Journal of Copyright in Education and Librarianship

ISSN 2473-8336 | jcel-pub.org

Volume 3, Issue 3

Cracking the Copyright Dilemma in Software Preservation: Protecting Digital Culture through Fair Use Consensus

Brandon Butler, Patricia Aufderheide, Peter Jaszi, and Krista Cox

Butler, B., Aufderheide, P., Jaszi, P., & Cox, K. (2019). Cracking the Copyright Dilemma in Software Preservation: Protecting Digital Culture through Fair Use Consensus? *Journal of Copyright in Education and Librarianship, 3*(3), 1-23. https://doi.org/10.17161/jcel.v3i3.10267

Cracking the Copyright Dilemma in Software Preservation: Protecting Digital Culture through Fair Use Consensus

Brandon Butler
University of Virginia Library

Patricia Aufderheide
American University

Peter Jaszi
American University Washington College of Law

Krista Cox
Association of Research Libraries

Correspondence regarding this article should be addressed to Brandon Butler, director of information policy, University of Virginia Library, bcb4y@virginia.edu; Patricia Aufderheide, university professor, American University; Peter Jaszi, professor of law emeritus, American University Washington College of Law; Krista Cox, formerly director of public policy initiatives, Association of Research Libraries

Abstract

Copyright problems may inhibit the crucially important work of preserving legacy software. Such software is worthy of study in its own right because it is critical to accessing digital culture and expression. Preservation work is essential for communicating across boundaries of the past and present in a digital era. Software preservationists in the United States have addressed their copyright problems by developing a code of best practices in employing fair use. Their work is an example of how collective action by users of law changes the norms and beliefs about law, which can in turn change the law itself insofar as the law takes account of community norms and practices. The work of creating the code involved facilitators who are communication, information sciences, and legal scholars and practitioners. Thus, the creation of the code is also an example of crossing the boundaries between technology and policy research.

Keywords: Software preservation, software, copyright, fair use, best practices

Copyright and Software Preservation

Software preservation serves two core purposes. It ensures the ongoing availability of the tools necessary to render digital artifacts, from word-processing files to computer-aided design (CAD) drawings. Software preservation also ensures that legacy software is available to researchers interested in studying it. The steps involved in preserving software for each of these purposes implicate copyright law.

Legacy software is becoming an essential tool for accessing digital artifacts for libraries, archives, museums, and other memory institutions. More and more, culture takes digital form, including images, documents, artworks, games, websites, and virtual worlds (Lessig, 2008; Meyerson, 2017). Preserving this work is also preserving our recent past for future research and teaching (Rinehart & Ippolito, 2014; Rothenberg, 1999).

The raw digital form of these artifacts is unreadable to humans; it must be "read" first by a machine running appropriate software, which renders the content in human-readable form. Maintaining the cultural record thus requires preservation not only of individual digital objects but also of operating systems, application programs, and other elements that make up the complex software environments that render digital files. Appropriate hardware or hardware emulators (software programs that simulate a hardware environment) are also required. While specialized software can sometimes extract relatively simple content (plain text, for example) from vintage formats, original software is typically required to ensure the faithful reproduction of a digital file as it appeared to its creator or its original intended audience. Important information can be lost when a digital file is rendered in a software environment other than the original one; software preservation is therefore a necessary feature of digital preservation strategy (Association of Research Libraries, 2018).

Software is also worthy of preservation as an object of study. Researchers have long been interested in the history of science and technology, either for its own sake or for the light it sheds on wider social phenomena, and scholars increasingly are investigating the history of programming and of the software tools that shape culture. Like other kinds of mass culture, software was not initially recognized as worthy of systematic inquiry. Today, however, scholars increasingly find studying legacy software an important part of their research. Fields such as code studies, software studies, and platform studies have grown as research areas in which software is central to the work. Many memory institutions have made software collecting, and with it preservation, an important part of their mission.

The resources necessary for software preservation—including expertise, specialized technology tools, and software itself—are unevenly distributed across memory institutions, and this problem is growing exponentially as software

burgeons. Assuring equitable and efficient access to software and software-dependent materials will increasingly require collaboration and resource sharing. Technologies such as emulation-as-a-service—which makes it possible to present users with an emulated version of a legacy hardware and software environment in a modern web browser running on modern hardware, allowing users to experience legacy software and digital content without access to legacy hardware—make achieving these goals cheaper, easier, and more secure. Community networks such as the Software Preservation Network are promoting interinstitutional cooperation.

Legal uncertainty has been one of the most persistent challenges to software preservation (Ayre & Muir, 2004). Core preservation activities almost inevitably trigger copyright concerns. Copyright law treats software as a "literary work" and protects it on the same basic terms it uses for books and movies. Almost every step in a typical preservation workflow is potentially regulated by copyright, starting with migration to a stable medium, during which software is reproduced—an activity reserved to copyright holders and their licensees unless a limitation, such as fair use, applies. Software may be modified in the process, perhaps creating a derivative work, another activity regulated by copyright. Providing copies of preserved software to researchers (or other institutions) is likely a form of distribution—another regulated activity. Making animated text or graphic elements viewable by the public may be a public performance—also regulated. Even showing a capture of the software in operation could constitute a public display or performance—yet another copyright-regulated act.[1]

The Copyright Act includes a laundry list of narrowly tailored provisions that favor specific groups or uses contemplated by Congress, including an exemption for libraries and archives. That provision, codified at 17 U.S.C. § 108, is a powerful tool that provides a great deal of certainty when it applies. However, it is widely understood that Section 108 falls short of covering the full range of actors and activities involved in modern cultural stewardship (Weston, 2017). For example, the provision excludes museums, which could give pause to institutions like the Computer History Museum that work to preserve a deep and diverse software collection. Other shortcomings of Section 108 include the seemingly arbitrary limit on the number of preservation copies an institution can make (no more than three) and the bar on public access to digital preservation copies beyond the physical premises of the institution.

