

## The Yale Medicine Thesis Digital Library

Arthur Belanger  
Charles J. Greenberg  
Gillian Mayman

Yale University Library & Information Technology Services, Yale School of Medicine  
charles.greenberg@yale.edu  
333 Cedar Street PO Box 208014  
New Haven, CT, 06520-8014  
ymtdl.med.yale.edu  
Keywords: MD, Medicine, Yale, Winternitz, Ingelfinger

### Abstract

*This paper presents the Yale Medicine Thesis Digital Library Project (YMTDL) including discussions of policy issues, project implementation and results. It also discusses how the presence of this project in one graduate school has stimulated thinking about Electronic Theses and Dissertations (ETD) in the broader academic community. The M.D. thesis at Yale, a tradition dating from 1839, has remained an essential part of the contemporary medical education curriculum. Students must create their own hypothesis, identify a faculty research mentor, develop a research protocol, be it literature review, laboratory study or clinical investigation, execute this protocol using current scientific standards and produce a printed thesis. Traditionally, print copies of the final thesis have been transferred to the Medical Library, and research results often appear in scholarly publishing. Unfortunately, locked shelving and skeletal cataloging for medical theses present barriers to access. The Office of Student Research and the Medical Library have begun to address these access barriers with this project. A repository of research findings will be available to a global audience, while respecting the student right to have their work published in high-impact peer-reviewed literature. With the Spring 2002 call for graduating medical student participation, the YMTDL Project team began processing digital copies of the theses and addressing institutional policy issues, with a goal of publicly launching the YMTDL in the early Spring of 2003 (<http://ymtdl.med.yale.edu>). The technology is based on the ETD-db project at Virginia Polytechnic Institute and State University (<http://scholar.lib.vt.edu/ETD-db/>).*

### Project Background

To understand and appreciate the circumstances that led to the Yale Medicine Thesis Digital Library ETD project at the Yale School of Medicine, a short history of the medical student thesis in the Yale medical school curriculum is in order.

The M.D. thesis at the Yale University School of Medicine is a tradition that dates from class of 1839. Most of the 1839 graduates, seventeen in number, presented dissertations on such subjects as dysentery, chlorosis, the color of the skin, or epilepsy. One future psychiatrist wrote on "Hope as a Remedial Agent." The 19<sup>th</sup> century history of the Yale Medical thesis was characterized as a well-intentioned strategy to keep students away from

non-academic distractions, an academic formality without methodological requirements.

A renewal of the scientific foundation of thesis activity occurred during the period 1920-1925, when a unique institutional program of medical education known as the "Yale Plan" was inaugurated. The Medical School Dean of that era, Milton C. Winternitz, featured in the "Yale Plan" the presentation of a dissertation, based on original scientific investigation, as one of the major requirements for graduation. The "Yale Plan" also stipulated that students could also decide whether or not to attend class. Class instruction, when attended, would not result in comparative grading. Students were given the maximum latitude to manage their own time.

Today, students continue to manage their own time and consider attendance at class optional, and all students continue to engage in thesis research, with the only possible exception being students who have already attained a Ph.D. degree. For all medical theses, there must be a specific hypothesis. A wide choice of subjects for research is permitted. Students may also perform a meta-analysis of existing research data in publications, using a new hypothesis to generate new data and evidence. Research must be designed by the students themselves, though each student works closely with a specific faculty research mentor. The production of medical student theses at Yale is a direct effect from the emphasis on the student research experience. The student-mentor collaborative experience at Yale also has a high probability of producing a scholarly research paper submission to a leading peer-reviewed biomedical journal.

The final approved and written thesis is presented to the Medical School Office of Student Research. An Awards Committee critiques and ranks all student theses submitted for honors. The highest ranked papers are presented in a Student Research Day program chaired by the Dean of the School, and graduation prizes are awarded for outstanding student research.

Traditionally, the medical library has received one printed copy of each student thesis. Due to the significant amount of time involved in cataloging, the library was forced to stop doing in-depth subject analysis. Consequently, the only current way to find a thesis in the online catalog is with author or title words. A separate project is

already underway to digitize thesis abstracts and add them to the catalog.

A further access barrier to print theses is their locked shelving location. Print medical theses are considered archival copies. Theses are retrieved only twice a day by library staff and may not be removed from the library. Additionally, all theses written between 1900 and 1974 are shelved in a facility about one mile away. Due to these significant barriers, the original student research contained in the print theses is hidden, inaccessible, and either undiscovered or overlooked.

The emergence of the concept and potential benefit of the ETD to the Yale Medicine student research community has occurred in our unprecedented 21<sup>st</sup> century era of electronic journal publishing. As widespread electronic journal access has reduced the perceived need to visit the library, there is also absolute certainty that locked shelving and skeletal cataloging for the thesis collection was rapidly relegating usage to an unprecedented level of insignificance. The path to obsolescence was taking shape.

## Project Proposal

In the spring of 2002 the Head Reference Librarian at the Medical Library, Charles Greenberg, initiated a dialogue with the Director of Medical Student Research, in order to share the prediction of print obsolescence. The Student Research Director, John Forest, Jr., M.D., reiterated his determination that a student thesis is part of an increasingly interconnected world of scientific knowledge and a rich source of inspiration for future researchers and students. The most formidable inhibition the Dr. Forest felt was the continued presence of the "Ingelfinger thinking."

