WORKSHOP ON THE PROTECTION OF INTELLECTUAL PROPERTY RIGHTS IN A DIGITAL LIBRARY SYSTEM

May 18 - 19, 1989

KNOWBOTS IN THE REAL WORLD

Corporation for National Research Initiatives

Reston, Virginia

TABLE OF CONTENTS

Preface

In the spring of 1989, the Corporation for National Research Initiatives (CNRI), conducted a workshop at its offices in Reston, Virginia, to discuss the protection of intellectual property rights in the proposed Digital Library System(s) (DLS).[1] Representatives from a variety of organizations in the information industry and legal community participated. The cross-section of individuals included those involved with publishers, libraries, universities, and others in the public and private sectors. Potential infrastructure providers, information providers, and institutional users were drawn together to share views on this important topic. Each attendee brought to the table a different perspective on the development of DLSs. A complete list of the participants is given in the Appendix.

This report attempts to capture the major topics which arose during these two days of discussions. It focuses on the impact of intellectual property considerations on the design and development of a Digital Library System. The report does not attempt to repeat these discussions verbatim, nor even to provide a chronological summation of the proceedings. Rather, it highlights the major discussion topics and questions raised.

1 See "THE DIGITAL LIBRARY PROJECT, VOLUME 1: The World of Knowbots (DRAFT)," Corporation for National Research Initiatives, March 1988.

1. A DEFINITIONAL FRAMEWORK

CREATING NEW TERMS, REDEFINING OTHERS,

DISCARDING EXCESS "DEFINITIONAL BAGGAGE"

At the outset of the workshop on intellectual property protection in a Digital Library System, it became clear that a common definitional framework was necessary to ensure that the participants were talking about the same things. The fundamental problem is that technology, law, and customary use of many terms useful in describing a DLS -- its purpose and components -- already have attached to them so much definitional baggage that either new terms must be developed or new definitions for existing terms must be decided upon lest that baggage mistakenly be carried forward. Thus, the participants spent several hours defining the major terms used in conjunction with a discussion of a DLS. The participants noted that the terms and their definitions may have to evolve over time, and indeed, many terms evolved over the course of the two days as the terms were applied subsequently to other terms and concepts. Below is a list of the major terms and the definitions adopted by the participants to be used in the workshop. When used with the definitions stated in this section, these terms will be capitalized. When they are used in conjunction with other meanings, including common usage, the terms shall appear in lower case.

a. **Object**: A piece of information, be it text, picture, sound, etc. As presently conceived, the Object would contain not only the object itself, but also critical information (e.g. Import Data) concerning that object (see Fig. 1).

b. **Creation:** The causing of a new object to appear in the "private space" of an author (either outside a DLS or in an area within a DLS to which only the author(s) has access).

c. **Rightsholder**: An owner of rights in the contents of an Object. The owner may be an author, his or her successor in title, or others who have rights or interests in the contents.

d. **Import**: The action by which a DLS is made aware of the existence and substance of an Object. Access may still be restricted, but at least a DLS knows that the Object exists. An Importer is the person or entity that makes an Object available for Import into the DLS.

e. **Annuli**: Conceptual partitions representing access levels to information within a DLS. As an information object traverses boundaries between annuli outward from the "private space" of a user, this represents increasing degrees of Access to the information object (see Fig. 2).

f. **Release**: The moving of an Object across an annulus boundary, thus increasing its availability. Terms discussed but rejected for this verb were: "create," "produce," "publish," "transition," "disseminate," "share," "transmit." A Release implies the application (or embedding) of Conditions to the Release.

g. **Access**: The range of uses which can be made of an Object, which are subject to Conditions. The larger the range of uses, the greater the "access level."

h. **Conditions**: The set of constraints placed upon a user's Access to Objects. These may be set at the time of Import by the Rightsholder or by the DLS itself.

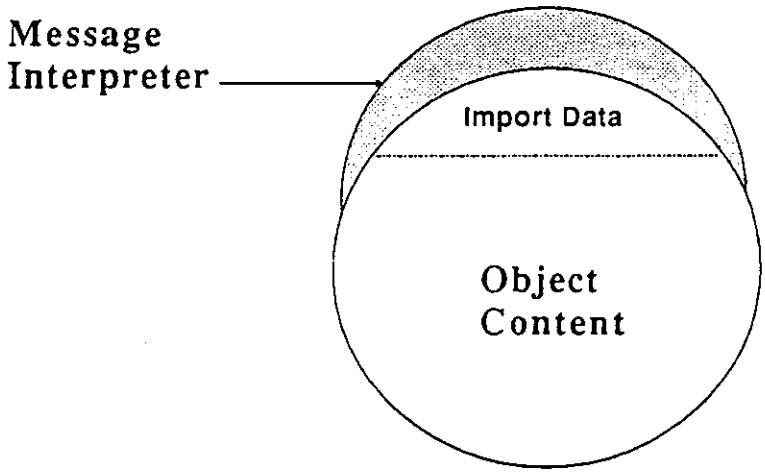
i. **Knowledge Robot (or Knowbot** ^{tm}**)**: An active intelligent program capable of exchanging Messages with other Knowbots and moving from one system to another carrying out the wishes of the user. Objects are manipulated in the system by Knowbots.

j. **Messages**: Information exchanged among Objects and/or Knowbots in a DLS. Examples include Object contents, audit or accounting information and search algorithms. Messages are interpreted upon receipt by Interpreters associated with Objects.

In addition, in discussing the actual workings of a Knowbot, as dictated by the computer software, two "slang" terms surfaced, and become part of the workshop vernacular.