Libraries and other cultural memory institutions have largely accepted the shortcomings of Section 108 and are increasingly comfortable relying on fair use to cover uses to which Section 108 does not clearly apply (Library Copyright Alliance, 2016). One of the chief advantages of fair use is the way that it can flexibly

[1] The language in this section borrows from the language of the code itself, which was crafted by the authors.

accommodate new technological uses and cultural practices that cannot be anticipated in the formulation of specific exceptions and limitations like Section 108. This makes fair use a perfect fit for software preservation, a practice that was certainly not on the radar as legislators and stakeholders developed the US Copyright Act.

Fair Use and Policy Change in Practice

Many of the legal problems associated with software preservation can be addressed by invoking fair use. Fair use is the US copyright doctrine that permits the unlicensed use of copyrighted material under some circumstances. While fair use is a broad and flexible doctrine that applies in a variety of situations, transformative use (that is, use for a new, socially beneficial purpose without intruding on the market for the copyrighted work) in appropriate amounts (which can include, as in the case of software, the entire work) has become an especially important category of fair use (Aufderheide & Jaszi, 2018). In its pivotal opinion in *Campbell v. Acuff Rose* (1994), the US Supreme Court declared that transformative uses "lie at the heart" of the fair use doctrine. Subsequent empirical scholarship has shown that transformative uses are highly favored by the courts, with the most recent and comprehensive survey concluding that "transformative use is eating the fair use world, and it is doing so more than we previously suspected" (Asay, in press, p. 6). Put quite simply, defendants that courts find to have engaged in transformative uses "nearly always win" and defendants that courts find not to have engaged in transformative uses "nearly always lose" (Asay, in press, p. 7). Software preservation will be well served, then, if the copyright-regulated acts that recur in the context of preservation work can be shown to be transformative. This showing is at the heart of the best practices process described below.

Extensive legal research has demonstrated the reliability and predictability of fair use, and especially of transformative use (Asay, in press; Netanel, 2011; Sag, 2012; Samuelson, 2009). But even when we focus on transformative uses, fair use is still a context-dependent doctrine that requires the assessment of whether an appropriate amount has been used in light of transformative purpose in each situation. In the past some critics have doubted that fair use is an expansive or reliable enough tool to address the many needs of a recombinant culture, especially in a digital era. Rather, they imagine change in law, such as new compulsory licensing or even a sweeping overhaul of copyright law, to be necessary to even begin to restore balance to the copyright system (Benkler, 2006; Lessig, 2004; Litman, 2001; McLeod, 2005). Despite making these dire predictions, even these same critics acknowledge the practical utility of fair use in specific situations (Lessig, 2008; McLeod, 2005, p. 155).

However, in many communities of practice fair use has often been at best a tool of last resort in spite of its potential utility. Many professionals have learned their field in a permissions-only culture, and express doubt and concern about employing fair use (Aufderheide, Milosevic, & Bello, 2016). Their doubt and concern has been deepened by media and software companies' aggressive antipiracy publicity, which often demonizes any act of copying (Gillespie, 2009; Patry, 2009; Rosenthal, 2015). Codes of best practices in fair use offer a bridge between community values and the seemingly abstract world of fair use, helping communities overcome fear.

The Code of Best Practices in Fair Use for Software Preservation is the latest in a series of best practices codes created in similar ways over the last 15 years. In 2004, Patricia Aufderheide and Peter Jaszi began working with communities of practice, starting with documentarians, to create codes of best practices in fair use that are tailored to individual user communities' needs and values. The success of the first code (available at cmsimpact.org/documentary) led to the creation of a dozen others (Aufderheide & Jaszi, 2018). All codes have been created with membership organizations in the relevant field and have often involved other facilitators as well, including the current authors.

Creating codes of best practices has been an important step in making it easier for any single professional to take action because they can do so knowing that their colleagues typically agree (Falzone & Urban, 2010). One critic has expressed doubts about the value of such consensus documents, believing, among other reasons, that the codes might be overcautious (Rothman, 2007). In fact, codes of best practices produced over the last 15 years have expanded access to fair use and made it easier for insurers, managers, lawyers, and others in positions of power to accept the legitimacy of fair use assertions by practitioners in the field (Weeramuni, 2019; Aufderheide & Jaszi, 2018, pp. 138–156; Donaldson, 2010).

This is to some an improbable result because codes of best practices are not legal documents, although they are reviewed by an advisory board of legal experts. These codes are summaries of major consensus points about interpretation of fair use in certain routine situations for the field. Their strength may partly lie in the fact that they are created within formal organizations of professionals. As legal scholar Michael Madison noted in an extensive study, judicial decision-making draws heavily on expectations for cultural practice (Madison, 2004). Best practices embraced by leading professional organizations are a powerful indicator of such expectations.

Recent codes of best practices in fair use, including the Code of Best Practices in Fair Use for Software Preservation, have followed a relatively consistent plan of development. The process is fourfold. First, a team of researchers interviews appropriate and experienced veterans in the field, using both social science in the

form of surveys and ethnographic methods in the form of observational and long-form interviews, to discover the kinds of routine practices in which fair use options might arise as well as to find out how the field currently manages those options. Second, after the results are circulated to the field facilitators meet with deliberative groups of professionals recommended by the field's organizations to discuss how the group believes fair use would appropriately be implemented in those situations, particularly with what limitations—that is, boundary points beyond which fair use would seem inappropriate to professionals. Third, the draft document is reviewed by a legal advisory board that confirms the norms in the document are consistent with existing law. Finally, sponsoring and partnering organizations circulate it to the field and solicit endorsements from related organizations (Aufderheide & Jaszi, 2018).

