INSTITUTIONAL REPOSITORIES: ESSENTIAL INFRASTRUCTURE FOR SCHOLARSHIP IN THE DIGITAL AGE

by Clifford A. Lynch, Executive Director, Coalition for Networked Information

Introduction

In the fall of 2002, something extraordinary occurred in the continuing networked information revolution, shifting the dynamic among individually driven innovation, institutional progress, and the evolution of disciplinary scholarly practices. The development of institutional repositories emerged as a new strategy that allows universities to apply serious, systematic leverage to accelerate changes taking place in scholarship and scholarly communication, both moving beyond their historic relatively passive role of supporting established publishers in modernizing scholarly publishing through the licensing of digital content, and also scaling up beyond ad-hoc alliances, partnerships, and support arrangements with a few select faculty pioneers exploring more transformative new uses of the digital medium.

Many technology trends and development efforts came together to make this strategy possible. Online storage costs have dropped significantly; repositories are now affordable. Standards like the open archives metadata harvesting protocol are now in place; some progress has also been made on the standards for the underlying metadata itself. The thinking about digital preservation over the past five years has advanced to the point where the needs are widely recognized and well defined, the technical approaches at least superficially mapped out, and the need for action is now clear. The development of free, publicly accessible journal article collections in disciplines such as high-energy physics has demonstrated ways in which the network can change scholarly communication by altering dissemination and access patterns; separately, the development of a series of extraordinary digital works had at least suggested the potential of creative authorship specifically for the digital medium to transform the presentation and transmission of scholarship.

The leadership of the Massachusetts Institute of Technology (MIT) in the development and deployment of the DSpace institutional repository system <http://www.dspace.org/>, created in collaboration with the Hewlett Packard Corporation, has been a model pointing the way forward for many other universities. In 2003, with funding from The Andrew W. Mellon Foundation and other sources, MIT’s DSpace is scheduled to be replicated at a number of additional institutions around the world; the software has also been released publicly under an open source arrangement, greatly lowering the cost and development barriers to implementing repositories for all institutions. The MIT software is not the only option available, although I believe it is the most general-purpose; for example, there is software from the University of Southampton in the U.K. <http://www.eprints.org/> designed more specifically for institutional or disciplinary repositories of papers, as opposed to arbitrary digital materials.

Over the past few months, I have had a number of opportunities to speak about the roles and significance of institutional repositories as a strategy for supporting the use of networked information to advance scholarship, notably at a workshop jointly sponsored by ARL, CNI, and SPARC in Washington, D.C., at the DSpace launch celebration at MIT, and at the University of Tennessee and the University of British Columbia. While video recordings of some of these events are available on
the Net, this article is an attempt to summarize and articulate the views I’ve expressed at these various events about the nature and functions of institutional repositories and their role in transforming scholarship.

**Defining Institutional Repositories**

In my view, a university-based institutional repository is a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organization and access or distribution. While operational responsibility for these services may reasonably be situated in different organizational units at different universities, an effective institutional repository of necessity represents a collaboration among librarians, information technologists, archives and records managers, faculty, and university administrators and policymakers. At any given point in time, an institutional repository will be supported by a set of information technologies, but a key part of the services that comprise an institutional repository is the management of technological changes, and the migration of digital content from one set of technologies to the next as part of the organizational commitment to providing repository services. An institutional repository is not simply a fixed set of software and hardware.

While early implementers of institutional repositories have chosen different paths to begin populating their repositories and to build campus community acceptance, support, and participation, I believe that a mature and fully realized institutional repository will contain the intellectual works of faculty and students—both research and teaching materials—and also documentation of the activities of the institution itself in the form of records of events and performance and of the ongoing intellectual life of the institution. It will also house experimental and observational data captured by members of the institution that support their scholarly activities.

At the most basic and fundamental level, an institutional repository is a recognition that the intellectual life and scholarship of our universities will increasingly be represented, documented, and shared in digital form, and that a primary responsibility of our universities is to exercise stewardship over these riches: both to make them available and to preserve them.

An institutional repository is the means by which our universities will address this responsibility both to the members of their communities and to the public. It is a new channel for structuring the university’s contribution to the broader world, and as such invites policy and cultural reassessment of this relationship.

I want to make the distinction between scholarly publishing as it is currently practiced and the broader, much more diverse, often less formal, and certainly more rapidly evolving set of practices that comprise scholarly communication; scholarly publishing is a very specific, circumscribed example of scholarly communication. I use the two terms “scholarly communication” and “scholarly publishing” distinctly and carefully in this paper. For example, the definition I propose for an institutional repository does not call for a new scholarly publishing role for universities, only one of dissemination of scholarly communication; scholarly publishing is much more than simple dissemination, and has typically been rather limited in the genres of communication that it does disseminate. I will have more to say about the relationships between repositories and publishing later.

For those organizations *within* the university concerned with stewardship—we think immediately of libraries, archives, and museums but should recognize there are also huge numbers of academic units that curate collections of information—it should be clear that institutional repositories raise complex and nuanced questions about organizational roles, responsibilities, resources, and strategies. Similar, but perhaps less complex, questions arise for all organizational units focused on dissemination of scholarly communication or more narrowly on scholarly publishing, such as university presses.

**The Strategic Importance of Institutional Repositories**

Scholarship and scholarly communication are changing. These changes start with risky and bold acts of individual creativity. They will extend slowly to cultural changes at the disciplinary level and ultimately to new interdisciplinary standards that are expressed in the decisions of institutional tenure and promotion practices.

Our institutions of higher education have overlooked an opportunity to support our most innovative and creative faculty for at least a decade now, to the detriment of both the faculty members and the institutions themselves. These faculty have been
exploring ways in which works of authorship in the new digital medium can enhance teaching and learning and the communication of scholarship; such innovations are essential to keeping scholarship vital and effective, and they must not only be supported but nurtured. Indeed nurturing these innovations reaches to the core mission of our universities, and to the core values of our universities. A much broader and generally more conservative group of faculty have exploited the Net as a vehicle for sharing their ideas worldwide, whether these ideas are expressed in relatively familiar forms such as digital versions of traditional journal articles or (less commonly) in entirely new forms that begin to map out the future evolution of, for example, the scholarly monograph in the digital medium. This embrace of new dissemination opportunities is also important for what it says about the roles of scholars and universities in society and in a global environment. Our universities have poorly served this broader group of scholars as well, though this may be less critical because faculty are well motivated to rise above the institutional failures to help them disseminate their works, because failures to effectively disseminate these works are less damaging than failures to legitimate nontraditional works, and because faculty concerned only with dissemination of traditional material are at less risk within their own disciplines.

