



New York University
A private university in the public service



Tisch School of the Arts
Moving Image Archiving and Preservation Program

721 Broadway, 6th Floor
New York, NY 10003-6807
Telephone: (212) 998-1618
Facsimile: (212) 995-4061
<http://www.nyu.edu/tisch/preservation/>

April 28, 2006

Mary Rasenberger
Policy Advisor for Special Programs
U.S. Copyright Office
By email: section108@copyright.gov

Dear Mary:

On behalf of the Moving Image Archiving & Preservation Program of New York University, I am submitting the attached written comments for the Section 108 Study Group. These are meant to supplement (not to replace) the oral comments I made at the March 16 hearings in DC. These comments incorporate more details about the relevant positions of international professional organizations on preservation copying. These comments also further explain some topics I outlined in our public comments.

Sincerely,

A handwritten signature in purple ink that reads 'Howard Besser'.

Howard Besser, Director
NYU Moving Image Archiving & Preservation Program
<http://www.nyu.edu/tisch/preservation>

General Issues

Underlying Rights and Incidental Capture

Archives engaged in digital preservation need indemnification against copyright violation of underlying rights or incidental capture. A creator of content may either rely on “fair use” or go through the steps of obtaining underlying rights to small components of a work (a small background music clip, a photo hung on the wall behind someone, ...). Or in the broadcast word, they may get 3-year or 10-year broadcast rights to use this underlying material. For example, we are relying upon fair use rights to include the below photograph, taken the day before the DC Section 108 Hearings, showing President Bush greeting Jason McElwain.



President Bush greets Jason McElwain, Associated Press, March 15, 2006

But this photograph also includes Bush holding the image of a mask of McElwain’s face. McElwain has sold the reproduction rights of his face and his story. And copying of the mask’s image violates rightsholder’s copyright. Associated Press probably felt quite comfortable relying upon fair use to distribute this photograph, particularly since it was a newsworthy item. We feel comfortable relying upon fair use to include it in our testimony. We’re not sure whether the Copyright Office feels comfortable relying on fair use to include both the mask and the photograph as part of the publication of our testimony. And could a library or archive that was engaged in preservation copying of the Section 108 Testimony feel comfortable that they would not get sued by the rightsholder of the mask or of the larger photo?

This kind of question comes up all the time in organizations trying to preserve media work. Copyright audits conducted on public television shows by NYU students as part of the NDIIPP grant showed hundreds of underlying rights questions that could impede preservation in just one episode of a program produced by the station trying to preserve the material.

Underlying rights and incidental capture do not necessarily carry on to the organization preserving the full work (which, in effect, becomes a compilation of all the underlying material). So a public TV station may have the legal right to show a work, and their preservation repository may have been given formal ok from the rightsholder of the program to make preservation copies of the work, but the preservation repository be on shaky grounds in making preservation copies because not all the thousands of underlying IP have been explicitly cleared for preservation.

Some European countries apparently grant exemptions to television production companies for what they term “incidental capture” in the course of production. This includes situations such as a poster in the background as the camera follows someone walking down a street, or background music playing in a public spot as a filmmaker tries to capture a street scene or interview someone. We need some similar way of indemnifying preservation repositories for cases of incidental capture or underlying rights.

Amendments to Current Subsections 108(b) and 108(c)

Off-Premises Access to Digital Copies

Though we believe that a limited degree of off-premises access is necessary for cultural institutions to continue to exert their fiduciary responsibility to their funders (usually taxpayers), the comments here solely address issues of a multiplicity of on-premises use. Here we are only addressing the issue of cross-institutional access rather than direct user access.

International audiovisual archiving professional societies have spent many years trying to craft public policy statements that balance rightsholders needs with an individual archive’s needs to handle preservation problems in conjunction with other archives. Public policy statements have been carefully developed by both the International Association of Sound and Audiovisual Archives (IASA)¹ and the UNESCO-based Coordinating Council of Audiovisual Archive Associations (CCAAA).² The International Federation of Film Archives (FIAF) spent half a day (April 26, 2006) at this year’s annual congress trying to finalize a 10-year discussion over details for a balanced policy statement on intellectual property.

One of the most critical issues for all these audiovisual archiving association is the need of each archive to share the results of their preservation work with other archives who would be able to promote the preservation work as part of a public education process. This is crucial for reasons ranging from sharing techniques, to inspiring good preservation in other archives, to exposing other rightsholders to the value of cooperating with preservation efforts, to performing the educative role that public archives must engage in as part of their

¹ IASA Policy guidelines on copyright and other intellectual property rights, available at <http://www.ccaaa.org/membpolicies.shtml>

² CCAAA statement on copyright and neighboring rights, available at <http://www.ccaaa.org/copyright.shtml>

responsibility to the taxpayers. In the analog world, this sharing involves the following types of activities within a sister archive: public screening of a restored print as part of an educative process on cinema (or audiovisual) history and preservation, the inclusion of short excerpts of preservation/restoration work done as part of an exhibition on preservation issues.

