

The imperative for digital libraries *

Chair:

Clifford Lynch, Executive Director, Coalition for Networked Information

Lynch leads a discussion with executives who are actively engaged in building collections of digital information. What are the issues? How are we managing them?

Speakers:

Peter Bolman	President Academic Press
Stephen Rhind-Tutt	President Chadwyck-Healey
Leslie Singer	President and COO, Institute for Scientific Information

The Digital Library and Licensing Imperatives

The Road to Digital Libraries

The road to digital libraries is paved with both good intentions and tricky issues, among them coherent access to collections, archiving issues, disintermediation of roles, new forms of publication, and lack of standards. Cliff Lynch, Executive Director of the Coalition for Networked Information (CNI) invites a well-rounded panel of industry authorities to debate some “today and tomorrow questions” regarding the clear need for digital libraries.

The panel – consisting of Pieter Bolman, President, Academic Press; Leslie Singer, President & COO, Institute for Scientific Information; and Stephen Rhind-Tutt, President, Chadwyck-Healey – engages in a roundtable discussion of Lynch’s questions.

Question One: How are we going to provide readers and researchers with coherent access to collections, and what are your roles?

Bolman believes that: (1) The journal will continue to exist, (2) Peer review will continue to be used, (3) Publish or perish will remain a driving force for scientists (to justify funding), and (4) Scientists must give credit to previous achievements (through citations). Citations can be linked directly to the full-text source, and that is the exciting part of electronic publishing right now. The coherence is given by the links and by the accessibility to the links and citations in any given article, and additionally by searching. This calls for more agreements between primary and secondary publishers, and more reliance on A&I services.

“I think today’s marketplace is demanding integrated information solutions”, says Singer, “and those who feel they can stand alone on great content will have great difficulties”. Web searching, filtering,

*In cooperation with the Coalition for Networked Information.

and linking capabilities is the way forward in the value-add business. She advises more links – between secondaries and secondaries, to patent organizations, to desktop tools, to the NCBI databases The vision for ISI is to be a “preferred entry point” for the end user.

From a humanist’s standpoint, there are some universal truths to consider first, says Rhind-Tutt: (1) Users are not prepared to go through lots of screens to find what they want, (2) If you look at the total number of links, it is companies like Yahoo and Alta Vista that are linking to everything they can. For the sake of coherence, “within the next two to three years, I would expect that everybody will be linking to each other”. This will inevitably lead to the rise of business arrangements between publishers to share linkages. With effective links in place, the next level of value add will be returning to editorial value.

Question Two: What are the business arrangements, if any, surrounding linking?

Rhind-Tutt says that, with Chadwick-Healey’s business arrangement with the Gale Corporation, the customer must be a subscriber to both services to use the links. Singer says that the links are just a necessary value-add, and subscribers do not pay anything extra for the links. Bolman says that, while he favors business arrangements to enable linking, the absence of generally accepted standards makes it more difficult to carry off.

Question Three: When is linking going to get easier? What standards initiatives will succeed to help linking?

Bolman is optimistic that standards will emerge sooner rather than later. The DOI (Digital Object Identifier) is necessary, but it’s not sufficient. There needs to be a standard in place as to how links are established, with something in the middle – a link processor – to maintain current links.

ISI has established a functional and precise linking infrastructure. Migrating to the DOI for ISI, however, would pose some difficulties because ISI would need to acquire the DOI and the accompanying metadata from the primary publisher and then ISI would have to go to another site to find the appropriate URL for the DOI. The speed of throughput with the DOI is another issue.

The linking issue is one reason that Rhind-Tutt and his colleagues believe that having discipline-specific Web sites is going to be a reality in the long term. Journals are less important to Chadwyck-Healey, but it is concerned about establishing a standardized way of describing text. Good linking still doesn’t solve some of the basic problems, such as how to organize the fields and editorial concerns about which resource you present first.

Question Four: What are your strategies for integrating information outside of the traditional canon of scholarly publishing – e.g., preprint servers, Web sites, self-published material, technical reports – with the more traditional literature systems?

ISI does include electronic-only material, and plans to integrate Web sites into its products. Chadwyck-Healey indexes sites on the Web at the document level through a Web trawler. In April, it will be reindexing 300,000 of its texts and the 16,000 Web sites it links to. The Web sites range from home pages through discussion and listserv groups.

Bolman refers to the “linking imbalance” between the formal literature and preprints. “It would seem to me that if you want to keep your system clean, that you only link to the formally accepted versions of the literature.” Technical reports, while not necessarily linked, should still be searchable.

Question Five: How are we going to get this electronic material archived?

The archive needs to be part of the refreshment (updating) cycle, says Bolman, and it is a publisher's responsibility. However, to make it worthwhile for publishers, they may have to license archives separately.

