CODE OF BEST PRACTICES IN FAIR USE FOR

SOFTWARE PRESERVATION

SEPTEMBER 2018

CMSI
CENTER FOR MEDIA & SOCIAL IMPACT
SCHOOL of COMMUNICATION
AMERICAN UNIVERSITY • WASHINGTON, DC

AMERICAN UNIVERSITY
WASHINGTON
COLLEGE OF LAW
PROGRAM ON
INFORMATION JUSTICE AND
INTELLECTUAL PROPERTY

ASSOCIATION
OF RESEARCH
LIBRARIES

CODE OF BEST PRACTICES IN FAIR USE FOR SOFTWARE PRESERVATION

SEPTEMBER 2018

COORDINATED BY

Association of Research Libraries
Center for Media & Social Impact
Program on Information Justice and Intellectual Property

FACILITATION BY

Patricia Aufderheide, University Professor
Research Fellow, Center for Media & Social Impact
School of Communication, American University

Brandon Butler, Director of Information Policy
University of Virginia Library

Krista Cox, Director of Public Policy Initiatives
Association of Research Libraries

Peter Jaszi, Professor Emeritus
Program on Information Justice and Intellectual Property
Washington College of Law, American University

FUNDED BY

The Alfred P. Sloan Foundation

AN AFFILIATED PROJECT OF

The Software Preservation Network

TABLE OF CONTENTS

September 2018

OVERVIEW

This is a code of best practices in fair use, describing the ways that fair use can be useful to software preservation in common, recurring contexts.

Fair use is the right given in U.S. copyright law to use copyrighted material without payment or permission, under some circumstances. A long pattern of judicial decisions applying Supreme Court precedent shows that an assessment of fair use typically depends on the answers to two questions:

- Is the use transformative—is the purpose for which preexisting copyrighted material is reused different from that for which it was originally created?
- Is the amount of material used appropriate to the purpose of the new use?

If so, it is likely that fair use applies. A fuller explanation of fair use law is in Appendix One.

This Code was made by and for the software preservation community, with the help of legal and technical experts. It provides librarians, archivists, curators, and others who work to preserve software with a tool to guide their reasoning about when and how to employ fair use, in the most common situations they currently face. It does not provide shortcuts in the form of prescriptive "guidelines" or rules of thumb. Nor does it seek to address all the possible situations in which software preservation professionals might employ fair use, now or in the future.

SOFTWARE PRESERVATION AND COPYRIGHT

For libraries, archives, museums and other memory institutions, legacy software is becoming an essential tool for accessing digital artifacts. More and more, our culture takes digital form, including images, documents, artworks, games, websites, and virtual worlds. The raw digital form of these artifacts (the "1s and 0s") is unreadable to human readers; it must be "read" first by a machine running appropriate software, which renders the content in human-readable form. Maintaining the accuracy and authenticity of 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 full and 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.

Utilitarian software tools also are worthy of preservation in their own right, as objects 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. Computer science researchers are also interested in studying software in order to understand previous programming design patterns and methods. Like other kinds of mass culture, software was not initially recognized as worthy of systematic inquiry. Today, however, scholars in diverse fields find studying legacy software an important part of their research, and 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. Most cannot support a comprehensive library of software, expert staff, and related tools. 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 could make achieving these goals cheaper, easier, and more secure. And professional groups such as the Software Preservation Network are promoting inter-institutional cooperation.

One of the most persistent challenges to software preservation has been legal uncertainty. Practitioners fear that legal structures developed to regulate

software in the commercial marketplace (like restrictive licenses and so-called "anti-circumvention" rules) somehow may impinge on their work. They also know that core preservation activities almost inevitably do trigger copyright concerns. 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." Software may be modified in the process, perhaps creating a "derivative work," while providing copies of preserved software to researchers (or other institutions) can represent a "distribution." Making animated textual or graphic elements viewable by the public may be a "public performance," and showing a screen capture could constitute a "public display." Each of the words and phrases in quotation marks appears in the list of exclusive legal rights of a copyright holder, raising the prospect that engaging in these activities without permission may be copyright infringement.

While courts have broadly blessed preservation activities in a handful of cases, no court decision so far has provided useful, detailed guidance for preservation undertaken in support of teaching and research. Nor do special provisions in the Copyright Act addressing archiving in some way, such as Sections 108 and 117, begin to reach the range of activities that make up preservation practice.

