Politicians give us many reasons to worry, and I don’t usually hold them up for public praise. But there is one thing politicians often do extremely well—they describe things in simple terms that everyone can understand. Al Gore, for example, when describing the virtues of technology, painted a verbal picture of a day when the contents of the Library of Congress would be available online to every school child in America.

Librarians across the country winced at that notion, particularly when they heard Gore’s description paraphrased by local administrators and trustees. Collectively, librarians protested—“No, no, that image is too simple. We can’t put everything online. We don’t have enough money. We don’t have all the legally required rights and permissions.” Also, perhaps most vehemently, we librarians protested that not everything that could or should be digitized is in the Library of Congress.

Our reaction was, to be blunt, a mistake. We missed an opportunity to capitalize on a politician’s vision. We lapsed into our usual professional concerns and institutional territoriality. Instead, we should have given the answer that I will give today to Karen’s question—“What do we have to do to realize the potential of digital libraries?” My answer is simple: we must build massive, comprehensive digital collections that scholars, students, and other researchers can use even more easily than they use the book-based collections we have built up over the centuries.

I do not say that building massive, widely accessible digital collections, in which I include library-acquired as well as library-created materials, will be easy. I do not say that it will be inexpensive. I do not say that libraries will magnanimously rush to collaborate to achieve it. And I do not say that we can readily unravel the complexities of linking libraries for users worldwide. But I do say that making much more content much more accessible is a great and worthy goal. If we could make it possible for every student in America—and for everyone else, in less favored countries, in repressed countries, indeed, anywhere in the world with access to computers—to use those computers as doorways to the libraries of the world, that would be an achievement truly worthy of the ingenious technologies that are beginning to make it possible.

To achieve such a goal, I believe that the digital library of the future will develop three overall characteristics.

- It will be a comprehensive collection of resources important for scholarship, teaching, and learning.
- It will be readily accessible to all types of users, novices as well as the experienced.
- It will be managed and maintained by professionals who see their role as stewards of the intellectual and cultural heritages of the world.

Also, I believe that achieving these things will require a fundamental transformation of the institutions we now call libraries.

If we are to bring about this transformation, if we are to adopt as a goal the provision of great quantities of digital library resources that can readily and widely be accessed and used by scholars, teachers, and students, how specifically would we begin? What must we as a community do?

One major requirement is obvious—launch a massive digitization effort to put our libraries online. The reason for doing so is that online is where faculty and students increasingly go to look for scholarly resources. We can transfer much more of our analog collections to digital so that the resources we have invested in...
developing all these years will not be lost from sight as scholars and students make digital the preferred mode. And we can also collect and provide wider access to “born-digital” material that scholars and students already are using in electronic form.

If there was any doubt about how extensively the use of digital information has permeated our campuses, it was dispelled by findings in a recently published major study, *The Council on Library and Information Resources and the Digital Library Federation* commissioned a survey of the information-use behavior of 3,234 faculty members, graduate students, and undergraduates in 392 American colleges and universities. High percentages of respondents of all kinds in all fields and types of institutions expressed comfort with electronic information, and the comfort level overall was almost as high as with print. Nearly 95% of respondents were comfortable with their institutions’ Web sites, and nearly three-fourths were participating to some degree in their institutions’ distance-learning opportunities. Desktop and laptop computers were readily available to most respondents, and laser printers and scanners to many.

Use of printed materials remains high, as does respect for libraries. But more than one-third of the survey’s respondents said they use the library less now than they did two years ago, apparently meaning the physical library. Significant proportions of students and faculty are using e-journals and even e-books. More than 80% indicated that they were finding more relevant information on the Internet than they did two years ago and that the Internet had changed the way that they used libraries. The survey does not indicate that print is going out of use or libraries out of business, but it does show how extensively students and their professors alike are going to the open Internet and library Web sites for the information they need.

And why not? The convenience of clicking on Web sites to find information rather than going to libraries is obvious to us all. Online library catalogs are easier to search than card catalogs. Finding bibliographic references through OCLC is a lot easier than tracking them down in individual libraries. Research findings can be brought to light more rapidly through e-journals, listservs, and e-mail than through printed journals and conference papers. Students can get material required for classes more readily from courseware offerings than from reserve shelves. And such conveniences are only some basic ways in which digital information facilitates academic work.