Rhind-Tutt says that, in English literature, the essence of electronic publishing is its dynamic nature – it is a performance that varies each time. Archives can only selectively draw out the more permanent aspects of the performance to preserve.

Singer says that ISI now archives back to 1973, with the goal to archive back to 1945 by the year 2000. The experience is difficult because, every few years, ISI changes policy and changes the ways in which it captures and stores information. Publishers have to make critical decisions about what is the most relevant information to archive.

Question Six: How do you see roles evolving vis-a-vis the role of libraries and universities vis-a-vis the role of scholarly societies and authors?

Rhind-Tutt acknowledges that there is a race on towards a “phenomenal” all-encompassing information system. The library systems believe that they will be able to do it. The A&I community says they have the expertise in making searches efficient. Publishers feel that they can reach the critical mass. The popular search engine outfits (e.g., Yahoo, Alta Vista) believe that they provide better access to collections. And Chadwyck-Healey believes that it has a niche, and that specificity is the way forward.

Technology gives rise to a change in roles, adds Bolman. The library will be the first screen, so the library will become a bundle of access rights. Libraries will also become very distributed on campuses.

Perspectives on Licensing

The evolution to digital libraries raises questions about the scalability of licensing. To be more efficient and cost-effective, the current licensing needs to scale up, says Ann Okerson, Associate University Librarian at Yale. She is cautiously optimistic that a number of initiatives on the horizon will improve licensing of digital information. They include:

- E-journal Consolidators (not aggregators): The benefits for e-journal users are one search engine, one easy source for articles, the ability to customize, and the possibility of archive services (e.g., OCLC and Stanford's High Wire). Consolidators do not, however, offer one single license for all journals.
- Consortia: Consortia are growing and becoming multi-type. Many consortia negotiate licenses for dozens or hundreds of libraries, and the consortia usually negotiate with the largest providers. These efforts represent efficiency of scale in licensing. The International Coalition of Library Consortia (ICOLC) is a virtual organization with more than 90 consortia members worldwide which is a voice of activism for consortia.
- Standard Licenses (Templates): PA/JISC (and NESLI) are developing a model electronic license. At Yale, the LIBLICENSE Project links to publisher e-journal licenses. Uniformity among these licenses indicate some convergence within the publisher community. The LIBLICENSE software is a downloadable license template that can be customized by others.
- Agreed Upon Guidelines: Copyright law may simplify licensing, but nothing in recent copyright law revisions resolves licensing. Guidelines surrounding copyright are more likely to have the potential of simplifying licensing negotiations. National and regional licensing helps licensing scale, but the impact in the United States is limited to date.

Chemical Abstracts Service (CAS) aims to provide scientists with the widest possible dissemination of chemical information – and the most serviceable options for access and use. At the same time, CAS must preserve the integrity of its databases. Michael Dennis, Special Assistant to the Director of CAS, says that copyright law, technology, and licensing are the best tools for balancing interests.

Dennis forecasts that, based on recent bills introduced in the House and Senate (H.R. 354 and the Database Fair Competition & Research Promotion Act of 1999), some form of database protection legislation is likely to pass in 1999.

CAS uses technology to ensure the availability, integrity, accuracy, and confidentiality of the CAS systems and databases. Specifically, CAS uses ID and password, digital certificates, encryption, fire walls, system alarms, and private lines. In 1999, there will be the first commercial uses of electronic paper, electronic books, a © chip (copyright box), and the DOI.

CAS licenses define authorized users (including occasional walk-ins), authorized use, and security requirements. Dennis' take away message is that licenses are here to stay: "In negotiating licenses, publishers should take a balanced approach: listen, articulate the reason behind your desired changes, do your homework, carefully evaluate the risks, and be open to compromise."

In 1999, Dennis says to watch for developments of UCC (Uniform Commercial Code) Article 2B (in its 19th draft) that has direct impact on licensing. There will be a key vote in July.

The American Association for the Advancement of Science (AAAS) is drawing from its print experience with *Science* to incorporate the best features of that subscription in its Science Online service, says Michael Spinella, Director of Membership & Meetings for AAAS. His philosophy of online subscription agreements is "less is better".

In his view, the disadvantages of negotiated licenses in bilateral agreements between publishers and libraries outweigh the advantages. Such licenses can be complex, costly (in negotiating time), and they can slow market penetration for publishers. What should publishers look for in agreements? Equitable terms, comprehensible and enforceable provisions, a fair market price, remedies commensurate to risk, and limited liability of users/institutions.

Web look-ups: ICOLC: www.library.yale.edu/consortia
LIBLICENSE PROJECT: www.library.yale.edu/~license
UCC Article 2B: www.law.upenn.edu; www.2BGuide.com