

2012

Developing an Electronic Repository for Undergraduate Theses

Foster Levy

East Tennessee State University, levyf@etsu.edu

Rebecca Pyles

East Tennessee State University

Celia Szarejko

East Tennessee State University

Linda Wyatt

East Tennessee State University

Follow this and additional works at: <http://digitalcommons.unl.edu/nchchip>

 Part of the [Higher Education Administration Commons](#)

Levy, Foster; Pyles, Rebecca; Szarejko, Celia; and Wyatt, Linda, "Developing an Electronic Repository for Undergraduate Theses" (2012). *Honors in Practice -- Online Archive*. 160.

<http://digitalcommons.unl.edu/nchchip/160>

This Article is brought to you for free and open access by the National Collegiate Honors Council at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Honors in Practice -- Online Archive by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Developing an Electronic Repository for Undergraduate Theses

FOSTER LEVY, REBECCA PYLES, CELIA SZAREJKO,
AND LINDA WYATT

EAST TENNESSEE STATE UNIVERSITY

INTRODUCTION

Undergraduate honors theses represent an intellectual asset that a university should recognize and manage as such. However, when theses were submitted exclusively in print copies, the work often faded into obscurity, forgotten by all but the student and mentor. While theses for advanced degrees have been accessible for many years via interlibrary loan or abstract services, similar access options have been unavailable for undergraduate theses because these works are most often associated with and maintained by the institutional honors program without involvement or support from the institution's library system. At best, an index of undergraduate theses might be available to the public, but print copies—often the only versions of theses—are traditionally housed in honors and are difficult, if not impossible, to obtain.

As undergraduate honors theses have become more commonplace and as online access to research has become virtually universal, honors programs do a disservice to their students, faculty, and the public if they do not provide access to the original scholarship produced by undergraduates. Furthermore, introducing undergraduates to electronic publication provides educational value by exposing them to the publishing demands they will likely encounter in their future education and professional careers. Two additional benefits to students are that an electronic publication saves the student time and expense in producing their final work and, most importantly, makes their work available worldwide to potential research partners and employers.

An electronic thesis repository provides several instruction-based opportunities to advance learning. For example, courses within an honors college often enroll students from a variety of disciplines. In research methods courses, students can easily examine examples of theses to familiarize themselves with the writing styles and formatting in their discipline. Similarly, in a colloquium-style course that addresses diverse topics, students can use the repository to become familiar with research approaches and writing styles outside of their discipline; this is a particular advantage for undergraduates because, as they progress in

their course of study and certainly once they begin graduate or professional programs, they rarely have time for this type of cross-disciplinary interaction. In addition, faculty in diverse disciplines can assign repository reviews, either in courses or when mentoring honors students, to demonstrate the level of scholarship expected in honors thesis work. Although these goals could be accomplished using hard copies, our current students are more comfortable with and more likely to access electronically available materials.

East Tennessee State University (ETSU) has made undergraduate theses available to the campus community and general public through an electronic repository and catalog maintained by the university library. The electronic system we implemented and others of similar quality provide a user interface developed for manuscript submission, review, and approvals. This type of system takes students' research experience to a final level of completion and assures that they learn how to navigate a process analogous to manuscript submission. Because adapting our institution's system for graduate theses and dissertations was not a feasible option and we could find no general guidelines to direct our efforts, we developed our own undergraduate honors thesis repository, and we hope that our efforts in this process will provide insights and guidelines for other institutions.

SURVEY OF USAGE AND PRACTICES

Interest in developing an electronic repository for undergraduate theses led us to conduct two assessments to gather knowledge of current practices. The first assessment, completed during the winter of 2010, was a survey of public access practices for undergraduate honors theses at seventeen ETSU peer institutions <<http://www.etsu.edu/iep/09FB/09TOC.htm>> and eight non-peer institutions. The non-peer sample comprised all institutions we could locate that had accessible (website) submission details about their honors theses and that provided contact information for relevant personnel. These non-peer institutions included honors programs at Boston College, Cornell University, Massachusetts Institute of Technology, Tufts University, University of Colorado at Boulder, University of Delaware, West Virginia University, and the Swedish University of Agricultural Science. ETSU, its seventeen peer institutions, and three of the selected institutions are listed in the online membership directory of the National Collegiate Honors Council. We also found that ETSU, eleven peer institutions, and two of the selected institutions are included in the online membership directory of the Council on Undergraduate Research.