But consider the plight of a faculty member seeking only broader dissemination and availability of his or her traditional journal articles, book chapters, or perhaps even monographs through use of the network, working in parallel with the traditional scholarly publishing system. Such a faculty member faces several time-consuming problems. He or she must exercise stewardship over the actual content and its metadata: migrating the content to new formats as they evolve over time, creating metadata describing the content, and ensuring the metadata is available in the appropriate schemas and formats and through appropriate protocol interfaces such as open archives metadata harvesting. Faculty are typically best at creating new knowledge, not maintaining the record of this process of creation. Worse still, this faculty member must not only manage content but must manage a dissemination system such as a personal Web site, playing the role of system administrator (or the manager of someone serving as a system administrator). Over the past few years, this has ceased to be a reasonable activity for most amateurs; software complexity, security risks, backup requirements, and other problems have generally relegated effective operation of Web sites to professionals who can exploit economies of scale, and who can begin each day with a review of recently issued security patches. Today, our faculty time is being wasted, and expended ineffectively, on system administration activities and content curation. And, because system administration is ineffective, it places our institutions at risk: because faculty are generally not capable of responding to the endless series of security exposures and patches, our university networks are riddled with vulnerable faculty machines intended to serve as points of distribution for scholarly works. And faculty create content at risk because they typically do not back it up appropriately, ensure its integrity (in part by hosting it on secure systems), and curate it properly.

For those faculty who are concerned not just with distribution opportunities through the network but with deeper questions of how to exploit the nature of the digital medium for new works of authorship, the situation is even worse. This is not just about more effective public access to recognizable and familiar genres of work such as journal articles which can, in the worst case, be reduced to printed forms for distribution to a tenure and promotion committee. These faculty take on a heavy burden in arguing for the legitimacy of investing their time in works of digital scholarship, and in making the case for the value of such creations in comparison to more traditional scholarly output. This is a cultural problem that must be played out discipline by discipline, and which must be worked out also in the evaluation, tenure, and promotion practices in place at an institutional level. However, preservability is an essential prerequisite to any claims to scholarly legitimacy for authoring in the new medium; without being able to claim such works are a permanent part of the scholarly record, it's very hard to argue that they not only deserve but demand full consideration as contributions to scholarship. Most individual faculty lack the time, resources, or expertise to ensure preservation of their own scholarly work even in the short term, and clearly can't do it in the long term that extends beyond their careers; the long term can only be addressed by an organizationally based strategy. Institutional repositories
can address both the near-term questions about continuity of access by providing an environment in which such new works of scholarship can be managed and disseminated—including such basic things as professionally managed systems and systematic backup procedures—and also the longer-term questions about preservation by creating an institutional commitment to such preservation.

The revolution in scholarly communications is not limited to the development of new genres of scholarly works that are enabled by the digital medium; even traditional forms such as journal articles now frequently include supplementary datasets and analysis tools. Scholarship has become data intensive; it is supported and documented by data and tools that complement interpretive works of authorship. For the sciences, these changes have been well documented in the recent National Science Foundation report of the Advisory Committee for Cyberinfrastructure chaired by Dan Atkins; while the report is focused on cyberinfrastructure to support the conduct of science, most of the discussion is in fact applicable beyond the sciences to the broader scholarly enterprise, including the humanities. Most scientific journals are now accepting what they characterize as “supplementary” materials as part of the publication of traditional journal articles, but it is much less clear what commitments these journals are making to actually integrating these supplementary materials into the permanent record of scholarship in the same way that they maintain the journal articles themselves as a part of that record. While it is clear that for some types of scholarly work we will see the continued evolution of disciplinary data repositories (consider, for example, molecular biology) and community norms that journal publication is complemented by deposit of data in these disciplinary repositories, it is equally clear that the scholarly enterprise is sufficiently diverse that these disciplinary repositories will never be fully comprehensive. Only an institutionally based approach to managing these data resources, which operates in alignment with what the faculty at each individual institution are actually doing, can provide a comprehensive dissemination and preservation mechanism for the data that supports the new scholarship for the digital world. Journals will move too slowly and too unevenly to manage these resources, and disciplinary data repositories cannot be comprehensive. Institutional repositories can maintain data in addition to authored scholarly works. In this sense, the institutional repository is a complement and a supplement, rather than a substitute, for traditional scholarly publication venues.

Institutional repositories also have roles beyond disseminating and managing the works of individual scholars that are part of the dialog of scholarly communications. I have argued that research libraries must establish new collection development strategies for the digital world, taking stewardship responsibility for content that will be of future scholarly importance. Institutional repositories are a place where they can put much of the material that research libraries identify as worth collecting. Finally, at least a few institutions themselves are changing their culture and are making commitments to globally disseminate extensive teaching and learning materials through the Net (for example, the OpenCourseWare initiative at MIT <http://ocw.mit.edu/>, or, at a less systematic but still important level, to digitally capture and preserve the many of the events of campus life—symposia, performances, lectures. Institutional repositories offer a framework for organized stewardship and accessibility of these materials.

To summarize, institutional repositories can facilitate greatly enhanced access to traditional scholarly content by empowering faculty to effectively use the new dissemination capabilities offered by the network. This is also occurring on a disciplinary basis through the development of e-print and preprint servers, at least in some disciplines. In cases where the disciplinary practice is steady, institutional repositories can feed disciplinary repositories directly. In cases where the disciplinary culture is more conservative, where scholarly societies or key journals choose to hold back change, institutional repositories can help individual faculty take the lead in initiating shifts in disciplinary practice.