But while preservation professionals' copyright concerns are not misplaced, their fears may be overstated. As this document shows, better understanding of the copyright doctrine of fair use can empower librarians, curators, and others to move forward confidently to accomplish their preservation mission. Such understanding also can help professionals to see copyright-related concerns such as licensing and anti-circumvention in a broader and less threatening perspective.

HOW THIS DOCUMENT WAS MADE

This document represents the consensus judgment of experienced professionals working with legacy software at a variety of institutions across the U.S. It is organized by common situations in which fair use is available to enable core preservation practice. It was created using a three-part process, which has been used successfully by over a dozen other communities of practice since 2004.

First, 40 seasoned practitioners were interviewed at length about copyright problems and concerns encountered in their efforts. Those interviews provided useful information about how the shared disciplinary values around preservation inform the work of librarians, archivists, curators, scholars and others. They also revealed an overarching theme of frustration with the growth of a "permissions culture." Many professionals mistakenly assume that the only safe path through the copyright thicket is to obtain express permission from a copyright holder for virtually all preservation activities. At the same time, they recognize that permission would often be impossible to get. This and related conclusions were documented in a white paper, *Copyright and Permissions Culture in Software Preservation and Its Implications for the Cultural Record*, published by the Association of Research Libraries and available at their website.

Second, eight discussion groups of professionals, convened in six cities and on two national video conference calls, to deliberate together at length about scenarios drawn from the first-stage interviews. A consensus built up over the course of these meetings was then distilled into a set of principles and limitations for the responsible exercise of fair use in software preservation.

Finally, a draft of these principles and limitations was circulated to an advisory board of legal experts for vetting, to confirm this consensus was within the realm of reasonable legal interpretation. The result is this Code of Best Practices in Fair Use for Software Preservation.

WHAT THIS CODE IS AND IS NOT

This Code addresses only activities undertaken to preserve and provide access to software for teaching and research, including work in the burgeoning field of software studies. These activities are most commonly supported by research libraries, archives, and museums, and are often associated with universities, although some of the institutions involved may be collecting software in service of another primary goal. The professionals with whom we spoke were unanimous that fair use rights should be equally available to all good-faith practitioners who share the same core values, including government agencies and private firms.

This document does not map the entire scope of fair use in software preservation. Fair use is available far beyond the situations in this document. But because it is a consensus document, it addresses only the applications of fair use that are the most common and about which there is substantial agreement in the field, rather than all possible or plausible ones. The principles below are as technology-neutral as possible, but new technology will inevitably expand the possibilities for software preservation. Practitioners should feel free to use their own judgment as they apply the broad principles of fair use beyond the scope of the Code or in emergent circumstances.

The Code does not provide specific guidance on challenges related to software licensing, or to the clauses in the Digital Millennium Copyright Act that forbid unauthorized decryption. These two issues have to be dealt with separately, but a solid understanding of fair use will be a powerful resource in addressing them. Appendices to this Code describe these issues and suggest directions for future action.

Nor does the Code address situations where software that is the focus of preservation effort is not subject to ordinary copyright restrictions, because the issue of fair use doesn't typically arise in those cases. These include government software, open source software (which may be freely shared and reused under appropriate circumstances) and software deposited by its creator under statutory or institutional mandates (e.g. for purposes of transparency or reproducibility). Where a particular open license does not authorize a preservation activity, however, that activity may still be permissible under fair use according to the principles below.

The software preservation community reached consensus around fair use principles for any software, whether or not the copyright owners are known or findable. Fair use applies equally to works where the owner is known and unknown. But if you are particularly concerned with "orphan works" (where the owner cannot be identified or located), you can consult the *Statement of Best Practices in Fair Use of Collections Containing Orphan Works for Libraries, Archives, and other Memory Institutions*, available at the Center for Media & Social Impact website among other places.

Software preservation is often closely intertwined with questions about digital preservation of content such as electronic literature, art, and games. This Code focuses on considerations unique to preserving utilitarian software and software environments. A paradigm case is preserving software applications and associated operating system and plug-in elements necessary for creating and viewing text documents, or computer-aided design files. The Code's guidance does not apply directly to digital content such as interactive art and videogames. The principles in this Code are often grounded in expert opinions about the characteristics of markets for software, including typical commercial lifespan and the likely substitutional value of older titles vis-a-vis currently commercially available ones. These considerations may not apply in the same way to cultural expression. That said, where such works do share relevant characteristics with utilitarian software, the principles below may provide guidance.

