intellectual property project

PROCEEDINGS

Technological Strategies for Protecting Intellectual Property in the Networked Multimedia Environment

COALITION FOR NETWORKED INFORMATION

INTERACTIVE MULTIMEDIA ASSOCIATION

JOHN F. KENNEDY SCHOOL OF GOVERNMENT
SCIENCE, TECHNOLOGY & PUBLIC POLICY PROGRAM

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROGRAM ON DIGITAL OPEN HIGH-RESOLUTION SYSTEMS

VOLUME 1 ISSUE 1 JANUARY 1994

The Journal of the Interactive Multimedia Association Intellectual Property Project

PROCEEDINGS

January 1994
Volume 1, Issue 1

The Journal of the Interactive Multimedia Association Intellectual Property Project

Published by:
Interactive Multimedia Association
Intellectual Property Project
3 Church Circle
Suite 800
Annapolis, MD 21401-1933
Phone: (410) 626-1380
FAX: (410) 263-0590

Copyright and Information Services in the Context of the National Research and Education Network[1]

by R. J. (Jerry) Linn

Abstract The High Performance Computing Act (HPCA) of 1991 (P.L. 102-194) places unenforceable requirements to protect copyrights and intellectual property rights on the National Research and Education Network (NREN). This paper discusses the roles and responsibilities of the NREN and associated information services; technical approaches to authentication, redistribution and authorization of use of electronic documents over the NREN; and an amendment to the High Performance Computing Act.

Introduction

It is clear that when the High Performance Computing Act of 1991 was written the notion of digital libraries was a consideration of the authors. It is also clear that the Congress intended that copyrighted materials be distributed over the Network. The legislative history of the Act affirms this position. The Act, as drafted in the 100th Congress (S.1067 and H.R. 3131, 1990), included provisions for authorization of appropriations for the National Science Foundation to establish digital libraries. Other bills introduced into Congress which provide for similar authorization of appropriations include S.2937, introduced in 1992, and S.4, introduced in 1993. Indeed, prior to 1991, digital libraries were integral to the thinking related to "information services" and were the stimulus for language incorporated into the HPC Act of 1991 with respect to protection of copyright.[2] The language employed in the Act of 1991 assumes information services are embedded in the Network, as part of a network infrastructure. However, the term "network" has very specific and narrow connotations when used by professionals in the computer and communications communities versus the broad definition of the term in the Act. Furthermore, recent papers and reports from a workshop focused on the National Research and Education Network (NREN) reflect the common understanding that the NREN is *only an access medium* to application services.[3] Therein lies the weakness of the legislation: the definition of the "Network" is too broad to assign responsibility for protection of copyright and intellectual property rights. Furthermore, the professional community to whom the courts would turn for expert witnesses to aid in interpretation of the law is not likely to agree with the reasonableness of the requirements that the Act places on operators of the Network or the

ability to enforce its provisions except in the computers attached to the Network which offer information services.

Specifically, the Act defines the "Network" in Sec. 4 as follows:

> (4) 'Network' means a computer network referred to as the National Research and Education Network established under section 102;

and Sec. 102 (c) "Network Characteristics" states:

> The Network shall —
>
> ....
>
> (5) be designed and operated so as to ensure the continued application of laws that provide network and information resources security measures, including those that protect copyright and other intellectual property rights, and those that control access to data bases and protect national security;
>
> (6) have accounting mechanisms which allow users or groups of users to be charged for their usage of copyrighted materials available over the Network and, where appropriate and technically feasible, for their usage of the Network;

There are several important things to note because they become "first premises" for a discussion. First, the NREN is a concept (the Act never defines who owns and operates it; the Act authorizes appropriation of Federal funding to agencies to implement the concept). Second, the NREN is a logical entity derived from a network of networks (an internet). And third, the NREN is a part of the *Internet*—that network of networks whose span is global and whose common denominator is a shared name and address space.