In our initial assessment, we attempted to address three questions (Table 1) through an examination of each institution's website and personal contacts via email or telephone. While all institutions surveyed provided some manner of public access to undergraduate theses, relatively few of our peer institutions offered either electronic access or cataloging within their university library system. The majority of the selected non-peer institutions provided both access and cataloging (Table 1). Our selection of non-peer institutions was far from

Table 1: Responses to Three Questions from ETSU Peer and Non-Peer Institutions

Question	Peer (17)		Non-Peer (8)	
	Yes	No	Yes	No
Does your institution use an electronic undergraduate thesis system for cataloging honors theses?	4	14	5	1
Are undergraduate theses/projects available for public access?	18	0	8	0
Are your honors theses available through an integrated library system?	5	13	7	1

random and designed specifically to find those institutions actively engaged. From our perspective, we felt the results from our peers were probably more indicative of the common situation: the absence of an electronic repository for undergraduate theses that is integrated into the institutional library.

Our second assessment was a survey of systems in use at institutions that have a functional electronic repository. Sampling for this survey consisted of an Internet search using the search terms “undergraduate” or “honors” plus “thesis” and “electronic.” We found that electronic undergraduate thesis repositories can be found in schools that include small, private liberal arts colleges (e.g., Washington and Lee University), regional universities (e.g., ETSU), state land grant institutions (e.g., Texas A&M University), and major international universities (e.g., Swedish University of Agricultural Sciences). Almost all electronic repositories have been recently established, a few before 2005 but most after 2008. Many repositories offer full text access to the general public, but some restrict access to campus users (i.e., those with university accounts). As of September 2011, the number of theses in electronic repositories ranged from a few in newly established systems to more than two thousand at the Swedish University of Agricultural Sciences. Most electronic repositories contain from twenty to three hundred theses at this time.

Little has been written about implementing an undergraduate electronic repository. A search of ERIC, Google Scholar, and Wilson OMNI returned no published journal articles. A search of the proceedings from the 2007 through 2011 International Symposia on Electronic Theses and Dissertations revealed only two presentations: one focused primarily on the process of instituting a thesis requirement for the West Virginia University Honors College (Garbutt & Simis); the second outlined problems related to integrating electronic undergraduate theses with theses and dissertations for advanced degrees (Fonseca).

In recent years, outlets for publication of undergraduate research have proliferated (for listings, see: <<http://www.jyi.org/resources/ugradPubs.html>>; <<http://urca.msu.edu/publishing>>). Some electronic journals are discipline-

specific while some are broad-ranging; most are refereed to some extent (some use in-house referees); and many are hosted by an academic institution but accept submissions from outside. Regardless of alternative avenues to publication, an electronic thesis repository serves an additional function in that it is an open record of all thesis works completed at a specific institution.

Undergraduate electronic journals are not a substitute for eThesis repositories. Students might prefer, for instance, to publish their work in a professional journal with a wider audience, and students submitting theses that include works of art or that are based on performances may seek alternative types of presentation venues. Furthermore, especially in the sciences, an undergraduate project is part of a larger research agenda where the results may be included in a more comprehensive publication. Undergraduate publication venues, unlike eThesis repositories, might not permit subsequent publication.

SYSTEM SELECTION AND IMPLEMENTATION

The ETSU Honors College launched its initiative to publish undergraduate honors theses online in 2009. Because staff of the ETSU Charles C. Sherrod Library had experience supporting electronic publication of master's theses and doctoral dissertations from our graduate school, we immediately sought to collaborate with our librarians in our endeavor.

