Enabling Access in Digital Libraries

A Report on a Workshop on Access Management

February 1999

Edited by Caroline Arms

with Judith Klavans

and Donald J. Waters

ISBN 1-887334-64-5

Published by:

The Digital Library Federation
Council on Library and Information Resources
1755 Massachusetts Avenue, NW, Suite 500
Washington, DC 20036

Additional copies are available for $15.00 from the above address. Orders must be prepaid, with checks made payable to the Council on Library and Information Resources.

The paper in this publication meets the minimum requirements of the American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials ANSI Z39.48-1984.

The Digital Library Federation

On May 1, 1995, 16 institutions created the Digital Library Federation (additional partners have since joined the original 16). The DLF partners have committed themselves to "bring together—from across the nation and beyond—digitized materials that will be made accessible to students, scholars, and citizens everywhere." If they are to succeed in reaching their goals, all DLF participants realize that they must act quickly to build the infrastructure and the institutional capacity to sustain digital libraries. In support of DLF participants' efforts to these ends, DLF launched this publication series in 1999 to highlight and disseminate critical work.

Donald J. Waters
Director
Digital Library Federation

Contents

Executive Summary

With digital information rapidly increasing in amount and availability, the information management community finds itself facing a wide-reaching and complex set of challenges. One of the primary challenges is how to manage access to information that is sensitive, proprietary, or protected by copyright. Addressing this question requires the attention of

- policy makers concerned with questions of privacy and protection of data,
- legal experts who draft contracts and licenses whose terms must be implemented through automated systems for authenticating users and authorizing access,
- technologists who design software for controlling electronic use and misuse, and
- publishers and librarians, who, as major providers of information, play a central role in striking a balance between protecting copyright and enabling access to the record of knowledge.

The workshop described in this report focused on the management of access to published information resources through research libraries. Topics discussed include privacy, protection of rights, authorization, and authentication. These are, in fact, important issues of concern to all citizens whenever access to information they seek is controlled automatically.

Among the groups seeking to meet the challenge of access management are the Digital Library Federation (DLF), which consists of major research libraries and archives in the United States, the Center for Research on Information Access (CRIA) at Columbia University, and the Information and Intelligent Systems Division of the Computers, Information Sciences and Engineering Directorate of the National Science Foundation (NSF). On April 6, 1998, they brought together expert practitioners and researchers from several disciplines at a workshop, held at the Brookings Institution in Washington, D.C., to explore some of the more pressing questions for research libraries, including:

> *How can members of a university that has subscribed to an electronic journal prove that they are authorized to access an article? How is a system to confirm that the staff member, professor, or student is not someone else? Are there ways to screen out impostors?*

How finely can information providers discriminate among potential users when making their materials available? What criteria should universities and public libraries, among other organizations, use to determine who should have access to a database of published information, such as the online version of The New York Times*? What options do public libraries have to be able to authorize the use of licensed materials to the general citizenry that they serve?*

How can authors and other creators of information resources be protected from digital thievery? Is Garrison Keillor correct in predicting that authors on the information superhighway will become "the deer in the headlights" of a vast traffic they cannot control?[1] *What means do custodians have to ensure that the cultural record is accessible but that the proprietary rights of authors and creators are protected against widespread copying and redistribution?*

Should digital data be fitted with a digital lock that can be opened only by users with matching keys? How does such a mechanism accord with constitutional and legislative mandates requiring that a balance be struck between the rights of authors and creators and citizens' accessibility to the cultural record?

Such questions and the discussions they stimulated led participants to identify five key properties for the design and adoption of systems that enable access for users while respecting the rights and interests of authors and publishers.

1. *Simplicity*. The less complex a system of access management, the more readily it can be adopted technologically and organizationally, and the more acceptable it is to all involved in its implementation.
2. *Privacy*. Systems that manage access to the cultural record must protect the privacy of users from detailed tracking and disclosure of use. User privacy must not be compromised.
3. *Good faith*. Agreements on access to scholarly information rely on trust among the parties involved. Users and providers would each prefer to depend, in an access management system that im-

[1] Garrison Keillor, remarks at a panel discussion, Session III, Conference on Intellectual Property Rights and the Arts: The Impact of New Technologies, sponsored by the New York International Festival of the Arts, December 13, 1994 (transcript on file with the Columbia Law Review).

plements these agreements, on reasonable barriers against abuse rather than complex restrictions that inhibit use.

4. *Trusted intermediaries.* Intermediaries play an essential role in providing access to the cultural record as parties trusted by both users and providers and as efficient aggregators of distribution and usage. System design must take the role of intermediaries into account.
5. *Reasonable terms.* Access management systems and license agreements must recognize the distinction between access and use. Overly tight control of access to a resource may impose inappropriate constraints on its use, especially in teaching and research contexts. The most useful system will not limit access to specific user groups known in advance to be interested in a resource, but will be reasonably open to serving unlikely users whose curiosity and research interests may lead them in directions not predicted by those responsible for making the agreements or designing the systems.

Workshop participants also recommended research and project evaluation in two key areas: *system usability* and *economic models.* First, an effort must be made to understand the ways in which users interact with systems, their needs in relation to new information types, and the functionality of these types in the emerging digital environment. Second, new standards of measure must be found to assess the usage of digital resources and thereby to develop alternative pricing schemes and payment mechanisms.

Although the conclusions reached at this workshop relate specifically to the problems of managing access to the cultural record in digital form for research and teaching purposes, they apply to other realms as well, including business, medicine, insurance, credit card transactions, and logfiles from Web browsers, all of which involve more sensitive information. Enabling appropriate access to digital information depends on the efforts and talents of many stakeholders: information specialists, librarians, publishers, computer scientists, lawyers, scientists, and policymakers, and the general citizenry.

Introduction

On April 6, 1998, the Digital Library Federation (DLF) and the National Science Foundation (NSF) sponsored a one-day workshop on ways to improve systems of managing access to digital information. The workshop was an outgrowth of a two-day meeting sponsored by NSF in September 1996 exploring the technology of the terms and conditions for access.[2] The consensus there was that input from a variety of user communities was required to develop formal mechanisms for implementing terms and conditions within digital libraries. This DLF-NSF workshop was convened to provide input from research libraries with a focus on requirements for access management systems that can be designed and deployed in today's technical, legal, and economic environment.

Workshop conveners Judith Klavans, director of Columbia University's Center for Research on Information Access, and Donald Waters, director of the Digital Library Federation, invited experts from the fields of computer science, library technology, publishing, information technology, and to exchange ideas on managing access law (see appendix A for workshop participants). Prior to the workshop, they gave the participants a list of suggested readings and asked them to consider two typical scenarios faced by research libraries (see appendix B for readings and figure 1 for scenarios). In the first scenario, libraries provide digital works to other academic institutions and the public; in the second, they license digital works from publishers or publishers' agents. The conveners asked participants to discuss the scenarios from the perspective of both users and providers of information as a basis for developing requirements for access management systems.

The workshop's objectives were

- to provide input to the development of the Coalition for Networked Information (CNI) White Paper on interorganizational access management,[3]
- to identify key research problems for programs such as the NSF's Digital Libraries Initiative and Knowledge and Distributed Intelligence program, and
- to provide a springboard for implementation projects.

---

[2] James R. Davis and Judith L. Klavans, "Workshop Report: The Technology of Terms and Conditions," *D-Lib Magazine*, June 1997. Available at http://www.dlib.org/dlib/june97/06davis.html

[3] Clifford Lynch (ed.), *A White Paper on Authentication and Access Management Issues in Cross-Organizational Use of Networked Information Resources*, Coalition for Networked Information, Spring 1998. Available at http://www.cni.org/projects/authentication/authentication-wp.html.

SCENARIO I:

Libraries as providers of digital works to the public and to other academic institutions

Consider that the Digital Library Federation seeks to integrate digitized works on the theme of "Making of America" from a variety of its participating institutions. Some of the DLF institutions, such as the Library of Congress and the New York Public Library, seek to provide their works generally to the public. Other DLF institutions have narrower goals, and aim to provide their works mainly to an academic constituency. Regardless of whom they regard as their main audience, all seek to distribute the work as broadly as possible while protecting the works they provide digitally from misuse. How can the roles of the expected user populations and the differing conditions under which the institutions operate as providers best be defined and matched without compromising the larger goal of effectively integrating the distributed collections of materials?

SCENARIO II:

Libraries as licensees of digital works from publishers or publishers' agents

A library licenses access to a set of journals on a Web site housed by a publisher or publishers' agent. The journals are not of general interest to the community served by the library, but only to a subset of users. Moreover, the license is only affordable if it is limited to a subset of users. What sets of user roles might the licensing agents on campuses plausibly want to differentiate? What conditions would a publisher need to provide in order to support such differentiated access?

Specific objectives desired from breakout groups:

- Define authorization requirements from user and provider perspectives
- Affect emerging technologies and their implementation
- Identify key research problems

Consider the following questions:

- What kinds of role distinctions are necessary?
- What rights and duties are expected?
- What are the privacy issues?
- How strong must the security controls be?
- What kind of accountability is necessary and what kinds of management data are needed?
- How do we evaluate effectiveness of the system from user and provider perspectives?

Figure 1. *Workshop Handout: Scenarios and Instructions for Breakout Groups*

Summary of the Day's Activities

In the opening statements, Judith Klavans explained the design and goals of the workshop, noting that its primary concern was *access* management rather than rights management, as originally announced. Donald Waters explained that systems for access management include two key technical components, *authentication* and *authorization*. He noted that the focus of this workshop was on mechanisms for authorization. He outlined issues and options identified through a parallel initiative at CNI, the development of a White Paper on Authentication and Access Management Issues in Cross-organizational Use of Networked Information Resources. Waters reminded participants that the two scenarios presented in the workshop handout should stimulate discussions later in the day.

