10-2011

Building a Digital Archive: The William & Mary Law School Scholarship Repository

Lauren Seney
lauren.seney@colorado.edu

Volume 21

Trends

IN LAW LIBRARY MANAGEMENT AND TECHNOLOGY

Edited by Philip C. Berwick ✦ *For academic, firm, corporate, and government law librarians*

Building a Digital Archive: The William & Mary Law School Scholarship Repository[I]

By LAUREN SENEY, Access/Technical Services Librarian, Wolf Law Library, College of William and Mary, Marshall-Wythe School of Law, Williamsburg, Virginia

In early 2010 the Scholarship Repository was merely an idea at the William & Mary Law School. Little did we know that a year later we would boast eighteen collections containing more than 6,000 items.

As ideas about establishing a repository began to swirl, we knew an additional librarian needed to be brought into the equation. Several recent retirements factored into the decision to restructure, add a new professional position, and institute a repository. The new librarian position—the Access/Technical Services Librarian—was created in the spring of 2010 to assist both the Access Services and Technical Services departments, and also to be primarily responsible for creating the William & Mary Law School Scholarship Repository (http://scholarship.law.wm.edu/).

Getting Started

Repository preparation began that spring by purchasing the back files of our law reviews from William S. Hein & Co., Inc, hosting a mini-symposium on institutional repositories to inform ourselves and others about best (and worst) practices, deciding to use Berkeley Electronic Press' (bepress) Digital Commons platform, selecting hardware and software to manage digital files, and working through the initial site design. A team consisting of the Head of Access Services, the Head of Technical Services, and the Access/Technical Services Librarian managed the project and worked with the director on the repository's blueprint. As the design neared completion, we decided how to store the files, began manipulating the law review articles, and used faculty members' resumes to initiate searches for their publications. By late July, with the help of staff members throughout the library, we had primed a portion of our flagship journal and began loading three articles per faculty member.

Locating and Preparing Content

We used a two-pronged approach to populating the repository. The library worked on our faculty scholarship and *William and Mary Law Review* collections simultaneously. The law review staffs were excluded from the process so we could preserve a cohesive repository design and maintain a consistent metadata set in the journals. This was an easy decision for us; we saw the repository as an extension of the library's print archive, and felt that digital content should be maintained in the same way.

The bulk of the work in preparing content for the repository fell to the new Access/Technical Services Librarian, though both department heads and several other staff members pitched in during the quieter summer months. As we approached the new school year, with work proceeding on both of these collections, we needed to document the most efficient workflows in preparation for a large student workforce. However, the procedures remained in flux throughout the school year as tasks were added and more efficient processes identified.

Our faculty scholarship was the inaugural content in the repository on July 20, 2010, followed by the most recent volumes of the *William and Mary Law Review* in early August. As school began, management of the repository project was assumed by the Access/Technical Services Librarian.

The Workforce

The repository's nine student workers, all first year law students, started on August 23 and we set the goal of loading all 51 volumes of the *William and Mary Law Review* and the majority of our faculty scholarship by the end of December. With almost a month of staff labor behind the project, we really weren't sure how we would achieve it. Despite that, we pushed forward with the repository students scheduled 36 hours a week (4 hours each).

We received two DVDs of files from Hein that contained our journals separated at the volume level. We could not incorporate them into the repository in this format, so we used the students to reduce them to individual articles before loading them. They used the procedures that had been streamlined during the summer in which individual articles provided the data for pre-determined metadata fields. We quickly found that we had underestimated how productive our students could be; their pace couldn't be matched by their librarian supervisors. There was also an unanticipated bottleneck in the workflow because we could most effectively back-check student work when none of them were working on a collection. Nine students working on two collections and the need to review all work before it went live on the site, left us with a small window to publish items to the web. To solve this, we initiated the design process on six additional journal collections, so in a few short weeks we were working on eight collections simultaneously.

It was also necessary to expand the students' assignments with a focus on tasks that were originally slated to be retrospective. This included expanding the metadata that we were associating with all publications. Our goal was to increase the repository's visibility in online search engines by associating keywords with each article. We had the students search for terms in electronic and print legal indexes to associate with the articles and pull them into a spread sheet which was reviewed by a librarian before they were loaded. During the fall semester the students also used online journal indexes to expand searches for faculty members' publications and used Adobe Acrobat Pro to edit and perform optical character recognition (OCR) on documents that required it.

The students' enthusiastic pace overwhelmed our quality control procedures in our inaugural months. The repository was not anyone's full-time job; however, validating student work, which included verifying keywords and providing missing metadata, before publishing content to the web was often more than a full-time task.

Results

In mid-November we achieved what none of us had dreamed possible six months earlier. Our entire journal archive of 142 volumes and almost 4,000 articles had been added to the repository, as well as nearly all our faculty scholarship. When we ended the calendar year just shy of 5,000 items, we started looking for additional collections to incorporate into the repository.

To prepare for the students' return in January we purchased a small, book edge scanner (a Plustek OpticBook 3600 Plus) and began looking for better OCR software; ultimately purchasing OmniPage Professional 17 in March. This allowed us to begin scanning documents on the history of the William & Mary Law School when the students were back on campus. We knew from the previous semester that the students were going to outpace us, so the collections were prioritized and cleanup work on the documents was based on those priorities. By the end of the spring semester we had digital copies of the Law School's admissions brochures, annual reports, graduation documents and several other historical collections. Many of these have a presence in the repository, but the process of cleaning up and reducing file sizes of the documents is complex, and continues to this day.

What Now?

Our thoughts turned to additional types of media, and we started adding photographs to the repository in the spring. We are now beginning to determine how we will create a video archive that includes presentations, graduation ceremonies, and the like. We hope to draw on what we learned during our first year of development, implementation, and assessment. Careful analysis during all of these stages supports our robust repository. We've learned that no matter how much you know—or think you know—your repository will be unique to your institution's goals, the content, and most importantly, to your staff. You can learn from others, but should plan on confronting the unexpected.

Notes

1. This article only breaks the surface on the implementation of the William & Mary Law School Scholarship Repository. More information can be found in the slides from Lauren Seney and Linda Tesar's presentation "Digitally Archiving Your Law Reviews" (available at http://scholarship.law.wm.edu/libpubs/21/).

Lauren Seney is Access/Technical Services Librarian, Wolf Law Library, College of William and Mary, Marshall-Wythe School of Law, Williamsburg, Virginia. Email <lpsene@wm.edu>.

Trends is published by
William S. Hein & Co., Inc.
1285 Main St. / Buffalo, NY 14209
Phone: (800) 828-7571; (716) 882-2600
Fax: (716) 883-8100
Email: mail@wshein.com
ISSN 0893-6773
The complete collection of articles published in ***Trends*** is available in HeinOnline (http://heinonline.org)

Editor: Philip C. Berwick, Associate Dean
for Information Resources
Washington University School of Law
Law Library
Campus Box 1171
St. Louis, MO 63130
Phone: (314) 935-4042; Fax: (314) 935-7125
Email: <BERWICK@WULAW.WUSTL.EDU>