

Brewster Kahle
Alexander Macgillivray
on behalf of The Internet Archive
P.O. Box 29244, The Presidio of San Francisco
116 Sheridan Avenue, San Francisco, CA 94129
<<http://archive.org/>>

Lawrence Lessig
Professor, Stanford Law School
Chairman, Creative Commons

Wendy Seltzer
Fellow, Berkman Center for Internet & Society

Office of the General Counsel
Copyright Office, L-M 403, James Madison Memorial Building
101 Independence Avenue, SE., Washington, DC 20024-0400

December 18, 2002

Re: RM 2002-4 – 17 USC § 1201 Exemptions Notice of Inquiry

1) Literary and audiovisual works embodied in software whose access control systems prohibit access to replicas of the works

Thank you for this opportunity to propose exemptions from Section 1201 of the Digital Millennium Copyright Act. Since the time of the first rulemaking we have directly experienced many additional problems with the Section 1201's prohibition on access control circumvention. We appreciate the chance to explain those problems and our proposed solution.

The signatories of this comment are concerned about the effects of Section 1201 on preservation of works controlled by access devices. The Internet Archive, this comment's principal author, is a non-profit library that started by archiving the web in order to provide an historical record to future generations, including those that visit its interactive sculpture in the Library of Congress's lobby. The Archive now provides free access to an enormous and wide-ranging collection of web pages, movies, books, sound recordings and software. We work every day to ensure that the explosion of digital creativity is not lost to history.

By this comment, we propose that “Literary and audiovisual works embodied in software whose access control systems prohibit access to replicas of the works” be exempted from the anti-circumvention provision of Section 1201(a)(1).

1. Class of Works: *Literary and Audiovisual Works embodied in software whose access control systems prohibit access to replicas of the works*

1.1. Summary of Argument

We have been given collections of copyrighted work in formats or on media that are degrading and becoming obsolete with time. To preserve these works, it is imperative that they be migrated to modern storage systems and that the reproductions be accessed to verify the accuracy and completeness of the transfer. This must be done before the formats or media become obsolete or damaged. This use is a non-infringing use protected by Sections 107, 108 and 117 of the Copyright Act. The proposed exemption is important to preserving the availability of the works and enabling archiving, future scholarship, and commentary. Furthermore, the proposed exemption is narrowly drawn and does not damage the market for the works. Without the ability to circumvent access controls in order to access archived reproductions, we have no way of ensuring this essential non-infringing use, and these works will vanish from the historical record of the 21st century.

1.2. Introduction

Preservation of digital works poses a unique challenge because of the quick deterioration of digital media and the rapid obsolescence of digital formats. Though a cave painting in pigment on rock may survive millions of years without any action on the part of archivists, the same is not true of digital works. The same image, created in a digital format such as Portable Network Graphics (“PNG”), on a computer storage device such as a CD-ROM, must be transferred to new media before the CD-ROM degrades and translated to new formats before the

PNG format becomes unreadable by available technology.¹ In this way, the digital record is endangered by any passage of time without its active maintenance. Mere neglect of the proper transfer and translation of these works over time destroys them for all of history.

The Internet Archive attempts to address this problem by soliciting donations, archiving, preserving, and providing access to a wide variety of digital works. Our collection includes more than ten billion web pages by date of access; movies and television programs including the Prelinger Archives and programs that air on the Public Broadcasting System; more than ten thousand texts, including *The Adventures of Huckleberry Finn*, *The Koran*, *A Christmas Carol* and *The Odyssey*; and more than one thousand five hundred recorded musical concerts. The Archive makes these materials available to researchers, historians, and scholars over the Internet.

We frequently receive new donations to the Archive, so our principal limitation is our ability to accept and process new works, not the number of works we are given. For those software works with measures that bar access to replicas of the software, Section 1201(a)(1)'s prohibition against the circumvention of access controls blocks our preservation of the donated software. We have already felt the effects of this prohibition and expect it to continue to prevent us from fully archiving a significant portion of our collection. We therefore respectfully submit that "Literary and audiovisual works embodied in software whose access control systems prohibit access to replicas of the works" should be exempted from Section 1201(a)(1).