Two invited presentations set the stage for the discussions. Russell Vaught, director of Academic Computing at Penn State University, described the enterprise-wide authorization database in use at Penn State. John Erickson, Vice President for Systems Development at Yankee Book Peddler, Inc., described the goal of the Rights Metadata Working Group established as part of the joint activities of the National Information Standards Organization (NISO) and the International DOI (Digital Object Identifier) Foundation. In particular, he presented the conceptual model developed by the group to represent rights transactions and a proposed schema for rights operations.

Three groups were formed for two breakout sessions. Each group reflected a balance of expertise and was instructed to use the same approach, outlined in the handout, to examine the two scenarios. In the first breakout session, they were to consider both scenarios from the perspective of users. For the second, their task was to concentrate on the provider's perspective. In the event, it proved hard to draw such lines. Perhaps because the challenge is to find a balance between the perspectives, the natural instinct was to consider both sides of issues such as security and privacy. At the same time, it was immediately apparent that the perspective of the library or institutional user differs markedly from that of the individual end user. In view of this complexity, the report summarizes the discussions from a thematic rather than a chronological standpoint.

Each breakout was followed by a plenary session at which a representative from each group summarized the discussions. Before the final discussion, Peter Jaszi of the Washington College of Law, American University, presented an update on legislative activities concerning copyright and related intellectual property rights (see appendix C).

Opening Statements

Judith Klavans, Director, Center for Research on Information Access, Columbia University

As a backdrop for discussion, Judith Klavans highlighted the findings of two previous workshops on topics related to access management. The first, a two-day workshop on Technology of Terms and

Conditions, was held in September 1996. This workshop, which Klavans chaired with Jim Davis of Xerox PARC, was also funded by NSF. Roughly 30 participants explored issues from multidisciplinary perspectives. Four breakout groups focused on different aspects of the overall problem, covering the following topics: infrastructure requirements and the factors that encourage or inhibit acceptance of systems for managing terms and conditions; the technical, political, and social uncertainties that prevent the formulation of descriptions of terms and conditions; issues of scale; and ways to express conditions of use.

Of the conclusions reached, three in particular gave stimulus to the present workshop:

- publishers vary in their approaches to licensing and the degree of control they wish to retain,
- user communities must be involved in design and testing, and
- community attitudes and acceptance are of prime importance.

Other important points raised in September 1996 were that technology must accommodate vagueness and ambiguity; ambiguity may be intentional, as a consequence of the legal needs for flexible interpretation; economic pressures push publishers and libraries in opposite directions, and legal and technological developments will affect the economic balance; and international perspectives must be considered.

In December 1996, the Digital Library Federation (then the National Digital Library Federation) and researchers from the six projects funded through the first phase of the Digital Libraries Initiative held a joint meeting at Stanford University. A discussion of terms and conditions focused on points of disagreement and other issues preventing progress in building systems to manage access to information in digital libraries. Participants explored requirements from three perspectives: publishing, libraries, and technology. Publishing needs included a link to systems for digital commerce, a legal infrastructure that offered protection for digital contents, and technical mechanisms for controlling and describing digital content objects. Libraries needed mechanisms for authenticating users and roles, the association of new metadata elements with digital content objects to support self-management, and systems that allow for third-party rights. From a technological perspective, it was argued that progress would be made by ignoring some complexities, partitioning the problem, and discarding (or deferring) intractable parts.

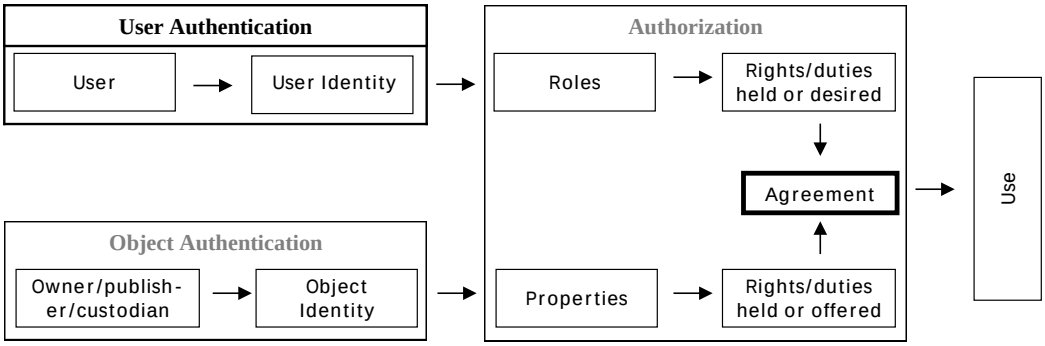
The present DLF-NSF workshop, Klavans continued, would undoubtedly contribute to the development of the CNI White Paper. In the longer term, the findings would influence ongoing collaborative projects in which participants are engaged and would guide the planning and evaluation of access management components of other digital library projects.

Donald Waters, Director, Digital Library Federation

Donald Waters opened with a brief description of the Digital Library Federation, a consortium composed of the Library of Congress, the National Archives, the New York Public Library, and sixteen of the nation's large research libraries. This organization was formed in 1995 to take the lead in identifying and lowering the barriers to federating digital libraries. One such barrier, said Waters, is the lack of adequate systems for access management, particularly for authorization. He emphasized that access management entails both authentication and authorization and presented a diagram to illustrate the elements of access management (see figure 2). Authentication refers to two distinct processes: verifying the identity of a user and ensuring that content is what it purports to be. Authorization ensures that terms and conditions in an agreement are being met by relating roles associated with a user to properties of an object.

Waters asked participants to focus on authorization within the context of access management. Participants could evaluate specific scenarios, he suggested, by drawing on the draft CNI White Paper edited by Clifford Lynch, which recommends that the following factors be considered in evaluating approaches to access management in universities: *granularity* or degree of role distinction required, privacy, strength of security, manageability with respect to accountability and ability to collect management data, technical feasibility, and affordability.

The White Paper identifies three approaches used on campuses to support authentication and facilitate authorization by remote information services or resources.



- User perspective: Associate roles of an identity with rights and duties in respect of the properties of an object to enable use
- Provider perspective: Link the properties of an information object with a set of terms and conditions for rightful use.

Figure 2. *Elements of Access Management*

1. IP source filtering: the institution warrants that traffic from a given set of Internet addresses is legitimate.
2. Proxies: the institution provides a specific machine through which all traffic to and from a service is routed and ensures that only legitimate traffic is permitted.
3. Credentials: each user presents a credential (such as a user ID and password or a digital certificate) to warrant legitimacy.

Waters then introduced the two scenarios and specific questions to be addressed in the breakout sessions (see handout reproduced as figure 1). The terms *authentication*, *authorization*, and others associated with access management are defined in appendix D, compiled from definitions used by Waters and the other presenters at this workshop, by the DOI Rights Metadata Working Group, and by Clifford Lynch in the draft CNI White Paper.

Invited Presentations

Creation of an Authorization Database

Russell S. Vaught, Director, Center for Academic Computing, Pennsylvania State University

Russell Vaught described the large multiyear effort at Penn State, a DLF institution, to build central authentication and authorization services for a large university with many campuses. The complexities, he pointed out, reach beyond the technical aspects of the project to issues of university policy and cost-benefit tradeoffs that affect far more than just the computing service organization.

Vaught sought to clarify the meaning of *authentication* and *authorization*. He pointed out that they are often confused because they are frequently employed together. Users, he noted, can be authenticated by something they know (such as an identity code or user ID and password), something they have (such as a SecureID card), or something they are (which can be verified, for example, using a retina scan). Authorization grants a user the right to use a system or data and usually presupposes authentication of users.

In 1992 the centralized Computer and Information Systems (CIS) service installed a distributed file system known as AFS, which relied on Kerberos, developed by the Massachusetts Institute of Technology (MIT), for authentication. AFS was developed at Carnegie Mellon University (CMU) and in its early stages was known as the Andrew File System (AFS). AFS and Kerberos both emerged in the 1980s, the byproducts of large projects dedicated to building campus networks and distributed systems. Kerberos provides authentication based on user ID and a password ("something you know"). In the summer of 1993, CIS decided to build a central authentication service based on version 4 of Kerberos to support all core computing systems (such as e-mail, dialup access, and the use of microcomputer labs). Kerberos is used in conjunction with SecureID cards for admin-

istrative applications in which the high cost of a security breach justifies the cost of the card (which generates passwords for one-time use). The Kerberos database includes 114,500 active principals (user identities); Penn State has 80,000 students and 30,000 faculty and staff at 24 locations.

In mid-1996, CIS decided to provide authorization services that could support more applications. The new system is based on the Distributed Computing Environment (DCE) Security Services, which uses version 5 of Kerberos for authentication. A cross-organizational task group was formed to develop an initial database to control authorization. The new system saw its first application in the summer of 1997. It is now being used to support many applications, including a proxy system for access to the remote JSTOR archive of scholarly journal literature. Vaught hopes that all systems using the old authentication service will be converted by the summer of 1999.

Although the system is complex, Vaught finds other options, such as a Public Key Infrastructure, equally complex and perhaps less cost effective. Performance and scalability, though still a concern, are expected to improve with the planned enhancements to the DCE directory component (and increased network capacity and processing power).

Reflections on the NISO DOI Rights Metadata Working Group

John S. Erickson, Vice President for Rights Technologies, Yankee Rights Management

John Erickson began by pointing out that copyright serves both as enabler and as inhibitor, establishing a balance to facilitate creativity for the overall benefit of society. His presentation described the current state of thinking of the NISO DOI Rights Metadata Working Group, chaired by Sally Morris, of Wiley, U. K., a working group formed to establish a standard rights metadata schema to facilitate electronic commerce for information objects (whether in digital or nondigital form). This working group, which has very active participation from U.K. publishers, is one of several emerging from a series of joint workshops organized by the National Information Standards Organization (NISO) and the International Digital Object Identifier (DOI) Foundation.[4]

The DOI system and related activities have developed within the publishing community and until recently, the focus has been on mak-

[4] The Digital Object Identifier (DOI) system is a mechanism for marking digital objects in order to facilitate electronic commerce and enable copyright management in a digital environment. The system emerged from activities of the Association of American Publishers, which is a charter member of the International DOI Foundation. As indicated on its Web site (http://www.doi.org/), the foundation is dedicated to supporting the needs of the intellectual property community in the digital environment, by establishing and governing the DOI system, setting policies for the system, choosing service providers for the system, and overseeing its successful operation.