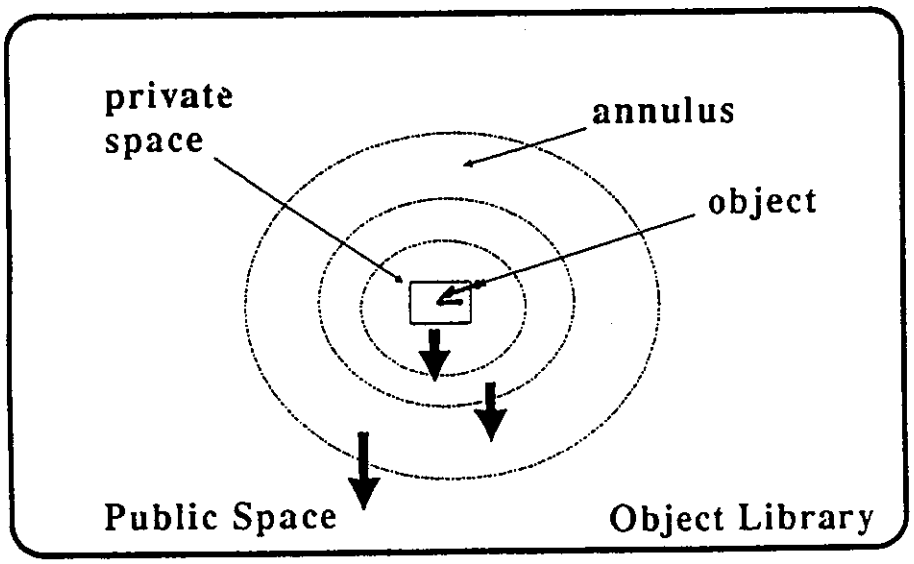
k. **Buttons**: Buttons represent functions or operations on an Object which are accessible to a given user. Conditions determine under what circumstances the buttons are applicable.

l. **Spigots**: Upon manipulation of an Object by a Knowbot, certain Messages will be passed from the Object to other parts of a DLS, including audit trails to the accounting functions of a DLS. These are emitted from the Spigots.



Object Model

Figure 1



The Annulus Concept

Figure 2

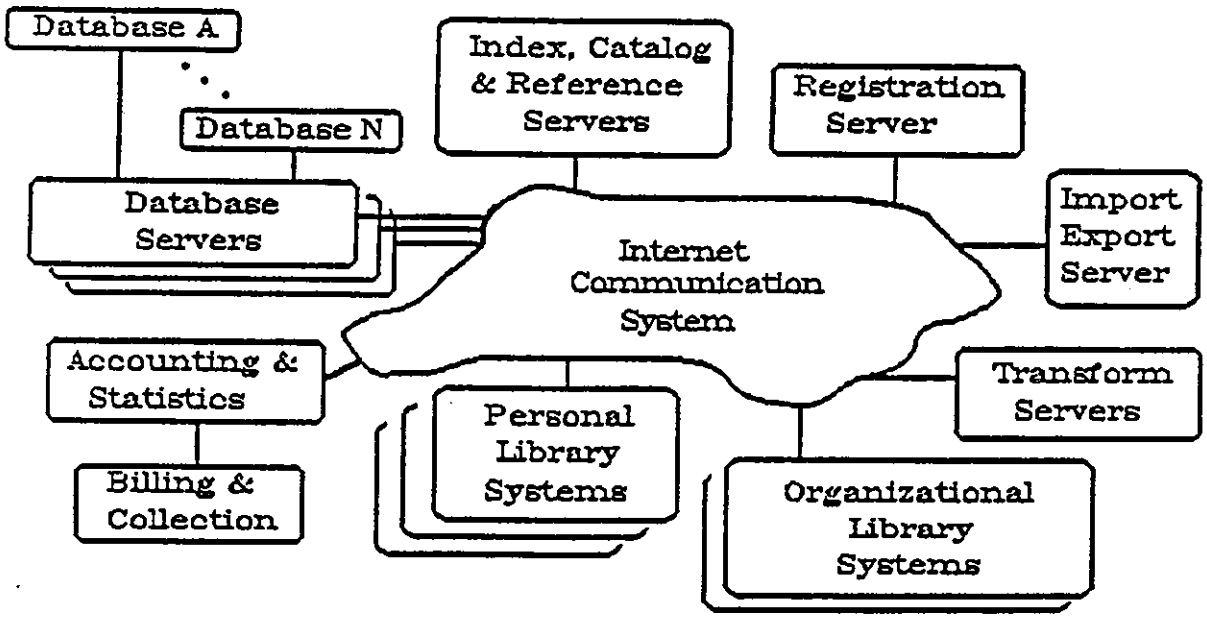
2. OPERATION OF A DIGITAL LIBRARY SYSTEM

A GUIDE TO ITS ARCHITECTURE AND FUNCTIONS

Conceptual understanding of a Digital Library System and its operation is necessary before one can discuss how existing legal regimes would apply to it. A schematic description of a Digital Library System is shown in Figure 3. Its components are 1) personal library systems for the users, 2) organizational library systems for serving groups of individuals or activities, 3) new as well as existing local or distant databases, 4) database servers to handle remote requests, and 5) a variety of system functions to coordinate and manage the entry and retrieval of data. The system components are assumed to be linked by means of one or more interconnected computer networks.

Local requests for information, if not satisfiable by the local personal library, are dispatched to other, larger or possibly more specialized sources of information available through the network. A single inquiry may spawn tens to thousands of exchanges among various parts of the full DLS. This could easily happen if the system must first query several databases before responding to a particular inquiry.

A Knowbot is typically constructed on behalf of a user at his personal library system and dispatched to a database server which interfaces the network to a particular database or set of databases. To accommodate existing database systems which are not capable of direct interaction with Knowbots, these servers can assist Knowbots in translating their information requests into terms which are compatible with the existing database's access methods. In the future, it is expected that database systems will be developed with built-in mechanisms for housing and catering to resident or transient Knowbots. It is possible, and even likely, that more than one Knowbot may be dispatched either directly from a personal library system or indirectly as a result of actions taken at a particular database server. These various Knowbots may rendezvous at a common server or all return to the originating workstation for assembly of the results.



The Digital Library Architecture

Figure 3

2.1 Import/Registration Function

In order for information to become part of a DLS, it must come with certain information which can be operated on by a DLS and appended to the Object it becomes. Minimal information which must be associated with the Object should include:

a. Name of Object
b. Submission Format (and format validity)
c. Date of Import (and possibly date of creation)
d. Name of Importer
e. Identity of Rightsholders (including author(s) if known)
f. Intellectual Property Rights in Object Contents
g. Conditions for Use
h. Certificate of Authenticity (if obtainable)

Some of this information must be provided by the Rightsholder(s), and some of it may be generated by a DLS (such as date of Import). The information above, combined with the content of the Object itself, can be used to produce a "Certificate of Integrity." While the DLS would normally protect the original Object against modification or destruction (except perhaps by a Rightsholder), technical means are available to detect such attempts. If any of the information contained in the body of the Object, or associated with the Object is ever changed, the Certificate of Integrity would automatically be violated. Such a violation is detectable and a Message can be passed to an appropriate part of a DLS to indicate that the Object has been manipulated in such a way that the original Object is no longer exactly the same as when it was Imported into a DLS (of course, a Rightsholder could always provide an updated version for Import into the DLS at a later time).

