We’ve worked with several law reviews and libraries over the last year to post the full law review archives, and to publish going forward. I’ve found there are really two “ins” to bringing your law reviews into the repository. One is by publishing going forward, the other is by starting with the archives. Ultimately, you’ll probably be aiming to do both but you can start with either.

Why think about the archives?
--Archival management and organization is the purview of the library.
--The archives are easily accessible (more on that in a minute).
--If you’re having concerns about getting your law reviews on board, bringing up the archives lays good groundwork. It establishes the history of the law review in digital, branded form, and helps hesitant law reviews see their place in that.
--Law review editors are generally persuaded by precedent, and a digital, comprehensive archive of the law review helps them see their place in that.

I’d recommend Kathleen Cowan’s presentation, “Owning the Online Future of Your Law Review,” to understand the main reasons a law review would want to begin to publish going forward in the repository. I’ve used it myself with several groups of law review editors and it has helped greatly to guide the conversation.

I said that we’d worked with several law reviews over the last year to bring the archives up. What follows are a couple slides showing some of those law reviews, some of which we’ve worked very closely with (Boston College and Marquette), and others of which we’ve learned from after or during the process.
Purchased digital files from Hein, received those in volume level PDFs.

Worked with library staff and students to break those into article level PDFs and capture the metadata.

Batch imported via XML.

Began to publish forthcoming papers as well as adding conferences/symposia.
We worked with BC Law on a digitization project.
They sent us their print archives and we coordinated with a company in California to scan and capture metadata for the full fifty years.
Built the journal and imported metadata and PDFs by batch upload via Excel.
Purchased archives from Hein.
Cleaned up PDFs.
Had metadata in digital form already.
Posted archives.
Has been posting law reviews to the IR for several years.
All three law reviews and archives are up.
Worked with them this year to convert to a journal structure for improved display and increased publishing tools.
As I see it, we can break the steps of creating your open access law review archives down into these five steps.

1) Get the archives in digital form
2) Prepare your PDFs
3) Capture the right metadata
4) Design and build the e-journal
5) Import and post your content
I'll be discussing the first three steps today in detail. I'll refer you to the presentation, “Owning the Online Future of Your Law Review” (http://works.bepress.com/kathleen_cowan/2) or to bepress Client Services to discuss the final two steps.
Both are reasonable options and have benefits and drawbacks.

**Purchase your archives**

--So far, everyone we’ve worked with who has purchased their digital files has purchased them from Hein.
--Cheaper cost at the outset
--Some have received archives in volume format, others in article format. Important to ask for what you need.
--You only get the files, so you’ll have to capture the metadata. Western New England for instance already had that data on file. If you don’t have the data and need to capture it, this could increase the cost of the project. Sometimes library staff takes this on, other times, students do the work. Or, you could consider outsourcing to your editorial staff or to a data entry company

**Digitize your print archives.** We worked with Boston College Law to do this.
--More costly at the outset.
--You can pay for OCRed scans, and well worth it to do so.
--Ask for files in both article AND issue/volume format.
--Metadata captured during scanning process.
--Total cost of scans plus metadata capture is known up front.
What do you need to know about your PDFs?

**Post PDFs by article.** You may have considered posting PDFs by volume or issue. I’d highly recommend posting by article. This allows for:

-- Targeted readership: readers download only what they’re looking to read. This means you get better stats on what is most popular. Also, title is important in Google search results. Titles are much more compelling to click on than “Vol 30, Issue 4.”

-- Tracking usage: You know what topics, titles and authors are most popular.

-- Usage reports targeted to every author. You can include an author’s email address with a submission. Each month, that author will receive an email detailing the downloads for all of his or her articles in the bepress system.

**OCR.** PDFs should always be OCRed. This means that they are not stored just as an image, but have a text layer embedded within them in order to facilitate full text indexing and searchability. If a PDF is not OCRed, you can do so within Adobe Professional. Ideally, your PDF provider will ensure that the PDFs are OCRed.