We considered and quickly rejected incorporating undergraduate honors theses into the existing system because configuration to accommodate a different audience and workflow would have been difficult and disruptive to existing users. Instead, our library pursued electronic publication of undergraduate honors theses at ETSU as an opportunity to identify and evaluate alternatives to replace the existing system for graduate works.

We were looking for an electronic publishing system that would manage the process of professional publication and would provide for discovery of and access to the final work via the web. Selecting a system that serves both purposes is a complex undertaking, especially since the final choice also depends on the resources available. Features that we desired included:

- System-supplied, unique ID/access URL for each work.
- Compliance with Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH) so that the contents of the collection could be automatically indexed by search engines.
- Customizable web interface that could inform visitors of policies on access and use; incorporate external links to integrate with other university web-sites; and provide guidelines and help to authors.
 - Search and browse with multiple browse options (author, year, discipline, department, or program).
 - RSS feed so visitors who discover the collection and want to be notified when additions are made can subscribe.

- Ability to use multiple files and content file types, including documents, presentations, images, audio, and video, each with appropriate data fields to describe content and/or access restrictions.
 - A file conversion utility to convert document formats like Microsoft Word to Portable Document Format (PDF) in the submission workflow.
- Structure that would provide for a professional publication process, including author submission, editorial review, revision, final disposition, and communication between author and reviewers.
 - A customizable submission interface that would provide the organization for students to describe their work, supply index terms for subject access, choose access restrictions if needed, acknowledge submission agreement(s), and submit all necessary files.
 - Ability to support and tailor multiple publication workflows so that the submission, review, and approval processes might be used by different groups for different publication types.
- Variable time embargoes for individual files (text or other) to address those disciplines with publisher restrictions about open access prior to formal publication and to provide options for protection of audio and video files. We wished to provide author, title, and abstract data openly but restrict access to the full text and apply access restrictions at the file level on both the main content and supporting files.
 - Structure that provided a summary page for each work that could include content description and specify any restrictions on access and use.
- User authentication that could be integrated with the existing campus computing environment. We wanted our students and faculty to log in to the honors thesis repository using the same username and password assigned for campus services. At our institution, we needed the system to support Windows Active Directory domain authentication via Lightweight Directory Access Protocol (LDAP), and SSL, in order to maintain secure user login via a web browser interface.
- Access/use statistics collection and reporting that would detail frequency of access by types of users and could be used by both the institution and authors to assess exposure of research to the outside world.

In addition to these features, we also had to evaluate systems (Table 2) based on two resource-related issues: (1) a hosted or local installation environment, which determines who is responsible for on-going computing maintenance and support; and (2) commercial or open source software, which would directly affect our budget. Commercial software developed and supported by a vendor would represent a large one-time immediate cost with smaller annual maintenance fees or a subscription service for an on-going annual charge but with maintenance included. Open-source software, on the other hand, requires

Table 2: Alternatives for Undergraduate Honors Thesis Electronic Publication Based on Resource Impact

Installation Environment	Software	
	Commercial	Open-Source
Local	<p><i>Content Pro</i> Innovative Interfaces Inc. <http://www.iii.com/products/content_pro.shtml></p> <p><i>Digitool</i> Ex Libris <http://www.exlibrisgroup.com/category/DigiToolOverview></p>	<p><i>EPrints</i> United Kingdom, University of Southampton School of Electronic & Computer Science <http://www.eprints.org></p> <p><i>DSpace</i> DuraSpace <http://www.dspace.org></p>
Hosted	<p><i>Digital Commons</i> Berkeley Electronic Press <http://digitalcommons.bepress.com></p>	Not considered

neither upfront cash nor the competitive procurement process that is often required in state-supported organizations, but it requires technical staff and computing resources to install, configure, maintain, and support it. Each university must evaluate its budget, staff time, and campus expertise to determine the relative cost-effectiveness of open-source versus commercial options.