1.2.1. Access control

One means of access control that prevents the Internet Archive from preserving works is the "original-only" access controls utilized by many works distributed on removable media,

¹ See TASK FORCE ON ARCHIVING DIGITAL INFORMATION, PRESERVING DIGITAL INFORMATION, May 1, 1996 (describing the problem, providing a good discussion of the processes required to preserve digital works and arguing for heightened use and reproduction rights for certified archives).

including floppy disks and CD-ROMs.² Typically these works do not employ copy protection. Instead they use a form of access control to deny access to any replica of the work in the ordinary course of operation. That is their function. They are not malfunctioning, obsolete or damaged. However, in denying access to replicas, “original-only” access controls and the prohibitions of Section 1201(a)(1) effectively prevent the Internet Archive from accessing its archival replicas and preserving the works.

For example, the Internet Archive has received a very generous donation from Macromedia, Inc. that cannot be fully archived because of “original-only” access controls.³ Macromedia donated over 10,000 software packages containing CD-ROMs and floppy disks storing copies of literary and audiovisual works embodied in software. The collected CD-ROMs contain a wide variety of materials, ranging from a compilation of aviation laws to a cultural guide to the Berkshires, college guides, children’s games and programs with which to create art. Our preliminary efforts to archive these titles has revealed that many are protected by “original-only” access controls. The impaired works include works illustrating the cultural significance of Totem Poles, giving a guide to colleges in 1996, providing study guides, for playing an interactive haunted house game, and for creating of video montages based on the popular television and movie characters “The Mighty Morphin’ Power Rangers.”⁴ The Internet Archive needs to circumvent the “original-only” access controls in order to complete its archiving of these works but is prevented from doing so by Section 1201(a)(1).

² “Original-only” access controls and their circumvention, are discussed at length in Vault Corp. v. Quaid Software Limited, 847 F.2d 255 (1988). In Vault, the Fifth Circuit found that there was no copyright infringement in the distribution of software to circumvent such access controls.

³ See *Macromedia Archive*, available at <<http://www.archive.org/cdroms/cdroms.php>> (providing access to an index of the collection and the ability to browse the collection).

⁴ The titles of these examples are: “Totem Poles - Myths, Magic and Monumental Art on the Pacific Northwest Coast,” “On Campus '96,” “Team SAT,” “I Spy Spooky Mansion” and “Mighty Morphin Power Rangers Create-A-Movie.”

1.2.2. Preservation requires copying **and** access

In order to archive a donated work, the Internet Archive must copy the work onto a different storage medium (typically a computer hard drive) and then verify that the work has been accurately copied by accessing the work in its new location.⁵ As many who have attempted to restore files from un-accessed backups can attest, merely copying a work does not ensure its preservation. The Internet Archive learned this lesson firsthand in archiving government web sites at the request of the Library of Congress in 2000. The Internet Archive had taken great pains to copy each web site accurately with an automated tool, yet when we first gained access to the archive, we found that some sites had not been correctly archived or preserved. Fortunately, because the replicated pages did not contain access controls prohibiting their access, the Internet Archive was able to discover the problem and change its procedures for future pages. We were able to capture later versions of these pages, although we cannot know whether some pages were altered between these two preservation sweeps. Had these works been protected by access controls, however, we might not have discovered the error until too late to make a quality archival record at all.

Access to the archived work is also important to periodic verification that the new medium on which the work is stored has not degenerated and that its format has not become obsolete. Over time, works must be migrated to new media without end. As one of the leading task forces to address digital archiving wrote:

The purpose of migration is to preserve the integrity of digital objects and to retain the ability for clients to retrieve, display, and otherwise use them in the face of constantly changing technology. The Task Force regards migration as an essential function of digital archives.⁶

⁵ This comment is not addressed towards the copy controls identified in Section 1201(b). In some cases, these copy controls may prevent the preservation of a work by preventing its migration to new media. However, even when no copy control prevents preservation, access controls such as the “original-only” access controls described above prevent preservation by preventing the verification of the archival copy.

⁶ *Supra note 1* at iii.