Institutional repositories can encourage the exploration and adoption of new forms of scholarly communication that exploit the digital medium in fundamental ways. This, to me, is perhaps the most important and exciting payoff: facilitating change not so much in the existing system of scholarly publishing but by opening up entire new forms of scholarly communication that will need to be legitimized and nurtured with guarantees of both short- and long-term accessibility. Institutional repositories can support new practices of scholarship that emphasize data as an integral part of the record and discourse of scholarship. They can structure and make effective otherwise diffuse efforts to capture and disseminate learning and teaching materials, symposia and performances, and related documentation of the intellectual life of universities.

**Cautions about Institutional Repositories**

There are at least three areas in which I am concerned attempts to develop institutional repositories could go seriously astray and become counterproductive.

The first potential danger is that institutional repositories are cast as tools of institutional (administrative) strategies to exercise control over what has typically been faculty controlled intellectual work.
I believe that any institutional repository approach that requires deposit of faculty or student works and/or uses the institutional repository as a means of asserting control or ownership over these works will likely fail, and probably deserves to fail. Institutional repositories will succeed precisely because they are responsive to the needs of campus communities, and advance the interests of campus communities and of scholarship broadly. To the extent that they try to enforce behavioral or cultural changes—and particularly controversial ones—within the campus community they will and should fail. The theme is accepting responsibility, not exerting new levels of control. This is not to say that policies mandating the deposit of materials that are broadly recognized as part of the institutional record (and recognized as being owned by the institution itself) are inappropriate. But institutions should move very conservatively down this path.

My second concern is somewhat similar to the first, that we respect institutional repositories as infrastructure and not overload this infrastructure with distracting and irrelevant policy baggage, but from a very different perspective.

We must not lose the crucial distinction between the role of institutions in establishing institutional repositories and the roles of scholarly communities within the institution's organizational units or within disciplines in creating and managing scholarly communication mechanisms that may build upon an institutional repository infrastructure. Campus administrators, librarians, and faculty members wishing to challenge existing systems of scholarly publishing (specifically their economic models and their creation of barriers to access through intellectual property control and licensing arrangements) may try to link their efforts too directly to institutional repositories by imposing inappropriate policy constraints upon the repository services.

Institutional repositories may legitimately serve as infrastructure to advance some of these interests—for example, groups might construct a peer-review process that certifies selected works that are accessible in various institutional repositories and even develop overlay systems that span a complex of institutional repositories and create a “virtual” journal. Note that such an effort would have to be extra-institutional and cross-institutional to have much scholarly credibility, for the same reason that university presses aren’t simply publishing outlets for the faculty at their parent institutions, and the editorial boards of institutionally hosted journals are drawn from beyond the host institution. Its extra-institutional nature should help to clarify that it shouldn’t be confused with the development of individual institutional repositories.

But this is not, to my mind, the primary point of institutional repositories. Indeed, it dramatically underestimates the importance of institutional repositories to characterize them as instruments for restructuring the current economics of scholarly publishing rather than as vehicles to advance, support, and legitimate a much broader spectrum of new scholarly communications. Further, I would argue that complex, cumbersome “gate keeping” policies for admitting materials to institutional repositories—particularly those that emulate practices from traditional scholarly publication such as the use of peer reviewers—are highly counterproductive; this will prevent institutional repositories from supporting and empowering faculty innovators and leaders. Membership in the campus community—certainly, if nothing else, membership in the campus faculty—should be sufficient credential to place materials in the institutional repository. To be sure, there are practical resource constraints that each institution will have to work out; some faculty have truly enormous datasets or multimedia collections that may be hard to accommodate. But recognize that the institutional repository isn’t a journal, or a collection of journals, and should not be managed like one. That’s not the point or the purpose of an institutional repository.

This does not preclude erecting superstructures on top of an institutional repository that implement elaborate gate-keeping mechanisms (the “community” mechanisms in DSpace, for example, allow the devolution of policies to specific groups and also sub-branding of areas within the repository as being under the policy control of specific groups) but the key point is that the basic repository service is an infrastructure service that should be kept divorced from policies imposed by such overlays. Such overlays might represent new journals, as already discussed; they might also represent archives, complete with appraisal systems and record-retention schedules, for example. My
argument is simply that it’s important to maintain a simple, low-barrier-to-submission, basic repository service as well, and that this service is much of the point of setting up the repository in the first place.

Institutional repositories are not a challenge or alternative to disciplinary repositories; rather, they complement them, just as they can complement existing venues of scholarly publication. The Open Archive Initiative Metadata Harvesting Protocol <http://www.openarchives.org/> gives us the tools for an institutional repository to act as an entry point for redistributing works to systems of disciplinary repositories. It is desirable to make this as simple as possible, and to insulate faculty from having to deal with the details of a constantly evolving multiplicity of disciplinary services. Better to present the faculty with a simple and stable submission interface to the institutional repository. In this sense institutional repositories can be an infrastructure upon which disciplinary services and repositories can build.

I have a third, rather different, concern about institutional repositories. We are now seeing a substantial number of leading institutions making commitments to implement them. In the near future, many campus communities may expect and demand that such services be made available rapidly; creating institutional repositories may also become fashionable in some administrative circles. My fear is that, at some institutions, repositories will be offered hastily and without much real institutional commitment.

It’s vital that institutions recognize institutional repositories as a serious and long-lasting commitment to the campus community (and to the scholarly world, and the public at large) that should not be made lightly. In establishing institutional repositories, institutions are both accepting risks and making promises; they are creating new expectations. In a budget crunch, the institutional repository may be one of the last things that can be cut, given the way that digital preservation demands steady and consistent attention and hence funding. Faculty who choose to rely on institutional repositories to disseminate and preserve their work are placing a great deal of trust in their institution and in the integrity, wisdom, and competence of the people who manage it. We need to ensure that our institutional repositories are worthy of this trust.