2.1.1 Liability for misregistration, misattribution, data input errors, etc.

Who would assume, or would be assigned, the liability for errors and misuse in the Importing and Access functions? DLS operators would obviously desire to minimize their liability, but Rightsholders also need a certain "confidence level" that their works will not be misused and that the integrity of their works will be maintained. Insurers may play a critical role here.

2.1.2 Avoiding value judgments when Importing Objects

Should all material be allowed to be Imported into a DLS, presuming the Importer has the necessary property rights to Import? Should a DLS refuse to perform Import functions of undesirable materials? What level of undesirability? Indecent material? Obscene material (defined under what standard)? Two countervailing forces were discussed: The desire to have only "quality" material in a DLS so that users would not have to wade through low utility material (however that is defined), versus the relative inability of a DLS to make such value determinations. What is of low value to most may nonetheless be very useful to others. There was a consensus that a DLS should not make such value judgments at the time of Import. Left unresolved was the issue of continued storage within a DLS of low utility material. The problem was temporarily assumed away by positing that storage capacity should not be a problem and that a DLS should be able to keep such material on file at little cost.

2.1.3 How will traditional publishers continue their role of quality control?

The participants agreed that in today's print media, publishers of text material, especially journals, play a critical role in quality control. Only those articles which meet the requirements of the publishers, editors or referees of a particular journal are actually printed. This quality control function is important to many of today's readers. Whereas a doctor might not normally read unsolicited papers sent to him in the mail, he probably would at least browse the New England Journal of Medicine each month, because he has come to expect that that journal will carry one or more articles of interest.

The question arises in a DLS environment, however, as to whether traditional publishers will continue to play that same role or whether others would undertake it. It was assumed that, in the near future, publishers would probably make electronic replicas of their paper journals available in a DLS. Looking to the future, however, it is difficult to predict whether the paper version and electronic version of journals will remain similar, or whether, for a variety of reasons authors may choose to bypass traditional paper journals altogether, and merely Import their articles into a DLS without the "stamp of approval" of journal publishers.

Consensus was reached that the quality control functions traditionally played by hard-copy publishers will continue to be valuable. A DLS, to the extent possible, should encourage continuation of this editorial function by working closely with publishers and encouraging publishers to accept the role of Importer. Such an infrastructure also mitigates against the operator of a DLS having to make value judgments.

2.2 Exporting Objects from a DLS

The ability to export material from a DLS presents a number of significant technical and legal problems. As a whole, publishers and authors may readily allow their works to be viewed in electronic form on a DLS by users for a fee, but most wish to impose conditions on the ability to reproduce or redistribute the work. From a technical standpoint, however, it may be difficult to keep a user who has viewed a document from copying it. If the document comes across a computer screen in its entirety in standard text format (e.g. ASCII), then present technology allows a user to save the text to a file as it appeared on the user's screen, thus creating a copy of the work.

The ability of a DLS to control selectively such copying by technical, legal, or other means is a critical issue in DLS development. Without a minimum degree of confidence that a DLS can enforce conditions on the actual export of Objects from a DLS, publishers and authors may be hesitant to place valuable material in a DLS.

2.3 Purging Documents from a DLS

The discussion on exporting documents from a DLS evolved into a longer discussion on the issue of whether an "owner" of intellectual property rights in an Object should have the ability to purge the Object from a DLS. Definitions of "purge" include the ability to deny future Access to the information ("de-release" it from the public Annulus), or wiping out all reference to the Object's existence (this could create substantial problems, as references to the Object could remain in a DLS, and Knowbots may contain pointers to the Object which lead nowhere). The traditional analog is the ability of an author to stop the printing of additional copies of a work, with the concomitant inability of the author to physically remove each existing copy from the shelves of all public libraries of the world and from individual homes and offices. Full purge

capabilities in a DLS, however, might be akin to destroying all known copies of a book, and deleting all references to the book in library card catalogs.

Scenarios were developed showing the benefits of purge, and showing its absolute need in certain instances. In cases such as incorrect medical information, the ability to correct the error, and thus purge the original Object from a DLS, may be critical. If not abused, the ability to correct errors in news stories was also cited as a beneficial application of the concept of purge (although a further analog would be the ability of some governments to "correct errors" in the news in such a way as to distort history).

It was pointed out in discussions that there may be economic reasons for restricting Import or discarding Imported material. Thus, we cannot and should not rule out the purging of some material for reasons having nothing to do with accuracy, only economic value. There was no real consensus on what should be done about purging. The participants agreed that this issue requires additional thought and discussion.

2.4 Catalog Index

Participants discussed the degree to which value added functions such as cataloging and indexing should be built into an operational DLS, and how much should be left to third-party development as a separate business. Obviously, some cataloging functions must be built into a DLS to allow it to function at all. After Import, an Object must be cataloged in some fashion in order for a DLS to access and otherwise act upon it. There appears ample room for third-party indexing, nevertheless, through either static compilations of particularized material (e.g. an index of a particular database, or an index of all materials in a DLS dealing with a single subject), or the development of search algorithms which can be tailored to individual needs (e.g., natural language search-and-compare software routines).

2.5 Accounting Functions

Participants first focused on the extent to which audit information should be collected. Privacy problems arise if each manipulation of an Object by a Knowbot produces an audit entry which is passed to accounting. To the extent the information is collected, it was assumed that it could be accessed by government or possibly others, no matter how

it is "protected." Keeping such an audit trail, however, would certainly aid in copyright infringement cases where access to infringed material is at issue. Consensus was that privacy should be a paramount concern for the DLS, even if it would entail omitting otherwise advantageous data collection mechanisms, which might be helpful but are subject to abuse.

Two completely separate accounting trails were suggested. One would provide information on what objects were accessed when (but not by whom); the other would track the number and billing class of Objects Accessed by a particular user, but not which Objects were Accessed. The two trails could not be combined so that the identity of both the Object and the person Accessing it could be obtained. Such an audit system would appear to provide sufficient data for accounting needs for establishing charges, but would also protect the privacy of individuals.