Of course, a number of libraries have already begun to move major holdings of texts and images from their shelves to our computer screens. Numerous valuable collections already are there. We’re all familiar with the American Memory collections on U.S. history and culture, accessible electronically from the Library of Congress—7 million digital items from more than 100 historical collections. Similarly, we know about the Making of America digital collection consisting of 267 monographs and 100,000 journal articles from 19th century imprints, pioneered by Cornell and the University of Michigan. We’re equally familiar with the extensive repository of digitized journals established by JSTOR—which, not incidentally, has discovered that digitization can bring resources to light in ways that greatly increase their use. Also well known is the National Science Digital Library’s aggregation of collections and services in support of science education, to which more than 100 institutions of learning are contributing, with financial help from the National Science Foundation.

Additionally and individually, the nation’s major research libraries have digitized hundreds of collections in numerous fields and are adding more, while also leasing digital resources from publishers. Statistics published for the year ending in June 2001 by the Association of Research Libraries (ARL) show that its members spent more than $132 million on electronic resources, which accounted for more than 16% of their library materials budgets. Moreover, the ARL said, and I quote: “In every year since 1992-93, average expenditures on electronic resources have increased at least twice as fast, and in some cases up to six times faster, than average library materials expenditures.”

Demand, however, may be rising even faster. Students who can’t find what they want on our library Web sites may go Googling for whatever seems relevant on the open Internet or more likely will go there in the first place. As David Seaman, Director of the Digital Library Federation, wryly observed about Internet search engines at a recent Federation meeting, “Our students have found another way not to come to the library.” Last summer, at a leadership development institute that CLIR sponsors with EDUCAUSE and Emory University for mid-level faculty and staff in higher education, participants expressed concern that failure to keep up with digital demand could put their institutions at competitive risk because students and even faculty increasingly expect fast and easy computer access to multiple kinds of digitized materials for coursework and research.

But our students now are going to a fragmentary digital library. Despite the growing body of digitized material and demand for it, collections that scholars and students can currently access electronically still remain a fraction of the holdings of libraries. What stands in the way of massively boosting library digitization?

Copyright, for one thing. You’ve all read of the Supreme Court decision that upheld the Sonny Bono Copyright Extension Act, which prevents a copyrighted publication from entering the public domain until 70 years after its author’s death and protects certain corporate properties for 95 years. What constitutes “fair use” is at issue as well, and libraries, on the advice of lawyers, have to be extremely careful about what they reproduce digitally. A legal expert on public-domain encroachments who recently met with my staff told of trying to track down the original source of a centuries-old poem by innocently making inquiries to libraries whose online collections contained it. Fearing that he intended to enforce some intellectual property restriction, many librarians responded not with information but with promises to remove the poem from their sites. The legal differences of opinion between publishers and librarians are real. Yet, it seems counterproductive to focus on these differences. One area of common interest is making new digital works accessible. Publishers and libraries are at least talking to each other about mutually useful solutions to potentially divisive issues. This happened at Columbia University in a workshop for educational publishers and representatives of the National Science Digital Library. It continues in meetings of a Working Group of Librarians and Publishers that the Council on Library and Information Resources set up in collaboration with the Professional and Scholarly Division of the Association of American Publishers. Resolution of copyright issues seems still far off, but people of good will are giving us hope. Both groups recognize that each side has much to learn about financial models and incentives for access.
Another big obstacle to mass digitization is, of course, money. Like most big goals, putting libraries online will require big bucks. But recognition is rising that digital development benefits would amply warrant more substantial investments. A foundation-funded coalition called the Digital Promise Project is calling on the Congress to recognize that digital developments can revitalize American education as much as did the provision of the GI Bill for veterans’ education, the Morrill Act for land-grant colleges, and the Northwest Ordinance for public schools. Bills have been introduced in the Senate and the House to finance a Digital Opportunity Investment Trust. The Trust would support innovative uses of digital technologies to improve education, in part through extended access to expanded digital libraries. Funding would come from government auctions of licenses to the publicly owned electromagnetic spectrum, which, over the next several years, could produce an estimated $18 billion.