This project also intersects with the academic discussion around public interest activism, policy, and scholarship (Napoli & Aslama, 2011). A central question in this area is how scholarship appropriately intersects with advocacy. Policy is often construed as focused on government action, but as the range of private actors in policy processes grows, governance often shifts to other domains (Braman, 2003, pp. 6, 14; DeNardis, 2014). The domain in this instance is community practice. Scholars often find it difficult to do research relevant to policy because, as Braman noted, they need to understand how decisions are actually made, know the relevant points of entry, and understand the pertinent technologies (Braman, 2003, p. 25). In this case, researchers investigated the actual conditions of software preservation in relation to copyright, thus ensuring they did understand the process of decision-making within the community. They understood, principally through previous research, the relationship of copyright exceptions to the problems that were discovered. Finally, they were able to match solutions to problems with the help of software preservationists who best understood the relevant technologies and the mission of their institutions and their field.

Methodology

Creating the code first involved grounding the project in research to determine current practices around copyright challenges in software preservation and their implications for software preservation's goals. This research was conducted by researchers from the Association of Research Libraries (Cox), the University of Virginia (Butler), and American University (Aufderheide and Jaszi). Jessica Meyerson, research program officer at the Educopia Institute and community cultivation advisor to the Software Preservation Network (SPN), a professional membership organization of 20 institutions whose website describes its mission as "advanc[ing] software preservation through collective action," advised on issues and literature (Meyerson, 2018), arranged for interview contacts, and commented on

results. However, as is typical in this process, she had no control over the researchers' conclusions.

Research prior to convening community members to deliberate on the terms of the code was conducted in two stages: first, a literature review informed by parallel efforts at SPN (Meyerson 2018) and second, interviews with 41 members of the software preservation community. The interviewees were selected by a combination of contacts from SPN and snowball sampling. Interviewees were either professionals directly involved in software preservation or those whose work involved software preservation. All had more than 10 years of experience in the field.

The interviews were hour-long, open-ended telephone, Skype, or Zoom conversations that were guided—but only guided—by a simple protocol:

1) Please describe your job's responsibilities.
2) Please take us through the process of executing one (or more) of those responsibilities relating to software.
3) Please tell us how you address copyright challenges at relevant points in this process.
4) Please tell us how the way you address these challenges accords, or not, with your expectations or standards for the best way to do software preservation in this (these) case(s).
5) Who else would you recommend we speak with?

In the next phase the researchers became facilitators of confidential, deliberative discussions among community members. The facilitators convened six groups in physical locations and two groups virtually via Zoom video-conferencing software. Each group ranged between 7 and 10 members.

Each group was presented with a version of a complex scenario:

> Your institution is new to digital preservation, but you have been exhorted by your Board to make "big splash" in the field. So you recently have made your first two major acquisitions. One consists of papers, digital media, and computers that belonged to the late Bella Brown, a visionary architect and hypertext author active in the late eighties and nineties. These records, which include all remaining documentation relating to Brown's famous project "Imaginary Cities" (unseen since a gallery installation in 1993), and the complete working files (including variants) for her 1991 work of "interactive electronic speculative fiction," *A Place in the Void*, are subject to a comprehensive deed of gift from Brown's estate, giving your institution "all rights." The software environments in which Brown did her path-breaking

work appear to have been an amalgam of off-the-shelf operating systems, rendering programs, and other specialized software tools and additional custom code written by her and various collaborators. The other acquisition was an eBay purchase: a large private collection (at least a 1,000 items, though there has been no systematic inventory) of commercial business software (including shareware) from 1980-1990 on manufacturers' floppy disks – many previously undocumented. A majority of firms that originally produced this software long ago disappeared from the industry, many without a trace. Only some retain their original packaging; your sampling suggests that many been sold with licenses authorizing only "individual use by the purchaser, on 1-5 computers located at a single residence or business premises."

You would like to do the following with these collections.

1. Produce faithful versions of the contents of all the digital media, along with images of all associated documentation to your servers as soon as possible,
    a. Does it matter if there is a clearly defined relationship between the collections and your institution's stated mission?
    b. Are there any other conditions you would impose on this activity, assuming the financial wherewithal were available?
2. Place a permanent preservation copy of each such record in the new secure National Digital Depository maintained by a Digital Preservation Consortium of trusted institutions.
3. Analyze the retained preservation copies as thoroughly as possible to assess the nature of the collections and their contents more completely, and to generate metadata or other cataloging information.
4. Run as many as possible of the program files you have acquired, using the original media and available legacy hardware, to create archival videos representing of the screen displays that they generate and illustrating the human-machine interactions characteristic of them.
5. Make these videos available to scholars and/or the general public.
. How broadly (scholars v. general public)?
    a. How openly (on request v. on line)?
    b. How extensively (part v. whole)?
    c. With what redactions (if any?)
6. Make available on site server-based environments using modern hardware configured so that the various programs in the collections can be run as realistically as technologically possible.

   a. Hardware emulation v. emulation with OS and other versatile softwre installed?
   b. Limits on access (qualified persons v. all comers)?
   c. User agreements?
7. Participate in an online network managed by the Digital Preservation Consortium, subject to appropriate security measures, in which participating institutions will share the environments they have created?
   Limits on access (qualified persons v. all comers)?
   a. User agreements?
   b. Other security measures
8. Enable access to these environments for unaffiliated scholars and the general, on site and/or on-line.
9. Distribute virtualized versions of some of Bella Brown's art works, including "authoritative" versions of "Imaginary Worlds" and "A Place in the Void."