3. CONDITIONS AND ENFORCEMENT

"BUTTONS AND SPIGOTS"

Attached to every Object entering a DLS must be a set of Conditions for Access (including a "no condition" option) in order for a DLS to be able to make use of the Object. These conditions will generally be set by the Rightsholders. The various uses which may be made of an Object must be delimited by specifying and attaching relatively specific Conditions to the Object. "Rules" adopted for the use of an Object may be enforced in the following three ways:

a. Technical enforcement (real-time system limitations on use imposed by a DLS);
b. Legal enforcement (post hoc challenges of uses);
c. Societal enforcement (including public education on norms for use of Objects).

3.1 Why Do We Need Conditions On Objects?

To understand the need for Conditions associated with operations on Objects, and to facilitate their uniform development, it is necessary to understand what a typical user might want to do with Objects in a DLS.

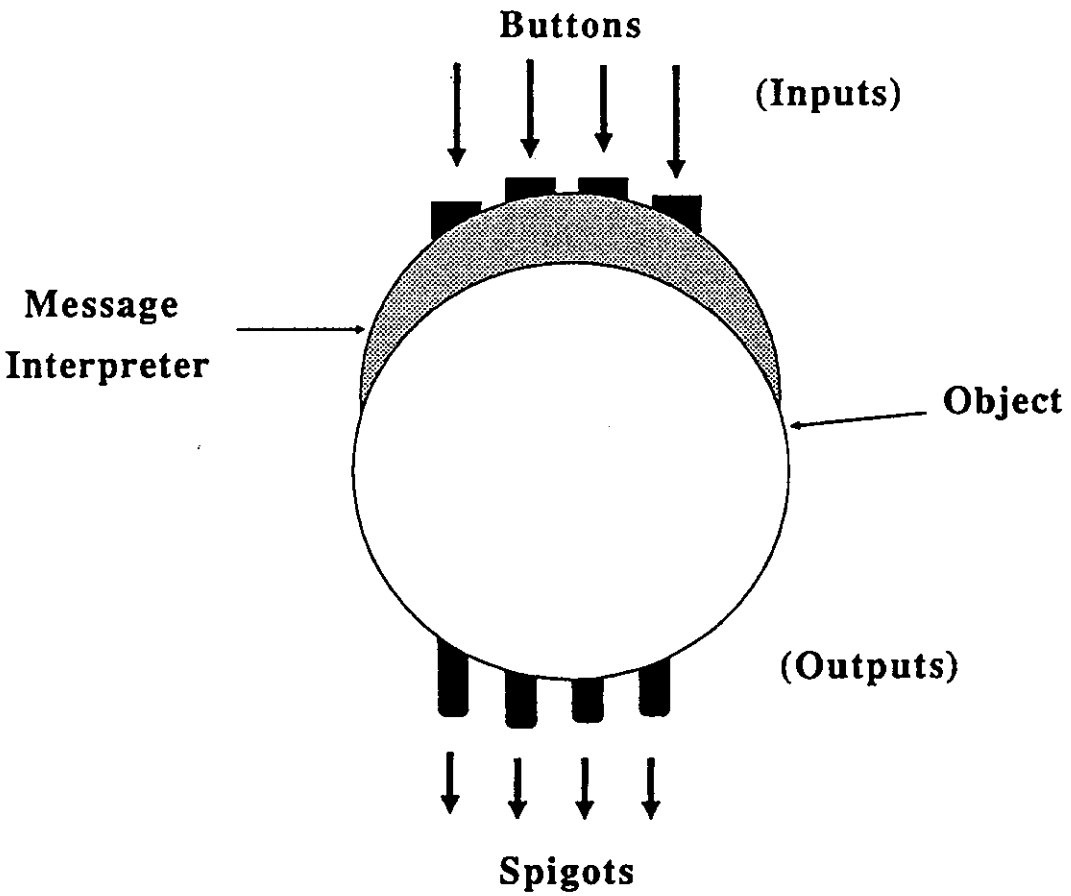
Participants divided the operations into roughly three categories, based on the likely need to enforce conditions on their application to an Object. These are *basic operations, special operations and prohibited operations*. Basic operations generally would be allowed on all Objects within a DLS, prohibited operations, with a few well-defined exceptions, would not be allowed on any Objects in a DLS, and Rightsholders would decide under what conditions and at what price any or all of the special operations would be allowed. Once these desired operations are defined, they can be implemented in such a way that conditions on their use can be enforced. A partial list of "user desires" includes:

Basic Operations	1. Search Object 2. Read or Display Objects 3. Extract part 4. Cite to Object
Special Operations	5. Copy All 6. Communicate with Rightsholder 7. Manipulate Objects (become an author of a derivative work) 8. Resell Original
Prohibited Operations	9. Modify Original 10. Destroy Original 11. Change Conditions

Graphically, operations can be thought of as "Buttons" on an Object that can be enabled or disabled, depending on the conditions set by the Rightsholder. "Spigots" are the output ports by which messages are emitted from Objects. Figure 4 illustrates this notion.

In order to feel comfortable providing Objects for a DLS, Rightsholders will want a certain "confidence level" that issues of compensation and protection will be properly addressed by a DLS. In particular, the association of Conditions with the application of operations on Objects through the three classes of enforcement mechanisms listed above should allow the following to be accomplished:

a. Compensation for viewing/downloading/using;
b. Protection against modification and/or destruction of original Objects (Integrity);
c. Provision for creation of derivative works while protecting against unauthorized repackaging of Objects (e.g. plagiarism);
d. Protection against aggregate replication;
e. Attribution (credit);
f. Level of Access/Release (no traversal of an Annulus unless authorized);
g. Protection from unfair competition (ability to detect and prevent misrepresentation of Objects as one's own);
h. Protection against reverse engineering (e.g., preclude decompilers from being applied to computer software object code to produce readable, modifiable source code);



Buttons and Spigots

Figure 4

i. Protection against destruction;
j. Protection against unauthorized alteration of conditions.

3.2 Some Case Examples: How Will Conditions Operate in Practice?

The participants attempted to apply the concepts mentioned above to some "easy" case scenarios. It quickly became apparent, however, that few cases were in fact easy. The group identified a number of example Objects to which the Condition set listed above should be tested. These example types included:

a. A Ph.D. thesis;
b. Public working paper;
c. Government Census (available only electronically in 1990);
d. VLSI chip designs;
e. Works of Shakespeare;
f. E-mail;
g. Natural Language search algorithms;
h. The Electric Cadaver (Hypertext/Hypervideo);
i. Film/Video Library.

Quickly, it became apparent that there are countless possible Conditions which could be made to apply to the works listed above, especially where the Object is not a traditional text-based work.

