February 17, 2000

Eric D. Scheirer
Leonard N. Foner
Media Laboratory
Massachusetts Institute of Technology
E15-401 (Scheirer), E15-305 (Foner)
Cambridge MA 02139-4307

This is a comment in response to the Library of Congress 37 CFR Part 201, treating the Exemption to Prohibition on Circumvention of Copyright Protection Systems for Access Control Technologies, as required in the Digital Millennium Copyright Act, 17 US 1201 ("DMCA").

We are researchers at the Media Laboratory of the Massachusetts Institute of Technology. The Media Laboratory ("Media Lab") is an interdisciplinary not-for-profit research institute with an international reputation. We conduct basic research in the development, distribution, and use of advanced multimedia technology and content. Our charter includes many areas of research that fall under the jurisdiction of the DMCA, including the study of media protection, media analysis, consumer privacy, encryption technology, and the design and development of content delivery systems.

The Media Lab also participates actively in voluntary, consensus-based standards-setting processs at the national and international levels, including the National Council on Information Technology Standards, the World-Wide-Web Consortium (W3C), the Internet Engineering Task Force (IETF), and ISO/IEC SC29 WG11 (the Moving Pictures Experts Group or "MPEG"). Our research and these standardization activities give us unique expertise regarding the impact of the DMCA on the multimedia research community.

Our comment is divided into two portions. In the first portion, we demonstrate how one particular field of study pursued by part of the Media Lab, namely the creation and analysis of technologies for the authoring, transmission, and reproduction of digital audio and video, is adversely impacted by the DMCA and in what ways the Librarian of Congress may mitigate this impact. In the second portion, we discuss more general problems, which plague those in the Media Lab and elsewhere who do research in systems that involves computer security and cryptanalysis, and discuss the chilling effect that the DMCA may bring to their attempts to publish the results of that research.

A. Comments on digital audio/video processing

This part of the comment will address several of the numbered questions specifically. We appeal to the Librarian of Congress to mitigate the effects of the DMCA on academic research in digital audio/video processing by exempting hardware and software tools that can be used to conduct such research from the purview of the Act.

1. What technological measures that effectively control access to copyrighted works exist today?

The most important such measure is the DVD Content Scrambling System (CSS), which controls access to movies and other video works distributed on DVD. Also, a set of measures to control access to music and other audio works distributed on compact disc (CD) and via digital delivery is presently being developed through the Secure Digital Music Initiative (SDMI). The eventual capabilities of the SDMI specifications may be speculated upon through reference to the technology developed by companies participating in the SDMI process, such as Liquid Audio, Intertrust, and Reciprocal Technology.

3. How has the use of technological measures that effectively control access to copyrighted works affected the availability of such works to persons who are or desire to be lawful users of such works?

One such lawful use of copyrighted works is research into content-based analysis of multimedia data (that is, the development of intelligent "media indexing" technology that can scan through multimedia content and determine features of human interest). Much of this research is conducted on advanced, high-performance computer systems such as Alpha systems sold by Compaq Computer Corporation, video workstations sold by Silicon Graphics, Inc., and networked clusters of computers running the open-source Linux operating system.

Since CSS-compliant video players and descramblers for these advanced computer platforms are not available, it is impossible to use such platforms to conduct research on works thereby protected. Further, as the DMCA has been used recently to justify prior restraint of third-party development and distribution of such players and descramblers (Universal Studios et al vs Remierdes et al, 00 Civ 0277, SDNY 2000), it will have a chilling effect on the future development of same. The restriction (through trade secret law) of the development of CSS players and descramblers to a small consortium of the multimedia industry means that it is unlikely that such will become available for these advanced platforms, as there is no market imperative for this consortium to develop them.

This lack of tools has a strong adverse effect on the availability of high-quality works of content to researchers who wish to lawfully use them for research and study.

6. If there are works that are available both in formats to which technological measures have been applied and in formats to which technological measures have not been applied, to what extent can the works in the latter formats substitute for the works in the formats to which technological measures have been applied?

For the purposes of research into multimedia content analysis as described in the response to (3), above, digital content is greatly preferable to analog content. This is because of the difficulty of converting analog media to digital representation for analysis, and because the lower quality of analog media makes the content works so stored less suitable for research purposes. Thus, it is only to a limited extent that unprotected analog works of content can be substituted for digital works protected by technological measures.

7. Are there works or classes of works that are available only electronically and only in formats to which such technological measures have been applied? If so, what are they?

As yet, we are not aware of any such works that would be suitable for research into multimedia content analysis. In our opinion, it is likely that, if the SDMI initiative is a success, some musical works will fall into this category within the next two years.

13. What impact has the use of technological measures that effectively control access to copyrighted works had on the ability of interested persons to engage in noninfringing uses of such works, including fair use and activities permitted by exemptions prescribed by law?