Our review of some of the leading literature in the field demonstrated that the software preservation community had already identified copyright as the primary area of legal uncertainty affecting their work (Rosenthal, 2015). Practitioners showed a relatively sophisticated understanding of the scope of protection afforded to software; they typically recognized not only that software is subject to copyright protection but also that preservation activities likely implicate copyright. Many writers expressed concern about software licenses, and some also flagged the Digital Millennium Copyright Act (DMCA) provisions relating to digital rights management (DRM) as a barrier to preservation. Though one scholar suggested civil disobedience, contract reform (which in some sense amounts to seeking permission), and resisting the permissions culture (Kraus, 2011), most solutions proposed in the literature revolved around seeking permission, perhaps through a blanket license collectively negotiated by libraries with a group of major software publishers (Engle, 2016; Rosenthal, 2015). Some expressed hope that Section 108 might somehow protect digital activity that resembles traditional library uses, such as a "virtual" reading room (Engle, 2016)—a hope that cannot be realized given the narrow confines of Section 108.

Our interviews included professionals in the following areas: digital preservation, archives, software studies, museum administration, digital humanities, computer science, and library information technology. Most informants were practitioners working in cultural memory institutions, but some were academics who rely on access to software for their research and others were administrators of institutions or departments where software preservation takes

place. There was some geographic concentration on the coasts due to the relatively high specialization involved in software collection and preservation, but there was a great diversity in kinds of institutions, from large R1 libraries to small independent museums. Although they may be essential to the software preservation ecosystem, we deliberately did not include amateur collectors and others who may have, and even preserve, old software but are not clearly focused on supporting research and learning. This choice was driven by the conviction, drawn from previous experiences in creating codes of best practices, that the community of practice was most usefully defined by reference to this shared professional context and mission, which helps underwrite the fair use reasoning process.

We found some commonalities across diversity of place, position, and mission. Software preservation professionals unanimously told us that copyright was a significant barrier to the fulfillment of the preservation mission. They usually understood copyright issues within licensing challenges; they rarely considered fair use as an option in their work. Their confusion about copyright led them to choices that limited their ability to fulfill that mission.

Generally, interviewees sought security in licensing. Most did not know that the "shrinkwrap" license terms for software did not necessarily apply to their institution, and they often did not know the licensing terms at all. They assumed that the harshest terms they were familiar with were probably the default for all licenses, although they rarely actually investigated the terms of licenses in their collections. Interviewees reported problems with "orphan works," legacy software programs whose owners cannot be identified or found. Even when putative owners could be found, the owners were often unsure of their own rights and were unwilling to grant permissions. Interviewees also worried about violating the anticircumvention provisions in the Digital Millennium Copyright Act, which bar tampering with "technological protection measures" (TPMs) that may prevent the accessing of software. Legacy software can often be encumbered by TPMs, such as encryption that requires a license key or the use of authentication servers. These keys and servers are frequently missing or out of commission for older software.

Interviewees often did not have anyone to turn to for good advice. Even when they had access to general counsel at their institutions, they found that attorneys in that office are unlikely to be experts in copyright law and are therefore most likely to urge caution.

The approaches interviewees used to manage copyright risk included:

- Limiting themselves, in a few cases, to acquiring and storing original media and documenting the behavior of the software when run on legacy hardware.
- Avoiding the need to preserve legacy software for access by migrating the contents of legacy files to new formats compatible with contemporary off-the-shelf systems. However, this was rarely feasible.

- Making copies for selection and description as well as for storage in a dark archive, but not for patron use or access.
- Letting a few vetted scholars study software on-site using original media running on donated or purchased legacy hardware, but limiting how they reference those materials.
- Providing on-site access via emulation run at a workstation but not offsite access via emulation-as-a-service.
- Making archived legacy software openly available online using emulation (this strategy was used only outside mainstream institutions).

As a result of these choices, which were driven by copyright anxiety, mission failures understandably often ensued, such as the following:

- Items could not be inventoried fully because the required software tools could not be accessed due to copyright concerns.
- Archivists could not diagnose errors or understand what may be missing or wrong in a digital archive without access to legacy software.
- Museums hesitated to collect digital artwork that relied on third-party software because of the concern that the work would become inaccessible once the required software was no longer available or supported on contemporary machines.
- Collections remained stuck in one physical site, useless to remote researchers.
- Migrating digital objects compromised their integrity, in some ways rendering them unfaithful to their original appearance or behavior.
- Funding for preservation was denied because of copyright concerns.
- Even on-site, scholars were frequently denied effective access.
- The software preservation community was poorly situated to innovate (or adopt) the next generation of new technological approaches.

Thus, the software preservation community was experiencing a high level of frustration in their activities due to perceived copyright barriers. This result was reported to the field and to the general public (Aufderheide, Butler, Cox, & Jaszi, 2018).

Code of Best Practices

After conducting eight discussion groups of librarians, archivists, curators, and related professionals in six cities and two national video conference calls, the authors distilled the consensus into a set of principles and limitations for the responsible exercise of fair use in software preservation.

Many preservation professionals have had or continue to have close relationships with commercial software production and distribution. Perhaps because of this deep cultural relationship, preservation professionals' reasoning reflected a double concern: to do their jobs well and to avoid interfering with developers' and publishers' current markets. However, preservationists clearly understood how different their work was from the activities of the commercial marketplace, and they were confident in the transformative nature of that work.