In the discussion of the treatment of a Ph.D. thesis, a number of issues and observations were made. It was thought that the author would be most interested in protecting the integrity of his work (i.e., would not want it modified by a third party) and making it easily citable. Monetary compensation was not thought to be a motive for placing a doctoral thesis in a Digital Library System, unless the thesis were supplied by a third party expecting to receive revenue for supplying it (e.g., University Microfilms). It was also unclear who would set the price and who would be paid if charges were levied for use.

Obtaining permission to make various uses of (apply various operations on) an Object was raised as a possible problem since the author typically retains copyright and may be difficult to track down. Of course, this is neither easier nor more difficult than the

situation which exists today. The case of a Ph.D. candidate who wants to publish his thesis for profit as a book was raised as a counter-example to the more typical case of registration with University Microfilms.

Participants agreed that some degree of flexibility in defining and applying Conditions must be incorporated into the Import/Registration portion of a DLS operation. Otherwise, at Import, a Rightsholder may be dissuaded from completing the Import if the desired use constraints are not part of the "Condition Menu" displayed. This will become more important as non-text Objects such as films, video, and sound recordings, and other multi-media, multi-work Objects[2] (or multi-Object works!) begin to appear in a DLS. A preliminary conclusion was reached that DLS programmers probably will have to create a "Condition and Operation description language" to allow Rightsholders to define operations and declare conditions of use for individual Objects.

2 As an example, a theatrical movie may contain many separately protectible works, including the video images, the script, the musical score, the musical recording, and the package as a totality. The Import function should allow each Rightsholder to define Conditions which will apply to the use of the separable parts of the work.

4. BUSINESS STRUCTURES

REACHING CRITICAL MASS WITH MAXIMUM FLEXIBILITY

Assuming that the technical barriers to creating a DLS can be overcome to give Rightsholders the necessary "confidence level" that continued integrity of their works will be maintained (i.e. that the originals cannot be modified or destroyed and that unauthorized copying can be constrained), the next step is to determine the appropriate business structure for a DLS. What economic incentives will publishers of traditional print material have to introduce it into a DLS? How will Rightsholders be paid for providing access to their works? How will end-users be billed for using a DLS? Can a DLS operate with multiple royalty mechanisms or with multiple pricing mechanisms or with both? What business form will a DLS operator take, and how involved will it be in negotiating with Rightsholders and end-users to establish pricing mechanisms?

The DLS is designed to augment the traditional intellectual property protection mechanisms with a technological assist to enforce restrictions that Rightsholders may wish to impose on others, either as access rights or as actions related to use. The DLS model assumes that, when entering an Object into the system, a Rightsholder can impose restrictions (e.g., by restricting the annuli to which it can be distributed and then by further imposing conditions on what a user can do with it) and that the DLS will enforce them. While this approach may be appropriate for publishers who can use it to charge for uses that limitations of the print media would not permit in practice, it puts control of intellectual property usage in the hands of the Rightsholders. The ensuing discussion addressed, without resolution, how public policy concerns could be injected into the process.

It was also noted that there already exist many on-line services which distribute information from a multitude of sources. Some of them do not differentiate between display and downloading of information. A DLS which provides the capability to

distinguish between display and downloading for reuse will have to accommodate these existing systems.

The issues above generated much debate among the participants, who generally agreed that the numerous business and "clearance of rights" issues were probably the most important to be resolved before a DLS can become a reality. Without a "critical mass" of Rightsholders who are satisfied that they will receive appropriate compensation for their works, and a "critical mass" of users willing to pay for Access to those works, a DLS will remain a laboratory experiment or at best a narrowly based capability only marginally evolved from that available through today's computer networks.

4.1. The Transactional Framework: Clearance of Rights, Billing of Users, and Distribution of Royalties/Payments to Rightsholders

Prior to the actual Import of a work into a DLS, some framework for making payments to Rightsholders in exchange for their agreement to make their works electronically available (the clearance of rights in the works) must be developed. How a DLS will accomplish this in an efficient and satisfactory manner to all parties provided fertile ground for discussion.

An efficient and economic billing structure to end-users also must be developed. Most notably, participants from different sectors of the information industry espoused markedly different views of how rights should be cleared and how billing and payments procedures should be implemented. From the discussion emerged two types of pricing structures: Block Payments and Per Access charging.

4.1.1 "Block" Payments (or Site Licenses)

Under such a plan, end-users would be charged a flat fee for use of a DLS, no matter what materials would be accessed or the amount of computer resources used within it. This fee might vary among types of users (e.g., academic, corporate, individuals), among DLSs, between DLSs and Rightsholders and with the size of the user group. Those from the academic community felt that a "block" fee end-user payment was most appropriate so as to encourage use by students and researchers and provide long-term budgetary tools.

Participants from the publishing community generally endorsed this scheme, noting that more specialized journals similarly would not prefer a "per Access" fee structure, as the resulting distribution (royalty payment) would not support continued availability on a DLS or possibly even print publication of the journal. There was a concern that a "per Access" charge could inject disequilibrium into the academic print publishing industry (it was noted that the market is already skewed, but no one wished to tamper with its workings).

4.1.2 "Per Access" Charges

The other transactional framework discussed was the "per Access" scenario. Here, a user would pay for each Object Accessed and function performed (such as searching, displaying, copying, or excerpting) and royalty payments would be made to Rightsholders on this same basis. Unique items such as patents, computer chip designs, and certain time-sensitive items (such as financial information) could be economically Imported into a DLS under this pricing scheme. "Per Access" charges could also include the pricing structures of many current on-line databases such as Lexis/Nexis, Westlaw, GENIE, Dow-Jones, Dialog and Compuserve, which charge users by the hour for access to all contents within the database.

4.1.3 Impact of Pricing Models on Traditional Avenues of Information Dissemination

A major concern to the participants was that the traditional market pricing mechanisms for information dissemination not be disrupted by the introduction of DLS services so as to result in a net loss of creation of new literary and other works. One specific example brought this issue into sharp focus. Assume publisher A produces two journals. Journal 1 is quite popular, has a large print subscriber base, and subscription price of $100. Journal 2, although of equal quality as Journal 1, is a "niche journal" which caters to a very small audience. It is not widely read. The subscription price of Journal 2 is set at $1000 so that the publisher can recoup at least most of its publishing costs. Now in a DLS environment, Publisher A makes both Journals available. Under a "per Access" pricing scheme, Publisher A will receive substantial revenues from Journal 1, but very little from Journal 2, for any reasonable price, since few people will Access it. Unless Publisher A chooses to subsidize Journal 2 from the receipts of Journal 1, Journal 2

may not participate, resulting in less material available in a DLS or even in a net loss of new creation of works, if Journal 2 cannot survive outside a DLS.