To address questions about preserving digital content not directly or fully addressed here, as well as questions about how digital objects (once preserved) can be employed in cultural practice, it may be helpful to consult other codes of best practices. For example, the *Code of Best Practices in Fair Use for Academic and Research Libraries* (cmsimpact.org/libraries) addresses preservation of a wide range of content types stored on deteriorating and near-obsolete formats, creating digital exhibits, and teaching with primary materials. Other documents, available from the Center for Media and Social Impact (cmsimpact.org), provide guidance to scholars, filmmakers, journalists, teachers, visual artists and other creators interested in quoting from in-copyright works. The reasoning in these documents applies fully to electronic literature, software art and video games, among other kinds of digital content.

Finally, we note that, like all fair use best practices codes based on community consensus, this one reflects the mission and values of a particular community, as carried out through its distinctive practices and techniques. Although the document does not represent a negotiated agreement between representatives of different stakeholders, the discussion among librarians, archivists, curators, and scholars that shaped this Code was pervaded with concern for safeguarding the legitimate interests of all, including software

developers and publishers. Indeed, many of the limitations to each principle are grounded in these professionals' awareness of the legitimate market interests of software creators and publishers. They were concerned to conduct preservation and access activities in a way that supplements and supports, rather than supplants, copyright holders' important role in the digital ecosystem.

PRINCIPLES AND LIMITATIONS

This section describes a series of situations in which librarians, archivists, curators and others working to preserve software can employ fair use. It describes the activities, states the principle informing the choice to employ fair use, and makes clear the limitations of such use—that is, the outer bounds of the community consensus at this time. Overall, the situations are organized according to the extent of access to preserved materials that they entail.

Throughout, professionals' reasoning reflects a double concern: to fulfill their own mission and to avoid interfering with developers' and publishers' current markets. Preservation professionals are confident their work is not a substitute for the market activity of copyright holders, and that it serves a new social purpose. Indeed, it is the unique province of these professionals to stabilize and document legacy software and the digital materials that depend on software for access.

Software preservationists therefore understand the uses of in-copyright software in the following situations to be transformative—to serve a different purpose than the works' original market purpose. They also believe that taking reasonable measures to protect against substitutional uses in their work is important.

As to the second question that courts typically ask—is the amount used quantitatively and qualitatively appropriate in light of the transformative purpose? — the appropriate unit for software preservation uses generally will be the whole work, although there may be exceptions, as detailed below. While literal limits on the amount may be inappropriate in the preservation context, many of the principles below include limitations on time, location, nature of use, and user. All these limitations have the effect of tailoring the use so that it is proportional to its transformative purpose.

SITUATION ONE: ACCESSIONING, STABILIZING, EVALUATING, AND DESCRIBING DIGITAL OBJECTS.

Software often comes to collecting institutions on unique, fragile, or obsolete storage media. Creating multiple copies is a well-established best practice in digital preservation. Therefore, software preservationists typically transfer the contents of fragile media by creating disk images. They also reproduce original associated materials such as packaging and documentation so that they can be stored alongside the associated software in stable digital formats. This intake process also typically involves evaluation and description, which may require running the software; it also may require use of legacy software environments (operating systems, drivers, and other elements) on which preserved software depends. At the end of this process, stable images are typically stored in multiple, diverse locations, which may include third-party digital repositories. Fair use protects these most basic activities, as they are essential to the core mission of software preservation, with appropriate limitations.

PRINCIPLE:

Fair use protects the internal preservation and documentation activities of software collections, subject to the following limitations.

LIMITATIONS:

a. Preservation activities should be related to the overall institutional mission.

b. Where materials have been donated, their preservation should be undertaken in light of the terms of donor agreements, which may limit reuse and access.

c. Reasonable care should be taken at this stage to identify software objects with sensitive content such as personal data or national security issues, as these issues trigger legal and ethical obligations that are not overcome by fair use.

d. Descriptions of preserved objects should be created, expressed, and shared to facilitate discovery by interested researchers within and, where possible, beyond the institution.

e. At this stage of processing, 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) who oversee or are engaged in the intake, description, and long-term preservation process, either on premises or in secure off-site environments.

SITUATION TWO: DOCUMENTING SOFTWARE IN OPERATION, AND MAKING THAT DOCUMENTATION AVAILABLE.