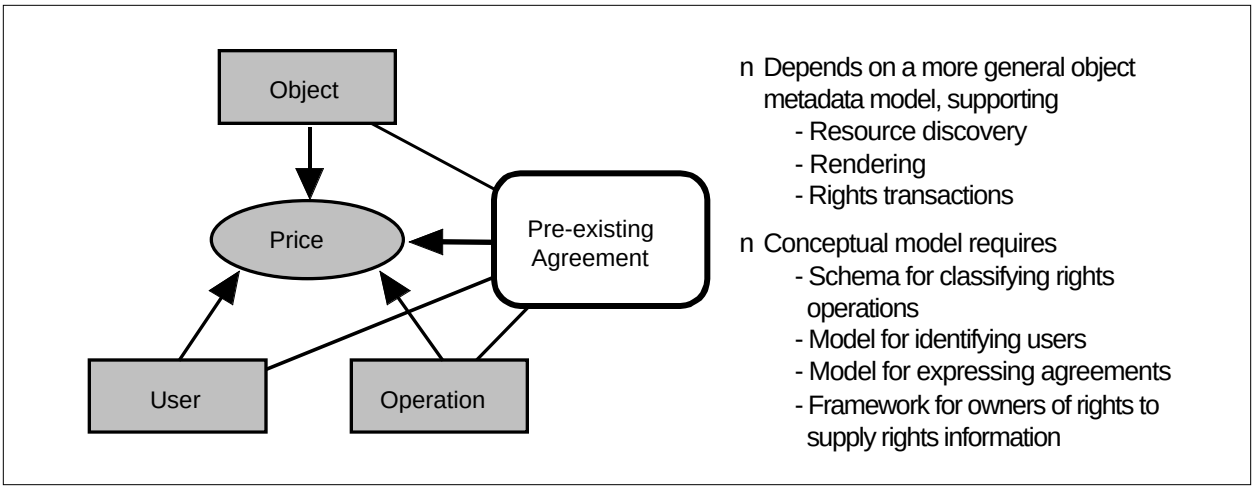


Figure 3. *Elements of the Conceptual Model for Rights Transactions*

ing money through the enforcement of rights. Erickson believes that the group is beginning to tolerate some degree of fair use and the related ambiguity. The joint activities with NISO signal recognition that discussion must be opened up to a broader community.

The group's stated objective is to develop "a consensus rights transaction model through very active, highly visible public discussions and information sharing." The resulting conceptual model is shown in figure 3.[5]

In the group's opinion, certain digital property rights languages, such as that proposed by Mark Stefik of Xerox PARC, have both advantages and disadvantages and hence alternative models are needed for purposes of comparison and practical evaluation. In particular, the group sees a need for a model to express agreements. Their current thinking borrows from approaches used by stock photography agencies and is based on the use of decision trees for evaluating permissions. A basic assumption here is that any use has a price, even if the price is zero. The model can accommodate a default agreement with standard prices for all users for a limited set of operations. Agreements could relate users or classes of users to certain operations on (uses of) classes of objects. Owners and administrators of agreements would have to be able to apply templates of operations and prices to groups of objects and users.

In Erickson's view, it is essential that the gathering of appropriate metadata become part of the publishing workflow. Two other important issues have been raised. Who would be accountable for codifying a license agreement and maintaining the data that supports access management? And would rights metadata for content be made available to third-party services along with descriptive metadata?

After Erickson's presentation, Clifford Lynch (CNI) provided some additional context for the activity of the NISO DOI working

[5] Erickson's full set of slides is available on the Yankee Book Peddler, Inc., Web site (http://www.ybp.com/yrm/presentations/DLF_CRIAShow/).

group. In a new approach to standards setting, NISO has sponsored exploratory workshops encouraging broad participation. In 1997 and early 1998, NISO and the DOI Foundation sponsored a series of joint meetings that addressed the question of whether DOI activities should be brought into the regular process for national and international standards. The meetings, the related electronic forum, and the five or six working groups they spawned have no formal standing within the national or international standards process. They are not intended to be exclusionary and have served a valuable educational role.

Discussion of Scenarios

The scope of workshop discussions was shaped partly by the facts in the scenarios and the questions asked, but perhaps more significantly by the balance of viewpoints inherent in the list of invitees and the experience of the individual participants. Most had considerable insight into the perspectives and legal responsibilities of both the users (or institutions acting as agents for users) and the providers of information resources. Many had also been involved in the parallel development of the CNI white paper and knew that some technical components were being explored in depth through that exercise. At this DLF-NSF workshop, they tended to focus on high-level requirements and policy issues, rather than on the technical details of automating the terms and conditions of use, which had been the major concern of an earlier NSF workshop. Instead, they concentrated on the problems universities and research libraries face today in their capacity as publishers of digital content created at their institution and as intermediaries licensing access from publishers and publishers' agents.

Participants brought substantial real-world experience to the discussion of the scenarios. Many had participated in the negotiation of licenses between libraries and publishers and were familiar with the economic realities that underlie such negotiations and with the practical problems of compliance. Others had struggled to establish what rights might pertain to materials in archival collections being converted to digital form and fully recognize that converting and making such materials accessible entail high costs.

The discussion focused on traditional scholarly resources and relations between academic institutions and publishers of scholarly materials. The market for such content is limited; little new money will be entering the system in the near term. The challenge is to take advantage of the opportunities offered by the electronic environment without "rocking the boat." The market for other classes of material, such as works aimed at the business or consumer market, might present very different issues. As pointed out in the earlier workshop, however, whatever the framework for managing access to digital works and balancing the rights and privileges of user and provider, its success depends on user acceptance. Any system that manages

access to the growing body of scholarly journal literature that publishers are making available in digital form, for example, must be accepted by the higher education community represented at this workshop or be doomed to failure.

The scenarios prompted discussions on a wide range of topics beyond the specific questions posed in the instructions for the breakout groups. Despite the limitation inherent in a one-day meeting, common themes emerged in the three breakout groups. These themes can guide the design and development of prototype systems.

This report summarizes the workshop discussions under three thematic headings rather than following the day's agenda. Although the discussions did not focus on technical matters, they were certainly built on some assumptions about the technical infrastructure. The first section below describes some of these implicit assumptions; they derive primarily from the CNI White Paper. The factors that affect user acceptance are drawn together in the second section. The third section extracts points that address the specific questions posed for consideration during the breakout sessions. Unanticipated issues that do not fit into these categories are described in a fourth and final section.

Technical Assumptions

A common framework for distributed access management is needed to avoid the proliferation of incompatible mechanisms developed to support specific arrangements. This framework must be general enough to support different mechanisms for authenticating users and must meet global requirements. It must permit access to be controlled at the level of individual objects (such as articles or books), not just at the entrance to a system or service that provides access to a large body of materials.

Today the most common method of controlling access is to filter by source address as defined by the Internet Protocol (IP). This mechanism is not adequate for the longer term. A limitation of particular concern to participants in this workshop is the exclusion of authorized users when they are away from an authorized site. In addition, IP source filtering cannot be applied when providing services to the general public or small organizations, such as schools, which may not have permanent IP addresses.

Universities need to develop campus-based authentication and authorization schemes for purposes other than access to licensed information resources. Authorization systems, such as that described by Russell Vaught in his presentation, are needed to control access to grades and other personal records, to charge for dining services or bookstore purchases, to permit entrance to libraries and sports facilities, and so on. In many cases, university libraries will be able to build on these capabilities to authenticate users and provide credentials acceptable to an access management system. In some cases, a

library may take a leading role in developing a campus-based authorization scheme.

As Donald Waters pointed out in his introduction, the CNI White Paper has identified three approaches to campus-based authentication and authorization that can interface with remote access management systems. The first approach is IP source filtering. The second is the provision of a gateway or proxy server to which each user must authenticate (typically using an ID and password) and through which all interactions with the remote system are transmitted. In the third approach, a campus-based authentication or authorization system issues credentials acceptable to the remote system. An important example of a credential is a digital certificate, having a data format compatible with Web-based security protocols and used for the distribution of secure information over the Internet according to a standard known as X.509. An acceptable access management framework must interface with all three mechanisms, since no one solution will be able to serve all campuses.

Issues Affecting User Acceptance

Where is the balance between two utopian visions?

Workshop participants observed that two contrasting utopian visions of the future of scholarly communication motivate developments in electronic publishing. For many publishers, the technology provides an opportunity to make more money by charging for every information use. For researchers, the technology holds the promise of free access to information for all. Librarians recognize that the information will not be free, but they seek to provide unfettered access to their users to the extent possible within existing budgets.

Over the last two or three years, interested parties have attempted to establish new rules of business for digital works with voluntary guidelines for fair use. These efforts, such as the Conference on Fair Use,[6] have foundered on the conflicting utopian visions of the parties involved. They have, nevertheless, served to educate all communities about the nature of the differences that divide them. Meanwhile, those engaged in private license negotiations and consortial arrangements have discovered the powers of the marketplace to forge workable solutions. As one participant observed, the longer the library and publishing communities are engaged in these activities, the more rational are the business models they adopt and the licenses they negotiate. The requirement for access management systems to allow for the ambiguity inherent in the related law also becomes clearer, as does the need for access management even when no fees are charged (for instance, to comply with the terms of a gift).

