In 1994 the National Science Foundation launched its Digital Libraries Initiative (DLI). The choice of combining the word *digital* with *library* immediately defined three interested parties: librarians, computer scientists, and publishers. The eventual impact of the Initiative reached far beyond these three groups. The Google search engine emerged from the funded work and has changed working styles for virtually all professions and private activities that involve a computer. The improved methods for searching video, which the Initiative funded, have led to new approaches for detecting longitudinal trends in nursing home occupant health (*CareMedia*). Similarly, the Initiative has touched the work of historians, anthropologists, political science and law, and many other professions.

But one of the more intriguing aspects of the DLI was its matchmaking coup of uniting librarians and computer scientists. Publishers were and are on the scene, but triangles are complicated, and just contemplating the binary union between academic librarians and computer scientists is rich enough for one article. Between the two, are there clear winners and losers after 10 years? Many would answer this question with a vigorous nod.

**Computer Scientists’ Expectations**

For computer scientists NSF’s DL Initiative provided a framework for exciting new work that was to be informed by the centuries-old discipline and values of librarianship. The scientists had been trained to use libraries since their years of secondary education. They could see, or at least imagine how current library functions would be moved forward by an injection of computing insight.

Digital library projects were for many computer scientists the perfect relief from the tension between conducting 'pure' research and impacting day-to-day society. Computing sciences are called on to continually generate novelty. On the other hand, they experience both their own desire, as well as funders' calls for deep
impact on society and neighboring scientific fields. Work on digital libraries promised a perfect resolution of that tension.

**Librarians' Expectations**

For librarians the new Initiative was promising from two perspectives. They had observed over the years that the natural sciences were beneficiaries of large grants, while library operations were much more difficult to fund and maintain. The Initiative would finally be a conduit for much needed funds.

Aside from the monetary issues, librarians who involved themselves in the Initiative understood that information technologies were indeed important to ensure libraries' continued impact on scholarly work. Obvious opportunities lay in novel search capabilities, holdings management, and instant access. Online Public Access Catalogs (OPACS) constituted the entirety of digital facilities for many libraries. The partnership with computer science would contribute the expertise that was not yet widely available in the library community.

**The Cuckoo's Egg Surprise**

The DLI proposals had architected what looked like a reasonably comfortable nest for the emerging union between the two disciplines. But just as the happy couple set out to work, the World Wide Web pecked its way into the scene and demanded attention while growing with astonishing rapidity.

The Web's advent significantly changed many plans. The new phenomenon's rapid spread propelled computer scientists and libraries into unforeseen directions. Both partners suddenly had a somewhat undisciplined teenager on their hands without the benefit of prior toddler-level co-parenting.

The coalition between the computing and library communities had been anchored in a tacit understanding that even in the 'new' world there would be coherent collections that one would operate on to search, organize, and browse. The collections would include multiple media; they would be larger than current holdings; and access methods would change. But the scene would still include information consumers, producers, and collections. Some strutting computer scientists predicted the end of collection gatekeeping and mediation between collections and their consumers; librarians in response clarified for their sometimes naive computing partners just how much key information is revealed in a reference interview. But other than these maybe occasionally testy exchanges, the common vision of better and more complete holdings prevailed.

The Web not only blurred the distinction between consumers and producers of information, but it dispersed most items that in the aggregate should have been collections across the world and under diverse ownership. This change undermined the common ground that had brought the two disciplines together.

For computer scientists this shift did not question the foundation of their work. On the contrary, early results within the Initiative projects had demonstrated a nagging downside of the existing DLI research environment. That environment was bound into special deals with publishers: the community could not truly share their results, such as new user interfaces to collections. Sharing the results would have required sharing access to the underlying collections, which were freely accessible only to one individual university that had partnered with a particular publisher. These restrictions were serious impediments, because computer scientists traditionally make their entire systems public. Personal fame is made by building and freely distributing entire operating systems (Linux), storage facilities (Berkeley DB), and user interface toolkits (Piccolo).

DLI researchers whose results were bound by per-project agreements with publishers realized that they could...
only share teasers with their colleagues. Copyright restrictions prevented access by a wider audience. While well-circumscribed holdings are a source of pride in the library world, the same concept therefore constricted the traditions of computer scientists. Working on the Web removed these restrictions.

The embrace of the Web by computer scientists was natural also because linkage of information is a much employed concept in computer programming. The notion is taught in every second-year programming class. The leap from linkage that is local to one machine, to linkage across continents is therefore not so foreign to students of computer science.