In early discussions informants spent substantial time grappling with the concept of transformative use and articulating the differences in purpose between preservation (including enabling access for research and teaching purposes) and normal commercial exploitation. The key realization for these early groups was that while the uses may be "the same" at the most literal level, this is true for *all* fair uses. For example, both an author and the critic of the author's work reproduce the author's words so that others might read and understand them, but the critic's purpose is more complex: the words are to be read *as evidence in a critical analysis*. The transformative use analysis must take into account an appropriately rich description of the secondary user's purpose, or else no use would be transformative.

Once the early groups had this realization, they saw that while the ordinary consumer and the software preservationist (or researcher) may both use Word 95 to open Word 95 documents, the preservationist/researcher's purpose is more complex: the documents are being opened *as part of accessing the digital past for research or teaching.* And like the critic, the archivist (or librarian, or museum curator) has a special social role distinct from the role of the original software publisher that requires the use of copyrighted materials in ways that are not, and typically cannot be, served by the ordinary commercial market. As we have seen, that is why these uses are "at the heart" of fair use. After the first two groups independently arrived at this conclusion, facilitators introduced the idea early on in other groups, who quickly embraced it and moved on to more specific applications of the concept.

The groups were particularly conscious of the lack of incentives for preservation in the marketplace. They also strongly believed in taking reasonable measures to protect against substitutional uses (where a user accesses preserved software in order to use it for its original purpose—to create new computer-aided designs, for example—rather than for research and study). They felt strongly that this would both help ensure their uses were transformative and protect the overall software ecosystem.

Confusion about the intricacies of intellectual property law continued into the group discussions. Frequently facilitators had to disentangle fair use issues from trademark, contract, and privacy law. This demonstrated not only non–legal experts'

understandable confusion about a confusing area of law but also a high degree of concern not to undertake unacceptable risk.

Further, preservationists frequently worried about taking any action that could enable a patron's inappropriate use of materials. This concern appeared to draw from both risk-avoidance and a deep understanding of the commercial field of software development. In fact, a patron's choice to use materials inappropriately is as much their own responsibility in an archive as it is anywhere else. But collections professionals typically take measures of various kinds—including labels, warnings, approval processes for access, limited physical access, and technical protection measures—to ensure that users access collections materials only for the purposes for which the archive is dedicated. This demonstrates their good faith. Software preservationists' deliberations also reflected this concern and approach.

Preservationists eventually identified five situations in which fair use might be relevant, drawing in part on the research previously conducted:

- Accessioning, stabilizing, evaluating, and describing digital objects.
- Documenting software in operation and making that documentation available.
- Providing access to software for use in research, teaching, and learning.
- Enabling broader networked access to software maintained and shared across multiple collections or institutions.
- Preserving files expressed in source code and other precompiled human-readable formats.

In each case preservationists proposed a set of limitations that effectively defined how fair use could be applied in typical situations. For example, in the first situation about intake of materials, the full set of limitations included the following:

- Preservation activities should be related to the overall institutional mission.
- The preservation of donated materials should be undertaken in light of the terms of donor agreements, which may limit reuse and access to donated materials.
- Reasonable care should be taken at this stage to identify software objects with sensitive content such as personal data, trade secrets, or national security issues, as these matters trigger legal and ethical obligations that are not overcome by fair use.
- Descriptions of preserved objects should be created, expressed, and shared in ways reasonably designed to facilitate discovery by interested researchers within and, where possible, beyond the institution.
- Access to software (including disk images) for preservation purposes should be limited to personnel (including staff, volunteers, and contractors or vendors, whether at the collection's home institution or at a partner

> institution or entity) engaged in the intake, description, and long-term preservation storage process, either on premises or in secure off-site environments.

In its appendices the code addresses legal areas adjacent to copyright and fair use, which may make preservation more difficult. For instance, many software preservationists worry that licenses, including the "shrinkwrap" licenses similar to terms of service, could prevent them from preserving software. These licenses often appear to prohibit someone from making a copy, or from making uses that are not "personal." However, licenses do not necessarily have the dire consequences many practitioners fear. For example, a license is only binding for those who are party to the contract (courts use the term "in privity" to describe this group), and users who obtain software through donations or other secondary transfers may not be in privity to the contract, depending on the circumstances. The stakes may also be lower because copyright's notoriously high statutory damages (fines as high as $150,000 per work if the court finds infringement was "willful") are not necessarily available to plaintiffs who merely claim breach of contract. In some cases breaching the terms of a license does bring statutory damages into play because it nullifies the license and unlicensed use may constitute infringement (*MDY v. Blizzard*, 2010). However, if the underlying use is fair, then a license is not required. As Section 107 tells us, fair use is "not an infringement" of copyright; permitting use without a license is the purpose of fair use. The standard remedy for a breach of contract is an award of actual damages, that is, compensation for the harm caused by the breach. Given the dim commercial prospects for obsolete software, the actual damages associated with breach of an old license should approach zero. Rights holders who realize the modesty of the expected remedy should hesitate to embark on costly litigation.