---

[6] *Report to the Commissioner on the Conclusion of the First Phase of the Conference on Fair Use*, U.S. Patents and Trademarks Office, September 1997. Available at http://www1.uspto.gov/web/offices/dcom/olia/confu/conclutoc.html.

Libraries are learning that they can respond to terms and conditions they see as unreasonable with the assistance of market forces. As they gain more experience in providing access to electronic resources, they are discovering that most users will accept less than the utopian ideal of free access to everything when they understand the underlying business model and find it reasonable. Publishers, too, are beginning to see that they are unlikely to earn more revenue from their traditional customer base for scholarly journals; furthermore, they recognize that libraries are making an honest attempt to comply with reasonable terms but cannot be expected to control or monitor their users' behavior closely. Nor can the lack of a robust means to enforce copyright still be blamed for holding back electronic publication of scholarly journals. If anything, their volume is increasing so rapidly that academic libraries are having difficulty absorbing them. Although license terms still vary from publisher to publisher, the agreements are growing more similar with experience.

Workshop participants argued that simple, liberal license agreements of the kind used by JSTOR should be the model for the future.[7] Such agreements are clearly less costly to negotiate and implement than others. Licensee institutions should be free to define their user community in the agreement and should take responsibility for authenticating their users (and providing credentials that certify roles, if necessary). Participants also urged publishers of specialized journals not to limit access to specific categories of academic users, as laid out in scenario 2. This practice was seen as a poor business model, since use by community members outside the specific group for whom access was purchased was unlikely to be significant. Prototype access management systems should operationalize simple agreements, to avoid raising expectations that unreasonable conditions could be enforced. Publishers who use simpler business models and offer reasonable license agreements, argued Vicky Reich of HighWire Press and Stanford University, are better able to expand into new markets.

At the same time, some participants strongly recommended that the academic community continue to press for free access for all to the scholarly literature. Such "blue-sky" talk, they said, would offset the commercial pressure for pay-per-view access and dampen the ability of publishers to implement access management systems based on that model. They believed other business models could be constructed to support the costs of managing and providing access to high-quality scholarly information.

What perspectives are needed?

Participants were instructed to consider the two scenarios from two perspectives: user and provider. All three groups, however, soon reached the conclusion that the discussions, particularly of scenario 2, must take into account the perspectives of three entities: publish-

[7] JSTOR Library License Agreement. Available at http://www.jstor.org/about/license.html.

ers, libraries as intermediaries or institutional users, and individual end users. To be successful, an access management system must be acceptable to both the end users and the intermediary library. The difficulty is, these parties differ greatly in their goals, economic motivations, legal responsibilities (particularly regarding liability), and in the different values or utilities they attribute to particular publications or works.

Licenses typically represent agreements made between libraries (or parent academic institutions or consortia) and publishers (or publishers' agents). Access management systems that operationalize those agreements, however, will limit what operations an end user can perform on a digital work. Systems must permit an end user to negotiate different terms and conditions for use of a work by establishing a different role through a separate or additional agreement. For example, enhanced terms might be based on an individual society membership or subscription, or on the acceptance of a charge. Some lawyers pointed out that for the end user to be able to exercise some of the privileges afforded by law or take advantage of ambiguities, he or she should be able to make an informed decision to ignore clauses in an agreement made between a library and a publisher. Agreements, suggested one participant, should incorporate formal loopholes permitting a wider range of operations from a special location or through an additional level of authorization.

The fact that all three discussion groups found it necessary to make a distinction between the end user and the institutional user acting as intermediary suggests that access management systems should be designed with such a distinction in mind. One group report included the observation that there will sometimes be a chain of obligation through several intermediaries from users through libraries to publishers (possibly through third-party aggregators) and eventually to authors.

Will there be slow evolution or a revolution?

For libraries, publishers, and the communities they serve, networked access to scholarly information is not a completely new business, but an extension of an existing portfolio of services in an existing economic structure, with staff and customers familiar with old practices. Although there are hopes for long-term efficiencies in replacing paper-based information products with electronic equivalents, and the transformation of the process of scholarly communication has begun, libraries and publishers must deal with both for many years to come. The continuing availability of well-managed, high-quality bodies of scholarly information will depend on professionals who must be rewarded for their efforts. The economic balance among authors, publishers, aggregators and other service providers, libraries, and users may adjust over time, but unless the adjustment is gradual, existing products and services are likely to suffer.

Although participants were not asked to consider the interests of authors, it was clearly assumed that access management systems

must be acceptable to authors who wish to disseminate their work to the scholarly and scientific communities. Simple business models reflected in effective access management systems would go a long way towards satisfying the needs of authors as well as other interested parties.

Academic users expect predictability and continuity. They expect the electronic environment to offer the functional equivalent of privileges that exist under current copyright law as applied to physical works, including but not limited to fair use. Conditions on the use of electronic versions of articles that create new impediments to research, to teaching practice, or to collaboration across disciplines or between faculty and students are cause for substantial complaint. The academic user expects to have access to information and be able to use it for scholarly purposes at a reasonable price, preferably, but not necessarily, zero. Charges for photocopying and photographic reproductions are common in academic libraries, which may also charge for or limit use of other services. However, users expect free access to the information traditionally found on the library shelves, such as journals to which the library subscribes.

Publishers and learned societies must find ways to reallocate resources and adjust their business model without destroying their short-term financial viability. When negotiating licenses for electronic versions of print publications, they clearly need to maintain revenue (or increase it to cover the new costs associated with the networked dissemination). Learned societies, in particular, are likely to have no cash reserves to invest in the hope of future cost savings; preserving cash flow is a matter of survival.

Libraries, too, are facing serious budget constraints. Many must make do with flat or shrinking budgets not only to maintain existing collections and services while struggling to keep up with prices for serials that are outstripping the rate of inflation, but also to meet the demand for new online services. They hesitate to accept pricing models that do not guarantee control over acquisition budgets.

Will economics govern acceptance?

The acceptance by libraries and end users of electronic publications and associated access management systems will be determined, at least in part, by economic factors. Libraries regularly look for cheaper ways to provide the same services or ways to provide enhanced services at costs they can justify or recover. Users will pay for services and academic administrations will increase budgets only if they expect to receive value for the expenditure.

Transaction costs associated with managing and providing access to scholarly information must be reasonable, whether incurred by users, libraries, or publishers. Arrangements between libraries for free interlibrary loan are common, in part to avoid costly accounting or payment procedures. Recent years have seen a growth in library consortia and third-party services that allow institutions to share the fixed costs associated with negotiating licenses and supporting co-

herent access to a variety of online resources. It is unlikely that information providers, whether publishers (as in scenario 2) or libraries (as in scenario 1) benefit by limiting access to a resource to subsets of users from an academic institution. The transaction costs of ensuring compliance with such limits will almost certainly exceed any loss of revenue sustained in granting more general access.

The pricing of institutional licenses for electronic resources remains a complex issue that requires further research, perhaps along the lines of the University of Michigan's PEAK project. Existing models are not wholly satisfactory. Pay-per-view is not acceptable as the standard pricing scheme for libraries acting as intermediaries. Similarly, it may make sense to base prices on a maximum number of simultaneous users when information is accessed via terminal sessions, but not when it is accessed via stateless Web interactions. For large and heterogeneous user populations (such as the entire population of a state), pricing by size of community makes no sense either. Alternative measures of volume are needed as a basis for subscription prices. But what metrics are appropriate and acceptable? Some participants suggested that fruitful analogies might be drawn from pricing schemes for network connections.

Simplicity pays

Perhaps the strongest message that emerged from this workshop was that whatever the system for managing access, it must be simple. It must be comprehensible and convenient for intermediaries and end users. The emphasis should be on finding ways to reasonably limit abuses and punish abusers rather than complicating life for every user. The system need not be designed to handle every special case but should be able to inform users of nonstandard provisions (such as the complex terms of a gift) without attempting to enforce them. Prototype systems should be developed to handle the majority of routine needs effectively. Publishers appear willing to tolerate a little leakage, if it does not turn into wholesale hemorrhage.

Complexity should be hidden from users, but those who want to know the full details of a complex deed of gift or the reason why access to an item is restricted should be able to find that information. Participants agreed that it is incumbent upon intermediaries (libraries and third-party aggregators) to negotiate simple licenses, with a view to making the management system simple to implement and to explain to users. Several argued that simple licenses benefit providers too, since they are less costly to negotiate and acceptable to a wider range of customers than are more complex licenses.

Systems that are straightforward to implement and easy to use will encourage compliance. Participants argued that is not absolutely necessary for systems designed to manage access to scholarly resources also to handle materials to which access must be limited for reasons of security. In the short term, the aim should be to build a system that operationalizes a few different, simple agreements. The design should be modular, flexible, and have the capacity for

growth. Extensions can be made later, on the basis of practical experience.

Convenors' Questions

What kinds of role distinctions are necessary?

Users of the resources in these scenarios may play many different roles, using the term *role* in a general sense. A faculty member may, for example, act as teacher, author or creator, researcher, consultant, or private individual. It would be impossible, participants argued, for an individual to declare that access to a particular article was being sought in conjunction with only one such role. Some expressed a fear that the mere technical ability to introduce and enforce distinctions among roles would lead to the adoption of practices that would discourage the general pursuit of knowledge and so would not be in the best of academic interests. One librarian recalled a case in which access to a licensed resource was permitted to faculty only during semesters in which they were teaching particular courses, regardless of whether the resource was useful to research or even to the preparation of the courses.

In the context of automated authorization and access management schemes, the term *role* has a related but more specific sense. It describes recorded characteristics of an individual user, such as membership in a group. Rules within the access management system determine whether a user with a particular role is able to access a resource and what operations he or she can perform on it. A user's role or roles might be established or negotiated in different ways, for example, through a campus-based proxy service or authorization scheme supported by a directory database, by membership in a professional society, or by acceptance of a charge to a credit card. Where institutional licensing of published journals is being considered, roles may be divided into those for which the institution can issue credentials and those that must be negotiated by the individual. Participants agreed that any access management scheme should allow an individual user to negotiate privileges beyond those afforded by institutional credentials or offered to the general public.

