By David Seaman

Deep Sharing: A Case for the Federated Digital Library

Libraries have reacted quickly to dizzying changes in the information landscape over the past two decades. Indeed, the rapid acceptance of relevant technologies is a library tradition, from the early adoption of typewriters, mimeograph machines, and photocopiers to the use of the fax for document delivery and the mainframe computer for online catalogs. In particular, libraries seized on networked delivery as a transformational new conduit to advance a core library mission: the dissemination of information in the pursuit of knowledge.

Every time one walks into a university or college library, one finds more evidence of the degree to which digital technology benefits library users: wireless laptops are loaned out in library lobbies; e-books and e-journals appear on handheld devices for quick review in the classroom or on the bus to class; e-reserves and digital interlibrary loans are common; and students move easily from virtual reference chat rooms to collaborative digital spaces in the physical library. Behind the scenes in a Special Collections Department, digital camera operators work alongside book conservators and catalogers, cheerfully converting rare treasures to rich facsimiles for high school, general public, and scholarly use. We librarians revel in the number of online visits these materials receive from users who will never set foot in our physical reading rooms.

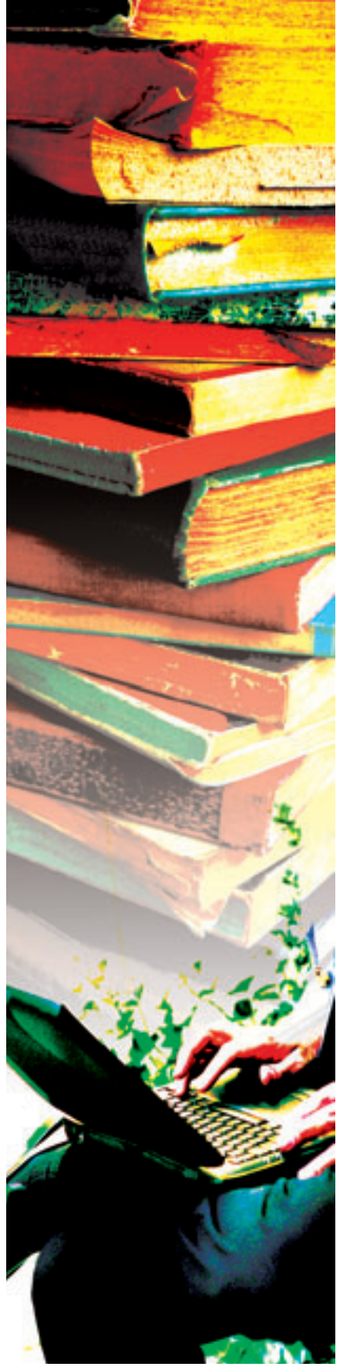
Libraries further demonstrate sophistication in the digital arena by the attention and thought expended on complex elements of the electronic revolution. Preservation, persistence, standards, archiving, training, and metadata pose difficult challenges in the brave new world of the disk and the tape. Such challenges in creating digital library services have led libraries to share knowledge, resources, and efforts in partnership organizations such as the Digital Library Federation.

Libraries are collaborative by nature—we freely share expertise, staff, ideas, and information about holdings for our collective good. Shared cataloging is a striking example: a cataloguer in one library creates a record about a book to share in a central database rather than in his or her local system, and all others who contribute to the collaborative can download that bibliographic record into their local systems rather than re-creating it at innumerable institutions. Librarians are talking about extending such interdependence and "deep sharing" to digital *content* by creating a Distributed Online Digital Library (DODL), which would depart from the status quo in terms of function, service, reuse of content, and library interdependency.

One might suppose that a shared, unified virtual library of content would add an unnecessary layer of cost and complexity with little benefit. Libraries already allow one another to look at (and link to) content on Web pages in local systems, achieving access to materials across the Web. But the current situation differs from the potential of a federated digital library. When creating digital library content today, librarians typically allow others to browse and search the content only on their respective Web sites, resulting in a bead-chain of collections barely coordinated in content or function. The current information landscape is dotted with rich content silos that can be visited but that resist innovative local recombination and reuse. As librarians become more ambitious in digitizing content from their local holdings, the more striking this data-isolation appears. This situation repeatedly thwarts the integration of remotely held content into local library services, courseware systems, and desktop data-analysis tools.