Under a "block payment" scheme, however, Publisher A would likely receive more for Journal 2 than under a "per Access" scheme, since royalty payments need not be tied directly to the number of users who Access Journal 2 and might be related to subscription prices. The net benefit of this scheme is that Journal 2 continues in existence, and its presence in a DLS helps compile the "critical mass" of materials necessary to attract both users and additional Importers. Additionally, Journal 2 may attract specific users to a particular DLS simply because of the existence of Journal 2 within it.

The consensus was that a DLS pricing model should be compatible as much as possible with the cost model. Most participants agreed that a DLS is characterized by high fixed costs, although a number of participants noted that communications fees and/or CPU demands to process Knowbots could result in high per Access costs ("the finger that launched a thousand Knowbots"). The DLS may need to implement ceilings on charges per transaction to prohibit "runaway searches".

All parties seemed to agree that flexibility in the cost/distribution model is desirable, thus making individual negotiations and contracts critical to the development of a practical DLS. The discussion did not resolve where existing on-line database providers would fit within this transactional framework. Further left unresolved was whether the two royalty payment models are compatible with multiple end-user pricing mechanisms. For example, does a DLS limit access of "block" users to items for which payments must be made to owners on a "per Access" basis? Or can a DLS calculate with a degree of confidence the usage behavior of "block" users so that the cost of their access to "per Access" materials is built into the "block" fee? In light of the success of multiple pricing mechanisms in the information industry of today, participants were confident that a transactional framework for a DLS is possible which combines block access and per access charging.

4.1.4 The Public Library Gateway Problem.

The question of public libraries and "information have-nots" presented a major problem in conceptualizing a transactional framework. There certainly will be a major effort to make DLS services available to as many people as possible, and those tradi-

tionally assigned the task of acting as repository of print information (i.e. libraries) will most likely identify a prominent role for themselves in DLS operations. If access is given free or at highly reduced rates to traditional library patrons (as is the current print model), then the possibility exists that the entire pricing scheme could be defeated. If an end user can access a DLS through a public library gateway, he or she could conceivably perform the same search and retrieve tasks at no cost that otherwise could cost hundreds of dollars an hour. Moreover, Rightsholders may hesitate to place them in a DLS if they may be accessed through a public library gateway.

One possible approach would be to limit the number of terminals (or gateways) available to the public library system on the theory that a commercial user will pay for the ability to retrieve information much sooner. A counter-scenario was suggested using the analogy of long distance telephone "least cost routing" businesses -- a business that does nothing more than monitor various "free" gateways into a DLS and assigns a search to a gateway as soon as it opens up.

It was pointed out, however, that public libraries are typically subsidized by tax-paying residents, so there is a pricing scheme which is independent of use. Thus, if the library pays for public use of Digital Library services, the pricing scheme for a DLS need not be destroyed by making DLS services accessible through a public library. It is possible that the public library might collect fees for use of DLS services from patrons. Participants agreed that this issue needs further thought.

A second related issue was how to bring the benefits of a DLS to those now not using today's print or electronic resources. The question was posed as to whether a DLS, with its graphical user interfaces, could offer a means of providing more information to less educated or to foreign-speaking individuals. It was unclear what transactional framework should apply to such individuals, as those who are illiterate also tend to be in lower economic brackets. The goals of education must thus be measured against the cost to a DLS operator, and the desire to funnel public funds from the government into a DLS to underwrite usage by the underprivileged.

4.2 Business Models

As the description of Digital Libraries in Section 2 indicated, a DLS may be monolithic or made up of smaller systems, each with the technical ability to interface. Given these two paths of evolution, the participants considered various business goals and organizational structures for an operational DLS. Such goals might include:

a. Developing cost-effective technical standards (common protocols for information representation and exchange);
b. Simplifying contract standards
c. Maximizing access;
d. Promoting fair access;
e. Minimizing costs to the end-user;
f. Minimizing the need for government regulation.

Taken together, these goals would produce a common infrastructure for DLS operations which would apply whether a single monolithic DLS were to develop, or many smaller DLSs were to evolve.

Goal (b), simplifying of contract standards, resulted in prolonged discussion as participants weighed such factors as the need for freedom to negotiate, need of educational institutions for a different pricing mechanism, and the monopoly/tariff model which would ensure access and no price discrimination. From this starting point a number of structural models for DLS operation were discussed.

4.2.1 DLS Model A: The Multiple DLS Environment

Most participants favored a decentralized DLS with links between various smaller DLSs. This would be the easiest transitional DLS, building upon existing on-line databases and early library projects such as Project Mercury at Carnegie Mellon University.

The major stumbling block to this approach is the problem of developing adequate *technical standards* for interoperability of multiple DLSs: how are we to ensure that the various DLSs will be able to pass data (Objects and Knowbots) and accounting information? Further, the communications costs of such a system are not insubstantial, although many references were made to existing government and privately supported networks. Transfer prices, fee collection, and other variable costs are substantial in this model, as each DLS would have its own transactional framework.

For example, suppose DLS-1 is operated within a "block" payment framework. Rightsholders are paid a flat royalty for Import, and users pay a specified amount for unlimited access to DLS-1. Now suppose DLS-2 is based on a "per Access" pricing method. Rightsholders are paid a certain amount for each Access and function performed on their Objects, and users are charged for each Access or function performed. Next suppose DLS-1 and DLS-2 are linked by a gateway. How are users of DLS-1 going to be charged for use of DLS-2? In addition to their "block" payment, are DLS-1 users going to receive a bill for their use of DLS-2?

If we envision a decentralized DLS with several hundred individual DLSs, accounting will be a major challenge. Under this decentralized model, therefore, much effort would be required at the outset to ensure that the varying pricing and royalty mechanisms between the various DLSs could allow full interfacing between them with minimal gateway roadblocks.

4.2.2 DLS Model B: The Monolithic DLS as a Regulated Monopoly

After discussing the transaction and transportation costs of multiple DLSs, the possibility was discussed that the U.S. information economy might be able to afford only one DLS. A single DLS would provide instant uniformity and most if not all of the common infrastructure goals would be achieved. Further, a single DLS would create numerous efficiencies from both a technical and business standpoint. A monolithic DLS, however, also would be a candidate for Federal regulation as a natural monopoly, akin to a public utility such as the power company or local telephone service. Governments have traditionally attempted to regulate natural monopolies in order to provide caps on prices and ensure quality of service that competitive factors generally provide in non-monopolistic industries.