We believe that there are some valuable lessons that the Section 108 Study Group can learn from the policy work by these audiovisual archive associations. First, there is an absolute need to share the activities and results of some preservation efforts with other institutions and the general public. In some cases fair use exemptions may allow use of limited content to achieve these goals, but we don't think that fair use will suffice for all such efforts. Second, these archives see themselves as part of what is essentially a closed network, and would like "on-site premises" to extend to members throughout their network. Their network is governed by Codes of Ethics³ that recognize the importance of the rights of rightsholders, and any breach of responsibility would be cause for expulsion from the professional society (and hence, the closed network).

Preservation-Only Exemptions

New exception for preservation of "at risk" digital works

This is a hugely important topic. The Section 108 Study Group has astutely pointed out that the inherent instability of digital materials necessitates up-front preservation activities, prior to deterioration. A major finding of the InterPARES preservation research project is that digital works cannot wait until a later "preservation" stage of their life-cycle for preservation activities to start; if preservation does not start early enough in the life-cycle of a digital work, it will not survive.⁴

Preservation of digital works is highly complex. Even when one is acutely aware of preservation problems at the point of digital creation, it is enormously difficult to preserve a digital work. Even with important works handled by curators aware of preservation problems, a significant portion of a digital work often becomes unretrievable within a decade or two, as these examples illustrate:

- The 1975 Viking Missions to Mars was almost all lost; it had to be re-typed from paper⁵
- A Feb 2006 survey by the UK's Digital Preservation Coalition⁶ found that

³ See, for example, the FIAF Code of Ethics, available at <http://www.fiafnet.org/uk/members/ethics.cfm>

⁴ See InterPARES website (<http://www.interpares.org/>), as well as Section 108 written testimony submitted by InterPARES 2.

⁵ Associate Press. "Coming Coming Soon: A Digital Dark Age?" on CBS News, Jan 21, 2003, <http://www.cbsnews.com/stories/2003/01/21/tech/main537308.shtml>

⁶ Digital Preservation Coalition. **Mind the gap: assessing digital preservation needs in the UK**, Feb 15, 2006, <http://www.dpconline.org/docs/reports/uknamindthegap.pdf>

- more than 70% of respondents said that data had been lost in their organization
- 7% recognised that corporate memory or key cultural material could be lost
- 60% said that their organization could lose out financially
- In a recent judgment, Morgan Stanley had more than \$1 billion awarded against them as a result of their failure to preserve and hand over some documents required by the courts. The Securities and Exchange Commission is also looking at fining the same bank over \$10 million – specifically for failing to preserve email documents.^{7,8}
- From 1993-2001 I kept all the materials for my “Impact of New Information Environments” course in digital form on the web. This included many different kinds of material (from digital videos of pioneering visionaries who spoke to my students, to student papers and powerpoint presentations, to threaded email discussions, ...). Though throughout the 1990s I made every effort to preserve this material in viewable form, it now serves as an important exercise in “digital archeology” for my Digital Preservation course. Many scores of important files from this course can no longer be accessed, and hundreds more cannot be revived without extensive and costly preservation efforts.⁹

Failure to preserve “at risk” material will have a downstream effect on all kinds of research and publishing. Commercial publishers will no longer have articles to publish without upstream scientific data. Entertainment industry rightsholders will have no creative workers if my School can’t teach the next generation of filmmakers using films and other media we have preserved from the past.

What is an “at risk” work?

All digital works are “at risk;” particularly difficult is guaranteeing the accuracy, reliability, and authenticity of a digital work over time. In empirical research involving dozens of multi-year case-studies, InterPares has concluded that preserving a digital work and guaranteeing that the preservation copy is accurate, reliable, and authentic is a hugely complex process (that requires involvement of a preservation archivist throughout the life-cycle of a work).¹⁰

⁷ cfo.com. **Missing Email May Cost Morgan \$10M, Aug 31, 2005,**

http://www.cfo.com/article.cfm/4340197/c_4340354

⁸ Reuters. “SEC may fine Morgan Stanley over e-mail”, via MSNBC online, Aug 30, 2005, <http://www.msnbc.msn.com/id/9128782/from/RL.1/>

⁹ The remnants of this content are at <http://besser.tsoa.nyu.edu/impact/>. A forthcoming article will explain the preservation problems. An earlier article covered some of the initial preservation problems – Besser, Howard. “Difficulties of implementing and maintaining current databases of WorldWide Web Information”, **Revue Informatique et Statistique dans les sciences humaines** 32 (1-4), 1996, pages 11-28.