A key characteristic of digital preservation is that, as described above, inaction leads to the loss of the work. If a work is not periodically migrated to new media and formats, the work will cease to be accessible and its accessibility will be unrecoverable. The Internet Archive, for example, has migrated its digital collection twice in just six years. Concurrently, unless access is allowed each time a work is migrated, some works will be lost on each migration. Over time, the cumulative effect of migration losses could leave a gaping hole in the digital record of history that could have been prevented.

This problem has no counterpart in the analog realm because a donor can give the physical work to an archive along with all the rights necessary to preserve it. By contrast, Section 1201(a)(1) reserves to the copyright owner, and not to a mere owner of a copy, the right to authorize access to a work protected by access controls. Thus, in the analog realm, the recipient of J.D. Salinger's letters may transfer them against the author's express wishes to an archivist for preservation, and Salinger is powerless to stop it. In the analog realm of slowly degrading media, that is enough to preserve the work through the eventual expiration of its copyright.⁷ In the digital realm, while a J.D. Salinger CD-ROM, with "original-only" access control, may be donated to the archive by any owner of a copy, it can be archived only with the authorization of the copyright holder. For practical purposes, therefore, archives cannot use donations from legitimate owners of works who do not also hold their copyrights. Section 1201(a)(1) forces us to rely for our historical record on every individual copyright holder, many of whom may be unreliable archivists, such as reclusive authors, defunct companies, squabbling heirs, or politicians shirking accountability for earlier promises. The proposed exemption restores the proper balance by enabling lawful preservation of what an archive has lawfully

obtained. If the exemption is granted, both the Salinger letters and CD-ROM may be archived, if not, absent Salinger's permission, only the letters will survive.

1.3. The Proposed Exemption Allows Non-Infringing Use

The Internet Archives' reproduction of and access to the works in its collection does not infringe the works' copyrights.. The Internet Archive's preservation use is explicitly exempted from copyright liability under Sections 108 and 117 of the Copyright Act governing archiving and libraries. Furthermore, even without Sections 108 and 117, the Internet Archive's use is protected by Section 107's exemption for fair uses.

1.3.1. Section 108

Under Section 108(a), it is not a copyright infringement for libraries and archives to reproduce and distribute a copyrighted work provided that (1) the reproduction is not made for financial gain; (2) the collections are open to the public or available to people doing research in a specialized field; and, (3) a copyright notice is provided. The Internet Archive's reproduction for preservation meets all three of these criteria. The Internet Archive is a 501(c)(3) non-profit organization that does nothing for financial gain, including the reproduction necessary to preserve works. The Internet Archive and all of its collections are open to the public either directly or through a straightforward registration process. The Internet Archive does not alter copyright notices available on works, and will provide explicit notice that works in its archives may be copyrighted. Section 108 was specifically adopted to address these types of archival reproductions and exempt them from liability under the Copyright Act.

⁷ Transfer of physical media allows for archival copying without violation of copyright through the practices outlined in 17 U.S.C. § 108.

1.3.2. Section 117

Section 117(a)(2) allows software owners to make reproductions for archival use. These reproductions do not infringe copyright as long as they are destroyed if the owner loses rightful possession of the original software. Because the Internet Archive is the owner of the software in its collection, makes its reproductions for archival use, and preserves rightful possession of the originals at all times, the Internet Archives archival reproductions do not infringe.⁸ The balance recommended by the Commission on New Technological Uses of Copyrighted Works for computer software explicitly accepted a right to make archival reproductions as valid and non-infringing when it recommended the adoption of Section 117.⁹

1.3.3. Fair Use

Even if Sections 108 and 117 did not apply, the Internet Archive's use is a fair use under the four factor analysis given in Section 107. As the House of Representatives reported in passing the Copyright Act of 1976 with regard to motion picture archivists:

The efforts of the Library of Congress, the American Film Institute, and other organizations to rescue and preserve this irreplaceable contribution to our cultural life are to be applauded, and the making of duplicate copies for purposes of archival preservation certainly falls within the scope of 'fair use.'¹⁰

No less presumption of fair use should be given to the Internet Archive's work in archiving software titles.