The Digital Millennium Copyright Act of 1998 includes provisions, codified at 17 U.S.C. 1201, barring circumvention of technological protection measures (such as encryption), and a violation of these provisions may not be subject to a fair use defense. One circuit has argued that a "nexus to infringement" is required in order to find liability for circumvention under Section 1201 (*Chamberlain v. Skylink*, 2004), while others have held that fair use is not a defense to a claim of unlawful circumvention (*MDY v. Blizzard*, 2010; *Universal Studios, Inc. v. Corley*, 2001). In the face of uncertainty, software preservationists often erred on the side of caution, assuming fair use would not be available as a defense. Obeying the decryption ban can have crippling effects on software preservation attempts (McDonough et al., 2010). Luckily, and as a result of the software community's heightened awareness of the importance of exceptions and limitations to their work, software preservationists petitioned the Copyright Office for relevant exemptions. In 2018,

the Library of Congress granted exemptions permitting circumvention when it is reasonably necessary for software preservation. The rules are fairly complex, but the Cyberlaw Clinic at Harvard Law School has written a helpful guide to the rules for lay practitioners (Lee & Albert, 2018). In addition to permitting circumvention in covered circumstances, the rule has another salutary effect; under Section 1201(a)(1)(D), the Librarian of Congress can grant an exemption only if the underlying behavior is "lawful." The grant of an exemption therefore gives credence to the underlying arguments that software preservation is not copyright infringement.

Discussion

The Code of Best Practices in Fair Use for Software Preservation creates conditions to change law in practice by drawing upon community consensus. It also demonstrates the success, for a particular kind of project, of collaboration between academic researchers and professionals.

The code provides extensive help to the preservationist by stating reasons that can guide decision-making without restricting the preservationist to rigid guidelines or checklists. Such rigidity would be antithetical to the logic of fair use because it strictly confines choice and freezes that choice within a particular historical moment. Rigidity is likewise not reflected in the statute, nor is it evident in case law (*Campbell v. Acuff-Rose*, 1994). It would undermine the proven distinctive value of the fair use approach over more specific limitations and exceptions to copyright. For instance, neither the VCR nor internet search was imagined when the doctrine of fair use was first articulated in the United States in 1841, but fair use enabled both innovations. In many countries without fair use, implementation of both was either slowed or made more difficult by lack of this doctrine (Aufderheide, Pappalardo, Suzor, & Stevens, 2018).

The guidance provided by software preservationists in the code is strongly marked by their concern to recognize the rights and business positions within the commercial industry. It also contains clauses that go beyond the requirements of fair use case law. For instance, preservationists repeatedly called for attribution or credit in their limitations. The fair use statute at 17 U.S.C. 107 makes no mention of attribution, and the issue has never been a particularly important factor in the case law. The US version of moral rights embodied in the Visual Artists Rights Act (codified at 17 U.S.C. 106A) is the closest thing to an attribution right in US law, but it falls far short of other countries' laws and has no effect on fair use. However, attribution is often a useful demonstration of good faith, which is an intangible but valuable element in any lawsuit and therefore valuable in deterring legal action.

The preservationists also called for attention to donor agreements. Donor agreements, as contracts, can create obligations that override fair use

considerations (though savvy institutions are careful to avoid this). Note that donor agreements, unlike consumer software licenses, are negotiated specifically between donors and collecting institutions, so the privity and harm arguments that mitigate risk in the case of software licensing are unavailable in the context of donor agreements. Uses permitted by fair use can still constitute breach, with serious legal and political consequences. Donor agreements give a clear indication of donors' wishes; donors are often software developers or software companies themselves, and collecting institutions value their relationships with these communities. Finally, like the librarians, archivists, and others who have developed previous codes of best practices involving collections (Association of Research Libraries, 2012), software preservationists showed concern about and referenced in their limitations noncopyright concerns such as privacy, trade secrets, and national security (as shown in the aforementioned example).

Thus, the code takes great care to stay well within the accepted terms of fair use rather than stretch it to its limits as well as to engage in broadly lawful and defensible practices. At the same time, software preservationists created in this code a clear articulation of the unique value of preserving the enabling elements of a digital culture. They asserted the legitimacy of their social role and practice relative to commercial exploitation of software.

Relationships between the academic researchers and the professionals were uniformly cordial and productive, as the professionals perceived the researchers throughout to be working in their interest. Academics faced no challenge to ethical standards in conducting research because the professionals did not establish the research design or conduct the research. Since the first part of the research—identifying patterns in current behavior in the field—was informed by a comprehensive literature review and preliminary interviews with practitioners, the researchers were fairly confident they would not encounter results that could be potentially negative to their partner. Since the strength of the second part of the research—creating the terms of a code of best practices in fair use—depended on the opinions of the professionals about appropriate limits within the law, researchers did not experience a conflict between the interests of the professionals and the academics in either methods or results.

The policies of software preservationists, already relevant to several fields of research, will only grow in importance. Software preservationists hold the keys to the memory of digital culture.

Limitations of the Code

The Code of Best Practices in Fair Use for Software Preservation has clear limits, even in applying fair use to software preservation problems. First, it covers only those areas where there was clear consensus in the groups. For instance, there

was disagreement about how to provide access to games given that many users simply want to replay them. (While the research team anticipated the difficulty of treating games with the same transformative rationale as utilitarian software, they explored the issue with early discussion groups to confirm this suspicion.) Second, while the specificity of the limitations describes generally the logic of fair use, it applies only to the most common situations in which current fair use principles are likely to apply. Fair use may well apply far beyond these situations, especially as both law and preservation continue to develop, but facilitators in discussion groups asked only about common situations, and they restricted the formation of principles and limitations to those where consensus existed.

The code also does not apply to expressive works that are built with software, such as games, electronic literature, and digital artworks. The logic of the limitations tied to the five situations described in the code applies to utilitarian software that ceases to have its ordinary commercial utility as technology advances. It might or might not apply to a particular expressive work. However, digital creators may well find other codes—such as those for filmmakers, nonfiction authors, and visual arts professionals[2]—relevant to digital expression, since fair use is platform-agnostic.

The Code of Best Practices in Fair Use for Software Preservation is also frozen in time in one crucial way. Its situations are grounded in the assumption that physical or downloadable copies and/or static iterations of code (e.g., 1.0, 2.0) exist. However, software vendors' business model is changing to a cloud computing or software-as-a-service model, with more and more customers purchasing access to software that runs on servers maintained by the provider rather than on local hardware.