A review of open-source electronic publishing systems produced at Johns Hopkins University (Cyzyk & Choudhury, 2008 and n.d.) brought *EPrints* to our attention. *EPrints* is free software developed for open-access publishing in the United Kingdom at the University of Southampton's School of Electronics and Computer Science (Gutteridge, Miles-Board, & Brody). *EPrints* uses software components with which we were already familiar, and a search of the Registry of Open Access Repositories (ROAR) revealed several institutions that had been using *EPrints* for master's theses and dissertations for several years. Three exemplars included Glasgow Theses Service <<http://theses.gla.ac.uk>>, University of Birmingham <<http://etheses.bham.ac.uk>>, and University of Nottingham eTheses <<http://etheses.nottingham.ac.uk>>. The Glasgow Theses Service illustrates the exceptional extent to which the *EPrints* user interface can be customized. Further investigation of resources available for documentation, training, and support convinced us to choose *EPrints* as the open-source software alternative that could be implemented using existing resources.

At the time of our system selection, *Digital Commons* (Berkeley Electronic Press) was the only commercial alternative that satisfied most of our functional requirements and required no local information technology (IT) resources to

implement. *Digital Commons* is a software suite designed to showcase an institution's scholarly work and includes an electronic publishing system (EdiKit) that supports editorial and peer review. *Digital Commons* is offered as hosted "Software as a Service" (SaaS) for an annual cost roughly equivalent to one half-time programmer/analyst. Exemplars currently using this software for graduate theses and dissertations include the University of Massachusetts at Amherst, University of Iowa, University of Tennessee at Knoxville, and University of Kentucky. We selected *Digital Commons* as our preferred option if we could find a budget source for the on-going subscription costs.

In 2009, we decided to install *EPrints* locally in order to move the project forward using existing resources while seeking support for *Digital Commons*. The *EPrints* ETSU Honors Thesis repository can be found at: <<http://honors.epub.etsu.edu>>. In 2012, we will be moving undergraduate honors theses as well as graduate theses and dissertations to *Digital Commons*.

IMPLEMENTATION OF THE *EPRINTS* SYSTEM

ACCESSIONING AND CATALOGING AN eTHESIS

The first thesis was accessioned into the ETSU Honors Thesis repository in 2009. After an initial voluntary trial period, we gradually expanded the requirement for use of ETSU eThesis across different programs within the ETSU Honors College. The first year, students within each honors program were offered eThesis as an option, and use of the system became mandatory the following year.

Support for students consists of written instructions available online <<http://www.etsu.edu/honors/Thesis/eThesis.asp>> and yearly in-person training sessions. During our search for repository examples, we also reviewed repository policies, copyright statements, and forms for open-access publishing permission to use as models for our needs at ETSU. Duke University's *DukeSpace* <<http://dukespace.lib.duke.edu/dspace>> emerged as the model we chose to emulate, not only for its breadth, depth, and organization, but also for its presentation of information about open-access publishing, rights and responsibilities of authors, explanations of benefits, and policies and procedures for managing the publication process. We modeled our Undergraduate Honors Thesis Availability Agreement and Non-Exclusive Distribution License form on the one used for electronic theses and dissertations by the Duke University Graduate School (Duke 22).

After two years, we have found the process works quite seamlessly and without any recurrent problems other than occasional assistance for some students to convert files to PDF format. Students and faculty find the system easy to use and fast; the entire submission process can be completed in about fifteen minutes. As of November 2011, there were forty-two theses in the repository comprised of one from 2009, nine from 2010, and thirty-two from 2011. We expect at least fifty in the coming year and then approximately a hundred each subsequent year.

SUBMISSION PROCEDURES

Typically, honors students defend their thesis in a public presentation and then complete final edits suggested by thesis readers and others. These traditional roles have not changed. Upon approval by the thesis committee, students submit written text to the repository in PDF format. Audio files, images, and video materials may be submitted in any format, a feature that is particularly well-suited to theses from both our University Honors Scholars and Fine & Performing Arts Scholars programs, which often include videos of performances, still photographs of artwork, and audio of musical compositions. Artists may have concerns about placing original artwork on the web, so our system enables students to upload video files or photographs at resolutions that allow viewing but are too low to permit quality or merchandisable duplication. Students may also protect audio and video works by restricting access either to repository users (honors students only) or to authenticated campus users.