Much of the discussion in this area focused on the granularity (degree) of role distinctions required and perhaps transmitted through credentials or gateway services that an institution may provide for members of its community. Privacy, cost of implementation, and institutional requirements (associated with varying missions and policies) were seen as factors here. Some argued strongly that the granularity should be no finer than membership in a community as defined by the licensing institution, in other words, that all those affiliated with a university should have access to the same resources on the same terms. Finer distinctions by school or department within a university (such as those suggested in the second scenario) are like-

ly to inhibit cross-disciplinary research. Distinctions between faculty, undergraduates, and graduate students would cause problems for teaching. Others suggested that some distinctions might be necessary because of institutional policies relating to services for alumni, say, or, in the case of state universities, services for the general public. The consensus was that fine role distinctions should be avoided and that certification of any distinctions should be the responsibility of the user institution.

The technology, said participants, should allow libraries and publishers to make the business agreements they want, but both sides are more likely to benefit if the agreements do not rely on complex role distinctions. In the second scenario, licensed journals are only of interest to a subset of the community; in such a case, the licensor and licensee might avoid the transaction costs in enforcing special limitations on access by negotiating the subscription price on a different basis. One suggestion was to base the price on the size of the subset interested in the resource (though not limiting access to this group). Others stressed the value of developing a volume-based approach to pricing other than a pay-per-view model.

The purpose of use, observed two breakout groups, is often more relevant than any characteristic of the user. In the first scenario (public domain materials digitized by libraries), libraries would probably encourage any use for teaching or research but wish to control commercial re-use of digital reproductions made at substantial expense, in order to recover costs or fund future digitization projects. The privileges afforded by the fair use doctrine and exceptions granted in copyright laws are also primarily based on the nature (and effects) of use and not on characteristics of the users. On further reflection, participants concluded that requiring users to declare in advance how they intended to use materials was unrealistic and would be seen as an invasion of privacy.

What rights and duties are expected?

One issue raised by this question related to the use of the term *rights*. Under U.S. copyright law, observed Mary Levering of the U.S. Copyright Office, publishers and authors have *rights* in intellectual works but that users exercise *privileges* and duties. Furthermore, copyright owners and their agents generally manage rights in copyrighted works, whereas libraries generally manage access to those works.

As pointed out earlier (under the heading What perspectives are needed?), users and providers have different expectations. Rights, privileges, duties, and responsibilities are shaped not only by license agreements, but also by the overall legal, economic, and technical environment. They will be subject to change over time.

Legally, privileges and duties may be established through a chain of obligation from author or creator to publisher, to library (possibly via a consortium or third party aggregator) to end user. Not every link in this chain is associated with a formal agreement. In the first scenario, where unpublished materials may be involved,

there may be no way to follow the chain and establish unambiguously the rights associated with the original materials. Complex terms of gift may impose additional duties on the recipient library. After converting the material to digital form and becoming the provider of online access, the library may wish to assert rights in the digital reproductions in order to safeguard the potential for income or retain control over how the materials are used. Most users of converted archival materials would comply with reasonable terms, if it were easy to determine what the terms were. Automatic enforcement of all such terms is infeasible, since they often apply to subsequent use rather than to access or to specific operations that might be controlled by technical means. Both providers and users would benefit in this case from a mechanism that cautions users about special conditions and allows them to determine whether or how to proceed.

Academic users value their personal space highly. In the words of one librarian, users want the library to "make the connection and get out of the way." They expect to be allowed to exercise personal responsibility or, as one breakout group reported, to have the "right to do reasonable things and the responsibility not to do unreasonable things." They would expect any access management system to allow them access to all the information that they are entitled to have access to, inform them of their privileges and responsibilities, and explain how they can negotiate additional privileges. They expect patterns of use permitted for print publications to carry over into the electronic environment. They also expect that publishers will somehow guarantee that the content they are accessing has not been corrupted inadvertently or maliciously.

For their part, publishers hope to maintain revenue, whether to satisfy shareholders, subsidize other activities, or simply cover costs. To achieve this end, they expect to control distribution of works for which they hold rights. They expect that privileges given to users based on a reasonable business model can be implemented by technical means. To be acceptable to publishers, an access management scheme must be customizable to individual license agreements and flexible enough to incorporate new types of agreement and new technology for authentication and for delivery of content. Market forces will determine which technical barriers to access and usage protect revenue and which inhibit market expansion.

As intermediaries, libraries have the responsibility to negotiate reasonable agreements on behalf of their user communities and parent institutions. They cannot be responsible for the actions of end users, but they do have a duty to take reasonable efforts to inform users of terms and conditions for access and use and to ensure that institutional policies, as well as systems or data that support access controls are effective and valid. They will expect to understand how license agreements are encoded and enforced within an access management scheme, in order to fulfil these responsibilities. Libraries (or their parent institution or agent) must make reasonable assurances that proxy or gateway services exclude unauthorized users and that credentials offered for users are valid. In return, they will expect

publishers' access management schemes to honor the credentials provided and facilitate access through such proxies.

In the print environment, libraries have assumed the responsibility for archiving materials for posterity. Under section 108 of the U.S. Copyright Law, libraries and archives may reproduce materials in certain circumstances, for example, to replace "a copy or phonorecord that is damaged, deteriorating, lost, or stolen, if the library or archives has, after a reasonable effort, determined that an unused replacement cannot be obtained at a fair price."[8] In an electronic environment in which the publisher controls the master copy, after-the-fact preservation will be impossible. Archiving for preservation must be planned for in advance. Libraries, as custodial institutions, will expect license agreements, and access management schemes that implement them, to provide contingency provisions and fail-safe mechanisms that ensure the long-term accessibility of the information resource. The long-term archiving of information in digital form presents a formidable challenge. Information, concluded one breakout group, "will only be preserved if someone's job depends on preserving it." Although the archiving challenge was beyond the scope of the workshop, participants noted that a possible contribution to an eventual solution would be special access management provisions that allowed libraries or trusted agents to make archival copies.

What are the privacy issues?

Participants were unanimous in their view toward the privacy of individual users, an important issue in the discussions surrounding the development of the CNI White Paper: the metadata that establishes privileges, they argued, should be under the control of the licensing organization and closely guarded. Using the CNI's categories of identification (anonymous, pseudonymous, pseudonymous with demographics, and actual identities), they recommended that campus-based authentication services, gateways, or proxies should not relay actual identities to access management schemes run by publishers or aggregators. Anonymous access, they concluded, poses the least threat to privacy. Pseudonymous identifiers ensure accountability by allowing a publisher to identify abnormal volumes of use by one (unidentified) user and notify the licensing organization. The association of demographic information with pseudonymous identifiers should be limited; under no circumstances should it be detailed enough to identify an individual user. As librarians have found, some publishers request more details than they can usefully analyze. However, libraries require some tracking of demographics for acquisition decisions and resource allocation, while providers may need such information to adjust business models.

Participants stressed that no unnecessary information should be tracked by provider or licensing institution. Users should not be re-

8 Copyright Law of the United States, contained in Title 17 of the U.S. Code, Section 108: "Limitations on Exclusive Rights: Reproduction by Libraries and Archives." Available at http://lcweb.loc.gov/copyright/title17/1-108.html.

quired to indicate the purpose of use. In many states, library reader records are confidential, and the law prohibits libraries from tracking readers' behavior. The academic community, some participants argued, should lobby for more extensive legal protection for privacy, extending to transactions with publishers and bookstores. However, it is reasonable to allow users to reveal personal information voluntarily in order to secure additional privileges, if they are told how that information will be protected.

How strong must the security controls be?

The design of any access management scheme will balance the tightness of security against user inconvenience and even denial of access to valid users in some cases. The degree of security enforced should be commensurate with the provider's trust in the user community. Publishers, it was also pointed out, do recognize that libraries are basically honest and will try to comply with reasonable license agreements to the best of their ability. Existing arrangements suggest that they would honor credentials generated through campus-based authentication schemes. Where trust between libraries is concerned, as in the first scenario, libraries have already proved the benefits of mutual trust in many resource-sharing activities, such as interlibrary loan. Libraries will certainly trust each other's authorization procedures if technically compatible.

In neither of the scenarios examined by the workshop does the content call for very tight security. The limited market value of scholarly and archival information is unlikely to invite widespread abuse. Thus, in the case of a student dropout, say, it would not be essential for the system to be able to revoke privileges immediately. Other classes of information, however, such as current recreational literature or some reference materials, might require more robust controls because of the potential for publishers to lose revenue.

Legal experts reminded participants that no access management scheme exists in a vacuum and that the external environment must be taken into account. They recommended that access management systems emphasize the detection of inappropriate behavior rather than enforcement ahead of time, which is likely to prevent some valid use. Users, they added, need to know what their responsibilities are, and institutional policies need to include adequate sanctions for abusers and procedures for dealing with them. Abuse could be punished by revoking privileges within the system or within the external environment.

In considering how to balance accountability and privacy in the campus environment, participants found one technical approach that had emerged in the discussions relating to the CNI White Paper as promising. Campuses could issue short-term pseudonymous certificates to authenticated users. Certificates valid for a semester or a year could act as credentials for access to most information resources. For selected resources, certificates valid for a few hours might be more appropriate.

What kinds of accountability are necessary and what kinds of management data are needed?

