

Beyond Digital Repositories

The (sometimes) wild world of open online access to legal materials

By Mikhail Koulikov

At their heart, both conceptually and historically, libraries are about books and librarians are those who work with books. Except, of course, what underlies all of librarianship is not books but rather the distinction among data, information, knowledge, and wisdom. Books are a convenient package for knowledge and information, but there is nothing about books in themselves that requires them to be the only possible packages.

This may be a slightly radical notion for the library world in general, but it is also something that every law librarian learns a week—if not less—into his or her job. In other fields and disciplines, the book really may be the basic unit around which the entire discipline is organized. But in law, that is rarely, if ever, the case. Rather, the most important operative units of law librarianship are individual blocks of text—statutes and judicial opinions—that may be published in book form, but do not have to be. Sure, the literature of law (and lawmaking) is still structured around the book paradigm, but there is virtually no reason for anyone to ever read the entire volume of a reporter straight through.

Of course, there is a reason why reporters and statutes exist in the form in which they do. Until recently, there really was no viable information representation alternative to the book. And so, to access any legal document, the reader had to access the book or book set, figure out that book's organizational structure, and then locate the individual item. But this is also exactly why the law is structured around a complex and complicated system of citation that allows for direct print access not only to books but also to specific documents, sections, paragraphs, and even lines.

Information Access in a Digital World

Now, more than 30 years into a world where computer-assisted legal research is the norm, when law librarians look at what the Internet means for all of us, what do we see? On the one side of the spectrum are the familiar commercial databases—Westlaw, Lexis, HeinOnline—that are essentially heirs

to the legal publishing environment as it existed before computers. On the other is the chaos of Web 2.0—user-created and user-submitted content that may be of interest to library and information science scholars but generally is largely irrelevant to the law librarian or lawyer (although this is also certainly changing as courts make a move to publish cases online and become increasingly comfortable with citing Wikipedia articles in their decisions). But what—if anything—is between the extremes?

Trying to reinvent the wheel and come up with an entirely new legal database is certainly not impossible, as can be seen in efforts by the likes of Bloomberg Law and Fastcase to challenge the Westlaw/Lexis duopoly. Reinventing the wheel, though, tends to require at the very least a major commitment and significant expense.

But of course, probably the single greatest effect of the Internet has been a radical decrease in the cost of entry into the dissemination-of-information market. Thousands of people participate in sharing information online. At one extreme, this can take the form of outright unauthorized distribution of audio, video, and, as is being noted more and more frequently, textual content. But voluntary information sharing is also what allows Wikipedia to exist. The bottom line is that from a purely technical standpoint, there is nothing preventing a single person (or an organization) from identifying a particular information niche and deciding to fill it.

Introducing the Micro-Archive

Digital repositories—that is, collections of published materials (usually, scholarly articles and research papers) that are placed online for free, open-access download—are a fairly recent development in the field of information management, but they have quickly made a major impact on just how scholarly and professional communications can take place. Some of these repositories, such as SSRN and SelectedWorks, are inherently multi-disciplinary, focused on individual authors, and operated by third-party organizations. Another type includes the institutional repositories being established by many academic institutions to disseminate the work of their faculty.





Micro-Archives: Diving In

Definitions and discussions can only go so far to illuminate the meaning of “micro-archive.” The easiest way to really get an idea of what a micro-archive can be—namely, the law librarian’s best friend—is to look at several examples.

Virtual Museum and Archive of the History of Financial Regulation

<http://sechistorical.org/museum/papers>

The scope of this collection encompasses several thousand documents, each somehow related to the U.S. Securities and Exchange Commission. The earliest is from 1912; the most recent from a few months ago. Many are government documents and so conceivably available from a federal repository or the National Archives, but placing them online on an open-access website makes them accessible from anywhere, by anyone, and requires no more skill than the ability to run a Google search.

ERISA Advisory Opinions

www.erisaadvisoryopinions.com

ERISA Advisory Opinions are issued by the Department of Labor (DOL) for the purpose of interpreting specific provisions of the *Employee Retirement Income Security Act* of 1974. They are available on the DOL website at www.dol.gov/ebsa/Regs/AOs/main.html going back to 1992. This micro-archive, edited by an associate at the New Jersey law firm of Niedweske Barber, extends the coverage all the way to 1974. Compared to a Westlaw or Lexis database, it is primitive. But much as comparing Wikipedia to the Encyclopedia Britannica ignores the simple fact that one is free and one isn’t, for a user who does not have access to the relevant database, a resource like this may end up being nothing less than absolutely invaluable.