To facilitate richer understanding of legacy software, and in some cases to capture otherwise inaccessible aspects of software functionality, a collecting institution may wish to create records of a program in operation. Videos, screenshots, and other documents capture important aspects of software operation that may be difficult to convey in textual metadata or to experience in emulation. Examples include the appearance of software operating on original hardware, or its behavior in conjunction with original input devices or as controlled by an expert user. Such documentation may even be, in some cases, the only available option to adequately preserve the content and texture of a work. Accompanying commentary by experts, creators, and others can add more value and provide context researchers may not otherwise discover. Fair use ensures that software collections can be fully documented in a variety of formats and contexts to preserve information that would otherwise be lost.

PRINCIPLE:

Fair use applies to the production and circulation of documentation of software functioning, subject to the following limitations.

LIMITATIONS

a. Such documentation should provide appropriate historical, technical, and other context where available.

b. The extent of the documentation should be proportional to the documentary purpose.

c. Special care should be taken in providing broad access if it could depress public demand for documentation available from authorized sources.

SITUATION THREE: PROVIDING ACCESS TO SOFTWARE FOR USE IN RESEARCH, TEACHING, AND LEARNING.

Researchers need access to legacy software. Providing access using only original media and hardware could place valuable resources at risk and create needless limits on access. The mission of memory institutions extends to providing access, either on-site using physical terminals or (increasingly) by means of remotely accessible online technologies such as emulation. In both settings, multiple interoperable programs can be run together in complex software environments. By definition, legacy software cannot be obtained in the commercial marketplace, so its availability for study and use depends upon memory institutions. Likewise, commercially-available rendering tools may not faithfully represent digital objects originally created in now-obsolete formats. Accordingly, experts in software preservation have taken a leading role in developing technologies to make access possible, while including features that can prevent corruption, diversion, or other unauthorized uses of preserved software. Fair use can help to realize new opportunities for the study of software and digital content in controlled environments.

PRINCIPLE:

Fair use applies to providing controlled access to software in support of research, teaching, and learning, with the following limitations.

LIMITATIONS

a. Individuals granted access to collection software should be notified that access is provided for teaching or research purposes, and they personally are responsible for ensuring that any further uses are lawful.

b. Where a preservation institution intends to provide only controlled access, it should take appropriate measures to limit the possibility of users copying or otherwise retaining software.

c. Access to commercially-available software should be restricted to minimize impact on ordinary commercial sales. Access may, for instance, be enabled on a case-by-case basis for limited purposes not served by commercial offerings, such as data verification and reproducibility studies, subject to the user's affirmative agreement to reasonable terms and conditions. Another approach could be to limit access to commercially available software to local terminals that limit how the software can be used or copied.

SITUATION FOUR: PROVIDING BROADER NETWORKED ACCESS TO SOFTWARE MAINTAINED AND SHARED ACROSS MULTIPLE COLLECTIONS OR INSTITUTIONS.

All collecting organizations, whatever their size and resources, benefit by establishing cooperative arrangements through which limited resources can be shared to make collections more widely available. No one institution can maintain a collection of software, software environments, and supporting resources (such as expert staff or specialized technology,) sufficient to facilitate access to all of the digital resources that may be of interest to its researchers. Sharing resources, including materials, facilities, and expertise, is a core value and practice in cultural memory institutions, and digital preservation in particular has benefited substantially from collective efforts. Sharing resources online has the additional promise of freeing access to collections from the physical premises and facilitating research by remote users. In addition, it has the potential to create new opportunities for collaborative research. Currently, the most promising technology for facilitating cooperative use of legacy software is Emulation as a Service (EaaS), which allows collections and, potentially, cooperating organizations to make software available to remote users in their web browsers, to study the environments themselves or to render other legacy digital content.

Within a well-organized cooperative effort grounded in shared mission, such technologies have a clear transformative purpose. They serve a need that is unique to memory institutions, one that the ordinary consumer market has not addressed. This principle extends the reasoning of Situation Three: fair use supports individual institutions providing access to software environments for research, and so it supports institutions combining their resources to serve more researchers more effectively. Cooperatives can use EaaS and related technology tools to maintain a high level of control over what authorized users can do with emulated environments, imposed either by agreements (e.g., in a terms-of-use agreement binding on each user), or by technological means (e.g., configuring the platform to prevent the relevant actions). These limits can help ensure that the extent of the use is commensurate with the transformative purpose.