Participants reiterated that libraries cannot, in practice, be accountable for the actions of users. Realistically, they can only make reasonable efforts to ensure compliance with license terms and the law. Any license agreement between a publisher or publisher's agent and a library will include some clauses relating to accountability of either party for complying with terms of the agreement. The JSTOR Library License Agreement, repeatedly cited as a model, stipulates that libraries must inform JSTOR if they are using a proxy server to control access, must exert reasonable efforts and cooperate with JSTOR in the implementation of security procedures, must work with JSTOR to inform users of the User Rules, and must notify JSTOR if the library becomes aware of violations. The license allows either JSTOR or the licensee organization to terminate access in the case of unauthorized use. To the extent that access to licensed resources is supported by technical means, some degree of accountability for the effectiveness of those technical controls is to be expected. As mentioned in the discussion on security controls, the group favored after-the-fact accountability rather than automated enforcement that might prevent valid access.

In conjunction with the discussion on privacy, participants observed that libraries, even when objecting to licenses that limit access to subsets of users, may still wish to collect usage statistics aggregated by demographic categories in order to make acquisition decisions and allocate resources. As noted earlier, some publishers ask for access to more demographic details than they fully use. No specific suggestions emerged as to an appropriate level of detail. In this instance, it is possible that both publishers and libraries would like to gather more detail for management purposes than is consistent with protection of the user's privacy.

How do we evaluate effectiveness of the system from user and provider perspectives?

According to participants, the basic test for a general access management scheme will be whether it is adopted in the marketplace. Its success will depend at least in part on quantity and breadth of use and its viability on whether the various parties receive appropriate value in the bargains they strike. Not surprisingly, no short-term or formal measures of effectiveness were discussed, since there is still much uncertainty about how best to evaluate digital libraries. No better criteria have emerged than precision and recall, which have served heretofore to evaluate information retrieval systems of much more limited scope.

Unanticipated Issues

Where do authors fit in?

As pointed out in the final plenary discussion session, the instructions for the breakout discussions omitted an important topic: the rights and perspectives of authors. In the end, it was decided that the relationship between authors, publishers, and users was too complex to bring into the discussion, and that a separate session would be needed to represent the perspective of authors. During the discussion, several points were made. Some participants argued that authors are both the ultimate information providers and, at least for scholarly journals, also the ultimate users. Faculty researchers must be educated to think twice about assigning all rights for articles to publishers, given the costs university libraries must bear in buying back the right to access the content and other barriers to broad access that publishers might wish to impose.

Other participants pointed out that the apparent equivalence between users and authors mentioned above is simplistic and demonstrates an American viewpoint rather than a global perspective. Legally, the rights of authors are very different from the privileges of users. In many other countries, authors retain moral rights even when they assign copyright to publishers; in some countries, they may not waive those rights. In Europe, authors have made their voices heard in objections to attempted agreements between libraries and publishers. Groups representing authors are working to use the technology to enforce their own rights, for instance by supporting the development of digital watermarking technology. In the United Kingdom, the Authors' Licensing and Collecting Society (ALCS) is adamantly opposed to the U.S. concept of fair use.

What about unaffiliated individuals and small institutions?

Scenario 1 highlights the fact that most academic libraries consider their services to unaffiliated individuals—the general public—an important component of their mission. For the Library of Congress and the New York Public Library, the provision of unfettered access to digitized collections is essential. Likewise, many state university libraries are required by law to provide service to unaffiliated users.

Increasingly, consortial and outreach activities undertaken by universities require academic libraries to provide services to smaller, less well-endowed institutions, such as K-12 schools, that do not have the technical infrastructure to provide authentication services. Institutions may wish to provide access to resources for which the library may not legally provide general access, but for which rights holders have granted permission for educational use. To facilitate access to such materials from schools and public libraries, the Library of Congress has considered establishing a site-license arrangement (at little or no charge). It lacks the technology and resources,

however, to implement such a scheme nationwide, because these institutions lack the necessary technical infrastructure. Even the simple application of IP source filtering is not feasible, since many small organizations do not have permanent IP addresses, but obtain them dynamically from Internet Service Providers each time they establish a connection. Clearly, state or local government agencies, consortia, or other third-party organizations must ensure that basic, uniform authentication and authorization services are available for small institutions and unaffiliated users.

Accommodating change

As pointed out earlier, the external legal, economic, and social framework in which access management schemes operate will change over time. Systems must be able to adapt to such changes if they are to succeed. The need for such flexibility may determine how best to represent rights or responsibilities within metadata associated with digital material and as rules implemented within access management systems. System designers should take nothing for granted: even status transitions that appear to be predictable—as in the case of expiration dates for copyrights held by corporate bodies or by creators who have died—are not certain, as current legislative recommendations to extend the period of copyright demonstrate (see appendix C). Furthermore, they must be prepared for changes of global dimensions: when photocopiers were introduced, copyright laws around the world were modified. Similarly, electronic publishing and network technology will disrupt the delicate balance between enabling creativity and inhibiting theft of intellectual property. As this effect is better understood, more changes to copyright laws are inevitable.

Two other notable trends in the external environment in recent years are creating further challenges for the design and deployment of access management systems. The growing phenomenon of distance learning in many universities leads to pressure for remote access to more library services in order to serve the expanded student body. Libraries are also establishing consortia to share the costs of licensing or of mounting electronic resources. For multicampus institutions, new centralized organizations may be formed to serve this role. For libraries without campus-based authentication systems, the consortia may provide authentication and proxy services. Access management schemes must be adaptable to a variety of third-party intermediate arrangements and changes in license agreements.

Widespread adoption of technology will undoubtedly stimulate further change. Access management systems must adapt to rights regimes around the world as global access to information becomes feasible from more countries. The scope of publications considered during tenure review is also likely to change. One participant urged the community represented in the room to persuade university presidents that the tenure process is possible without paying exorbitant sums to publishers. Whether or not they are persuaded, the growing

importance of the "grey" literature (such as electronic preprints) in some disciplines will undoubtedly affect the perspective of both users and scholarly publishers.

Accommodating ambiguity

The 1996 workshop on Technology of Terms and Conditions revealed the ambiguities and uncertainties inherent in copyright and related law. Systems developers were surprised to learn that laws often allow for flexible interpretations (in other words, are intentionally ambiguous), with the expectation that different interpretations will be tested in the courts and evaluated in the light of practical experience and other laws. Workshop participants devoted considerable attention to the ambiguities that might affect the design of access management systems, most of which have already been noted in this report but merit repeating.

License agreements can clarify some ambiguities, but not others. The rights pertaining to the historical materials Digital Library Federation members may propose to digitize for the Making of America project, for instance, may be impossible to ascertain. The moral rights of authors (in countries that recognize such rights) are not usually reflected in license agreements between libraries and publishers. Economic constraints may prevent publishers from establishing unambiguously the rights associated with illustrations and other subcomponents in old publications that are now being made available in electronic form. Like the participants in the Making of America project, they must find a way to assess and manage the risk. The original creators may or may not be interested in asserting any rights, depending on the age and nature of the works, the purpose of the use, and the current commercial viability of the works.

Another ambiguity relates to exceptions and limitations to intellectual property rights afforded by laws. Exceptions are usually based on the purpose of use and its effects on any market for the protected work. At the point at which access is controlled, users may not even know how they propose to use the work (beyond ascertaining whether it is of interest at all). In the view of workshop participants, it would be unreasonable and an invasion of privacy to require users to declare why they were accessing a work. A case in point might be researchers who are unwilling to explain their interest in a particular research topic when preparing a grant proposal or in existing patents in an area in which they have developed patentable technology. Users searching MEDLINE to research a serious medical condition may not wish to disclose that they or a family member have health problems. There is no way to enforce the fair use provisions of U.S. copyright law on the basis of characteristics of users. Each case is judged on its merits.

Conclusions

The workshop on access management held in Washington, D. C., on April 6, 1998, yielded several conclusions worth highlighting. It identified the need for research and evaluation of prototype projects in two key areas: *system usability and economic models*. The design of access management systems should be based on a better understanding of how users interact with such systems, what new information types will meet user needs, and what function these types perform in the emerging digital environment. To establish a viable economic balance for publishers, libraries and other intermediaries and users in the academic community, new standards of measure must be found to assess the usage of digital resources and thereby to develop alternative pricing schemes and payment mechanisms.

In addition, workshop participants identified five key properties for access management systems that would make them acceptable to users and libraries while respecting the rights and interests of authors and publishers. [9]

1. *Simplicity*. The less complex a system of access management, the more readily it can be adopted technologically and organizationally, and the more acceptable it is to all involved in its implementation.
2. *Privacy*. Systems that manage access to the cultural record must protect the privacy of users from detailed tracking and disclosure of use. User privacy must not be compromised.
3. *Good faith*. Agreements on access to scholarly information rely on trust among the parties involved. Users and providers would each prefer to depend, in an access management system that implements these agreements, on reasonable barriers against abuse rather than complex restrictions that inhibit use.
4. *Trusted intermediaries*. Intermediaries play an essential role in providing access to the cultural record as parties trusted by both users and providers and as efficient aggregators of distribution and usage. System design must take the role of intermediaries into account.
5. *Reasonable terms*. Access management systems and license agreements must recognize the distinction between access and use. Overly tight control of access to a resource may impose inappropriate constraints on its use, especially in teaching and research contexts. The most useful system will not limit access to specific user groups known in advance to be interested in a resource but will be reasonably open to serving unlikely users whose curiosity and research interests may lead them in directions not predicted by those responsible for making the agreements or designing the systems.

The findings of this workshop are relevant to a wide range of interested parties:

---

[9] Gerry Bernbom provided this useful summary of design properties in correspondence with Donald Waters, July 29, 1998.

- policy makers involved in making decisions on managing digital data in relation to questions of privacy;
- legal experts who draft contracts and licenses which must be implemented through technical mechanisms for authentication and authorization;
- technologists designing new software for controlling electronic use and mis-use; and
- publishers and librarians, who, as major providers of information, play a central role in striking a balance between protecting copyright and providing access to the cultural record of knowledge.