Another fair use of such copyrighted works is their use as examples within the context of presenting research results. That is, once a research project is completed, and it is published and disseminated throughout the research community, short selections of copyrighted works can be valuable in describing the function and efficacy of the new techniques being presented.

Technological measures that effectively control access to copyrighted works can make such dissemination difficult. For example, some music on the Internet is distributed in the "RealAudio" format created by RealNetworks, Inc. A technological feature of this format is that it is difficult for users to capture and store the music data (this format is a "streaming" format, in which the data are transmitted continuously over the Internet, rather than stored in a fixed file). Because of this, the fair use of this music to serve as an example within a research presentation is adversely affected.

28. What other comments, if any, do you have?

An important aspect of the increasing protection of copyrighted works through technological means is the ability of the rights owner, through that protection, to limit or disable fair uses of these works by the public. The development of fair use law was originally intended to serve as a defense for individuals and organizations who would otherwise be breaking copyright statutes with their uses of copyrighted works. However, as the sophistication of content-protection technology increases, it is no longer the copyright owners who must be protected from public theft of their material. Rather, sophisticated copyright-protection technologies serve to swing the balance of power far toward the rights owners, enabling them to unilaterally determine what uses of their content are allowable.

We believe that the fair use of copyrighted works for criticism, scholarship, reporting, and teaching is a fundamentally important part of American copyright law. Fair use must be strongly protected, not only as a defense of individuals against copyright litigation, but as a limit to the controls that rights owners are allowed to place on their content. Rights owners should not be allowed to technologically prohibit fair uses of their copyrighted works, even if it is within their technological ability to do so.

29. Do you wish to testify at a hearing to be conducted by the Copyright Office in connection with this rulemaking?

Representatives of the Media Laboratory would be pleased to testify at a hearing on the relationship between copyright law, fair uses of copyrighted works, and the Digital Millennium Copyright Act.

B. Comments on cryptography and security

The DMCA criminalizes a large amount number of academic activities which were formerly completely legal. As a result, those activities should be labelled exempt from its influence so as not to irreparably damage academic conduct.

It should be pointed out that these comments are solely addressed as remedying the deficiencies of the DMCA within the ability of the Librarian of Congress to do so. The DMCA itself is a deeply-flawed law, which makes illegal many activities simply on the basis of their being carried out in a digital medium, and as such is very probably unconstitutional at its core. However, the Librarian of Congress does not have the power to overturn the law on this basis—this is more properly a matter for the Supreme Court—and hence these comments are more narrow than is our preference.

We concentrate on three major areas for which we appeal to the Librarian of Congress to moderate the deleterious effects of the DMCA:

- Reverse-engineering of security systems in general, and cryptographic systems in particular, as the way in which academic security and cryptographic research proceeds
- Reverse-engineeering in the pursuit of creating interoperable systems and standards
- Chilling of Constitutionally-protected speech in the dissemination of the results of the first two activities.

1. Security, Cryptography, and Academic Research

Security

The computer science and applied mathematics communities have invested several decades of research in the questions of how to make computer systems secure from unauthorized modification, and how to safeguard the confidentiality and authenticity of their users' communications and stored data. These pursuits generally are divided into two major areas:

- Theoretical analysis of existing and proposed security and cryptographic systems
- Case studies of fielded systems

This approach is used because purely theoretical concerns are never sufficient when evaluating or implementing real programs and real systems. The theory of how to make a system secure is not sufficient because mathematical proofs are not powerful enough to completely represent everything interest about a program longer than a few dozen lines of code, and are particularly useless when the issue of real users' behavior must also be considered. Instead, actual examination of deployed systems has historically yielded a wealth of real operational data about where their weaknesses are. These weaknesses are often a result of incorrect implementation, incorrect use by the intended users of the system, previously-unknown environmental influences of which the original systems designers' were unaware, and so forth.

The DMCA would make such case studies illegal, because doing such a case study might demonstrate how to compromise the security of a system designed to protect copyright content. This criminalizes academic research originally aimed at learning how to -protect- this content.

Cryptography

Cryptography is the science of developing, using, *and breaking* mathematical codes. These codes are used for a large variety of tasks in the modern world, including (a) preventing private communications from being overheard, (b) guaranteeing the authenticity of a communication which might be forged, (c) preventing and/or detecting unauthorized modifications to the stored software and/or user data of a computer system, and (d) controlling the uses to which copyrighted information may be put.

The DMCA is focussed on point (d), but it bears on all other points, as well. This occurs because use (d) is only possible because of the enormous literature accumulated, over the last few decades, in pursuit of points (a), (b), and (c). In other words, using cryptographic techniques for protecting copyrighted content owes a huge debt to the infrastructure—in mathematics, computer software, and individuals trained in their use—developed to aid in confidentiality and authenticity of both stored and transmitted data.