An institutional repository can fail over time for many reasons: policy (for example, the institution chooses to stop funding it), management failure or incompetence, or technical problems. Any of these failures can result in the disruption of access, or worse, total and permanent loss of material stored in the institutional repository. As we think about institutional repositories today, there is much less redundancy than we have had in our systems of print publication and libraries, so any single institutional failure can cause more damage. I worry a great deal about what the various impacts and implications of the first few major failures of institutional repositories—for whatever reasons—will be; I fear, for example, that they may greatly set back scholarly acceptance of authorship of digital works; they may have a corrosive effect on the trust that underpins campus communities; they may undermine broad social support for higher education. Sadly, I have little doubt that we will see such failures within the next decade or so. I hope I am wrong.

Stewardship is easy and inexpensive to claim; it is expensive and difficult to honor, and perhaps it will prove to be all too easy to later abdicate. Institutions need to think seriously before launching institutional repository programs.

Institutional Repositories and Networked Information Standards and Infrastructure

I believe that institutional repositories will promote progress in the development and deployment of infrastructure standards in a variety of difficult or neglected areas. Here I’ll mention only three.

Preserveable Formats. Institutional repositories make promises about stewardship and preservation. These promises are necessarily qualified. Institutions will make choices based on a balance of campus community demand and local assessments about technical feasibility, which will result in lists of file formats that they will commit to preserve in accessible forms (presumably through format migration); in other cases, they may preserve the bits that make up a file, but will offer no guarantees that these bits can be interpreted in the future without the development of specialized programs to read them. These choices can then be collected into broader community consensus within the higher education and research community as a form of bottom-up standards development that benefits from active work in curation and ongoing faculty involvement.

Identifiers. The ability to make persistent reference to materials in institutional repositories will clearly be
critical as these materials will form an important part of the scholarly dialogue and record. This will have to include provisions to deal with issues like versioning. Higher education and the library community have not been sufficiently active in this area, largely ceding the field to commercial interests and traditional publisher agendas. The deployment of institutional repositories will drive pragmatic solutions in this area.

Rights Documentation and Management. The management of rights for digital materials will be essential. The whole point of institutional repositories is to facilitate access, reuse, and stewardship (which may itself involve reformattin) of content, and we need methods of recording and documenting the rights and permissions associated with works that facilitate these goals of the research and education community. Part of this is a technical problem involving metadata structures; the other part is building consensus around a relatively small number of sets of terms and conditions that can cover the majority of the materials in practice. Working “standards” like the stock licenses under development by Creative Commons <http://creativecommons.org/> will be important here, and institutional repositories will be a way to make campus community members aware of these developments. Again, institutional repositories offer the opportunity for bottom-up, community-driven, consensus development about rights and permissions.

Future Developments in Institutional Repositories

I’ve described the current developments in institutional repositories and tried to explain why these are so deeply and strategically important to the enterprises of scholarship and higher education. The perspective has been largely a near-term one. In concluding this paper I want to at least sketch a few additional developments that may build upon an increasingly well established institutional repository model.

Not every higher education institution will need or want to run an institutional repository, though I think ultimately almost every such institution will want to offer some institutional repository services to its community. We will see various forms of consortial or cluster institutional repositories. Well designed institutional repositories will separate system operation from curatorial and policy control (e.g. submission, preservation, etc) of specific sets of content. Thus we can expect institutional repositories to be a basic part of the negotiations in the development of regional or disciplinary consortia among universities or libraries.

There is a clearly evolving idea of “federating” institutional repositories but as yet little concrete exploration of what this means—cross-repository search, swaps of storage between institutional repositories to gain geographic and systems diversity in pursuit of backup, preservation, and disaster recovery, or other capabilities. This will be a fruitful area for exploration and innovation. Another part of federation is that faculty often don’t stay at a single institution for their entire career, and they frequently disregard institutional boundaries when collaborating with other scholars. Federation of institutional repositories may also subsume the development of arrangements that recognize and facilitate faculty mobility and cross-institutional collaborations.

Finally, university institutional repositories have some very interesting and unexplored extensions to what we might think of as community or public repositories; this may in fact be another case of a concept developed within higher education moving more broadly into our society. Public libraries might join forces with local government, local historical societies, local museums and archives, and members of local communities to establish community repositories. Public broadcasting might also have a role here. In the long run it raises questions about “publishing” (and particularly nonprofit publishing) not in the scholarly context, but by members of arbitrary, perhaps but not necessarily geographically defined, communities or other interest groups. It is not inconceivable that we might also ultimately see commercial repository services for the public at large.

It is clear that the institutional repository is a very powerful idea that can serve as an engine of change for our institutions of higher education, and more broadly for the scholarly enterprises that they support. If properly developed, it advances a surprising number of goals, and addresses an impressive range of needs. Some of the results seem clear, though there are also likely to be any number of unexpected consequences. This is an area where I believe universities need to invest aggressively, but where they also need to implement thoughtfully and carefully, with broad consultation and collaboration across the campus community (with intellectual leadership from the faculty and the library working in partnership) and with a full understanding that if they succeed they will permanently change the landscape of scholarly communication.

—Copyright © 2003 Clifford A. Lynch

Framing the Issue: Open Access
by Mary M. Case, Director, Office of Scholarly Communication, and Judith Matz, ARL Communications Officer

In early 2002, an Association of Research Libraries task force recommended that the Association promote "open access to quality information in support of learning and scholarship." A key component of this effort is educating members of the research and academic communities about open access and its potential. ARL encourages discussions among library staff, campus administrators, university counselors, faculty, and policymakers about open access and how its application in research institutions can provide a cost-effective way to disseminate and use information. This resource guide will assist and inform these discussions. The guide highlights the key points to consider in thinking about and discussing open access, gives examples of open access implementation, and provides sources for more information.

Why Is Access to Information Important?

Society benefits from the open exchange of ideas. Access to information is essential in a democratic society. Public health, the economy, public policy all depend on access to and use of information, including copyrighted works.

Access to copyrighted materials inspires creativity and facilitates the development of new knowledge.
- Intellectual property is the life-blood of progress in the sciences and arts.
- New knowledge is developed from existing information. Authors build on the intellectual products of others to create new works

Copyright exists for the public good.
Copyright was intended to serve the public interest by encouraging the advancement of knowledge while protecting the rights of authors and copyright owners. It is meant to balance the competing interests of creators, publishers, and users, not stifle the free flow of information.