Although the workshop focused primarily on the means of managing access to published knowledge in digital form in the context of the research library, it also made clear the much larger dimensions of access management issues. With the enormous growth in digital records of every form, the issues of privacy, protection, authorization, and authentication are fast becoming a concern for all citizens.

Appendix A: Workshop Participants

Caroline Arms
NDLP Program Coordinator
National Digital Library Program
Library of Congress
101 Independence Avenue SE
Washington, DC 20540-9300
Phone: (202) 707-0105
Fax: (202) 707-0955
E-Mail: caar@loc.gov

William Y. Arms
Vice President
Corporation for National Research Initiatives
1895 Preston White Drive
Reston, Virginia 20191
Phone: (703) 620-8990
Fax: (703) 620-0913
E-mail: warms@cnri.reston.va.us

Ross Atkinson
Deputy University Librarian
Olin Library
Cornell University
Ithaca, NY 14853
Phone: (607) 255-3393
Fax: (607) 255-9346
E-mail: ra13@cornell.edu

Gerry Bernbom
Special Assistant for Digital Libraries and Distance Education
Office of the Vice President for Information Technology
Indiana University
2711 East Tenth Street
Bloomington, Indiana 47408
Phone: (812) 855-4624
Fax: (812) 855-3310
E-mail: bernbom@indiana.edu

Dennis Cromwell
Chief Scientist, Advanced Technology Lab
University Information Technology Systems
Indiana University
2711 East Tenth Street
Bloomington, Indiana 47408
Phone: (812) 855-7326
Fax: (812) 855-7868
E-mail: dcromwel@indiana.edu
URL: http://ezinfo.ucs.indiana.edu/~dcromwel/home.html

John Erickson, Ph.D.
VP-Rights Technologies
Yankee Rights Management
999 Maple Street
Contoocook, NH 03229
Phone: (802) 649-1847
Fax: (802) 649-2193
Email: jerickson@ybp.com
URL: http://www.ybp.com/yrm

Eric G. Ferrin
Director of Library Computer Services
8E Pattee Library
University Park, PA 16802
Phone: (814) 865-1818
Fax: (814) 863-3560
E-mail: egf@psu.edu

Les Gasser
Director, Computation & Social Systems
Division of Information and Intelligent Systems
National Science Foundation
4201 Wilson Boulevard, Room 1115
Arlington, VA 22230
Phone: (703) 306-1927
E-mail: lgasser@nsf.gov

Bernie Hurley
Chief Scientist
The UC Berkeley Library
Rm. 245, Doe Library
Berkeley, CA 94720-6000
Phone: (510) 642-3773
Fax: (510) 643-8179
E-mail: bernie@library.berkeley.edu

Peter Jaszi
Washington College of Law
American University
4801 Mass Avenue
Washington D.C.
Phone: (202) 274-4216
E-mail: pjaszi@american.edu

Andrea Keyhani
Manager, Publisher Relations
OCLC Online Computer Library Center
6565 Frantz Rd.Dublin, OH 43017
Phone: (614) 764-6474
Fax: (614) 764-1640
Email: keyhani@oclc.org

Dr. Judith L. Klavans
Director, Center for Research on Information Access
Research Scientist, Department of Computer Science
Columbia University
535 West 114th Street, MC 1103
New York, NY 10027
Phone: (212) 854-7443
Fax: (212) 854-9099
E-mail: klavans@cs.columbia.edu
URL: http://www.cs.columbia.edu/~klavans/home.html

Michael Lesk
Division Director
Division of Information and Intelligent Systems
National Science Foundation
4201 Wilson Boulevard, Room 1115
Arlington, VA 22230
Phone: (703) 306-1930
E-mail: mlesk@nsf.gov

Mary Berghaus Levering
Associate Register for National Copyright Programs
U.S. Copyright Office
Library of Congress
Washington, DC 20540-6007
Phone: (202) 707-8350
Fax: (202) 707-8366
E-mail: mlev@loc.gov

Melissa Smith Levine
Legal Advisor, National Digital Library Program
Library of Congress
Washington, D.C. 20540-1300
Phone: (202) 707-1783
Fax: (202) 707-0815
Email: mele@loc.gov

Wendy P. Lougee
Asst. Director, Digital Library Initiatives
University of Michigan Library
818 Hatcher South
Ann Arbor, MI 48109-1205
Phone: (313) 764-8016
Fax: (313) 763-5080
E-mail: wlougee@umich.edu
URL: http://www-personal.umich.edu/~wlougee/

Clifford Lynch
Executive Director
Coalition for Networked Information
21 Dupont Circle, Suite #800
Washington, DC 20036-1109
Phone: (202) 296-5098
E-mail: cliff@cni.org

Carol Mandel
Deputy University Librarian
510 Butler Library
Columbia University
535 West 114th Street
New York, NY 10027-7029
Phone: (212) 854-2226
Fax: (212) 854-9099
E-mail: mandel@columbia.edu

Charlene Mason
Associate University Librarian
499 Wilson Library
309 19th Avenue South
University of Minnesota Libraries
Minneapolis MN 55455
Phone: (612) 624-4520
Fax: (612) 626-9353
E-mail: c-maso@tc.umn.edu

David Millman
Manager, Research & Development
Academic Information Systems
Columbia University
612 West 115 St
New York, NY 10025
Phone: (212) 854-4284
Fax: (212) 662-6442
E-mail: dsm@columbia.edu

Ann Okerson
Associate University Librarian for Collections Development and Management
Yale University Library
New Haven, CT 06520-8240
Phone: (203) 432-1763
Fax: (203) 432-8527
Email: ann.okerson@yale.edu

Vicky Reich
Senior Librarian
Green Library
University Libraries and Information Resources
Stanford, CA 94305
Phone: (650) 725-1134
Fax: (650) 725-4902
E-mail: vreich@Sulmail.stanford.edu

Terry Ryan
Associate University Librarian for Systems
11334 University Research Library
University of California at Los Angeles
Box 951575
Los Angeles CA 90095-1575
Phone: (310) 825-1201
Email: tryan@library.ucla.edu

Sarah E. Sully
General Counsel, Director of Publisher Relations
JSTOR
188 Madison Avenue
New York, NY 10016
Phone: (212) 592-7345
Fax: (212) 592-7355
E-mail: ss@jstor.org
URL: http://www.jstor.org

Russell S. Vaught Director, Center for Academic Computing
229 Computer Building
Pennsylvania State University
University Park, PA 16802
Phone: (814) 863-0421
Fax: (814) 863-7049
Email: rsv@psu.edu
URL: http://www.personal.psu.edu/rsv

Donald Waters
Director, Digital Library Federation
Council on Library and Information Resources
205 Church Street, Third Floor
New Haven, CT 06510-1805
Phone: (203) 498-6076
Fax: (203) 498-6078
Email: dwaters@clir.org
URL: http://www.clir.org/diglib/dlfhomepage.htm

Appendix B: Suggested Readings

A report from a previous, related workshop

Davis, James R., and Judith L. Klavans. "Workshop Report: The Technology of Terms and Conditions." *D-Lib Magazine*, June 1997. Available from http://www.dlib.org/dlib/june97/06davis.html.

Core background materials for workshop on April 6

Arms, William Y. "Implementing Policies for Access Management." *D-Lib Magazine*, February 1998. Available from http://www.dlib.org/dlib/february98/arms/02arms.html.

Gladney, H. M., and J. B. Lotspiech. "Safeguarding Digital Library Contents and Users." *D-Lib Magazine*, May 1997. Available from http://www.dlib.org/dlib/may97/ibm/05gladney.html.

Stefik, Mark. "Trusted Systems." *Scientific American*, March 1997. Available from http://www.sciam.com/0397issue/0397stefik.html.

Wiseman, Norman. "Implementing a National Access Management System for Electronic Services: Technology Alone Is Not Enough." *D-Lib Magazine*, March 1998. Available from http://www.dlib.org/dlib/march98/wiseman/03wiseman.html.

Related background readings

Arms, William Y., Christophe Blanchi, and Edward A. Overly. "An Architecture for Information in Digital Libraries." *D-Lib Magazine*, February 1997. Available from http://www.dlib.org/dlib/february97/cnri/02arms1.html.

Bide, Mark. "In Search of the Unicorn: The Digital Object Identifier from a User Perspective." A report for the British National Bibliography Research Fund, November 1997. Available from http://www.britain.eu.net/~bic/bicinfo.html.

Cross-Industry Working Team. "Managing Access to Digital Information: An Approach Based on Digital Objects and Stated Operations." May 1997. Available from http://www.xiwt.org/documents/ManagAccess/ManagAccessTOC.html

Garrison, William V., and Gregory A. McClellan. "Authentication and Authorization, Part 2. Tao of Gateway: Providing Internet Access to Licensed Databases." *Library Hi Tech* vol. 15, no. 57-58 (1997): 39–54.

Kahn, Robert, and Robert Wilensky. "A Framework for Distributed Digital Object Services." May 1995. Available from http://www.cnri.reston.va.us/home/cstr/arch/k-w.html.

Lynch, Clifford A. “Authentication and Authorization, Part I. The Changing Role in a Networked Information Environment.” *Library Hi Tech*, vol. 15, no. 57-58, (1997): 30–38.

Machovec, George. “User Authentication and Authorization in a Networked Library Environment: Alliance Issues.” November 1997. Available from http://www.coalliance.org/reports/security.html.

Paskin, Norman. “Information Identifiers.” *Learned Publishing*, vol. 10, no. 2 (April 1997): 135–56. Available from http://www.elsevier.co.jp/inca/homepage/about/infoident/.

Riddle, Bob. “The ICAAP Project, Part Two: The Web Architecture.” *Library Hi Tech*, vol. 15, no. 57-58 (1997): 71–78.