Though the Internet Archive reproduces the entirety of a copyrighted work entitled to full copyright protection, both the character of the Archive's use and its effect on the market for the copyrighted work compel a finding that the Internet Archive's use is fair. The Internet Archive's

⁸ For a discussion of permissible archival uses and the requirements of Section 117, see Vault, 847 F.2d at 264-67 (finding that backing up software from media that could be damaged is a non-infringing use).

⁹ See NATIONAL COMMISSION ON NEW TECHNOLOGICAL USES OF COPYRIGHTED WORKS, FINAL REPORT, at 31.

¹⁰ H.R. REP. NO. 94-1476, at 73 (1976).

preservation of creative works has no commercial motive. We are a 501(c)(3) non-profit archive and we believe that archiving is a public service in exactly the way the House of Representatives described. Our copying of copyrighted works and our verification of those copies through access does not bring the Internet Archive any commercial gain. Each copy of a work we preserve is legitimately acquired by the Internet Archive. There is no market for vast majority of these works because they are either obsolete or near obsolescence. In the case of the Macromedia Collection, the CD-ROMs were published in the 1990s for use on 1990s machines and operating systems. Many are incompatible with current platforms. The reproduction of these titles for inclusion in our Archive and access to them for the purpose of verifying the archived replica does not substitute for the purchase of another copy even if one were available.

1.4. The Exemption Satisfies the Statutory Categories of 1201(a)(1)(C)

In passing the Digital Millennium Copyright Act, Congress directed that the Register of Copyrights recommend exemptions to Section 1201(a)(1) based on examination of the impact of any exemption on:

- (i) the availability for use of copyrighted works;
- (ii) the availability for use of works for nonprofit archival, preservation, and educational purposes;
- (iii) the impact that the prohibition on the circumvention of technological measures applied to copyrighted works has on criticism, comment, news reporting, teaching, scholarship, or research;
- (iv) the effect of circumvention of technological measures on the market for or value of copyrighted works; and
- (v) such other factors as the Librarian considers appropriate.¹¹

This exemption increases the availability of copyrighted works; ensures their availability for non-profit preservation, education, scholarship and research; allows a more accurate historical

¹¹ 17 USC § 1201(a)(1)(C).

record to be preserved; does not harm the market for copyrighted works; and, protects against the irreparable harm of their loss. For each of these reasons, it should be adopted.

1.4.1. Availability for use of copyrighted works

Many copyrighted software titles have extremely short time periods of availability. For example, the vast majority of the titles in the Macromedia Collection are not currently available from computer retailers.¹² For these titles, there is no possibility that granting any exemption will decrease their availability. Indeed any exemption that allows others, including the Internet Archive, to preserve these works can only increase their availability.

The platforms required to use copyrighted software have a similarly short life span. This compounds the problem of software availability because even if a software title were made available by its publisher after a long time period it might still be unavailable for use.

The British Broadcasting Corporation's experience with its digital version of the Domesday Book exemplifies the dangers of platform unavailability. The original Domesday Book was compiled in 1086 by King William of England as a comprehensive survey of lands, assets, taxes and ownership. In 1986, the BBC created a digital sequel providing a rich audiovisual record of life in the 1980s. The BBC's digital archive was stored on laser disk and designed to last at least as long as its predecessor. Just sixteen years later, however, the new digital Domesday had become obsolete and could not be read. Only through a very expensive long-term effort has a team from Leeds University and the University of Michigan been able to restore a means to play the digital version.¹³ Such an approach will not scale for the preservation of millions of other copyrighted works.

¹² This assertion is based on an informal sample of a portion of the Macromedia collection.

¹³ See *Digital Domesday book unlocked*, BBC News, Dec. 2, 2002, available at <<http://news.bbc.co.uk/2/hi/technology/2534391.stm>>.

This problem is compounded by the fact that many digital works created today are created in digital format and not produced in any analog format. For these works, the more gracefully decaying analog version is not available if the digital version should become lost or inaccessible. If a software platform for a work available only as embodied in software becomes obsolete, archivists will have no analog copy to which to return and the work will become permanently unavailable. For this reason, it is imperative to preservation of the availability of copyrighted works that their archiving not be thwarted by access controls acting in concert with Section 1201(a)(1)'s prohibition on their circumvention.