In the late 1960's, under the editorship of Franz Ingelfinger, the *New England Journal of Medicine (NEJM)* adopted a policy of declining to review or publish research that had been previously published elsewhere, seeking to protect the health of the public from non-peer reviewed research news prematurely distributed to the popular press. Other biomedical journals followed the lead of the *NEJM*. Why would any medical student or advisor intending to submit their research to a prestigious journal risk a rejection because of a prior ETD edition?

Mr. Greenberg related the recent history of open archiving initiatives. He also described how ETD implementations could allow student authors to either suppress access to the electronic thesis for a specified time period or limit access to the institutional campus. With assurance that such reasonable publishing delays and controls could be implemented, the Director of Student Research agreed to co-direct the pilot project and approved an immediate request for graduating student voluntary participation in the Yale Medicine Thesis Digital Library (YMT-

DL) in March 2002, less than two months before the thesis submission deadline for the next graduating class.

## Project Implementation

The software solution for serving theses was at that moment undecided, and the Yale Library at that moment was not a member of the Networked Digital Library of Theses and Dissertations (NDLTD). The "shock of the new" ETD format for thesis would be considerable at a Medical School with much reverence for both history and paper. To immediately build on the decision to launch the project, Medical students in the graduating class of 2002 received an email invitation to be a pioneer author represented in YMTDL. Twelve students, or about 15% of the graduating class responded by filling out and submitting a participation form, and eventually manuscripts in word processing format were received from 8 students. The participation form was modeled on the ETD project at Virginia Polytechnic Institute and State University (Virginia Tech). The low percentage of response was nevertheless considered significant, in that the status of the project at the moment of submission was primitive, consisting mainly of good intentions, a participation form, and an FAQ on the library web.

During the summer of 2002, in recognition that the urgency to launch the project may have avoided sufficient university review, the participation form was submitted to the Yale Office of the General Counsel for comment. In August, a member of the Office of General Counsel approved the format and wording of the participation agreement and recommended valuable changes. A new version of the form was prepared for use in 2003.

In October of 2002, the Associate University Librarian for digital acquisitions and licensing was asked for her opinion on joining the NDLTD. The benefits of membership included use of the ETD software from Virginia Tech, access to national and international experts for ETD publishing and standards, and participation in an international collaboration at a time when Yale was eager to expand its international profile. The Associate University Librarian readily agreed to the Yale Library's membership. However, because of scant interest in ETD publishing in other campus graduate schools, the Library membership would avoid delays in joining NDLTD. In her opinion, Medical student theses are unique enough to receive consideration independent of other dissertation treatment on campus.

In late October 2002, Information Technology Services-Medicine (ITS-Med) SYSTEM administrator Arthur Belanger installed the Virginia Tech ETD software on a secure server. Mr. Greenberg proceeded to convert student word processing manuscripts into PDF format and upload records. The Medical Library webmaster, Gillian Mayman, proceeded to create a prototype public inter-

face that possessed visual synergy with a recent comprehensive library web redesign. During the preparation phase for public debut, additional effort was made to contact student participants that never handed in a word processing file. Several students eventually responded.

## Project Public Debut

The public announcement of YMTDL (<http://ymtdl.med.yale.edu>) was issued in late February, 2003. Completing the work necessary for the public debut in time for the School of Medicine graduating class of 2003 participation was important in "making it real." Because a unique open URL identifies each thesis, the library online catalog records for both electronic and print versions of theses can be enhanced with a link to the electronic version. As the project continues to develop, we expect to participate in the NDLTD union catalog experiments.

Project leaders have discussed ways to improve the quantity of participants in the School of Medicine graduating class of 2003. Publicity was simplified and made more attractive. During library education activities with 4<sup>th</sup> year medical students, the YMTDL project was mentioned. Posters were also conspicuously placed in student mailboxes, campus bulletin boards, and near Medical Library computer labs where students normally work on the final version of their thesis.

The concept of electronic submission and approval through the YMTDL SYSTEM was initially deferred in precedence to developing the author permission and publishing facets of the project. The benefits of ETD submission and approval capability have been discussed with the Office of Student Research, and the development of a pilot demonstration of efficacy is anticipated for 2003-2004 academic year.

## Technical Specifications

The technology for YMTDL is based on the ETD-db project at Virginia Polytechnic Institute and State University (<http://scholar.lib.vt.edu/ETD-db/>). Yale has joined the VT ETD-db consortium. We have downloaded and installed Version 1.7c. Unlike Virginia Tech, who has the SYSTEM installed on a dedicated web server, we have installed their software on our institutional web server (<http://info.med.yale.edu>) as a virtual host (<http://ymtdl.med.yale.edu>). This web server consists of a clustered pair of Compaq Alpha DS20s, each with 1 GB of RAM and 170 GB of shared disk space. The operating SYSTEM is Compaq True64 5.1a. The web server software is Apache 1.3.2.7 with Perl 5.6 and MySQL3.23. We decided to use the existing web server to take advantage of the support structure already in place. This eliminates the need for additional hardware as well as the increased administrative overhead needed for a web server. The institutional machine is maintained by ITS-Med's Production Systems and Web Design and Development groups. The SYSTEM is backed up nightly by our Tivoli Storage Manager SYSTEM.

The software installation was straightforward. The changes we made were primarily in the configuration variables defined in the installation documentation. The only other changes we made were in the HTML headers and footers; we did this to give the site the look and feel of the Medical Library's site (<http://info.med.yale.edu/library>). It is clear that the developers at Virginia Tech spent a lot of time and careful thought in the design and implementation of this software. We would like to publicly thank them for making their code available to the rest of the Academic Community.