PERSONNEL

Management of our eThesis system is a collaboration between a librarian and members of our honors college, specifically the directors of various honors programs and the director of undergraduate research (DUR). Access to the system is limited to students expecting to complete a thesis. Each year, the directors of honors programs forward the names of thesis writers to be given access that year and our librarian enters those students into the system. This control ensures that the system does not get overloaded with superfluous submissions.

After a thesis is submitted, it becomes available for review by the DUR, who checks the submission for completeness and formatting. However, unlike dissertations that often have very specific format and page guidelines, we permit a wider variety of styles for undergraduate theses at ETSU. We encourage students to use a format identified with a journal that represents a potential publication outlet for their work. In addition to the electronic submission, we require hard copy of the cover page signed by the mentor and thesis readers as well as a signed release form that grants ETSU permission to post the thesis. The release form offers an option to embargo a thesis from public viewing for up to two years so that researchers can publish work that has not been previously published. On approval by the DUR, the files are moved into the repository and indexed, and an abstract page is generated. Catalog records for works added to the repository are generated automatically on a monthly basis but may be generated and loaded manually as needed.

ANTICIPATED QUESTIONS

HOW LONG DID IT TAKE TO INSTALL AND CUSTOMIZE *EPRINTS*?

Software installation is quick and relatively painless; it can be done by an experienced analyst/programmer in a matter of hours. Customization is more difficult, and the time depends on prior knowledge and the complexity of desired outcomes. The changes we made were few and simple, and the work was done in a matter of days. We found the documentation on the *EPrints* Wiki helpful and the technical support responsive.

WHAT KINDS OF SKILLS ARE REQUIRED?

At ETSU, an analyst/programmer and a librarian formed a team to implement *EPrints* for honors theses. The programmer was responsible for installing the software; SSL certificate; configuring the web server; and scripting backup and data-conversion tasks. The librarian completed the application configuration and customization in consultation with honors college staff. The same librarian currently helps manage honors repository user accounts; monitors the transfer of data from the repository into the library catalog; and corrects catalog records. We made no attempt to redesign the web interface beyond simple branding because we felt it was unnecessary; the “out of the box” interface is clean, functional, and intuitive.

HOW DO PEOPLE FIND ETSU HONORS THESES?

The content of ETSU honors theses is searchable using Internet search engines such as Google and Google Scholar. Within the honors repository, users can browse by year, subject, or program (labeled *Division*) and may search by keyword. The library catalog provides the ability to search the ETSU eTheses and Dissertations collection by author, title, subject, or keyword and to sort results by date, title, or relevance.

HOW DO YOU GET HONORS THESES INTO THE LIBRARY CATALOG?

A locally developed Perl program handles exporting catalog data from the honors repository in XML format and converting it to MARC format for loading into the library catalog. A separate program on the library system fetches and loads the data file. The process runs automatically on a regular schedule.

FUTURE DIRECTIONS

A comprehensive electronic repository of honors theses should include those produced before 2009 that exist only in hard copy. While hardware is available to scan these paper documents into electronic form, the cost of

personnel to complete the task is currently an issue. Staff would be needed to supervise the scanning process, annotate the resultant products, and secure requisite permissions from students and their faculty mentors. The latter is complicated because both our graduates and some faculty members have relocated. Getting permissions thus becomes progressively more difficult the farther back we go in time. However, we continue to seek ways to begin this task, knowing that the effort will better serve our programs, faculty, and graduates.

After implementing the ETSU eThesis system, we dropped the requirement for submission of a paper copy, requiring only the signed cover page. However, we now recognize that digital copies alone do not represent a sufficient long-term archive. We currently are planning to require students to submit one printed copy to accompany the digital copy, but we are also examining other archival options. Our particular concern is how to archive image, audio, and video files that are part of a thesis. A potential solution is the subscription service *OCLC Digital Archive* <<http://www.oclc.org/digitalarchive>> for long-term preservation of honors theses and associated media files.

