
by Jonathan Zittrain

In the late Nineteenth Century the Spanish Marquis de Olivart—a writer, ambassador, professor, and sometime foreign minister—had amassed an enviable collection of some fourteen thousand international law books. He then gave the collection to the Spanish government, moved, he said, “by a patriotism that was as ardent as it was sterile.”

The government didn’t stick to the terms of the gift in maintaining the collection, and the disillusioned Marquis managed to claw it back. Word got around that it might be for sale. The Harvard Law School Librarian lobbied to put in a bid—one that would cost the school nearly every spare cent it had. After contentious discussion, the faculty approved. Gold bullion was deposited into the Marquis’s bank, and the books were smuggled out under
cover of night, apparently to avoid inciting the Spanish government to ban their export, or perhaps to avoid the eye of the Marquis’s lenders.

Thus did a law school library score a coup of materials, and whet its appetite for more. The next year, acquisitions vaulted Harvard’s collections to over eighty percent of all the world’s English law books published before 1601.

Tales like these, shared in the law school’s own official reminiscences published in the 1950’s and 60’s, reinforce the notion of a library as a storehouse of rare and precious things. And with good reason. Libraries originated at a time when books were expensive, difficult to copy, and thus perhaps irreplaceable. To be able to pool them in one place had some real benefits:

- Materials could be shared among lots of people, whether for mere browsing or full reading. The cost could be borne by an institution on behalf of its patrons rather than individual collectors.

- Books could be indexed and grouped together, so that someone wanting to learn about a topic would have lots of possible books to consult in one visit. And they could be secured, both against the elements and against marauders. Many libraries don’t just look like fortresses—that’s what they are.

With those features, you can see why people thought more was more: a library’s greatness could be measured by how many volumes it possessed, and the prestige of an institution of learning could be gauged by the size of
its library. To this day, the Harvard Law School Library, which I direct, claims distinction as the largest academic law library in the world.

Volume of volumes was never the whole story, though. Libraries have, over time, inhabited the roles not only of guardians of knowledge, but of curators, and not merely for their owners or immediate communities, but for the world at large. The curatorial role became crucial as the range of possible things to read vastly exceeded the amount of time someone would have to read them—and finding something responsive to one’s query required mastering the baroque art of search. Librarians apprenticed to degrees in information science to know how to find things, and they coupled that skill with a professional commitment to neutrality, or at least absence of bias.

Fast forward past the year 2000, and both carefully accruing collections and expertly guiding people through them might seem quaint. The Web contains information about nearly anything, and search engines effortlessly index the Web. The only thing left, by this theory, is mopping up: retroactively digitizing materials that are sitting quietly in archives and depositories. Once that’s done, the only function a general-purpose library could serve is as a place to put terminals with Internet access and hold classes on Web browsing, so that anyone can use one. That may explain why the use of the Library folder in Mac OS, complete with Parthenon-columned icon, is as a repository of system files usually hidden to the user. Valuable documents and files go elsewhere.
Despite these appearance, libraries—real ones concerned with guarding and curating knowledge—remain crucial to free and open societies, and not simply because their traditional services within academia, from curation to preservation to research, remain in high demand by scholars. More broadly, they crucially complement the Web in its highest aspirations: to provide unfettered access to knowledge, and to link authors and readers in new ways. Here’s why.

First, information may be easy to copy, but it’s also easy to poison and destroy. The Web is a distributed marvel: click on any link on a page and you’ll instantly be able to see to what it refers, whether it’s offered by the author of the page you’re already reading, or somewhere on the other side of the world, by a different person writing at a different time for a different purpose. That the act of citation and linkage could be made so easy to forge and to follow, and accessible to anyone with a Web browser rather than special patron privileges, is revolutionary.

But the very characteristics that make the distributed Net so powerful overall also make it dicey in any given use. Links rot; sources evaporate. The anarchic Web loses some luster every time that something an author meant to share turns out to be a 404-not-found error.

I co-authored a study investigating link rot in legal scholarship and judicial opinions, and was shocked to find that, circa late 2013, nearly three out of four links found within all Harvard Law Review articles were dead. Half of the links in U.S. Supreme Court opinions were dead. Before the Web, the only common link was an analog: an author had to name with great precision a source, and a reader could nearly always take that citation to a library and expect to be able to access the source. Labor intensive, but the
barriers to publishing meant that most stuff linked was in books and other systematized formats that libraries were likely to store. Post-Web, much can be published without burdensome intermediaries, but if it vanishes, it vanishes.