The Network established under Sec. 102(a) does not imply that the Federal government installs or owns the physical assets of the NREN (e.g., optical fiber cables, routers) nor does it preclude the NREN from being derived from commercial, private sector sources and services. *This ambiguity is important.* The definition and ownership of the NREN are not cast in concrete (like highways); this omission allows the NREN (or parts of it) to transition from government provided and/or subsidized services to commercial for-profit services, or an evolving combination of both. Evolving Federal policy supports transition to commercial services as required services become commercial commodities.

Which networks comprise the NREN and who owns/subsidizes them are not as important as understanding that "ownership" of subnetworks, levels of subsidy and recipients of subsidy are all subject to change over time. Therefore, defensible answers for issues related to copyright, intellectual property rights and the NREN must take into account the diversity of the technology base in component subnetworks, of ownership, of agency missions and goals, and of those services accessed by the NREN versus common services provided by subnetworks comprising the NREN/Internet. This complexity suggests that it will be beneficial to partition the problem into smaller components for analysis and discussion.

Subsequent subsections present the "Network" as a set of services, establish both technical and pragmatic reasons for doing so, and discuss protection mechanisms appropriate for the decomposed services. Specific

technical mechanisms are outlined which may be employed to distribute and protect copyrighted materials by an information service. Finally, an argument is presented that the HPC Act of 1991 should be amended such that the protection of copyrights and intellectual property is properly the responsibility of information service providers and users. An amendment is offered which would realize the position presented.

Delineation of Services

A delineation of network services aligned with widely recognized technical boundaries and terms will aid in a dialogue because functions and responsibilities can be discussed within an established framework. Professionals familiar with network architectures associate specific functions and services with well-known named layers of a network architecture. The terms and concepts used below are recognized by an international community of computer and communications professionals.[4] Thus, it is unnecessary to define new terms and concepts in order to establish a framework to discuss issues.

The functions associated with the two lowest layers of a network architecture are *physical*, point-to-point connectivity and signaling, and data transmission via *data links* which interconnect computers or routers. Next in the hierarchy are *network-layer* functions which select routes and relay data packets enroute to their destination. These functions are the least common denominator of a "computer network" and are often implemented by routers which comprise or interconnect wide area subnetworks.

The *transport layer* establishes end-to-end connectivity and may provide for retransmission of data packets lost or corrupted by lower layers. Thus, the transport layer provides a reliable end-to-end communications medium for application programs and services. Note that the public switched network may also be used to provide an end-to-end communications path between computers; however, end-to-end communications is achieved by different technical means.

Information services are provided by *application-layer* programs and supporting protocols at the end points of a communications path. Examples of application-layer services are electronic mail and file transfer, which are implemented by application-layer protocols (e.g., Simple Mail Transfer Protocol (SMTP) and X.400 are electronic mail protocols).

Connectivity of subnetworks in the NREN/Internet functions at the network layer (see Figure 1). Each subnetwork serves as a switching fabric for a set of computers; i.e., the network layer software receives and relays packets of data from one node in the network to another node based only on its destination address. Note that the routing and relay (switching) functions assigned to the network layer are the least common denominator of the Internet (NREN). Specifically, a subnetwork (e.g., college campus, midlevel network or the NSFnet of the National Science Foundation) may use one set of technologies and another subnetwork may use another. However, the "glue" that interconnects them is a **common, minimal set of protocols necessary to provide the routing and relay functions**. Any additional set of functions is optional in the network layer and is only likely to be incorporated if actually required in a given environment (e.g., security, network management). Therefore,

any assumption that the "Network" is a uniform, ubiquitous environment is erroneous—particularly when the NREN is viewed as a set of interconnected autonomous subnetworks.

This is a greatly simplified sketch of a multi-layered network architecture. The sketch highlights crucial networking design concepts; i.e., specific functions are assigned to layers in a network to accommodate an array of lower-layer communications technologies and for design and maintenance purposes. However, we have sufficient information and a set of terms which is rich enough to pose questions about how and where the requirements of the Act might be implemented and to explain why they might or might not be reasonable requirements in the first place. We are also prepared to identify and discuss conflicting objectives, if proper design and engineering principles are not followed.