PRINCIPLE:

Fair use applies to institutions making software available on a cooperative basis to broaden research opportunities, including off-premises access using technology such as Emulation as a Service, subject to limitations below.

LIMITATIONS

a. To ensure collective activities are conducted within the scope of institutional missions, the roles and responsibilities of member institutions in a cooperative arrangement should be specified in a Memorandum of Understanding or other agreement.

b. Participating institutions should set policies about how and when to extend access to their own affiliated teachers and researchers, mindful of the need to safeguard the legitimate interests of software owners and vendors. For instance, they may adopt measures to discourage users from building private software libraries, or to prevent substitution for commercially-available tools (such as word processors or computer-aided design programs).

c. Likewise, any cooperative effort should develop policies to evaluate requests for access from non-affiliated researchers, and grant them subject to appropriate safeguards.

d. Participating institutions should establish and publicize a mechanism for registering and following up on concerns expressed by software developers, publishers, and other stakeholders about the availability of specific software programs or environments within the network.

SITUATION FIVE: PRESERVING FILES EXPRESSED IN SOURCE CODE AND OTHER HUMAN-READABLE FORMATS.

Research in fast-growing academic fields such as software studies, critical code studies and platform studies depends on access to source code, the human-readable format in which software is written. Source code can reveal important information, for example, about the process of software creation. Some developers release their code to the public for study and reuse, and some donate copyright or grant broad licenses along with code to collecting institutions, enabling public access and even reuse. These common items pose special problems, however, when special arrangements for public access and use have not been made. This is because, unlike compiled object code designed to run in particular hardware/software environments, legacy source code can be more easily adapted for use in new commercial software that may compete with a copyright holder's own offerings. Thus, providing unrestricted access to source code in collections might facilitate unfair appropriation, or contravene trade secrecy. Fair use ensures the longevity of source code, and access for study and teaching under appropriate conditions.

PRINCIPLE:

Fair use applies to the preservation of source code, and to making it available for research use, subject to the limitations below.

LIMITATIONS

a. Restrictions expressed in donor agreements should be strictly observed. Where the donor was the author or publisher, the agreement provides especially critical information about their wishes for future access and use.

b. In most cases, source code should not be made available to the general public online, and access should be treated similarly to requests for unpublished manuscripts.
c. As appropriate, researchers' access to proprietary content not related to the research inquiry should be limited, by redaction or otherwise.
d. As reasonably possible, attribution of authorship and ownership should be provided along with any files made available.

EPILOGUE: THE FUTURE OF SOFTWARE PRESERVATION

Since the beginnings of the software industry, programs have been commercialized through the distribution of physical copies, whether loaded on install media like floppy disks or CD-ROMs, or by means of downloads that result in a file present on the user's hard drive. Increasing numbers of memory institutions have been able to acquire such copies and build collections that incorporate them. The principles and limitations detailed above will continue to be relevant as long as such collections continue to exist. In the future, however, there will be important new challenges to the preservation functions around which this Code is organized, and to the goal of making the record of software development and digital artifacts with software dependencies accessible for teaching and research.

The business model of the software vendors is changing, with more and more customers purchasing access to software that runs on servers maintained by the provider, rather than on local hardware. The perceived commercial advantages of this "cloud computing" or "software-as-a-service" model include the abilities to update programs continuously rather than through the periodic release of new versions, to deal more effectively

with software vulnerabilities, and to exercise greater control over users' activities.
As a practical matter, without reliable access to complete 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. The Library of Congress will, of course, continue to be in a position to demand copies for its collection under 17 U.S.C. Sec. 407. But, the Library's mission does not extend to the broad dissemination of information about its holdings to researchers.

Other collecting organizations may therefore be well advised to negotiate with vendors to assure that this cultural heritage is safeguarded. As documented in the *Permissions Culture* white paper, however, experience so far suggests that such negotiations are ineffective when undertaken by individual institutions on an ad hoc basis. Prospective agreements between vendors and collections may offer the best hope of preserving software distributed in this way, but new approaches may be needed to make such agreements possible.

We hope that broader discussion of the issues addressed in this Code may spark a cooperative public-private discussion of why software preservation matters, and how it can be furthered by voluntary action—before it is too late.