OCRing also makes it possible to search within the document and across the archives. This is because it can be full text indexed. Documents in a Digital Commons repository are full text indexed in the system as well as in search engines like Google and Google Scholar. Full text indexing dramatically improves discovery.

Law review editors love that their reviews are full text indexed via Google and Google Scholar. It’s an additional big win that the full archives are aggregated and searchable in the system.
3) Capturing the Right Metadata

There are two components to any repository record:

1) Metadata
2) PDF

An optional component is supplemental material.

Let's take a step back. In order to talk about capturing the right metadata, I need to talk about what goes into a repository record. A repository record has two components: metadata, and a PDF (or other discrete file, but for the purposes of the law review, we'll think in terms of PDFs).

You can optionally add supplemental content to your records, or to your journals as a whole. I won't address this any further in this presentation, but you can always talk to the bepress Client Services team about adding supplemental content, multimedia, blog feeds, Twitter feeds, and capturing conferences or symposia.
This is a repository record from Boston College Law Review. It includes both metadata (like title, author, etc) and access to the PDF (by clicking on the download button).

You are looking at the metadata we’ve chosen to display to users. Additional metadata is captured for OAI purposes, as well as management and organizational purposes, but is not displayed. We can choose to display more or less at any given time.
For the Boston College Law Review, we captured two types of metadata: 1) metadata relevant to the journal as a whole; and, 2) metadata relevant to each article.

We used the Table of Contents in the law review as a guide. You can see here an early mark-up of one table of contents. Here, we’ve marked the metadata relevant to building the journal – volume, issue, title, special title, publication date.

This told us how many volumes and issues to build, with what publication dates and titles.
We also captured metadata relevant to each article. We did so with three goals to guide us:

1) We needed to construct a good table of contents for each issue.
2) We wanted to be able to generate a proper citation.
3) We wanted to increase searchability and discoverability.

Most of this metadata was also in the Table of Contents, but we had to go searching for the names of student authors for notes and casenotes. They never get front page billing.
Then we captured metadata relevant to each article. Here you see a spreadsheet where all the data was captured. The company we worked with compiled the data into massive spreadsheets that we provided to them.

The data that helped us to build the journal (Volume, issues, title) is over on the left and has blue headings. The data describing each article record (like what goes into the blue book citation, for instance) makes up the majority of the columns on the right. (I’ve hidden a couple author columns on this spreadsheet so it would fit into the slide.)
Here’s the Table of Contents for an issue of the Boston College Law Review.

One of our guiding principles and goals was to generate a table of contents. We wanted to make sure we had all the data we’d need to construct a table of contents, and to modify that table of contents later on. For instance, right now we display the following on the table of contents: Volume and issue; type of article (so article, remark, essay, casenote, note, etc); article title; author or authors. But we left it flexible so that if we want to display first page at some later point, we can do so.

The green arrows point to some but not all metadata elements.
Here’s an article metadata page from the Boston College Law Review. I’ll talk a bit about working toward our other goals in the context of this page.

Goal #2 was to ensure we could construct a proper citation. The bottom green arrow points to the citation.

Goal #3 was to improve search and discovery. First, keep in mind that we capture more metadata than what we choose to display on the user facing end. All of these fields are included in a search. Secondly, full-text indexing is the most important factor in improving search and discovery. However, full-text indexing can be moderately improved with focused keywords and subject headings, as well as abstracts. Generally, subject specialists are leveraged to provide keywording and subject headings.
Finally, to finish up, you'll work with Client Services to design and launch your journal, where you then import your content.

Contact Client Services to learn more about the journal design process and get a copy of the journal design forms.
And learn more about the batch import process here:
http://digitalcommons.bepress.com/reference/10/
Questions?

Courtney Smith
Outreach and Scholarly Communications

csmith@bepress.com

Bepress Client Services

support@dc.bepress.com

510.665.1200 x2