¹⁰ See InterPARES website (<http://www.interpares.org/>), as well as Section 108 written testimony submitted by InterPARES 2.

Any work that requires specific hardware or software to view it is “at risk”, particularly if that hardware/software environment changes over relatively short time periods. For example, in the 50 years since the beginning of videotape (April 1955), there have been more than 50 video formats that enjoyed a fair degree of market penetration.¹¹ Videotape, audiotape, and other technologies that decay quickly need to be included. These technologies are similar to digital in that they face both deterioration of base materials and the problem of fairly rapid technological obsolescence.

Exemptions should be made for any “fixed format” that has a short lifetime. I would submit that the criteria for judging what constitutes a “short” lifetime should be guided by the length of copyright term. Any format with an average life-expectancy shorter than copyright term (plus extensions) means that works in that format will likely not survive long enough to enter the public domain. The long-term effects of digital works not surviving into the public domain would be unacceptable for both the research and the artistic communities, which rely heavily upon public domain works. Shrinking numbers of public domain works would also affect the many businesses that rely upon public domain works. Not having works survive into the public domain is also constitutionally suspect.¹²

Copies Maintained in a Restricted Archive

Fully “dark archives” are a bad idea. Preserving digital works is very complex, and at no time in the foreseeable future will we be able to feel fully confident that our preservation work is successful. Just one seemingly minor system change 10 years from now could render vast quantities of digital information inaccessible. Without regular access, when a problem with the digital preservation is discovered, it may well be too late to make the works viewable. By then the viewing/rendering software may be unavailable (or not work on a current operating system), links may be corrupted, or files themselves may have lost sub-pieces.

There is much false confidence around preservation in dark archives. Anecdotal evidence suggests that attempts to preserve without access is not as successful as the preservation managers think. For example, Jim Grey of Microsoft Research approached Supercomputer Centers that claimed to have fail-safe storage systems for their astronomical data. When pressed to find things, 30% of the data could not be retrieved.¹³ The California Digital

¹¹ See Vidipax Corporation. Vidipax Videotape Format Guide
<http://www.vidipax.com/format.html>

¹² As Article 1, Section 8 of the US Constitution implies, Congress is permitted to pass laws granting monopoly status to a rightsholder (copyright laws) “for limited times”, after which the work is to become part of the public domain (so that, collectively, public domain works can “promote the Progress of Science and useful Arts”). If works cannot survive until the point at which they can enter the public domain, this would threaten the “progress of science and useful arts” and destroy the careful balance built into laws governing the copyright monopoly.

¹³ Jim Grey, as quoted by Brewster Kahle in the Section 108 Public Hearings, March 8, 2006, Topic 3, available at <http://www.loc.gov/section108/docs/0308-topic3.pdf>

Library trusted 2 major research institutions with a large corpus of preservation data that was corrupted approximately one year after the digital data was handed over to these institutions.¹⁴

We would suggest that the concept of “dim archives” replace the notion of “dark archives”. A dim archive would only allow access to a limited group of users. Enough users to generate the kind of queries that would alert the archive to failures while it’s still possible to correct them. But not enough users to pose a serious threat to rightsholders. A dim archive might limit access to a single university, or to a particular group of researchers (such as members of a couple of professional societies).

New Website Preservation Exemption

General Comments

The Web is part of our shared popular culture, and we need to preserve widespread representative samples of it in order for future generations to understand our current time period. For history & scholarship, everyday media images of the past are crucial. It is only many years later that society begins to recognize the importance of everyday seemingly ephemeral material. For example, the Prelinger collection of educational and industrial films sheds great insights on gender issues that affect everyone who grew up in the 1950s (with educational films on how to date, how to be a good homemaker, job counseling, etc.). And our students are working on restoring old Coca-Cola commercials that have been deemed worthy of entering the Library of Congress collection. Film and television production studios make extensive use of advertisements from 19th and 20th century newspapers to decide on proper clothing, atmosphere, and prices for period-piece films and television works. The Web offers that kind of insight into everyday life and culture that will be important for future individuals, researchers, and media companies who try to understand or portray life in our current time period.

But websites are incredibly difficult to preserve, even when you know what you are doing, and you make great efforts to preserve them (see discussion of Howard Besser’s website dating back to 1993 in section on “New exception for preservation of ‘at risk’ digital works” above). Few individuals or organizations have the know-how (let alone the resources) to attempt to preserve their websites. In addition, most websites contain material authored by someone other than the website owner (such as photographs, icons, short reproductions, blogs, user-contributed reviews, etc.), which raises questions over whether anyone can give the explicit right to an organization to copy the entire website.

