

*This paper was commissioned by the Boston Library Consortium with support from the Alfred P. Sloan Foundation for distribution at the Universal Access Digital Library Summit on September 25, 2008 in Boston.*

---

## **FREE OUR LIBRARIES!**

### Why We Need A New Approach to Putting Library Collections Online

By Richard K. Johnson

---

Long ago, mankind figured out that special kinds of public institutions are needed to preserve cultural heritage—libraries and museums. They’ve been a grand success, enabling successive generations to learn from and build on the record of man’s achievements and failures.

But a momentous, ill-considered shift is now afoot that threatens to limit the public rights in the collections assembled and maintained, often at public expense, in libraries around the globe.

Today Google and other businesses are scanning millions of books from the world’s great libraries and offering access to them on the Web. This conjures up the vision of a vast, free, Internet public library of accumulated knowledge. It seems like a marriage made in heaven—the union of corporate capital and enormous library collections, carrying knowledge into virtually every home and workplace.

Unfortunately, it’s not.

#### ***Paradise Lost?***

What is emerging is less than the vision of our dreams. Barriers to use of digital texts are popping up almost as fast as books are being scanned. The rights that readers enjoyed in the print world are being eroded as books are electronically transformed. In the process, we’re at risk of losing some of the rights we enjoyed under copyright.

The problem of commercial control is illustrated by the plight of public domain works. These works belong to the public. They are, by intention, not copyright protected and may be used by anyone for any purpose without asking permission. The public domain is a rich resource and an essential foundation for the Internet public library. It includes centuries of great literature and is a chronicle of civilization and learning.

Before the Internet, there was little argument over what people could do with public domain works. They could do *anything*. But technology makes it possible to impose new technical and contractual protections that can be applied willy-nilly to in-copyright and public domain works alike. The lawyers and economists call this the “enclosure” of the public domain and it looms large as the future of the Internet public library plays out.

For example, companies that are scanning library collections have required users to gain online access to books solely via proprietary search engines. They also have prohibited users from employing third-party computing tools such as screen readers for the visually impaired or scholarly text analysis tools. In effect, they are securing and enforcing a monopoly on the digital texts of works that are in the public domain. Of course, other businesses might also scan a library's collections, but this is a substantial undertaking and, as a practical matter, isn't apt to happen anytime soon.

The reason for such heavy-handedness isn't nefarious intent on the part of those undertaking digitization of library collections. After all, it's their responsibility to shareholders to profit from the investment and to build enduring competitive advantage. In the bargain, information sharing will dramatically expand.

But expanded access is a by-product not the objective. As long as commercial motives rule, critical decisions will be driven foremost by business imperatives that often are directly at odds with the public interest. This will inevitably reduce the rights of readers, especially with respect to out-of-copyright works.

Abandoning the public domain is also bad for the institutions that open their shelves to scanning. Most of these institutions are universities. Whether supported by tax dollars or private endowment, they usually express their mission in terms of advancing knowledge, promoting teaching and research, and addressing societal challenges. How can rolling back the public domain do anything but subvert this work?

### ***Business Basics***

It's true that profit-seeking businesses do many things better and more efficiently than the public sector could ever hope to do. But turning things over to business isn't the right answer to every question.

If it was, public libraries might be run on a for-profit basis. But they're not. Since the early 1800s communities across the US have invested their scarce tax revenue in public libraries in order to support community literacy and the economic and social benefits that spring from this. The investment paid off and today libraries are a vital part of most communities and among the most highly rated public services provided by municipalities.

Could they have achieved this if their every action was framed by concepts like market share, revenue, and profit? It is precisely because libraries are insulated from volatile market forces that they are able to be servants of library users and stewards of our cultural heritage. They are enduring institutions, guided by deep-seated values and beholden foremost to their users.

In contrast, insolvency and discontinuity are facts of life in business. Even mighty Microsoft is subject to changes in strategy. This past May, the company announced it is closing the books on its library-scanning venture to focus on services that "offer users cash back on their purchases from our advertisers."

Google, the first-mover in library scanning, could likewise change course. An investor advisory service recently estimated Google's book scanning program trims \$200 million off the company's market capitalization, likening such free services to a "hobby" instead of a business. "The market," they said, is concerned that these services "do the company little good if it cannot charge for them." As the economy slows and advertising spending shrinks, who knows what's ahead for Google Book?

### ***Open or Shut?***

Another fundamental problem is that businesses value confidentiality over openness. Their "secret sauce" enables them to get and keep competitive advantage. In contrast, most libraries' doors are open and their collections are "Free to All", as an inscription above the Boston Public Library entrance famously proclaims. They are built on a foundation of sharing and the recognition that this advances rather than impedes their success.

Companies that are scanning library collections explain it isn't feasible to allow other search engines to crawl the books they digitize since the impetus for the investment is to gain an edge for their service. From a commercial perspective, it seems pretty simple. If a company pays to scan works, shouldn't they get to decide who can see them and under what conditions?

No, they shouldn't be allowed to make such a decision. The works they're exploiting are available as a result of public funding or the tax advantages afforded non-profit institutions, so the public's interest should take precedence.

### ***Common Cents***

Scanning and hosting are cheap—less than \$30 for a typical book, according to one authority. Companies that are digitizing library collections already have the capacity to reliably store the bits and bites, find information in milliseconds with sophisticated search engines, and make it all readily available on the Internet for viewing or downloading by anyone, anywhere, anytime. So their added costs are relatively modest in comparison to the public outlay.

The harder, more expensive part is selecting, acquiring, cataloging, storing, and preserving books and other manifestations of our written heritage—the work libraries do.

Over a period of years a library may have invested many hundreds of dollars to buy, house, maintain, and provide access to a single book that costs a tiny fraction of that amount to scan. Cumulatively libraries have spent billions of dollars on their collections and, for most libraries, they've done so with public support of one sort or another. The value of their investment is far more than what private companies are spending to put it on the Web.

Yet companies are paying nothing for access to the crown jewels. And that's the way it should be. After all, libraries are in the business of sharing.

Just as timber companies don't get a free pass to exploit our National Forests, strong protections for the public interest must be part of all agreements for businesses to exploit library collections. If mass digitization efforts are to yield more than a series of well-guarded information vaults, certain key decisions need to be taken out of the hands of firms that are using library collections to create profit for themselves and value for their shareholders.

Publicly supported libraries and tax-advantaged, privately endowed institutions alike need to negotiate scanning deals with provisions that vigorously and explicitly defend the public's usage rights. To date, this has been new ground for libraries and some of the deals have been less than perfect. But, having seen the early outcomes of commercial efforts, they better understand today what needs protecting. And many recognize that the interests of individual institutions are best served when all seek to unlock the potential of networked digital libraries.

### *A Different Approach*

The Internet has given us a glimpse of what universal online access to library collections could mean to people from all walks of life—how it could advance education, strengthen economies, and improve lives. But if copyright exemptions and other public interests get pushed aside, that breathtaking potential will never be realized.

No matter how hard they try, libraries and businesses may never find a suitable middle ground where shareholders and the public interest are both winners in the commercial digitization of library collections. We may awaken one day to find our digital heritage has become private property instead of a public good.

To prevent that risk, we need new funding strategies, coordinated library action, and forward-looking principles to guide us. It's time to sort out the right roles and responsibilities for companies, libraries, governments, and private funders and to get about the work of building an Internet public library that puts the public first.

###

*Richard K. Johnson* is a scholarly communication consultant and senior advisor to the Association of Research Libraries. He was the founding Executive Director of SPARC (Scholarly Publishing and Academic Resources Coalition) from 1998-2005.