Federal investment in R&D is leveraged by access to research results.
The federal government spent close to $50 billion on non-defense-related R&D in 2002. The government depends on the dissemination of the results of that research as a stimulus to further economic, scientific, medical, and environmental development.

What Obstacles Limit Access?
Rapidly rising journal subscription prices have severely eroded the ability of libraries, universities, and scholars to purchase the publications necessary for research and education.
- Expenditures for serials by research libraries increased 210% between 1986–2001 while the CPI increased 62%. The typical library spent 3 times as much but purchased 5% fewer titles.
- Book purchases declined by 9% between 1986–2001 as libraries sought to sustain journals collections. Based on 1986 purchasing levels, the typical research library has foregone purchasing 90,000 monographs over the past 15 years.
- In the electronic environment, the model has changed from the purchase of physical copies to the licensing of access. In general, libraries do not own copies of electronic resources and must negotiate licenses (rather than depend on copyright law) to determine access and use.
- Large bundles of electronic journals offered by major commercial publishers will force smaller publishers out of business. Multiple-year licenses to large bundles of content that preclude cancellations will force libraries to cancel titles from smaller publishers to cover price increases of the bundles. This diminishes competition and increases the market control of the large publishers.
- Lack of corrective market forces has permitted large companies to reap high profits from publishing science journals. In 2001 Reed Elsevier’s STM division’s operating profit was 34%, while its legal division’s operating profit was 20%, its business division’s 15%, and education 23%.
- Mergers and acquisitions increase prices and eliminate competition. Research has shown that mergers exacerbate the already significant price increases of journals owned by the merging companies. While there were 13 major STM publishers in 1998, only 7 remained by the end of 2002.

Legislative and Legal Issues.
In addition to the use of licenses, several large commercial entities, anxious to protect their content and profits, have promoted legislation to extend limitations on access to copyrighted materials.
- DMCA (Digital Millennium Copyright Act) of 1998 is the most comprehensive reform of copyright law in a generation. It is a flawed attempt to take copyright principles into the digital age.
  - Fair use and other exemptions of the Copyright Act. The anti-circumvention provision of the DMCA could eliminate fair
use and other exemptions because it allows content owners to use technical protection measures to control access to digital works. Fair use is a provision in copyright law that allows the use of copyrighted works for such purposes as teaching and research without seeking permission if certain conditions are met. Content owners could use technical protection measures not only to restrict access, but also to restrict use of copyrighted electronic resources in libraries. This could jeopardize fair use and libraries' ability to preserve digital works. Several court cases have upheld the anti-circumvention provision.

- **First sale provision** allows libraries to loan copies of works and receive gifts from donors. The DMCA commissioned a study by the Copyright Office on whether first sale should apply to digital works. The Copyright Office agreed with the arguments of content owners who said that first sale should not cover digital transmissions.

- **Sonny Bono Copyright Term Extension Act (CTEA) (1998)**

  - CTEA added 20 years to copyright protection and eliminates adding any works to the public domain until 2018. It was designed to protect the economic interests of large content companies, such as Disney, and does not take into account the vast amount of copyrighted material shared by scholars once it enters the public domain.

  - The CTEA was challenged in a case brought before the Supreme Court (Eldred v. Ashcroft). The court upheld the law in its ruling in early 2002.

- **Database Legislation**

  - Would bypass copyright law and create a new form of intellectual property for databases. Proponents have been pushing database legislation for many years.

  - Could allow publishers to restrict access to and use of databases of factual or public domain information, which are not protected by copyright.

  - Combined with new technological protections, this legislation would give database owners unprecedented control over factual information.

- **UCITA (Uniform Computer Information Transactions Act)**

  - Proposed state law that would legalize shrink-wrap and click-on non-negotiated licenses for computer software and digital information, thus permitting content owners to use these licenses to override copyright exemptions.

  - Only passed in two states amid growing awareness of potentially harmful effects on consumers. American Bar Association Working Group recommended that entire legislation be redrafted.

  - Recent amendments to UCITA by NCCUSL (National Conference of Commissioners on Uniform State Laws), the organization drafting and promoting the legislation, still do not address library, consumer, and other-sector concerns.

### What Is Open Access?

**Open access is a cost-effective way to disseminate and use information.** It is an alternative to the traditional, subscription-based, publishing model made possible by new digital technologies and networked communications. As used by ARL, open access refers to works that are created with no expectation of direct monetary return and made available at no cost to the reader on the public Internet for purposes of education and research. Any user of open access works could read, download, copy, distribute, print, search, or link to the full texts of works, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the Internet itself. Open access does not apply to materials for which the authors expect to generate revenue.

**Open access operates within the current legal framework of copyright law.** Authors own the original copyright in their works. In the process of publishing, authors can transfer to publishers the right for publishers to post the work freely on the Web, or authors can retain the right to post their own work on institutional or disciplinary servers. Authors, however, retain control over the integrity of their work and have the right to be properly acknowledged and cited.

**Open access is intended to be free for readers, not free for producers.** The costs of producing digital open-access literature are believed much lower than the costs of producing print literature, but financial and human resources are required. Author or institutional fees for dissemination have been proposed as possible alternatives to the traditional library subscription model for funding the costs of open access.

**Open access focuses on academic research.** Open access is concerned with scientific and research texts that scholars give to the community without expectation
of direct monetary return, including peer-reviewed journal articles, preprints, preliminary findings, and data sets.

**Open access and peer review.** Open access does not mean that peer review is bypassed. Peer review is medium-independent, as necessary for online journals as for print journals, and no more difficult.

**Initiatives that Call for Open Access**

**Budapest Open Access Initiative (BOAI)** is a statement of principle, strategy, and commitment to making research articles in all academic fields publicly available on the Internet. The initiative has been signed by a growing number of researchers, universities, laboratories, libraries, foundations, journals, publishers, learned societies, and scholars from around the world. The BOAI recommends using two complementary strategies: self-archiving in institutional/disciplinary repositories and open access journals.