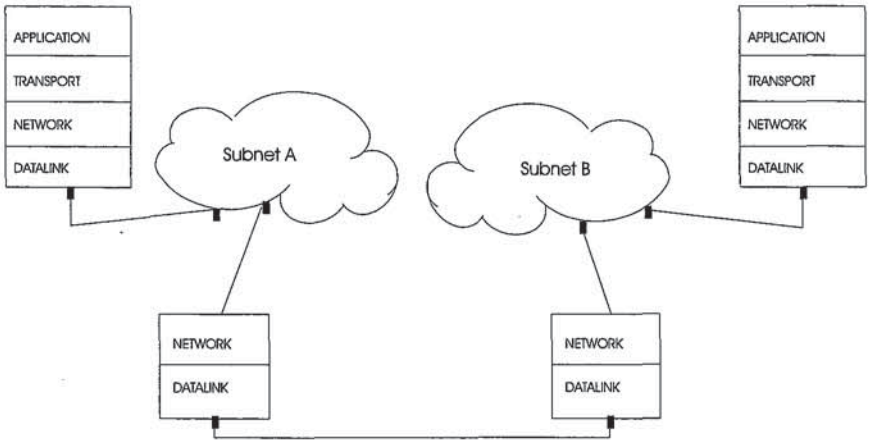

Figure 1: An internet -- a network and networks

Thus, when the Act states:

> the Network shall —
>
> ....
>
> (5) be designed and operated so as to ensure the continued application of laws that provide network and information resources security measures, including those that protect copyright and other intellectual property rights, and those that control access to data bases and protect national security;

we can ask: "What does this mean? What protection is required? How may required protection be achieved in the context of existing network architectures? And, who should be responsible?"

Clearly, the routing and relay functions of the network layer will not protect copyrighted materials. In fact, they do not even assure delivery of data packets. Therefore, the "Network" described in the Act requires more functions than those described for the network layer. So it is appropriate to ask: "What protection is inherent in a network; what additional protection is required; and where is it most appropriately offered?"

Under normal circumstances, network-layer software *does not* inspect the contents of data packets. There are at least two good reasons not to do so. First, inspection of packets for any purpose introduces unnecessary overhead and degrades the throughput of the network (a serious consideration in high-speed networks). Second, inspection of packets (or streams of data) jeopardizes the privacy of the information being transmitted. Also, recall that the network-layer software was described earlier as "least common denominator," with the implication that any additional functions were optional. Consequently, the network layer is not a viable candidate for uniform protection of copyrighted materials.

Data integrity protection against accidental changes is assured if specific transport protocols are employed at the end points of a connection. Specifically, the network can protect against accidental loss or corruption of data during transmission from one point to another. This is true for the Transmission Control Protocol (TCP) and the Organization for International Standardization (ISO) Transport Class 4 (TP4); both detect and retransmit lost and corrupted data. However, the transport protocols cannot protect against redistribution of materials obtained from a legitimate source, nor can they assure the authenticity of the materials transmitted over the network. The means to assure authenticity of materials and achieve protection from deliberate abuse by end users is to implement the required protection mechanisms in computer systems as part of the application programs which deliver services to users.

Technical Means for Protection of Copyrighted Materials

New protective services can be created for information dissemination which can also be applied to those materials that have a copyright. However, requirements for protection must be defined before describing how protection might be achieved. Below is a set of requirements which serve as a starting point for a discussion.

Protections and Features Required

Authentication: A mechanism is required to certify that any material received is a bona fide copy of the original (data authentication) and possibly who it came from (origin authentication). If the copy is not authentic, then this fact should be detectable and the copy discarded. Recall that the transport layer *may* provide for integrity protection against accidental changes, but authentication provides a means for protection against both accidental and intentional changes.