That’s why the HLS Library is proud to be a founding member of perma.cc, a consortium complementing the extraordinary Internet Archive, seeking to preserve copies of the sources that scholars and judges link to on the open Web. The preserved materials can be readily accessible for the ages, placed on the record within a formal, disinterested, distributed repository of the world’s great libraries. This is especially important as information might not only vanish, but be adulterated. When Barnes and Noble can offer a book as canonical as War and Peace with key changes quietly (if accidentally) made to its vocabulary, it’s a signal that our knowledge requires actual guardians ready to preserve and fight for its integrity, rather than, in the words of John Perry Barlow, merely vendors treating ideas as “another industrial product, no more noble than pig iron.”

Indeed, Barlow, writing in 1996, didn’t fully appreciate that ideas could become less industrial products than industrial services. That is, they could be rented rather than sold—at great profit. The world’s great research libraries have thus, in a time of greatly lessened barriers to the movement of information, oddly found themselves finding it more costly to acquire—at least when they want to gain access to articles published in prestigious academic journals. These journals are, in most fields, published by for-profit vendors. Before the rise of ubiquitous networks, profits were limited by the fact that vendors, like regular booksellers, could charge roughly only one price for their wares. Try to discriminate too much among consumers, charging more for those with larger wallets, and the used book market would intervene—why pay a lot for a book when you can borrow or buy it from someone who managed to get a preferential price?

That’s changed. Institutions, on behalf of their patrons, now can be charged special (and secret!) prices just for them, and only for access for an interval. Stop paying your fees and your “collections” disappear. With information now a service rather than a product, we’ve seen prices (those we can discern, at least) go through the roof.
The most bizarre thing about the serials crisis is that the works that have become so pricey are, by and large, generated by the very scholars and institutions getting soaked to access them. One might wonder why scholars keep publishing in journals that treat them that way, when it’s so easy to publish online. And the answer has been that publishing in prestigious journals is important for career advancement, and until other journals can come along ready to add more value to the scholarly ecosystem, we’re stuck. That’s what open access publishing has been about—allowing the contents of journals to be accessed and shared freely—and it may yet do the trick. But even there one can find complex incentives and micro- and nano-economic forces to account for: some so-called open access journals are spammy fronts for a quick buck made from authors who pay open access “fees” to get published, and even distinguished venues have found themselves publishing gibberish papers as they expand the scope of their offerings to profit from any author willing to pay a fee to present at a conference.

This points to a second role for libraries in addition to safeguarding otherwise-ephemeral online materials: libraries are natural places to help keep the markets for information honest, and they can even participate in them themselves. Once freed from the trap of spending every last penny on renting access to journal articles, they can sponsor their own journals and other proceedings of scholarly work, working with faculty to hone the best ways to select, edit, and publish scholarship. Indeed, they can help build discussions around it, including on work that has yet to be formally published. There is little reason now to separate the pre- and post-publication stages. The former need not be invisible, and the latter need
Law will be a particularly interesting area in which to experiment, if the public domain cases on which much scholarship is based can themselves be digitally freed for all to study. That’s because legal journals are not only non-profit, but also run by law students. That’s a feature, not a bug: those students can begin to work more closely with their professors to experiment with new publishing methods and styles, taking risks that entrenched, profitable publishers in other disciplines would have no reason to adopt. And we can re-imagine textbooks starting with legal casebooks, which cost hundreds of dollars each today, even though they comprise mostly public domain material: judicial opinions. Libraries can not only help produce standard course texts at low or no cost, but more important, make them remixable, so that the courses themselves can evolve as students and professors adopt and adapt others’ syllabi.

Such experiments in scholarship can draw upon the information science skills, and near-sworn neutrality of the modern librarian. In a world suffused by truthiness—where anyone can adopt the form of a think tank or a scholarly paper without actually adhering to fundamental norms of pursuing truth through empirical data and rigorous argument—we more than ever need a space where knowledge is not only generated but venerated, and debate is earnest and fierce but not bought and paid for. And where finding particular information amidst a Babel-like ocean of bits is not consigned to a handful of for-profit search engines and information-feed-generators who can tweak what people see for their own purposes—or at the behest of regulators.

That’s why it makes so much sense for us to be thinking about the revitalization rather than replacement of the library. In a world suffused with so much transient information as to inspire epistemic paralysis, we acutely need libraries’ power, independence, and ethos: institutions conceived to fight on behalf of their patrons, which is to say for the public and for the preservation and intelligibility of the public record.

Jonathan Zittrain is the George Bemis Professor of Law at Harvard Law School and the Harvard Kennedy School of Government, Professor of Computer Science at the Harvard School of Engineering and Applied Sciences, Co-founder of the
Update: Knight Foundation, in the words of its co-founder, seeks to “bestir the people into an awareness of their own condition, provide inspiration for their thoughts and rouse them to pursue their true interests.” Fittingly, the Foundation’s next $2.5 million news challenge solicits answers to the question, “How might we leverage libraries as a platform to build more knowledgeable communities?” I’ll be entering at least one proposal—and anyone working in this area should strongly consider developing an idea. Details are available here and here. Deadline is September 30, 2014.