**SPARC (the Scholarly Publishing and Academic Resources Coalition)** is actively promoting both open access journals and the development of institutional repositories. SPARC has a number of open access partners including: Algebraic and Geometric Topology, Documenta Mathematica, eScholarship, Geometry & Topology, Journal of Insect Science, Journal of Machine Learning Research, and New Journal of Physics.

**Public Library of Science (PLoS)** is a grassroots initiative signed by over 30,000 scientists to encourage publishers to deposit their journals in central archives, like PubMed Central, within six months of publication. These scientists believe that information from multiple sources stored in a common format in central repositories can significantly enhance their ability to search across collections, manipulate data, and develop tools to integrate the literature with a variety of other information resources. Having generated only modest response from publishers, the leaders of the PLoS are planning to develop their own set of open access journals.

**Open Access Implementation: Some Examples**

**PubMed Central.** A digital archive of life sciences journal literature developed, managed, and supported by the National Center for Biotechnology Information at the U.S. National Library of Medicine. Access to PubMed Central is free and unrestricted. Participation in PubMed Central is voluntary and publishers can deposit journal articles at any time. Copyright remains with the journal or author.

**BioMed Central (BMC).** An independent commercial publishing house committed to providing immediate free access to peer-reviewed biomedical research. BMC publishes more than 50 online journals in biology and medicine and uses authors’ fees and institutional memberships to fund its open access journals. Authors who publish in BMC journals retain copyright.

**arXiv.org e-Print archive.** Started in 1991, arXiv.org is a fully automated electronic archive and distribution server for preprints in physics and related disciplines, mathematics, computer science, and cognitive science. The service, formerly hosted by Los Alamos National Laboratory, was transferred to Cornell in September 2001.

**Institutional repositories** are digital archives of intellectual products created by the faculty, staff, and students of an institution and accessible to end users both within and without the institution, with few if any barriers to access. Institutions may act independently or within a state or regional consortium. A number of institutions, such as Caltech and the University of California, have already built such repositories. DSpace, a repository being developed at MIT, will host research material from professors at the institute and will allow researchers to select access levels for each item they contribute.

**Disciplinary repositories** are used by certain academic disciplines to facilitate the sharing and storage of research materials. These repositories (known as "e-print servers") have enjoyed high rates of participation within their respective fields. Repositories exist in disciplines such as classical literature, history of philosophy, economics, chemistry, cognitive sciences, mathematics, and physics.

**Self-archiving, generally within an institutional or disciplinary repository, refers to the depositing by the author of a digital document in a publicly accessible Web site. It includes articles and preprints by individual researchers.**

**Standards that Enable Open Access**

**The Open Archives Initiative (OAI)** develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content. The OAI Metadata Harvesting Protocol allows third-party services to gather standardized metadata from distributed repositories and conduct searches against the assembled metadata to identify and ultimately retrieve documents. While many proponents of OAI advocate open access, i.e., the free availability of works on the Internet, the fundamental technological framework and standards of the OAI are independent of the both the type of content offered and the economic models surrounding that content.
Online Resources for More Information

Promoting Open Access


Copyright and Intellectual Property


Obstacles to Open Access


Open Access Initiatives


Open Access Implementation


DSpace at MIT. <http://dspace.mit.edu/>.


Open Archives Initiative


For a more extensive list of resources, see the Web version of this guide <http://www.arl.org/scomm/open_access/framing.html>. 
THE END OF HISTORY?
REFLECTIONS ON A DECADE
by William J. Crowe, Spencer Librarian, Kenneth Spencer Research Library, University of Kansas

I am honored to speak at this special anniversary meeting. It was only three years ago that I took part in my last ARL meeting, in Kansas City and Lawrence, and I welcome the chance once again to be among many good friends and colleagues. I also recall listening, with great appreciation, at our 60th anniversary meeting to David Stam’s witty and revealing tour d’horizon, aptly entitled “Plus ça change.”

David’s work is a very important contribution to understanding the character of this Association.

This talk is meant to be different from David’s effort. It is not a history. As an aid to historians of the future, the ARL staff has done a very good job of recording for this meeting information about the many milestones of the last decade. However, I chose consciously not to try to take an historical perspective, because that is of course not possible. We are too close to the last decade. Many of the issues, and many of the people, are still with us. Indeed, this talk is very much a reflection. It is not based on research, but on my recollections of participating in the Association, some reading of secondary sources, and some conversations with ARL staff and others who have represented member institutions.

At the outset of the ‘90s there was much promise of change in the world of information and libraries, even if we did not always recognize where it might take us. In looking at ARL during that halcyon era, I am happy to observe that we did not fall prey to the techno-geeks or wild-eyed enthusiasts. However, neither did we shutter the windows. We worked hard to understand change and, as David’s analysis of our traditions makes plain, also to make change, always with a mind to sustaining our shared mission and promoting common values as stewards of the great majority of North America’s research libraries.

There were so many ARL initiatives, so many ventures, so many new partnerships that it is almost dizzying to recall them. Still, David was—and still would be—right to observe that there was much continuity of purpose, if not always of method; there was sustained strategy, if not always the same tactics.

Teasing out which of those trends might warrant more reflection was a challenge. In the end, I chose on the basis of what I think might capture the interests of the next generation and that could, I think should, continue to shape a large part of the future agenda for the Association. The four trends that I regard as most important coming out of the ARL experience in the ‘90s are:

- Attention to measurement/assessment (moving beyond a rearview-mirror approach)
- Changes in scholarly communication (not too large)
- Changes in the theories and practices of organizing and preserving knowledge objects (who knew “knowledge objects” or “metadata” 10 years ago?)
- Recognition of the need to prepare for a transformed workforce, including directors, in the research library community

In each of these areas, ARL made enormous strides in the ‘90s in mobilizing facts and opinions, seeking partners, and finding champions among member library directors and staff, and so was often in a position to promote wanted change.

Here is a core message for those of you new to ARL: Never doubt that one person or a small and determined band of people can make something happen. Consider the following examples.