Fundamental needs in the digital library will be better served by exploring the possibility of collaboratively built, jointly owned, centrally pooled library content—a federated digital library to which librarians contribute content that users will see as a unified collection and from which libraries can draw files into local collections for innovative reuse and rearticulation as the needs of local users dictate. This federated digital library could provide quite a different experience for users. First, it would allow a common interface (such as now exists for catalog records) for distributed collections for those who want a massive, unified digital library of content. Second, and more radically, it would enable librarians and end-users alike to download "digital master" files as malleable objects for local recombinations, to be enriched with context from librarians or teachers, crafted for specific audiences, and unified in appearance and function. A user could download, combine, search, annotate, and wrap the results into a seamless "digital library mix" for others to experience.

Imagine that you are teaching a class about the Civil War or that you are a librarian crafting a collection in support of a seminar on that subject. You quickly discover via the Internet many books, images, letters, and manuscripts scattered across dozens of institutions. At present, all you can do is scrutinize that data where it resides, in formats that the creator of the content determined. You can make a list of resources—a Web bibliography—or write a description with hypertext links to the scattered collections, and your students can visit these other collections, taking "digital fieldtrips." But that is all you can do—you can have a passive engagement with the content but not an active one. You cannot combine those scattered objects into something new, improved, and shaped for your local needs. Twenty books in twenty different locations cannot easily be searched together, or enriched with information and design elements of value for a local project, or dropped into desktop software that may allow annotation by the user, or subjected to linguistic or statistical analysis that the original Web site does not support, or delivered in a format (PalmPilot, e-book) that the producer did not think appropriate, or used in myriad other ways. Libraries create high-quality digital masters for long-term preservation and reuse but then typically expose only one view of a file to the user, in one particular search-and-display software package. This serves one type of need but underserves others, reducing options for the reshaping and combining of content by teachers, librarians, and other users.



Along with richer services, the economics of the shared digital library are compelling. Imagine that thirty libraries coordinate to digitize content out of their collections. They individually fund pieces of the endeavor, but all have access to the sum of their activity. For the cost of building one digital object at each institution and depositing it in the DODL, each library would gain thirty downloadable objects built to enduring archival standards. This equation becomes even more compelling as participation grows. Duplication of effort ceases. The ability to build something—to contribute to a growing shared collection—may be the incentive libraries need to direct more local funding and efforts to digitizing content.

Sharing digital content sometimes touches a negative emotional nerve in librarians—a knee-jerk response that I have observed in myself. For generations, we librarians have measured our worth in part by how much stuff we have, which is often the first thing anyone asks about a library. In the digital world, however, we can simultaneously have a book and (legal rights permitting) give it to others too. Sharing a treasure does not diminish our pile of content. Competing on locally held book hoards may not serve us well in the digital world. What should now distinguish libraries one from another is what we do with the federated digital library content to which we all have access—the tools, training, design, context, and services we craft around the content and the richer scholarship and pedagogy these things permit.

We need to exploit all available standards and technologies to infuse malleability, interoperability, and repurposing into the use of digital library objects. Innovative users need content that fosters discovery, engagement, and experiment. Users are too often invited to watch "content channels" whose aesthetics, services, and behaviors are dictated by terms of the creating institutions or publishers; they are rarely able to download, manipulate, morph, annotate, cross-search, and repurpose digital library content (in the "music mix" model). The hope is that the creation of a DODL will make seamless access to and rich reuse of content a "doddle" (to borrow a colloquial Britishism): something very easy to do, after the current panoply of digital library skills is brought to bear on the problem.

The networked computer has applications to daily work in every corner of the academic library. It revolutionized the way we create and use bibliographic records by replacing note cards in drawers with the computerized union catalog, which takes us far beyond the collections of any single library. Now, our offerings are being enriched with electronic versions of the content itself: journals, photographs, manuscripts, slide libraries, and books. This digital content has an inherent malleability that paper-bound information lacks. It is time for extending the benefits of federation to maximize the recombinant potentials of such content.

David Seaman is Director of the Digital Library Federation, a consortium of thirty-four major academic libraries and related organizations dedicated to using electronic information technologies to extend their collections and services.