International audiovisual archivist associations have crafted policy statements that both recognize rightsholders rights, but also maintain that they need to have the right to copy and preserve material that the public maintains for free. The CCAA presented one such

¹⁴ Because of institutional sensitivity, I cannot publicly divulge which institutions these were.

statement to WIPO this past November.¹⁵ (We would go further in stating that material meant for free public viewing, that passes over public right-of-ways [such as government-regulated spectrum, satellite feeds that pass through the public air, cable that passes through city-owned streets – all of these should enjoy preservation exemptions, because all are part of freely available daily life, a portion of which must be preserved for future generations.)

When a Website opposes crawling

Opt-in for web crawling just won't allow enough content to be harvested to be of use to future generations (and would also run the risk that only a highly skewed set of content survives). Opt-out configurations cannot rely on conventions created for other purposes (such as robot.txt, which was designed to avoid the system load of crawling, not necessarily to oppose web harvesting for preservation). A good example of this problem is whitehouse.gov, which indicates that it should not be crawled.

But even when rightsholders oppose it, saving wide swaths of the Web is crucial to understanding culture and history and to future research & scholarship. The 108 exceptions for audiovisual news recognizes that for certain types of material, preservation should be permitted even if opposed by the rightsholders.

From history, we know that rightsholders cannot be relied upon to preserve the material they own. Individual rightsholders seldom have the means to preserve works, and corporate rightsholders often do not consider this a priority. The purpose of a corporate entity is to generate profits, and there is no reason for that entity to engage in costly preservation unless it perceives that the cost of preservation will lead to increased profits. In the past, many corporate rightsholders have actively opposed public archives' efforts to preserve their content, fearing that those efforts set bad precedents over the monopoly control that copyright grants.

Much that remains of the early stages of each 20th century communications technology content only exists because archives operated at the borderlines of copyright legality, attempting to make preservation copies of media works in violation of the wishes of the rightsholders. Here are two examples:

- It is well known in the film world that for the first 75 years of cinema, rightsholders did little or nothing to preserve their films. In fact, many large rightsholders periodically threatened public archives who were trying to preserve their films. But with the emergence of a videotape aftermarket in the 1980s, many large cinema rightsholders came to public archives begging to borrow the only decent extant copy of their films. And it was not until the 1980s that the Hollywood Studios started their own internal programs to seriously maintain their films.

¹⁵ CCAA statement at the 13th session of the Standing Committee on Copyright and Related Rights, WIPO, Geneva, 21-23 November 2005, available at <http://www.ccaa.org/copyright.shtml>

- The only remaining copies we have of most pre-1980s television news broadcasts were recorded by the Vanderbilt Television News Archive (VTNA). Soon after Paul Simpson began recording nightly television news programs in the late 1960s, CBS objected, and eventually sued VTNA for copyright violation. In 1976, in the midst of a lengthy court battle between CBS and VTNA, Congress passed a law amending Section 108(f)(3) which legalized Vanderbilt's recordings. Many thousands of researchers have relied heavily on Vanderbilt's collection as the only publicly accessible comprehensive collection of US television news over the past 40 years. Paul Simpson's original idea for VTNA was to hold public officials and the news media accountable for past statements they had made, and VTNA has greatly served the public interest in this respect. In addition, both networks and local news channels have come to rely upon VTNA's extensive indexing of each nightly news program to help them find particular past stories. This great national treasure emerged in spite of fierce legal opposition from a corporate rightsholder.¹⁶

Rightsholders will only undertake costly preservation for high-value material. After all, why should they spend on this when the cost of preservation exceeds any potential financial return?) A rightsholder's valuable assets are likely to be a few particular high-value items on a website containing thousands of elements. From a cultural history standpoint, libraries and archives are concerned with preserving the larger context of those thousands of low monetary value items which we know that the rightsholder will not preserve.

Unless a Web preservation exemption similar to 108(f)(3) is crafted, few websites will survive until they reach the public domain, and much of present-day cultural history will disappear. And as the Vanderbilt exemption has shown: there was no significant market harm to the television industry; much of that industry actually benefited from the added value that Vanderbilt brought to them; and that this exemption provided a huge social benefit in terms of: preservation of our heritage, teaching and research into the past, holding public servants accountable for things that they claimed, etc.

¹⁶ Lucas Hilderbrand. *yet-unnamed doctoral dissertation*, New York University Cinema Studies Department, to be filed June 2006.