Limited redistribution: Publishers want to control distribution to those who have paid a fee for the use of copyrighted materials. Mechanisms should be implemented to restrict the number of copies printed to those paid for and to the individual who paid for them.

Protection against plagiarism and change: Authors and publishers do not want their materials used without appropriate attribution, nor do they want the materials excised, edited, or modified such that authenticity is jeopardized. Information should be stored in a form which makes it difficult, if not impossible, to remove the copyright mark, or excise or modify text.

Object form: Information should be stored and exchanged in standardized but device-independent forms. Processing software employed by a

user should display or print the materials in an appropriate form given the constraints of the user's video display and/or printer.[5]

To discourage plagiarism, excising parts of the text and other unauthorized uses of the information, an object could be put in a "sealed envelope" and distributed in one of several forms which are not easily read and modified by humans. These forms could include SGML, G4Fax, and PostScript or other useful forms. SGML denotes the Standard Graphics Markup Language. SGML text would require processing of the input text to render meaningful output on either a video display or printer. G4Fax denotes Group 4 Facsimile which is a compressed bit stream using an international standard for scanning and compressing facsimile images. It may be displayed or printed on raster scan output devices (video display or printer). G4Fax could readily be used for interlibrary exchange to avoid document handling and scanning. PostScript denotes the form used by PostScript printers. It is a page description language that is widely implemented, is useful for printing purposes only, and would not require significant processing if directed to a printer.

Appropriate remuneration: Remuneration could take the form of a subscription fee, license fee, contract, or fee for services rendered, as appropriate. Dissemination may be by an author, original publisher, information service, or library (hereafter called an authorized distribution source).

It is assumed that interlibrary loan and electronic redistribution of single copies of papers to individuals by libraries who have a subscription, license or contract with a publisher constitutes "fair use." It is also assumed that fees for services will be established (commercial, for-profit; and not-for-profit) and public access could be via public libraries. Specifically, an individual could ask for and get a copy of a paper or article as easily as he or she can reproduce it on a copier in a library (and at a comparable price). Remuneration by an individual patron could be at the time the material was obtained, if there was a fee.

Technical Mechanisms

A set of mechanisms may be combined to address the requirements outlined above. For discussion purposes, we consider a body of material (information) as an "object" with certain components and attributes. One attribute is an electronic "copyright" mark; the object forms noted earlier are another attribute. Object-oriented technology associates processing of objects with their attributes. For simplicity, however, we describe an object as an envelope and its contents. The information on the envelope is visible and the contents hidden and sealed with a digital signature. Examples of information on the envelope could include title, author(s), abstract, keywords (e.g. full bibliographic record) and attributes describing the form of the object, a digital signature, copyright status (yes/no), and date and timestamp associated with an authorized copy. Visibility of information on the envelope has other obvious advantages related to search and retrieval of information stored in digital libraries, but they are outside the scope of this paper. Figure 2 presents a graphic perspective of the concepts.