The New Jersey Digital Legal Library

<http://njlegallib.rutgers.edu>

With a state-specific focus, the scope of this micro-archive is on an ever-growing collection of highly specific administrative documents, such as the opinions of the state attorney general, the state governor’s executive orders, and departmental policies. Previously, most of these would only be available from the state archives or the state library, and even then many are thoroughly obscure and would probably be unfamiliar to all but the most expert legal researchers.

But there is also a third type. Websites gather together, index, and place for potential distribution documents and materials from a range of sources that are united by some sort of common theme. For lack of a better term, this type can be called the “micro-archive.”

What Micro-Archives Are and What They Do

A micro-archive is decidedly not Web 2.0. Rather, it is controlled tightly by whoever is actually in charge of uploading content. But from the user’s perspective, if it turns out that a micro-archive exists for the area or issue that you (the user) are interested in, the details of how or why it does are decidedly secondary. So, in a way, what micro-archives can be compared to most directly are traditional print subject bibliographies, except rather than listing only the citations to specific documents, the micro-archives will also contain the documents themselves.

Beyond this basic definition, the kinds of micro-archives that exist out there already—and that may be of interest to law librarians—vary widely. They can be organized and supported by a state library, by an academic institution, by a nonprofit, or even by a law firm or corporation. Some have explicit policies on how their contents can be used; others are either not that advanced, have simply decided not to bother, or index and archive documents that are by definition not copyrightable. Because these micro-archives are largely labors of love, they certainly will never compete with either commercial databases or with something like FDSys in terms of user-friendliness, search functions, or even metadata and annotation. Where their strength lies is not in breadth of coverage but in specific focus on materials of a very specialized, very specific type.

Micro-Archives in Practice

Accepting the idea of the micro-archive requires a bit of courage—not from the point of view of the end user, who simply needs access to a document and couldn’t care less how the document is acquired, but from the point of view of the librarian.

We as librarians are still taught to think in terms of books as whole structures, not in terms of what is actually in them. If nothing else, this is why the idea of content-neutral citation to court decisions is still so novel; it makes sense and it is certainly appealing, but going from an appealing idea to an effective execution clearly takes not just effort, but a thorough readjustment of our own understanding of the nature of how information can be organized. And even the technical ability to run full-text

searches on statutes and case law to retrieve only instances of particular terms has been criticized for making it too easy for researchers to ignore contexts.

The micro-archive concept, on the other hand, is a powerful reminder of the shift in the definition of what is “published” and “accessible.” Until recently, only a document that was formally included in a published compilation that was then in turn listed in some sort of index would be available for review. Now, though, any piece of writing that exists can potentially be located by a user. And so, for the law librarian, the paradigm has shifted. Before, when faced with a document title and an unfamiliar citation, the first step would be to decipher the citation, then search either the library’s collection or the databases it subscribed to and hopefully be able to locate whatever source the citation referred to. This set of responses will never disappear entirely. But the truly effective librarian will have the courage to accept the fact that the existence and growing importance of micro-archives is making “Google it” an entirely acceptable first step in retrieving legal and government information online.

What then, one may justifiably ask, is the purpose of the librarian? Why can’t the student, the professor, or the associate do the same thing? Well, the simplest answer is *time*. The best use for the time of the associate or the student is not research, but analysis and interpretation.

The more elaborate answer has to do with what is probably the best mark of an excellent librarian—not knowing an answer but rather knowing where to *find* the answer. Using a standard online search engine in a truly effective manner requires a great deal of familiarity with its conventions and protocol. And beyond that, the skilled librarian will know exactly what kinds of resources are available out there, what kinds of information and materials they can contain, and what are the relative advantages and disadvantages of each.

The micro-archive is not going to be the end-all and be-all of human knowledge. But every time one opens—and is identified by its potential users—the universe of just what kind of information can now be meaningfully changed into knowledge and wisdom expands a little bit more. The role and goal of the effective librarian is to know not about books, but about information sources, and the micro-archive concept is a powerful reminder of the fact that information sources, just like information itself, are never static. ■

Mikhail Koulikov (mkoulikov@nyli.org) is reference/research librarian at the New York Law Institute in New York City.