Would SPARC have arisen if not for the passionate advocacy of Ken Frazier and the hard work of Mary Case and others?

Where would advocacy for fair use and a reasoned balance in considerations of intellectual property be without Jim Neal and Prue Adler?

Where would the new measures initiative be without Carla Stoffle and Martha Kyrillidou?

Where would our very effective partnerships with the AAU, NASULGC, EDUCAUSE, CLIR, NSF, IFLA, and so many others be without a long line of ARL presidents, board members, and most especially Duane Webster?

In every one of the four areas I have cited we face new challenges. Who now will be the champions? Who will expand ARL’s reach to cement relationships with old and new partners? Who will steward the time and energy of the Association and its money, to target these always scarce commodities away from some still worthy purposes in order to meet new challenges?

I see four major and I hope transitory challenges in all of these areas:

Scarcity of money, tight budgets at home, foundations still giving but from constricted funding bases, government showing some promise, but in neither Canada nor the U.S. do the levels of funding for libraries approach those of an earlier halcyon era, just a generation ago.

A persistent belief by many leaders in higher education and government in the efficacy of the market, with a concomitant tendency to accept the commodification of information. Is there hope yet? Perhaps our frustration is like that felt by the environmental community of a generation ago, before Silent Spring (a book, isn’t?) brought the society around.

An aging population of librarians, notably in the ARL community, presaging at once the loss of much
experience (and some wisdom?) and, at the same time perhaps, the hanging on of too many of the Baby Boomer cohort, many because of suddenly flimsier retirement nests. Might some of us stay beyond our time, frustrating the younger, and sometimes more venturosome and energetic 20-, 30-, and 40-somethings? Will they be able to wait?

An erosion of the near reverence for libraries that many in earlier generations of scholars and decision makers in higher education and sometimes government had. We may have been over successful during the last half-century in professionalizing the work of the library, and so unwittingly have pushed to the sidelines some of our strongest allies and advocates. Here, I see the transformation of scholarly communication, which is their issue and ours, reviving the chances for a true partnership, which the last 50 years may too often have eroded. We need to sustain our shared work, recalling the lessons of our storied and sometimes colorful past.

It is for that reason that I can ask and answer so confidently the question: Were the ‘90s the end of history? Of course not. Only another fin de siècle (not only David Stam can recall apt French terms) in which each of the institutions represented here came together to sustain our responsibility to society.

Let me close not with another French phrase, but with an English one. It comes from my home library, the Boston Public Library [BPL], where I began work as a high school student. As those of you with ties to Boston may recall, the north facade of the McKim building of the BPL, visible where one emerges from the subway, bears a very New England admonition: The Commonwealth Requires the Education of the People As the Safeguard of Order and Liberty.

Not a bad sentiment from the ‘90s, the 1890s of course. This is a statement of purpose worthy of any great research library and of the Association that brings such libraries on this continent together in shared purpose.

—Copyright © 2002 William J. Crowe

The editors thank the author for permission to publish this article, excerpted from a paper presented at the ARL Membership Meeting, October 16, 2002. The full paper is available on the ARL web site <http://www.arl.org/arl/proceedings/141/crowe2.pdf>.

1 See <http://www.arl.org/arl/plus.ca.html>.
3 The acronyms stand for Association of American Universities, National Association of State Universities and Land-Grant Colleges, Council on Library and Information Resources, National Science Foundation, and International Federation of Library Associations and Institutions, respectively.
NEW SPEC KITS FROM ARL/OLMS

SPEC Kits Exchange Information
Designed to examine current research library practices and policies and serve as resource guides for libraries as they face ever-changing management problems, each SPEC Kit contains a summary analysis, survey questions with tallies, pertinent documentation from participating libraries, and a reading list and Web site references for further information on the topic. ARL/OLMS published the following six SPEC Kits in fall 2002. The tables of contents and executive summaries are available online at <http://www.arl.org/spec/specialist.html>.

Reference Service Statistics and Assessment
SPEC Kit 268
This survey documents how ARL member libraries are collecting, analyzing, and using their reference service statistics, revealing strong dissatisfaction with current assessment techniques. It is hoped that the survey will inspire libraries, librarians, and library associations to devote the necessary resources to developing innovative, efficient, and effective procedures that can be widely adopted. Eric Novotny. September 2002. ISSN 0160-3582. 142 pages. $45 ($35 ARL members).

Integrating Preservation Activities
SPEC Kit 269
This survey explores to what extent preservation activities have, in fact, been integrated into research library operations by examining the many facets of mature preservation programs and determining how much activity takes place within, and how much outside, the purview of preservation departments. Karen Brown and Emily Holmes. October 2002. ISSN 0160-3582. 144 pages. $45 ($35 ARL members).

Core Competencies
SPEC Kit 270
Core competencies are the skills, knowledge, abilities, and attributes that employees across an organization are expected to have or achieve so that they can contribute successfully within a particular organizational context. This survey examines whether research libraries are defining or adopting core competencies and how they are integrating them into recruitment, staff development, evaluation, and compensation programs. Beth McNeil. October 2002. ISSN 0160-3582. 128 pages. $45 ($35 ARL members).

Library Systems Office Organization
SPEC Kit 271
The past decade has seen dramatic changes in information technology, library services, staff responsibilities, and organizational efforts to manage these changes. Through this survey the investigators document the changes in research library systems operations since 1994—when ARL last surveyed its members about the organization of the library systems office—and identify future trends. Scott P. Muir and Adriene Lim. November 2002. ISSN 0160-3582. 128 pages. $45 ($35 ARL members).

Insuring and Valuing Research Library Collections
SPEC Kit 272
Risk management in research libraries has gained increasing attention in recent years. The role of the institutional risk manager has expanded, often impacting the business processes of the library. This survey investigates a variety of methods used to value and insure research library collections. It informs librarians for discussing these issues with administrative officers and risk management staff. Susan K. Martin. December 2002. ISSN 0160-3582. 88 pages. $45 ($35 ARL members).