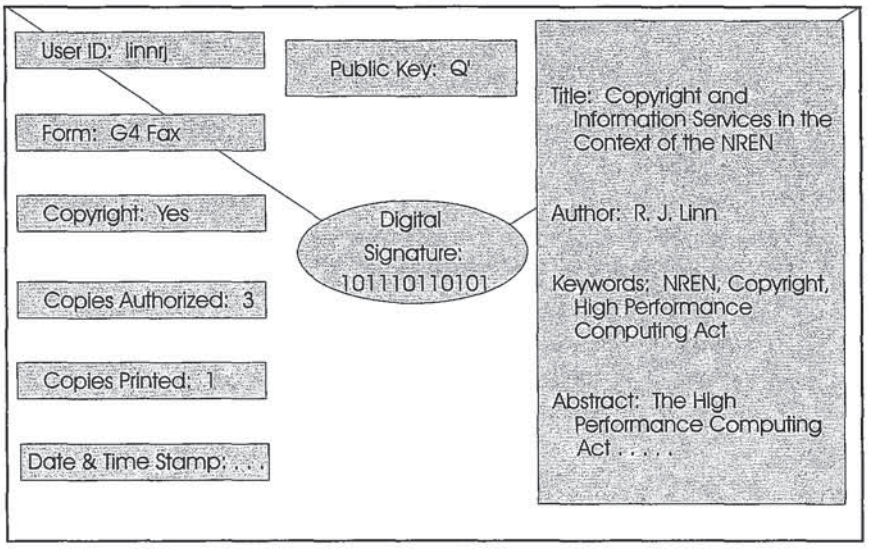

Figure 2: An Information Object -- Envelope plus Content

For our purposes we assume:

- an object is processed by standardized software (hereafter called *rendering software*);
- creating an original information object, file transfer over a network, and rendering of the information on a video display or printer are built-in functions of the rendering software;
- the rendering software is
  - inexpensive or free because it is in the interest of the public, and of publishers and authors to protect their intellectual property, and
  - widely available; e.g., distributed by publishers, information service providers, computer manufacturers;
- copies of objects are exchanged using the rendering software—a copy is obtained from an authorized distribution source (may be an individual if there is no fee for use); and
- the structure and exchange formats of objects are standardized (either de facto or de jure).

Active Protection Mechanisms

Two active mechanisms *implemented in the rendering software* will achieve the requirements for protection outlined in the previous section.

Authentication: Confirmation of authenticity of the source and contents of the envelope can be achieved by use of a public key, digital signature algorithm. The public key is provided by the author or publisher and is written on the envelope. The public key is used to verify the digital signature of the information written on the envelope and its contents. If either is changed, the digital signature verification algorithm detects and reports failure. If verification failure is detected when an object is being obtained from an information service, its retransmission

should be requested. (This might occur if data were lost or corrupted.) If a failure is detected when displaying or printing an object, further processing should be inhibited. This might indicate a bootleg copy, or a mismatch of user identification with that on the envelope, or it might indicate that the authorized number of copies have been printed. Optionally, the object could be destroyed by the rendering software when verification fails.

Limited redistribution: Identifying the holder of the copy on the envelope (e.g., user identification) and a copy counter can be employed to limit electronic redistribution. The user identification and initial value of the copy counter stored on the envelope are established when a copy is obtained from an authorized distribution source. The number of printed copies allowed is a function of the fee paid. The copy counter is used to restrict the number of copies rendered on a printer. As the copy counter is decremented, a residual copy count and new digital signature is computed and affixed to the envelope to prevent an unlimited number of copies from being printed.

Note that sending a copy of an object via electronic mail, redistribution by a bulletin board and other simple copying mechanisms will not update the contents of the envelope which contains the date and timestamp of an authorized copy. If the date and timestamp in the directory entry for a file containing an object do not match those in the envelope of the object, the rendering software considers the copy to be unauthorized. Consequently, the information contained in the envelope will not be presented to a user by the rendering software and unauthorized copies are useless.

Materials may be displayed on a video display an unlimited number of times by the user identified on the envelope. Other users are prohibited from displaying an object with the "copyright" attribute. However, unlimited rendering and redistribution is permitted if an authorized distribution source omits the "copyright" attribute on the envelope, or enters "unrestricted" in either the user identification or copies authorized fields.

Passive Protection Mechanisms

Object form: The object forms described above are not human interpretable forms (SGML, G4Fax, PostScript). Furthermore, an object is stored in a form which may not be displayed or printed without the rendering software unless it is extracted from within its envelope. Although this is a passive protection mechanism, significant technical information and expertise are required to defeat it.

Note that all the forms described above prevent easy redistribution by simply making a copy and mailing or printing it with utility software because the rendering software is required to display or print an object. These forms also inhibit using a simple editor to "cut and paste" text into another document because no form is human readable, and direct user access into the contents of the envelope is not allowed by the rendering software.