One other initial problem with this scenario is reaching critical mass. How can a single system smoothly evolve from the present patchwork of on-line databases with markedly different pricing structures? Moreover, how can this system evolve from the existing print-medium distribution system in an equitable fashion? For example, could a publisher, in effect, be forced to participate in a DLS under unfavorable terms and conditions simply because the DLS has become the primary means for locating and retrieving published material? Resolution of this critical mass question requires careful consideration of economic issues (including added revenues, potential for lost print sales, and overall profitability) as well as protection of intellectual property rights in a DLS.

Once an Object is placed into a monolithic DLS, can a Rightsholder be assured of protection and future control of the Object? One technical scenario discussed was the "neutral ground" Knowbot. Under this scenario Objects would never actually leave the owner's library (no Release to public access), but rather, a Knowbot with search instructions would pass into the owner's library or "nearby" neutral ground and perform its functions there. A counter-example based on viruses was also raised: the owner might feel even less safe if there were a possibility of a rogue Knowbot entering his work space and stealing or modifying everything. It was supposed that Knowbots would be subject to local controls on access to Objects.

Regulation as a utility would interject some degree of government intrusion into the operation of a DLS which some participants argued outweighed efficiency gains of a monolithic DLS. As discussed in Section 5, below, the interests of the government may not be totally consistent with the interests of a DLS operator. Obviously, the more control the government has over DLS development and operation, the more its interests will take precedence over the countervailing interests of the private sector.

An analogy to the public television broadcast system is relevant here. In this case, a variant of the monopoly model would utilize government funding in the establishment of multiple DLSs which would then compete. The major potential for dependence on the government led the participants to conclude that this model probably is not workable as a long-term solution.

4.2.3 The Role of a Testbed DLS in Aiding Infrastructure Development

Given these two very diverse potential evolutionary routes to DLS operation, the participants then considered the question of how the technical and business problems can be managed on a smaller scale. The following testbed model was discussed as a possible first step in developing a DLS prototype. The testbed model's parameters were :

a. Free of charge (at least initially)
b. Window-based display
c. Access via current R & D networks
d. Populated by the following free works:
    -- Technical reports
    -- Ph.D. dissertations
    -- Free software
    -- Patents and other government publications
    -- E-Mail
    -- Library Catalogues
e. Experiment with "access aids" -- services and mechanisms to help users find useful information.
f. Timing: prototype within one year, pilot DLS on-line three months later.

This testbed model generated extensive discussion. Many participants viewed the first goal of free access as counter-productive, since one of the major questions in a DLS is the pricing/royalty distribution structure necessary to generate "critical mass" of users and Object importation into a DLS. Members of the publishing community thought that this prototype could provide an interesting experiment in negotiating for rights to import Objects into a DLS. However, limiting the contents to free works would not give a realistic approximation of expected use or even of the eventual composition of the user community. In addition, using only public domain materials could create an expectation in the user community that any information in the system could be used freely without regard to intellectual property rights. This expectation could be hard to reverse if the system later included copyrighted materials. While the testbed model might generate some interest among the research community, that collection of users may not be the optimal test environment for a DLS. Participants therefore suggested broadening the types of Objects available and therefore the user community to provide a richer testbed for exploring the technical **and** business issues involved with DLS operations.

5. THE ROLE OF GOVERNMENT IN A DLS

BIG PLAYER OR "BIG BROTHER"?

As with many new technologies, one may expect that the Federal government will play some role in its development. There was a generalized reluctance on the part of many of the participants to readily encourage government participation. The source of this reluctance was explored. The starting point of the discussion was to try to list the interests the government might have in a DLS:

a. Access to government information;
b. Regulation of anticompetitive practices;
c. Universal access (rural America and information "have-nots");
d. Uniform Technical standards;
e. Social conditions imposed on information marketplace (obscenity);
f. Limiting extent of DLS operator liability;
g. Government gathering of personal information.

The government's potential ability to Access personal information about DLS users caused the greatest concern to participants. Although most participants agreed that providing the public with easier access to government documents would be a laudable goal, the reverse situation of allowing the government easier access to private sector information concerning its citizens would be much more troublesome. Most participants felt that there would be a direct correlation between the amount of private information the government could access (including possibly accounting and other internal DLS information) and the government's overall involvement in a DLS.

The degree to which universal access to a DLS is compatible with a viable business venture also was discussed. The more the government requires a DLS to make itself available for free or at reduced rates for underprivileged users, the greater the potential for abuse by users otherwise able to pay. There is the additional compatibility problem between "pay per Access" Objects and universal access. Presumably, Rightsholders may demand a payment for each Access to their Object because of the perceived high value

of that Object. An owner may hesitate to include the Object in a DLS if universal access requirement will allow some DLS users to Access that Object for no charge.

In discussing universal access, various analogies were made to telephone service, cable and broadcast television. Telephone service has become such a necessity that Congress has taken steps to ensure that the economically disadvantaged still can afford local telephone service as the industry evolves away from long distance charges subsidizing local exchange service. Similarly, the Communications Act of 1934 requires the FCC to allocate television stations on a "fair and equitable" basis so as to provide the greatest number of television signals to the greatest number of persons. With the emergence of cable television and direct broadcast satellite as direct competitors to free over-the-air television, they are beginning to bid directly for programming traditionally carried on free television. Congress is now beginning to look at whether alternative services may be evolving in such a way that universal access to them may be required by regulation or statute.

The question then arises whether and when DLS services will be perceived as a necessity to all persons in this country. If the government makes that determination, then universal access may become part of the regulatory environment. Many participants reiterated that the more government is involved early in the operation of a DLS, the sooner regulations requiring universal access may be imposed. The participants concluded that in the best of all possible worlds the government should provide research funding for DLS development and possibly involve itself in the development of technical standards (to make sure government publications were easily Importable), but otherwise not be involved with the operation of a DLS.

The other alternative for funding would be to try to develop a DLS with the financial support of the communications and information industries. This approach also has its down-side, as any company paying a substantial amount of the research and development costs is likely to want rights to the system software. It was noted that this proprietary view could seriously impede the ability to reach agreement on broadly-based standards required for a DLS to become a major piece of our national information infrastructure. A "matching grant" scenario involving a number of industry leaders each providing a smaller share of the funding appeared to be more attractive from this perspective.

