students themselves did not realize they were speaking out loud. The answer is to begin to build student awareness about these issues from day one, through small steps such as assigning ORCID iDs to new students and more complex tasks such as including

discussions of licensing and data sharing in methodology classes. Accepted graduate students should, from the outset, understand the role of their future research as a contribution to the scholarly dialogue. Advisors can ensure that the developing scholars under their charge become familiar with issues around the opening up of the research process, access to an ever growing number of artifacts created in and out of that process, and how to explore areas of interest within this increasingly exposed context. Deans, graduate divisions and libraries can work together more closely to better understand each other's goals, responsibilities, and workflows and thus identify complementary tasks and compatible practices to achieve the best outcomes for students, the institution and the public.

The scholarly communication revolution continues to touch all parts of the academy, including the time-honored exercise of producing a dissertation. As scholarly artefacts in general become more heterogeneous in concept, construction and format, so too will dissertations. And as research outputs of all sorts become more readily available to the higher education community and the world at large, so too will dissertations. Though legitimate reasons for restricted use will persist and should be supported, the clear benefits of surfacing the knowledge created in our institutions of learning add momentum to those same efforts, inspiring us to increase the availability of more categories of materials, including dissertations. IRs, instrumental in enabling the opening up of faculty material of both familiar and novel types, will continue to be a strategic and efficient venue for liberating dissertations from within the stacks and behind subscription databases. Though these student works retain their traditional and critical role as foundational, certifying documents, they act also as tremendous contributors to the advancement of the scholarly record by the academy's newest members. The dissertation, then, is not just a cornerstone in the higher education's edifice, but a window onto the new as well, one that is well framed and supported by the institutional repository.