Even if Congress fails to approve funding from that source, however, we should not give up on other prospects for providing expanded access to massive digital libraries. Instead, we need to remind ourselves of a similar situation in the 1970s and early 1980s when every library in the country needed to convert its card catalog to an online format. Long-term planning and goal setting, combined with strategic grants from funders and the creation of new services such as OCLC, resulted in the massive conversion of library catalogs.

Even if we overcome the financial and copyright obstacles to massive digitizing of library holdings, however, another major problem remains—how to organize the effort. Because libraries acquire many of the same books, significant overlap exists in their collections. Nothing is gained by digitizing every library’s copy of the same title. The Digital Library Federation is promoting creation of a registry of digital materials so that, among other things, duplicative digitization could be avoided. But this step, important as it is, would only allow libraries to see what others had digitized, not designate responsibilities for seeing that digitization goes forward.

No library alone can afford to digitize all that its patrons could use. Can libraries parcel out digitization responsibilities among themselves? Can institutions that have spent decades if not centuries competing to build rival collections of print materials agree to merge these collections electronically? Can they agree to digitize their unique holdings for universal access if other libraries agree to do the same with theirs? When the smallest college can offer its faculty and students electronic access to the library holdings of the largest research university, what happens to relative status in the competition for top scholars and students? It may be hard to work for universal benefits if the cost seems to be a reduction in one’s own prestige. All our energies could be exhausted just in trying to find acceptable terms for such collaborations. But if, in the digital era, libraries must continue to compete, it can cease to be about collections and become about services—the ingenuity with which individual libraries tailor resource access to particular needs of their user communities.

Actually, some libraries already have contributed to aggregations of digital resources, such as those I named earlier. And, as members of the Digital Library Federation noted in a recent discussion of strategic planning, technological tools are now available or under development that will make possible what members called “federating content.” Realizing that “crafting content into a whole,” as one member put it, could be of great significance for future teaching and learning, the Federation is already raising questions about whether and how access centers might be linked and responsibilities divided.

The responsibilities, I should add, extend to keeping digital resources accessible, as well as simply creating them. Because of the fragility of digital media, the rapidity with which systems for accessing digital information are superceded, and the uncertainty of the long-term efficacy of current digital maintenance techniques, preservation of digital information is a growing concern as well. Access to the kind of comprehensive digital library I envision must extend widely not only now but through future generations. Preservation, too, will be easier if collaborative.

Thus we need money, intellectual property agreements, and library collaborations to build the massive and accessible collections of enduringly valuable cultural resources that I am proposing. Success also will require some less obvious things. One is something that even traditional libraries seldom fully achieved—genuine collaboration with scholars.

In the digital era, faculty members are not just clicking on digital resources for research and education that libraries provide. They are creating their own. In research centers, learned societies, and other discipline-based organizations, scholars collaborate to build electronic databases, large and small, of use in their respective specialties. Some publish their own e-journals outside the purview of established publishers. Others maintain Web sites for exchanging reports on research results not yet ready for formal publication. Scholars even develop electronic tools for searching, comparing, recombining, and analyzing digital information. Teaching faculty set up Web sites for their classes through which to provide electronic access to assignments, syllabi, and course readings.

Sometimes faculty seek help from campus libraries in such endeavors, but more often they don’t, at least not until the reality dawns on scholars that what they have created may need a more permanent home. Much of what they create is, in fact, worth preserving for use by others now and in the future. Accordingly, some alert librarians now collect various manifestations of e-scholarship, as MIT’s D-Space project does, and even provide e-publishing assistance to scholars, as does Cornell University’s Project Euclid.

Building comprehensive collections of scholarly digital resources means incorporating resources developed by scholars and teachers. Librarians need to help them solve their problems while capturing content for digital libraries. Every library with sufficient technological capability can do this on its own campus, but libraries need not stop there. MIT’s librarians have identified several ways to persuade scholars of e-repository benefits. MIT is working with other institutions in the U.S., the U.K., and Canada to federate D-Space in hopes of building what D-Space developers call, in the January 2003 issue of D-Lib Magazine, “a critical corpus of content that represents the intellectual output of the world’s leading research universities.”