Without reliable access to static distributed copies of future software releases, memory institutions may be unable to create and maintain a stable record of them, regardless of their fair use rights. Libraries have relied on donations from collectors or their own purchased physical copies to build the software collections they currently hold. Capturing software stored on vendor servers raises legal and political issues that institutions have yet to address—most importantly the body of law that bars unauthorized access to servers, including the Computer Fraud and Abuse Act and torts like trespass to chattels. It is unclear how this situation will be addressed in the future; it is unlikely it can be addressed without some form of commercial-industry involvement. This creates even greater urgency around the ability of software preservationists to articulate their unique role in the software ecology.

The kind of relationship between academic and professional that drove this research is also limited. The findings in this research are applicable specifically to

[2] All these and other codes are available at cmsimpact.org/fair-use.

software preservation. Other organizations and professions may wish to replicate this process themselves, as a dozen others already have, in order to tailor the interpretation of fair use for their particular practices. Furthermore, the mere existence of the code itself does not improve practice—it only provides an opportunity. Barriers to implementing the code may be not only ignorance but also reluctance to be perceived as going against the grain of tradition or irritating a potential donor, an administrator, a general counsel, or another entity in a position of power. Thus, while this work is directly aligned with policy change, it in itself cannot change overall copyright policy. It can only encourage change in the behavior of the relevant professionals, which can change the meaning of existing policy for that group.

Effectiveness

Will developing the code change any behavior? In the past, such codes of best practices have changed insurance practices, created new publishing policies, enabled new artworks, and made available for remote access major archival collections (Aufderheide & Jaszi, 2018).

Certainly the code has been embraced by the organizations with widest reach in the software preservation community and the institutions that house it. The lead organization in facilitating it, the Association of Research Libraries, anchors the field of libraries. Endorsers include the Society of American Archivists, the lead professional association for archives; the Software Preservation Network, a young but vital organization of preservationists that focuses on institutional challenges; the American Library Association (ALA), the largest association of libraries in the United States; and the Association of College & Research Libraries, the division of ALA comprising higher-education libraries.

The field has shown engagement with the code. A seven-part webinar series immediately filled with a hundred participants was recorded and is being transcribed to become a general reference. Meanwhile, there has been a complete absence of pushback from any software-industry company or association. In the Copyright Office DMCA rulemaking process, industry groups generally endorsed software preservation and some ultimately supported the grant of an exemption (US Copyright Office, 2018).

If previous experience is any guide, the biggest challenge for adoption will be entrenched practices, policies, and habits within institutions that have routinely discounted, postponed, or delayed work that cannot be fully licensed. However, it was widespread frustration with achieving even basic goals that drove the field to search out and deliberate together to create a code of best practices in fair use.

Conclusions

The process of creating the Code of Best Practices in Fair Use for Software Preservation demonstrates the power of active relationships among communication scholars, legal scholars, and field practitioners. None of the participants had the full range of knowledge to be able to draw a relevant conclusion from the question, How does fair use apply to software preservation? Collaboration in good faith made possible a result that, historical precedent suggests, may be able to change preservationists' behavior in ways that permit better preservation without impinging on current markets. Those changes are arguably policy changes that occur outside traditional policy realms.

Without software preservation, culture in a digital age is simply made invisible. Software preservation's capacities are intimately linked with copyright policies. Fair use enables the fulfillment of software preservation's mission and also benefits the wider society in both the short and long run.

References

Asay, Clark D. (in press). Is transformative use eating the world? *Boston College Law Review.* Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3332682

Association of Research Libraries. (2012). *Code of best practices in fair use for academic and research libraries*. Retrieved from Center for Media & Social Impact website: https://cmsimpact.org/code/code-best-practices-fair-use-academic-research-libraries/

Association of Research Libraries. (2018). *Code of best practices in fair use for software preservation*. Retrieved from Center for Media & Social Impact website: cmsimpact.org/software

Aufderheide, P., Butler, B., Cox, K., & Jaszi, P. (2018). *The copyright permissions culture in software preservation and its implications for the cultural record*. Retrieved from Association of Research Libraries website: https://www.arl.org/wp-content/uploads/2018/02/2018.02.09_CopyrightPermissionsCulture.pdf

Aufderheide, P., & Jaszi, P. (2018). *Reclaiming fair use: How to put balance back in copyright* (Second edition). Chicago: University of Chicago Press.

Aufderheide, P., Milosevic, T., & Bello, B. (2016). The impact of copyright permissions culture on the US visual arts community: The consequences of fear of fair use. *New Media & Society*, *18*(9), 2012–2027 https://doi.org/10.1177%2F1461444815575018

Aufderheide, P., Pappalardo, K., Suzor, N., & Stevens, J. (2018). Calculating the consequences of narrow Australian copyright exceptions: Measurable, hidden and incalculable costs to creators. *Poetics, 69*, 15–26. https://doi.org/10.1016/j.poetic.2018.05.005

Ayre, C., & Muir, A. (2004). *Right to preserve? Copyright and licensing for digital preservation project: Final report*. Retrieved from Loughborough University website: https://dspace.lboro.ac.uk/2134/343

Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. New Haven, CT: Yale University Press.

Braman, S. (2003). *Communication researchers and policy-making*. Cambridge, MA: MIT Press.

Campbell v. Acuff-Rose, 510 U.S. 569 (1994).

Chamberlain v. Skylink, 381 F.3d 1178 (Fed. Cir. 2004).

Copyright Act of 1976, 17 U.S.C. §§ 101 et seq. (1976).