Write protection: Write protection is the first line of defense required to protect the authenticity of information disseminated by an information service. It restricts the privileges to create or modify stored information

to the rightful owner(s); these are called "write privileges" associated with a file. Restrictions are essential for any information service and must be implemented within the computer system offering the information service. Write protection is not a function of the "Network" but is a responsibility of the parties operating an information service.

In summary, two active forms of protection are proposed for intellectual property: *authentication* and *limited redistribution*. Two complementary, passive mechanisms are also identified, but are inadequate on their own (*object form* and *write protection)*. All the mechanisms suggested are implementable on computers accessed by a network, and are completely independent of the networking technology used to access an information service. All mechanisms are applicable to any information distributed over a computer network whether or not the information carries a copyright mark.

Summary Arguments

Separation of the roles and responsibilities of the "Network" and "information service providers" provides a logical and pragmatic framework for disentangling and discussing the legal and technical issues related to the NREN and copyright.

First, the NREN is a concept (or logical entity) rather than something physical with fixed boundaries. The present and future NREN will be part of the global Internet. As such, its owners are both public and private entities, and it is not uniform in the underlying technology deployed. Pragmatically, it is impossible to require any owner of part of the Internet (a subnetwork) to add new, optional network functions which do not serve the owner's immediate needs. Consequently, the Network as a whole can only provide the "least common denominator" services with respect to networking functions. These common functions are selection of routes and forwarding packets enroute to their destination; this is called "packet switching." Often, technical people think of the "Network" in terms of these limited functions; e.g., NSFnet provides the packet switching and routing functions to interconnect other networks.

Second, the language of Sec. 102 (c)(5) implies that operators of subnetworks which are part of the Network could be liable for the illegal actions of both the providers of information services accessed via the Network and the users of these information services; i.e., "must be designed and operated to ensure ... including those that protect copyright ..."

These requirements to protect copyrights and intellectual property rights are at odds with established protection for common carriers who also provide networks capable of providing access to information services which distribute copyrighted materials. Carriers are not liable for the illegal activities of their users. Surely, a telephone company would not be held legally liable if an information service used facsimile machines to illegally sell and distribute journal articles. Note that it is technically feasible for the NREN to become integrated with the public switched network in the near future (e.g., narrowband ISDN services (Integrated Services Digital Network) could be used to access the Internet). Using this situation as an example, there could be a dichotomy in terms of requirements and liabilities related to operators of subnet-

works with respect to a single illegal act; e.g., if part of the access path was via the public switched network and part via a midlevel network.

Third, consider that the "operator of the Network" is responsible for collecting and redistributing fees to the "appropriate entity" for use of copyrighted materials (c.f. Sec. 102 (c)(6) in the introduction). Is it likely that private sector providers of information services (e.g., a publisher) want an intermediary (Uncle Sam/Federal agencies) to collect and redistribute funds for services rendered? Even if an information service did want this service, which "network operator" is responsible (or would accept the responsibility)? Federal agencies operating a subnetwork do not want the responsibility of collecting and redistributing fees for private sector parties. Note that definitions of "operator of the Network" and "appropriate entity" (author, publisher, ...) are open questions. Particularly, when user access is granted via a sequence of subnetworks, who is the network operator? Is it the "operator" who provides the "user" access to the network, the operator who connects the information service, both, or some more complex combination?

Finally, a number of network-independent mechanisms may be employed by information service providers to limit redistribution and assure that copies remain unmodified. These include data compression, authorized use meter (copy counter), and public-key, digital-signature techniques. Digital signature can be employed as a tool to "seal an envelope" and verify the authenticity of copyrighted materials distributed over the Network. These mechanisms can be implemented to protect copyrights and the interests of publishers and authors completely independent of the network technology used to access the materials.

Definition of standardized technical practices to achieve the desired results and inexpensive software to distribute, protect and render copyrighted materials are all that are needed to protect the interests of publishers and achieve the intent of the High Performance Computing Act.