Roscheisen, Martin, and Terry Winograd. “The Stanford FIRM Framework for Interoperable Rights Management.” Forum on Technology-Based Intellectual Property Management. Interactive Media Association, White House Economic Council and White House Office of Science and Technology. Washington D.C., 1997. Available from http://mjosa.stanford.edu/~roscheis/IMA/index.html.

Appendix C: Update on Related Legislative Activity

At the April 1996 workshop on access management, Peter Jaszi (Washington College of Law, American University) gave a summary of current legislative activity relating to intellectual property rights, drawing attention to some potential problems for libraries. March 1998 had seen considerable activity in Congress, under pressure from the commercial sector, particularly from the motion picture and sound industries. According to Jaszi, the leading legislative proposals made no significant distinctions for scholarly communication or academic use. His impression was that there was an inclination to pass bills within the forty legislative days then left in the current session of Congress.

The proposals dealt with three issues, all of which had been under legislative consideration for more than two years: copyright term extension [HR 2589], database protection [HR2652], andWIPO implementation [HR 2281, Coble]. All three proposals had been reported out of the House Subcommittee on Courts and Intellectual Property and the House Judiciary Committee for floor action. HR 2589 had been referred to the Senate Judiciary Committee. Jaszi indicated that the Senate Judiciary Committee was considering passing the bills on for floor action without full consideration within the committee.

Term extension

HR 2589 would add 20 years of copyright protection for works currently under copyright protection. The extension would not apply to works for which copyright has already expired. The new copyright term would be the life of the author plus 70 years. The proposal included no concessions for academic use. According to Jaszi, the exemption for libraries and archives in relation to making copies of out-of-print works, for purposes of preservation, scholarship, or research, during the last 20 years of protection was weak.

Database protection

HR 2652 (the Collections of Information Antipiracy Act) proposed a new right for databases that are simply compilations but require significant investment. Based on principles of unfair competition rather than copyright, this bill was partly in response to the decision in Feist Publications v. Rural Telephone Service Co. Inc, [499 U.S. 340, 18 USPQ2d 1275(1991), 41 PTJC 443, 453]. Feist denied copyright protection to compilations, unless there is value added through selection, coordination, or arrangement. As written, this new protection for databases would apply to any information that can be organized systematically, including facts, numerical data, and works of authorship. Penalties would apply to those who damaged the actual or potential market for a protected database by extracting data. These penalties would apply to end users as well as to commercial re-users, although an amendment provides partial protection for nonprofit organizations. The period of protection is 15 years.

In response to a question, Jaszi confirmed that there is no associated requirement for deposit (as there is, under mandatory deposit regulations, for most published works protected by copyright). Jaszi reported that the exemptions for scientific and academic use have been described as ineffective.

WIPO implementation

HR 2281 would provide reasonable prohibitions and penalties relating to tampering with copyright management information. In addition, it included prohibitions against circumvention of copyright protection schemes, whatever the motivation. Jaszi regarded it as significant that the bill did not reaffirm the principle of fair use, offered no exemptions for digital preservation by libraries or distance education, and provided no preemption of contract terms by constitutional privileges or federal law. Jaszi noted that a pair of companion bills suggested an alternative approach to conformance with the World Intellectual Property Organization treaty that addressed many of these concerns. These are Senator John Aschroft's (R-MO) Digital Copyright Clarification and Technology Act (S 1146) in the Senate and the Digital Era Copyright Enhancement Act (HR 3048), introduced by Reps. Rick Boucher (D-VA) and Tom Campbell (R-CA) in the House.

Jaszi noted that the committee actions on HR 2652 and HR 2281 were very recent. For updates, he recommended consulting the Web site of the Digital Future Coalition, which represents many library organizations, at http://www.dfc.org/.

Appendix D: Definitions

The following definitions have been compiled from those used by Waters and the other presenters at this workshop, by the DOI Rights Metadata Working Group, and by Clifford Lynch in the draft CNI White Paper.

Terms with an asterisk are used in figure 4, which was drawn up following the workshop as a result of discussions among Donald Waters, John Erickson, William Arms, and Caroline Arms. Figure 4 combines elements of diagrams presented by Waters and Erickson at this workshop (figures 2 and 3) and in William Arms's February 1998 article in *D-Lib Magazine*, "Implementing Policies for Access Management." The grey boxes represent components of an automated authorization system, whereas the white boxes represent the interactions between users and providers in the external environment or between either party and the access management system.

Access Management

Access management is a process mediated by information managers by which users gain authorized access to the intellectual property of creators / owners / providers. Access management systems make use of authentication and authorization services to enable or control access to and use of a networked resource.

**Agreements*

In figure 4, the term *agreements* refers to licenses and other legal agreements, entered into by or on behalf of users with information providers. Agreements may be made in different ways, for example, through formal contract or acceptance of terms at time of access.

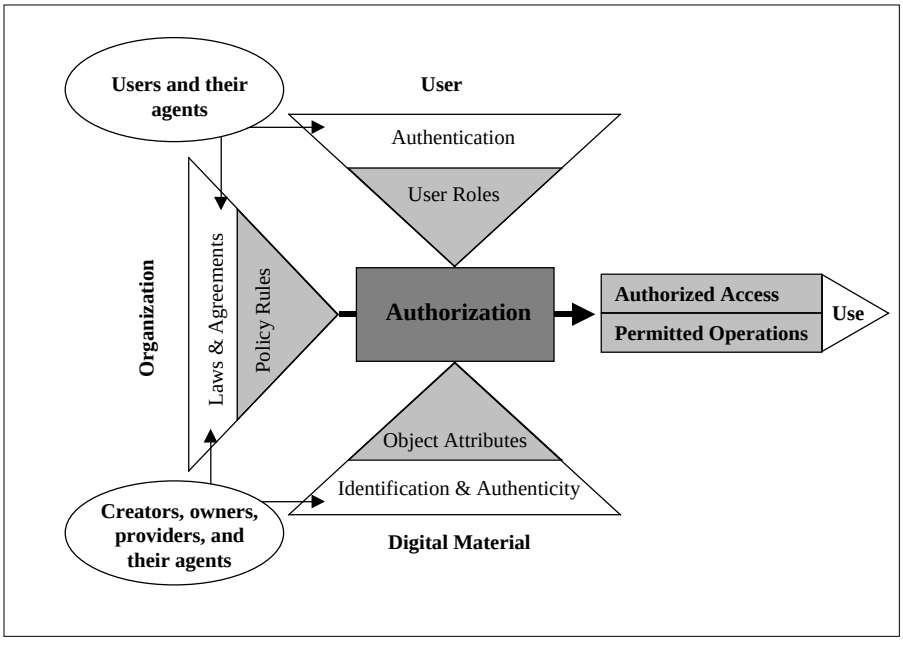


Figure 4. *Authorization in an Access Management System*

Within an automated access management system, the agreement and applicable laws will be implemented through a set of policy rules.

The DOI Rights Metadata Working Group uses the term both for a legal agreement between the parties and the corresponding representation of rules that control access within a system. For that group, "an Agreement is a statement of permitted operations, and applicable terms, for a given object or set of objects, and a given user or set of users. The default Agreement is that for 'all users', probably permitting a limited set of operations at standard fees. A site license that is already in place would be a specific Agreement."

*Authentication of users

Authentication is the process whereby a network user establishes a right to an identity or name (such as a user ID, or credit card number). A user can establish this right through

something he or she knows (user ID and password),
something he or she has (ID card), or
something he or she is (handprint or retina scan).

*Authenticity of content

The authenticity of digital content requires mechanisms to assure users that content has not been corrupted. Since digital content is easily manipulated in ways that cannot be easily detected, users will expect publishers and third party custodians of information to provide assurances of authenticity.

*Authorization

Authorization is the process of determining whether an identity (given a set of role attributes associated with that identity) is permitted to perform some action, such as accessing a resource. The identity may represent a particular individual or be anonymous or pseudonymous. In the context of access management, authorization will rely on access policy rules, role attributes of the user, and terms and conditions attributes of a digital object to determine whether the desired action is permitted and how to disseminate the material.

Credential

In the context of access management, a credential is something that a user can present to an authorization system operated by an information provider as evidence of legitimacy. One form of credential is a user ID together with password. Another is a digital certificate following the X.509 standard. Support for transmitting such certificates and using them to control access is being incorporated into the latest versions of Web browsers and Web servers. See the CNI White Paper for more discussion.

*Object attributes

If access to certain digital material is to be managed, metadata attributes must be associated with each digital content object to indicate what terms and conditions apply to that object. An attribute would typically indicate a class of material to which an object belongs, with a common set of terms and conditions applying to all material in that class.

Operation

In the context of access management, an operation is any act that can be done with an object, internal or external, to a given computer system. Examples: view, print, save; modify, redistribute. (This definition is from the DOI Rights Metadata Working Group.) The word *action* is sometimes used in roughly the same way.

*Policy rules

Within in an access management system, policy rules are encoded to determine whether a user is entitled to access the digital object being requested and whether the requested operation is permitted. Policy rules operationalize agreements, such as licenses, and applicable law.

Proxy

In the context of access management, a proxy is a special computer that acts as a gateway to one or more resources. The licensee organization (or its agent) typically deploys a proxy. The proxy relies on authentication services to establish the legitimacy of a user and then routes all traffic between that user and the licensed resource. See the CNI White Paper for more discussion.

*Role

In the context of access management, role is specified by attributes associated with a user's identity. Examples of role attributes are: membership in a university community, fulltime student, or individual subscriber to a scholarly journal. Authorization mechanisms use role attributes to determine whether a user is permitted to perform certain actions or operations on a resource or content object.

*Use

Use of a digital resource may extend beyond the operations performed online by a user when accessing a digital object. A faculty member who has printed a copy of an article may, for example, use it for personal reference or to make fifty copies for distribution to a class. Access management systems can limit access and have the potential to limit online operations; systems cannot fully control subsequent use.