Yet another factor that drew large numbers of computer scientists to the area was that as the Web grew, machine learning, statistical, and other heuristic approaches became applicable to information search and organization. This broadening of relevance to additional subdisciplines of computer science increased the number of researchers and made what began under the name Digital Libraries an area with "sex appeal" to many.

For librarians the intrusion of the Web into the work on digital libraries was much more difficult to integrate. Losing the notion of collection in visions of the future threatened a weight-carrying pillar of traditional librarianship. Now all that seemed left of the original partnership with computer scientists were theoretical computing algorithms without clear connection to recognizable, traditional library functions.

The disruption to the library community was greatly exacerbated by many journal publishers’ business decision to charge at a premium for digital content. This decision has been forcing academic libraries to cancel subscriptions, undermining their role as conduits to scholarly work.

In truth, however, both communities were fundamentally affected by a radical change in attitude that the World Wide Web introduced. Many computer scientists, most notably the database community, had shared with librarians the value of data completeness, reliable availability, and content integrity. Both communities considered predictable, repeatable collection access and retrieval to be a prime value. The Web's hyperabundance of duplication and content alternatives generated a popular culture of laissez-faire retrieval. A wide sector of the public was much less upset than one might have predicted by Web links that led nowhere. The massive size of the Web makes up for local failures. The implications of these unruly information behaviors still has both communities tumbling today.

Far from enjoying immunity from the restrictions that tradition imposes on older fields like library science, computer scientists struggle with ambivalence towards the fruits of their own digital revolution. At the same time that they herald online self-publishing as the breakage of shackles from the tyranny of publishers, tenure committees cannot wrestle themselves into granting the same respect to online-only work that they afford traditional journal articles.

**Mutual (Mis?)Conceptions**

Amidst the stress of this destabilized common ground, a number of complaints have been leveled among the partners. Some librarians had expected DLI money to flow into collection building. Instead, they perceive, computer scientists have hijacked the money and created an environment whose connection to librarianship is unclear. Some felt that their fast moving computing enthusiast partners too thoughtlessly dismissed important functions, like collection development, as quaint.

The impatient among computer scientists in turn could not understand why librarians are so annoyingly deliberate about metadata, spending years arguing about structures that the computer scientists felt would be replaceable by just another clever search algorithm improvement. Most of all, some computer scientists
couldn't understand why librarians couldn't be, well, normal computer scientists.

So, has computer science swept across the information landscape like the Vandals of old and left librarianship in ashes? This bleak prediction is unwarranted. A number of developments call for a brighter assessment.

While information accession now rests on a highly technical infrastructure, the core function of librarianship remains. The information must be organized, collated, and presented. Highwire Press, for example, provides not merely a single keyword search field, but also an overview of available resources, as well as browsing facilities.

The notion of collections is spontaneously re-emerging in the form of what computer scientists have named information 'hubs.' These new incarnations of collections are Web sites whose primary goal is to direct visitors to other Web sites that all specialize on the hub's topic focus. (The hubs' topic-specific target sites are called 'authorities' in the literature). Quality hubs require curation. This type of curation differs from its traditional cousin in that it must cope with sometimes ephemeral material, and in that it requires a different underlying skill set. The spirit of this function, however, remains.

Opportunities now arise for direct connections between librarians and scholarly authors. The production of computer science scholarly content has over the past ten years emerged into a prototype for how other disciplines may soon generate their publication output: Computer scientists are now expected to generate camera-ready copy of their work. Of course, no camera is involved in final production. The product instead moves directly online and into the hands of what should be a new breed of librarians who add value to the product and maintain hubs and online archives. Computer science authors and librarians have thus moved to be direct neighbors in the publication workflow (see, for example Dutch academics declare research free-for-all). This closeness opens new opportunities for librarians to expand their operations. For example, librarians might over time reach back into the workflow and help scholars who are not computer savvy produce output that works well in an online world.

The accomplishments of the Digital Libraries Initiative and many related activities external to its work have broadened opportunities for library science, rather than marginalizing the field. As models for new libraries, like the ACM Digital Library for computer science emerge, other scholarly fields will follow suit and require help in developing their own styles of online presence.

Now, if only ACM librarians would stop making computer scientists choose content descriptors from a controlled vocabulary of terms... Oh well, in any union both sides need something that recalls their old identity. In return, Web servers of flighty computer scientists get to return their irresponsible "404 Not Found," which makes any librarians worth their salt grit their teeth.

Copyright © 2005 Andreas Paepcke, Hector Garcia-Molina, and Rebecca Wesley