Conclusions

Sections 102 (c)(5) and (6) of the Act place unrealistic and unenforceable requirements on the "Network" and its operators (Federal, State or private sector parties) to (1) protect copyrights and intellectual property rights; and (2) account for use, collect fees and remunerate copyright holders. These should be the responsibility of the information service providers and users of information services. These are unrealistic burdens to place on Federal agencies or private sector operators of subnetworks which are part of the NREN (Internet).

While it is impossible to assure complete protection against malicious individuals, the appropriate remedy is to develop and deploy technical protections in the appropriate places, and apply the law in the same manner it is used to prevent bootleg copies of paper documents being reproduced on copiers.

The rationale developed in this paper could be used to interpret the existing law and develop regulations and rules aligned with the proposed amendment. If regulations and rules with the same intent were written, they would not clarify the intent of Congress[6] and would be more readily challenged in the courts. An amendment would clarify the intent of Congress and make the law enforceable. The author believes that clarity

on this issue is in the public interest as well as that of authors and publishers. To this end, an amendment is proposed as an appendix.

APPENDIX Proposed Amendment to the HPC Act of 1991

Insert the following definition at the end of Sec. 4.

"(6) "Information Service Provider" means an entity or individual who disseminates information, data, or copyrighted materials to others, for free or for fee as appropriate."

(Note that this definition is broad enough to include libraries, for-profit publishers, or individuals who want to participate in an "electronic press"—and is not restricted to the dissemination of copyrighted materials).

Substitute the following for Sec. 102 (c)(5) and (6):

The Network shall —

....

"(5) be designed and operated so as to enable the continued application of laws, regulations, directives and standards that prescribe security measures for network and information resources and those that control access to data bases and protect national security;

"(6) have accounting mechanisms which allow users or groups of users to be charged for their usage of the Network, where appropriate;"

and insert after Sec. 102 (e) —

"(f) Information services which distribute copyrighted information shall be designed and operated so as to enable the continued application of laws which protect copyrights and other intellectual property rights, including appropriate remuneration of copyright holders, while allowing for the `fair use' provisions of the copyright law."

and renumber Sec. 102 "(f)" and "(g)" as "(g)" and "(h)".

NOTES

1. This paper is a contribution of the National Institute of Standards and Technology. As such, it is not subject to copyright. The opinions expressed in this paper have not been endorsed by the Federal Networking Council, or any other federal working group.

2. These provisions were first specified in a draft of H.R. 3131, Title III, "Information Services," Sec. 302, "Copyrighted Materials," 1990.

3. Proceedings of the NREN Workshop, Monterey, CA, Sept. 16-18, 1992, EDUCOM.

4. The terminology employed is based upon the **"Open Systems Interconnection—Basic Reference Model,"** published by the Organization for Interna-

tional Standardization in 1984. A similar delineation of functions and terminology is used in the Internet architecture defined by the Internet Architecture Board/Internet Engineering Task Force.

5. The techniques described in this paper are equally applicable to output media other than video displays and printers. Thus, "object form" is intended to denote some machine-processable form of digital information which requires "rendering software" to present the content of the information in a human interpretable form—video, audio, printed text or some combination thereof.

6. In a conversation with the author, Mike Nelson ,who was on then Senator Gore's staff and now is in the Office of Science and Technology Policy in the White House, said,"Yes, we knew headers were required, but protection of copyright by the 'Network' is essential. Thus, the law reflects the intent of Congress."

Biography

R.J. (Jerry) Linn, a computer scientist, is Associate Director of the Computer Systems Laboratory at the National Institute of Standards and Technology (NIST) in Maryland. As a Commerce-Science Fellow in the U.S. House of Representatives, he worked on the High Performance Computing Act of 1990. His research activities include formal protocol design, specification and testing.

R.J. Linn
Associate Director for Program Implementation
Computer Systems Laboratory
B164 Technology Bldg.
National Institute of Standards and Technology
Gaithersburg, MD 20899
linnrj@osi.ncsl.nist.gov