By criminalizing conduct which might compromise use (d), the DMCA makes it all too easy to accidentally commit illegal acts while in pursuit of uses (a), (b), and (c).
In particular, no skilled cryptographer becomes that way only by creating codes and ciphers. Instead, the cryptographer must also have experience in the ways in which others may *break* his or her schemes, and this knowledge is in turn obtained by attempting to break the schems proposed by others. Additionally, the cryptographic community has learned that no new cryptographic system may be trusted unless it has been subject to years of scrutiny by others in the field. In other words, an unexamined system is virtually certain to be weak and easy to break. (The recent cases involving the cryptographic systems in the DVD system are a case in point—the system was developed in secret, and, because of this, was shown to be easy to break when its underlying algorithms were first made public. On the other hand, public algorithms such as the US Data Encryption Standard (DES) have withstood decades of scrutiny, and were easily repaired (by now using triple-DES) when that scrutiny showed that advances in computer power were threatening to undermine it.)

By criminalizing such activities, the DMCA promises that the only good cryptographers will be those who learned their craft outside of the United States. That has obvious implications for US national security.

Academic research

While many who study computer security or cryptography are members of academic institutions, a very large number are not. In particular, the computer security field is immense, and employs thousands of individuals at hundreds of companies which make computer security products. The same is true of cryptographers—many work as independent consultants, for example, and are not members of formally-recognized academic institutions.

This means that an explicit exemption only for people who can demonstrate a particular tie to a particular academic institution would disenfranchise many thousands of individuals from pursuing work that was formerly legal, solely because of their lack of institutional affiliation. Such a discriminatory law is surely not what is intended by the DMCA and is not supported by its legislative history. While the current DMCA is deleterious to academic research, as demonstrated above, exceptions to its reach should *not* therefore be limited only to academics, but rather to anyone who is engaged in computer security or cryptographic research and must therefore be able to both create and attempt to break cryptographic systems—including those which may, under some circumstances, be used to protect copyrighted content.

2. Creating interoperable computer systems

Computer systems must be able to understand each other's data, regardless of whether or not the same entity programmed them. This principle has stood at the heart of creating both computer applications and computer networks for decades.
In many cases, the file formats and network protocol standards which describe the format and structure of the data to be exchanged have been public documents. While this would appear to make life easy for the programmers tasked with making disparate implementations interoperate, it is very commonly the case that the documentation available is either incomplete or outdated. (It is the nature of computer documentation that it is rarely as complete or precise as the program(s) it describes, because, unlike programs—which must run on a particular computer and which do not work if they are incompletely written—documentation is both written and read by humans, and accidental or deliberate omissions are much more likely to go unnoticed until they are needed to instruct a much-less-forgiving computer.)

Because such documentation is often sparse or wrong, many computer programmers have extensive experience in reverse-engineering the function of a program or device written or built by someone else, in order to create a complete and correct description of its behavior so that they may build another program or device that may exchange data with it. The DMCA threatens to outlaw such a practice if the data involved may be copyrighted—regardless of whether the program or device actually *handles* any copyrighted data. All that must be true is that it *may be used* to handle such data *in addition* whatever other use it may have.

In addition to documented (albeit incompletely) interfaces, computer systems engineers also have a history of reverse-engineering undocumented interfaces, again for interoperability reasons. This behavior has been shown to be protected under law—see, for example, the case of Sega Enterp. Ltd. v. Accolade, Inc., 977 F.2d 1510 (1992), in which it was held that Accolade, Inc had the right to reverse-engineer the interface of particular Sega videogames so it could create additional game cartridges for the machine. The DMCA attempts to outlaw such behavior, with the clear result that the diversity of products in the marketplace suffers.

3. Chilled speech

The aforementioned activities are the bread and butter of both academic and industrial computer security and cryptography practitioners. The DMCA threatens to chill their speech, even if a particular practice might fall outside of its scope, because it is rarely the case that an individual researcher will publish—whether on a web page or in an academic journal—research that he or she thinks might lead to a lawsuit brought by the giant corporations that control the vast majority of all copyrighted content in the US today. That enormous resources can be brought to bear on a virtually global scope by such companies is undisputed; one need only notice the recent activities of the DVD Copy Control Association in the DeCSS brouhaha to assure oneself that even simple descriptions of algorithms, even if not personally reverse-engineered by the person responsible for their publications, may easily lead one into being named as a party to very expensive litigation.

Thus, the fear of being hauled into court by a powerful multinational corporation for pursuing what used to be legitimate research may well lead US-based researchers to abandon any research which might possibly arise the ire of such corporations. This has the effect of moving all such research overseas, to venues in which the DMCA does not apply, and seems hardly the outcome that the DMCA is intended to foster.

Respectfully yours,

Eric D. Scheirer
Research Assistant
MIT Laboratory

Leonard N. Foner
Research Assistant
MIT Media Laboratory