1.4.2. The availability for use of works for nonprofit archival, preservation, and educational purposes

As stated above, the proposed exemption of this class of works is directed at improving the availability of works for use for nonprofit archival and preservation purposes. Without access to replicas, no archive can be created in order to preserve literary and audiovisual works embodied in software. These works will not be preserved absent this exemption.

1.4.3. The impact that the prohibition on the circumvention of technological measures applied to copyrighted works has on criticism, comment, news reporting, teaching, scholarship, or research;

Lack of access to replicas of works embodied in software thwarts the preservation of such software. As a result, these source materials will be inaccessible to criticism, comment, news reporting, teaching, scholarship and research in the future. The historical record of these works will be limited to whatever contemporaneous commentary they inspired. Works that did not attract contemporary notice and comment will simply vanish from the historical record. Thus the negative impact on commentary, scholarship and research is strong and irreversible.

1.4.4. The effect of circumvention of technological measures on the market for or value of copyrighted works

This exemption will have no adverse effect on the market for copyrighted works because the Copyright Act will continue to prohibit the creation of illegitimate replicas to be accessed. Section 1201(a)(2) will still prohibit the distribution of devices to aid in accessing any such replicas. This exemption will not affect the rights of copyright owners or the public except for granting access to legitimate replicas under 1201(a)(1).

To the extent the value of a work to the public is considered under this criterion, the proposed exemption increases that value by allowing archivists to assure the availability of a work in perpetuity. Significantly, this public value is drastically increased upon a work's entry into the public domain at the expiration of its copyright. Without the proposed exemption, that value to the public, an essential part of copyright's bargain, may never be recouped for works protected by access controls that prohibit access to replicas.

1.4.5. Other factors

One other factor not directly addressed in Section 1201(a)(1)(C) that should be considered is the scope of harm done by 1201(a)(1) as opposed to the scope of potential harm done by any proposed exception. To the extent that Section 1201(a)(1) does irreparable harm that cannot be cured by the later granting of an exemption to it, that exemption should be granted. Similarly, to the extent that any potential harm a 1201(a)(1) exemption might cause can be reversed in a future rulemaking by declining to extend the exemption, the exemption should be granted.

The proposed exemption is an example of an exemption that meets both criteria. The damage done by the Internet Archive's inability to preserve these works is irreparable. Once a platform has become obsolete, a format no longer supported, or a medium degenerates so that it

is no longer accessible, the later grant of an exemption cannot rescue it. Whatever works have not been archived never will be.

On the other hand, any potential negative consequences of the proposed exemption can be ended by failure to extend its duration. This is true even for works that have been copied and accessed in replica form. Since each circumvention of access controls is a separate potential violation of Section 1201(a)(1), any access to the work after expiration of the exemption on a rulemaking that failed to extend it would be punishable under Section 1201(a)(1) to the same extent that it would have been before any exemption. The asymmetry of the long-term effect of grant or denial of the proposed exemption should lessen the strength of any arguments against its adoption that are not based on irreparable harm and give yet another reason for its adoption.

1.5. Conclusion

Section 1201(a)(1) is currently preventing archivists from archiving software whose access control systems prohibit access to replicas. The effect of Section 1201(a)(1)'s prohibition on circumvention is to cripple archivists and make these works inaccessible for all time. The proposed exemption offers a narrow solution to this problem that does not damage copyright holders, meets each of the established criteria for the rulemaking, and protects an important non-infringing use of these works. For these reasons, we respectfully request that an exemption from Section 1201(a)(1) be granted for "Literary and audiovisual works embodied in software whose access control systems prohibit access to replicas of the works."

A century from now, the forms of creative expression, the technology that transmits such expression, and the intellectual property laws that regulate them may be very different. Many digital formats will have faded from popularity to obsolescence and the vast majority of current

digital storage media will be unusable. As a result many creative digital works may only be preserved as inanimate manuals and unplayable disks – never to be experienced again. Still, digital archivists will struggle to ensure that the creations of the previous centuries are not lost for the next. Unfortunately, software protected by access controls that prohibit access to replicas will not exist to be preserved by next century’s archivist unless today’s archivists are allowed to preserve it. Section 1201(a)(1) prohibits that work. Our proposed exemption would allow it. That is the straightforward import of this comment.