If the time and effort are expended to create an electronic repository, it is desirable to know the extent to which items are accessed. Unfortunately, we do not have sufficient data to answer this question at present. To have reliable data, we need to distinguish between access to the abstract pages and file downloads and access related to the submission, review, and approval processes. We expect to be able to obtain this type of use data when honors theses are moved to *Digital Commons*.

We also want to know what influence our repository may have on our programs and students. Do honors students use the works in the repository to give them ideas for further research? Does the online availability of theses improve the quality of the work being done? Has the work in the repository generated new interest among faculty or students? Are there long-term benefits to our students who use the repository for their research? We plan to include assessments of these topics in our senior exit evaluations and alumni surveys. Together with the access/use data, we hope this information will detail the benefits of the ETSU eThesis repository for our students, faculty, and administration.

ADVICE AND RECOMMENDATIONS

Our experiences may not be typical of all institutions. However, knowing our own difficulties in locating guidelines and advice to assist in creation of our eThesis system, we are providing our most significant recommendations below. Be sure the system you select offers the following:

- Support for non-document object files (e.g., audio, video, and images) and the descriptive data associated with those objects;
- Processes available that enable images to be adjusted in resolution to ensure they cannot be downloaded in quality form and also options for protecting content of audio and video files;

- Support for variable time embargoes for individual files (text or other) to address those disciplines with publisher restrictions about open access prior to formal publication and to provide options for protection of audio and video files;
- Content that is discoverable on the open web and search capabilities in the local catalog that support identification of undergraduate theses separate from graduate theses and dissertations; and
- Easy conversion of catalog information to the main university library catalog, including checks for errors in any text conversion.

Our most important advice is to understand that undergraduate honors theses represent an intellectual asset of the university that should be recognized and managed as such. An institution's library is a natural for the role of custodial care given its historic responsibilities for managing collection, description, and accessibility of resources that support the teaching and learning mission of the university. You also must provide processes or options to protect any original creative property produced by students. Finally, the issue of long-term archival storage is always a concern with electronic documents because technologies are changing at such a rapid rate. Consideration of permanent archival versions of all documents is a necessity; at a minimum, use of paper copy or microfilm is recommended. Our current conundrum of finding an appropriate archive for video and audio recordings has yet to be resolved.

REFERENCES

- Czyzyk, M. & Choudhury, S. (2008, April 28). A survey and evaluation of open-source electronic publishing systems [PDF document]. Retrieved 11/2/2011 from <<https://wiki.library.jhu.edu/display/epubs/Home>>
- Czyzyk, M. & Choudhury, S. (n.d.). A survey and evaluation of open-source electronic publishing systems, summary table of data. Retrieved 11/2/2011 from <<https://wiki.library.jhu.edu/download/attachments/22964/index2.html?version=1>>
- Duke University Graduate School. (2011, June). Guide for the electronic submission of thesis and dissertation [PDF document]. Retrieved 11/2/2011 from <http://gradschool.duke.edu/documents/policies_and_forms/ElectronicThesisDissGuide2010-2011.pdf>
- Fonseca, A.J. (2010). Taking your four-year institution to the next level: Introducing the ETD concept at undergraduate universities. Abstract, 13th International Symposium on Electronic Theses and Dissertations, Austin, TX.
- Garbutt, K. & M. Simis. (2009). Going Digital: The Implementation of an electronic honors thesis (eHT) at WVU and its impact. Abstract, 12th International Symposium on Electronic Theses and Dissertations, Pittsburgh, PA.

DEVELOPING AN ELECTRONIC REPOSITORY FOR UNDERGRADUATE THESES

Gutteridge, C., Miles-Board, T., & Brody, Tim. (2009, May 13). *EPrints* 3.1.3 [software]. Available from <<http://www.eprints.org/software>>
Registry of Open Access Repositories (ROAR). Search for Repository Type matches any of "e-Theses" AND Software matches any of "EPrints."
Retrieved 11/2/2011 from <http://roar.EPrints.org/cgi/roar_search/advanced?location_country=&software=EPrints&type=theses&order=-recordcount%2F-date>

The authors may be contacted at
levyf@etsu.edu.