6. CONCLUSIONS OF THE WORKSHOP

WHERE DO WE GO FROM HERE?

6.1 Intellectual Property Rights Issues

After two days of discussions, it became apparent that the group had focused less on legal theories of intellectual property protection and more on how intellectual property rights might be applied in the development and operation of a DLS. A number of legal issues were considered, however, and provided a firm backdrop to the discussions of Conditions, Import, and transactional framework models.

Many of the Conditions discussed, and the functions which could be performed on Objects as a result of the existence or absence of particular Conditions, were firmly rooted in traditional intellectual property rights theory. Historically, copyright holders have certain divisible rights to control the dissemination of their works and what others can do to those works.

One of the major issues in the Import discussion revolved around ensuring that an alledged Rightsholder indeed had the right to Import an Object, and the liability of a DLS should an Object (for whatever reason) be placed in a DLS without the right to do so. The conclusion that a "Condition and Operation description language" may need to be developed evolved directly from the discussion of the limitations on Access that Rightsholders may require, and are entitled to enforce under existing intellectual property law, before they will be willing to Import a work into a DLS.

The discussions conducted at the workshop could be thought of as an attempt to overlay DLS development onto intellectual property law and practice, rather than attempting to apply intellectual property law to DLS development. The discussion of

transactional frameworks was not conducted in a vacuum. Participants, at least implicitly, drew from their experience in today's print world in which owners of copyrights and their agents negotiate for reproduction, transmission, the right to make a derivative work, and other rights. Discussions of the transactional framework of multi-media Objects had a clear analog in present-day standardized agreements between Rightsholders and users. Given that the participants came to the workshop with a working knowledge of many aspects of intellectual property law and how the present information industry works within that legal structure, the route taken may have been more productive than an attempt to shoehorn DLS development and operations into particular legal "pegholes."

One suggestion of a next step in the discussion of protection of intellectual property rights in a DLS was for participants from the legal community to prepare papers for presentation on a number of legal issues surrounding intellectual property rights. A possible topic could be a presentation on what current legal assumptions are called into question in a DLS environment. Will traditional concepts of "fair use" have to be altered with the technology? Does the technical ability of a DLS to completely purge an Object, and all references to it, call into question the rights of authors to control their works? Will a "public right" to information develop so as to impinge upon the number and types of Conditions a Rightsholder may specify? Will the ease of Access to Objects created by a DLS affect the traditional issue of access to documents in copyright infringement cases, and can audit trails be constructed so as to allow for legal proof of access to an alleged infringing work, yet still protect the privacy of individuals?

Another approach to assessing the role of intellectual property rights on DLS development would be to develop a number of case studies of different types of Objects, and track the legal issues confronted upon Import, registration, cataloging, indexing, manipulation, and export from the system of a protected Object.

Finally, it was noted that a need exists for coordination between the various library projects currently underway in the United States. Substantial support was expressed for CNRI to undertake this coordinating role.

6.2 Future Workshops/Conferences

Participants discussed the feasibility and desirability of future workshops, and whether one or more conferences should be conducted. At the close of the workshop, the participants were encouraged to submit a list of questions and issues which could be pursued at another meeting similar to this one.

One possible future meeting was discussed involving those who are actively involved in DLS research. Another workshop would include persons and groups that have a near-term interest in becoming involved in a DLS project. One other possible avenue for discussion is the White House Conference on Library and Information Services to be convened sometime in the Fall of 1991.

6.3 Conclusions

There was a clear feeling after two days of discussion by the participants that the workshop had barely scratched the surface of the legal, economic, and social issues involved in the development, implementation, and operation of a fully functional DLS. Although no participant could point to any insurmountable legal hurdles to DLS development, there was no agreement on what constitutes the best or even a good framework on which to build such complex systems. Participants came away from the workshop with a feeling that, with certain modifications, existing intellectual property and contract law would allow authors and other Rightsholders to continue to exercise their constitutional and statutory rights to control their works and receive compensation for the use of those works by others. That is not to say that the transition from print to electronic libraries will be a smooth and easy road. The extent to which these issues can be resolved efficiently and quickly remains to be seen. Nevertheless, whether through lengthy litigation, reactive legislation, or by contract, the intellectual property laws of the United States have proven to have sufficient flexibility, or amendability, to accommodate emerging technologies. So long as there is economic gain to be made from populating and using a DLS, it should be possible for the information industry to resolve the outstanding questions in such a fashion that legal issues need not be a major inhibitor to the development of a DLS.

ACKNOWLEDGMENTS

This report is based on the discussions which took place during May 18-19, 1989 at the workshop on Intellectual Property Rights in a Digital Library System and on comments made by the participants in reviewing an early draft copy of the report. The original draft of the workshop report was prepared by James E. Dunstan of Haley, Bader and Potts. The workshop was jointly sponsored by the Corporation for National Research Initiatives and the American Association for Artificial Intelligence.

APPENDIX

LIST OF PARTICIPANTS

William Y. Arms
Project Mercury
Carnegie Mellon University
Pittsburgh, PA 15217

Daniel G. Bobrow
American Association for
Artificial Intelligence
445 Burgess Drive
Palo Alto, CA 94025

David Brandin
A.T. Kearney Technology
One Lagoon Drive
Redwood City, CA 94065

Vinton G. Cerf
Corp. for National Research Initiatives
1895 Preston White Drive
Suite 100
Reston, VA 22091

James E. Dunstan
Haley, Bader & Potts
2000 M Street, NW
Washington, DC 20036

John Garrett
Copyright Clearance Center
27 Congress Street
Salem, MA 01970

Brian Kahin
254 Concord Avenue
Cambridge, MA 02138

Robert E. Kahn
Corp. for National Research Initiatives
1895 Preston White Drive
Suite 100
Reston, VA 22091

Patrice Lyons
Haley, Bader & Potts
2000 M Street, NW
Washington, DC 20036

Susan Martin
National Commission on Libraries
and Information Science
1111 18th St. NW, Suite 310
Washington DC 20036

Carol Risher
Assoc. of American Publishers
1718 Conn. Ave., NW, 7th Fl.
Washington, DC 20009

Pamela Samuelson
Emory Law School
Gambrell Hall
Atlanta, GA 30322

Keith W. Uncapher
University of Southern California
Olin Hall 330A
Los Angeles, CA 90089