At the same time, libraries need to work with learned societies and other discipline-based groups of scholars to understand their digital content needs. What are the most important works in the various academic fields to digitize? Which resources could teachers best use in digital form for classroom instruction? Which e-scholarship projects developed by scholars would most help others if inte-
grated into a digital library system for widening accessibility? Which e-projects are worthy of long-term preservation? All this requires much more from librarians than accepting recommendations from scholars on new books to buy. This requires thinking through with scholars the steps to take toward building a new kind of library.

Additionally, to build this library effectively, we need to enlarge our understanding of how users seek and employ digital materials. Much work has been done to gain insight into user behavior, capabilities, and preferences. Insight is coming from the extensive survey of users that I mentioned earlier, from JSTOR's revealing studies of e-journal use, and from the Scholarly Work in the Humanities Project, which studied how humanists operate in the evolving information environment. The need now is to pull together what we have learned from these studies for guidance as we work to create usefully accessible digital resources. Publishers and librarians alike must find ways to respond to users’ interests—even demands. Publishers are now building proprietary systems to provide access to their works. Librarians struggle to find ways to pull proprietary materials into a unified system of access for their patrons. Librarians and publishers both have an interest in access, which provides a common ground for working together on this problem. Both groups' digital futures are intertwined with users’ expectations.

We realize already that no such thing exists as "the user." Instead, users come in many categories with heterogeneous needs and requirements. Today’s technologies allow us to customize information resources for users, which means that the future digital library can and must be tailored to the needs of their communities. Those collections have also proven enormously valuable over time. There is every reason to think that the same will be true of digital resources. But because of the heavy technological dependency of digital information, and the ease with which almost anyone can "publish" it, effective stewardship will require collaboration among librarians, publishers, scholars, and information technologists. In projects funded by the Mellon Foundation to advance e-journal archiving by librarians and publishers, in work librarians already are doing with scholars on e-scholarship, and in consolidations on some campuses of IT departments with libraries, some of the needed collaborations are beginning. We must strengthen and extend such efforts if we are to be effective digital library stewards. In spite of bureaucratic boundaries, seemingly conflicting interests, and threats to professional and organizational identities, I believe that effective collaborations can be forged.

Massive digitization is, in many respects, a gigantic new publishing effort. Might it be possible for agreements to be reached whereby publishers give permission to libraries to digitize materials now protected by copyright in exchange for opportunities for publishers to sell products online through libraries? Also, could librarians and publishers work together on portal developments, through which researchers could search across library collections and publication lists among sources of scholarly information of relevance to their projects? Reconceptualizing the role librarians and publishers play may enable us, in the digital environment, to structure relationships that allow both to accomplish their important goals. To such relationships, libraries could bring such things as digitizing experience, high-level metadata, delivery systems, and large-scale storage, while publishers could bring their expertise in marketing, pricing, and customer support.

Similarly, I believe that we can integrate the work of scholars and librarians, particularly as librarians come to recognize that scholarly resources no longer are restricted to traditional publishing channels but now are based on digital and other media besides print, and as scholars recognize the value of pre-print publication and of digital library preservation of their work. In building the future digital library, professional separation is only a burden to all.

So, that concludes my picture of the future, an understandably simple one I hope, albeit a bit more filled out than Al Gore’s. That’s what I think the future digital library should be like and what it will take to achieve it. I have not shied from identifying some of the obstacles to achieving this vision, which seem substantial. But I take heart from something that Bill Frye of Emory University said when he was on the board of the Council on Library and Information Resources and agreed to take responsibility for a preservation planning committee that we had charged with nothing less than outlining a national program for preserving millions of books in danger of deterioration. Liking this situation to eating an elephant, he advised: “Start with a single bite.”

Bite by bite—or maybe I should say, byte by byte—we are advancing toward the things I argued at the outset were needed for the future digital library: a comprehensive collection of digitized resources, readily accessible to all types of users, and managed by professionals who see their role as stewards of the intellectual and cultural heritage of the world. I hope we will not betray the possibilities of the new technologies by settling for anything smaller.