Chat Reference
SPEC Kit 273
ARL reported on electronic (primarily e-mail) reference in SPEC Kit 251, Electronic Reference Service, in 1999. Since then, many libraries have adopted interactive chat reference as a service delivery mode. This new survey gathered data on chat reference service in ARL member libraries. It provides assistance to libraries that are exploring or planning to initiate such a service. Jana Ronan and Carol Turner. December 2002. ISSN 0160-3582. 141 pages. $45 ($35 ARL members).

Subscribe to SPEC
SPEC Kit Subscription (ISSN 0160-3582), 6 issues per year, shipping included. U.S. and Canada: $190 ARL member, $260 nonmember; International: $330. Payment by check, money order, or MasterCard or Visa is accepted. Make check or money order payable in U.S. funds to the Association of Research Libraries, Federal ID #52-0784198-N.

Shipping and handling for individual orders:
U.S.: UPS Ground, $10/publication
Canada: UPS Ground, $15/per publication
International and Rush Orders: Call (301) 362-8196 or e-mail <pubs@arl.org> for quote.

To subscribe or order individual SPEC Kits, contact:
ARL Publications Distribution Center
P.O. Box 531
Annapolis Junction, MD 20701-0531
Phone: (301) 362-8196
Fax: (301) 206-9789
E-mail: <pubs@arl.org>
Web: <http://www.arl.org/pubscat/order/>
TRANSITIONS

California, Los Angeles: Janice Koyama, Associate University Librarian for Public Services, was appointed Interim University Librarian effective January 1, 2003. She replaced Alison Bunting, who recently announced her retirement.

California, Riverside: Ruth Jackson, former Dean of Libraries at Wichita State University, was named University Librarian effective November 2002.

Colorado State: Catherine Murray-Rust was appointed Dean of the University Libraries effective July 1, 2003. She is currently Associate University Librarian for Public Services and Innovative Technology at Oregon State University.

Emory: Joan Gotwals, Vice Provost and Director of Libraries, announced her retirement effective fall 2003. Linda Matthews, Special Collections Division Leader, was appointed Interim Director for three years.

Kentucky: Carol Pitts Diedrichs, Assistant Director for Technology Services and Collections at Ohio State University Libraries, was named Dean of Libraries effective July 1, 2003.

Miami: Steven G. Ullmann, Vice Provost for Faculty Affairs and Dean of the Graduate School, is serving as the Interim University Librarian.

Other Transitions

Association of College and Research Libraries: ACRL contracted with Susan Martin to be the ALA division’s part-time visiting program officer for scholarly communication. She was previously University Librarian at Georgetown University and is President of SKM Associates, Inc., a consulting firm.

Canadian Library Association Don Butcher was named the Association’s new Executive Director. He was formerly Executive Director of Ontario’s real estate appraisers association.

OCLC: The OCLC Board of Trustees elected two new board members, and the board’s chair, William J. Crowe, Spencer Librarian, Kenneth Spencer Research Library, University of Kansas, was re-elected and will continue to serve as chair for a fourth consecutive year. The two new board members, who will serve four-year terms, are Edward W. Barry, former President of Oxford University Press, and Martin Gomez, Executive Director of the Friends & Foundation of the San Francisco Public Library. Also, two librarians elected to the board by the OCLC Members’ Council have begun their terms: Larry Alford, Deputy University Librarian, University of North Carolina, and Jerry Stephens, Librarian and Director, Mervyn H. Sterne Library, University of Alabama, Birmingham, will serve six-year terms on the board.

JAN MERRILL-OLDHAM HONORED BY ARL

During the October 2002 ARL Membership Meeting, ARL honored Jan Merrill-Oldham, Malloy-Rabinowitz Preservation Librarian, Harvard University, for her service to research libraries. Nancy Gwinn, Director of the Smithsonian Institution Libraries and Chair of the ARL Preservation of Research Libraries Materials Committee, presented Ms. Merrill-Oldham with a certificate that reads:

The Association of Research Libraries extends its gratitude to Jan Merrill-Oldham for her 15 years of service as a consultant to the Preservation of Research Libraries Materials Committee. With the passion of a practitioner, the vision of a leader, and the eloquence to convey both, Jan dedicated herself to helping ARL directors understand and appreciate the importance of preservation to the well-run research library. While building excellent preservation programs at the University of Connecticut and now at Harvard, Jan helped to shape a preservation agenda for ARL and guided the development of standards for mature preservation programs in ARL libraries. Quick to understand the implications of the digital revolution for preservation, she was also careful to remind the Committee of the continuing need to preserve brittle books and formats not yet addressed. Jan is a leader in the preservation profession and trained many of the preservation administrators now in ARL libraries. We commend and thank Jan Merrill-Oldham for her service to the Committee, the Association, the profession, and scholarship.

Pictured at the Membership Meeting, left to right, are Duane E. Webster, ARL Executive Director; Paula Kaufman, University of Illinois at Urbana-Champaign and ARL President; Ms. Merrill-Oldham; and Ms. Gwinn.
ARL Calendar 2003

April 10
New Ways of Listening to Users: LibQUAL+™
ARL Preconference at ACRL
Charlotte, North Carolina

April 10
Open Access 101: What, Why, and How You Can Make it a Reality
SPARC Preconference at ACRL
Charlotte, North Carolina

April 16–18
Advanced XML Workshop
Charlottesville, Virginia

April 28–29
CNI Spring Task Force Meeting
Washington, D.C.

May 12–16
Service Quality Evaluation Academy
San Antonio, Texas

May 13–16
ARL Board and Membership Meeting
Lexington, Kentucky

July 28–29
ARL Board Meeting
Washington, D.C.

September 16–18
Facilitation Skills Institute
Location to be announced

September 22–25
Library Leadership for New Managers Program
Washington, D.C.

October 7–8
Leading Change
Washington, D.C.

October 14–16
ARL Board and Membership Meeting
Washington, D.C.

October 17
Scholarly Tribes and Tribulations: How Traditional Practices in the Disciplines are Driving Technology in Different Ways
Washington, D.C.

November 4–6
Library Management Skills Institute I: The Manager
Location to be announced

December 8–9
CNI Fall Task Force Meeting
Portland, Oregon