DeNardis, L. (2014). *The global war for Internet governance*. New Haven, CT: Yale University Press.

Donaldson, M. (2010). Fair use: What a difference a decade makes. *Journal of the Copyright Society of the U.S.A.*, *57*(3), 331–335.

Engle, E. (2016, January 22). Intellectual property rights issues for software emulation: An interview with Euan Cochrane, Zach Vowell, and Jessica Meyerson [Blog post]. Retrieved from https://blogs.loc.gov/thesignal/2016/01/intellectual-property-rights-issues-for-software-emulation-an-interview-with-euan-cochrane-zach-vowell-and-jessica-meyerson/

Falzone, A., & Urban, J. (2010). Demystifying fair use: The gift of the Center for Social Media statements of best practices. *Journal of the Copyright Society of the U.S.A.*, *57*(3), 337–350.

Gillespie, T. (2009). Characterizing copyright in the classroom: The cultural work of antipiracy campaigns. *Communication, Culture & Critique*, *2*(3), 274–318. https://doi.org/10.1111/j.1753-9137.2009.01039.x

Kraus, K. (2011). "A counter-friction to the machine": What game scholars, librarians, and archivists can learn from machinima makers about user activism. *Journal of Visual Culture, 10*(1), 100–112. https://doi.org/10.1177/1470412910391576

Lee, K., & Albert, K. (2018). *A preservationist's guide to the DMCA exemption for software preservation*. Retrieved from Software Preservation Network website: http://softwarepn.webmasters21.com/1201-exemption-guide-for-software-preservationists/

Lessig, L. (2004). *Free culture: How big media uses technology and the law to lock down culture and control creativity*. New York: Penguin Press.

Lessig, L. (2008). *Remix: Making art and commerce thrive in the hybrid economy*. New York: Penguin Press.

Library Copyright Alliance. (2016). Statement of the Library Copyright Alliance on the Copyright Office's notice of inquiry concerning Section 108 of the Copyright Act. Retrieved from Library Copyright Alliance website: https://www.librarycopyrightalliance.org/documents/united-states-documents/section-108-of-the-copyright-act/lca-issues-statement-on-the-copyright-offices-notice-of-inquiry-concerning-section-108-of-the-copyright-act/

Litman, J. (2001). *Digital copyright: Protecting intellectual property on the Internet*. Amherst, NY: Prometheus Books.

Madison, M. J. (2004). A pattern-oriented approach to fair use. *William and Mary Law Review*, *45*(4), 1525–2425.

McDonough, J., Olendorf, R., Kirschenbaum, M., Kraus, K., Reside, D., Donahue, R., . . . Rojo, S. (2010). *Preserving virtual worlds: Final report*. Retrieved from Illinois Digital Environment for Access to Learning and Scholarship website: https://www.ideals.illinois.edu/handle/2142/17097

McLeod, K. (2005). *Freedom of expression: Overzealous copyright bozos and other enemies of creativity.* New York: Doubleday.

MDY v. Blizzard, 629 F.3d 928 (9th Cir. 2010).

Meyerson, J. (2017, November). *Software preservation network: The access breakdown and system level change*. Paper presented at the ADE Summit, Library of Congress, Washington, DC. Retrieved from http://digitalpreservation.gov/meetings/ade/slides/session4_speaker2_Meyerson.pdf

Meyerson, J. (2018). *Software preservation network lit review*. Retrieved from osf.io/qdsyc

Napoli, P. M., & Aslama, M. (2011). *Communications research in action: Scholar-activist collaborations for a democratic public sphere* (First edition). New York: Fordham University Press.

Netanel, N. (2011). Making sense of fair use. *Lewis & Clark Law Review*, *15*(3), 715–772.

Patry, W. (2009). *Moral panics and the copyright wars*. New York: Oxford University Press.

Rinehart, R., & Ippolito, J. (2014). *Re-collection: Art, new media, and social memory*. Cambridge, MA: MIT Press.

Rosenthal, D. (2015). *Emulation & virtualization as preservation strategies*. Retrieved from Andrew W. Mellon Foundation website: https://mellon.org/resources/news/articles/emulation-virtualization-preservation-strategies/

Rothenberg, J. (1999). *Ensuring the longevity of digital information*. Retrieved from CLIR website: www.clir.org/wp-content/uploads/sites/6/ensuring.pdf

Rothman, J. E. (2007). The questionable use of custom in intellectual property. *Virginia Law Review*, *93*(8), 1899–1982.

Sag, M. (2012). Predicting fair use. *Ohio State Law Journal*, *73*(1), 47–91.

Samuelson, P. (2009). *Unbundling fair uses*. Retrieved from Berkeley Law website: https://scholarship.law.berkeley.edu/facpubs/1521/

Universal City Studios, Inc. v. Corley, 273 F.3d 429 (2nd Cir. 2001).

US Copyright Office (2018, April). *Copyright Office Section 1201 roundtable hearing transcript April 12, 2018*. Roundtable held at the James Madison Building, Washington, DC. Transcript retrieved from copyright.gov website: https://www.copyright.gov/1201/2018/hearing-transcripts/1201-Rulemaking-Public-Roundtable-04-12-2018.pdf

Weeramuni, L. (2019). How to fight fair use fear, uncertainty, and doubt: The experience of one open educational resource. *Journal of Copyright in Education and Librarianship*, *3*(1), 1-21. https://doi.org/10.17161/jcel.v3i1.9751

Weston, C. (2017). *Section 108 of Title 17: A discussion document of the Register of Copyrights*. Retrieved from copyright.gov website: https://www.copyright.gov/policy/section108/discussion-document.pdf.