APPENDIX ONE: THE FAIR USE DOCTRINE AND PRESERVATION PRACTICE

The goal of US copyright law is to promote the progress of knowledge and culture. Its best-known feature is protection of copyright owners' rights, but the law includes protections for the public, too. Copying, quoting, recontextualizing, and reusing existing cultural material are critically important to creating, spreading, and preserving knowledge and culture, so the law strikes a balance between rightsholder control and public access and reuse.

This balance is part of the social bargain at the heart of our copyright law. Creators get some exclusive rights in new works, not as an end in itself but to encourage them to produce culture. At the same time, copyright protection is limited to reflect the interests of the law's primary intended beneficiary—the public. The public includes the current and future generations of creators, who may want to refer to or invoke copyrighted culture; scholars and students who need access to culture as part of their research and teaching; and the librarians, archivists, and curators who collect and preserve culture for current and future study.

The public interest limits on copyright begin with the fact that copyright lasts for a limited time. After that, works enter the public domain and are free for use by all. Even so, the duration of protection stretches for generations. So other limitations allow the use of works that are still protected by copyright without permission from or payment to the copyright owner. Fair use is the most flexible and widely applicable of these.

FAIR USE IS A RIGHT WITH A LONG HISTORY.

The "right of fair use" (as the Copyright Act describes it) is grounded in the

First Amendment, and it has been part of US copyright law since at least 1841. Section 107 of the current Copyright Act specifically provides that "fair use . . . is not an infringement of copyright." In litigation, fair use is invoked as an "affirmative defense," like other expressive liberties. In everyday life, fair use is exercised and experienced routinely as a right we all enjoy.

FAIR USE IS FLEXIBLE.

The law describes fair use in general terms, so that it can adjust to evolving circumstances. How can a judge tell when a use is fair? The law says judges must consider at least four factors: the purpose and character of the new use, the nature of the work used, the quantity and quality of the portion that was used, and the effect of the use on the market for the work. This balancing test is sometimes referred to as an "equitable rule of reason" because it uses all the facts and circumstances to evaluate whether each new use has social or cultural benefits that outweigh the cost imposed on the copyright owner.

FAIR USE IS CONTEXT SENSITIVE.

The balancing approach means that it is important to look at each use in light of its overall or ultimate purpose, rather than in isolation. This is why, in our conversations with practitioners about software preservation, we always explored why a particular copyright-related action (copying, distributing, and so on) was important as a part of the overall preservation and access mission, and what preservation goals each of the core practices described in the principles above helps achieve. A common mistake that can lead practitioners in all fields to underestimate their fair use rights is viewing their actions in isolation— as just "making a copy" or "running the software," rather than part of a larger process with a new, positive purpose. Courts have taken this contextual approach in declaring reverse engineering to be a fair use: Reverse engineering requires copying of protected works, but it yields new knowledge and often new products, to society's benefit. In judging fair use, the full story—especially the ultimate purpose—matters.

FAIR USE IS PREDICTABLE.

Choices about whether to exercise fair use always involve judgment, but over the past 25 years, some clear expectations have emerged. Courts have established that usually the most important question about the fairness of fair use is whether the use is "transformative"—whether it "adds something new, with a further purpose or different character," as the Supreme Court put it in the 1994 case *Campbell v. Acuff-Rose Music*. Since then, cases have shown that a "transformative" use does not have to—in fact usually does not—literally change or revise the original material. Using that material *in a new context* different from that of its ordinary market is what "adds something" and makes the use transformative. The opposite of a transformative use is a substitutional one—a use that merely offers consumers a copy, or a portion, or a version, of the work itself. Understanding the transformative use concept makes fair use much easier to understand and predict.

FAIR USE IS CONSISTENT.

Fair use is flexible and context-sensitive, not arbitrary. Fair use treats similar uses similarly. Once you have established that fair use applies to your use of software in a particular context, that same logic applies the next time you do it. In this way, fair use can become part of daily practice, and practitioners can rely on it to protect them consistently from case to case.

FAIR USE IS SHAPED BY "TRANSFORMATIVE USE."

This concept influences how all of the statutory fair use factors are applied. It is in relation to the "purpose and character" of the use (the first factor) that the "nature" of the copyright work is assessed. Because the preservation and study of software differs so profoundly from distribution to consumers, the second factor, the "nature" of the work—usually considered the least significant in isolation—has little relevance here. Analysis of the third factor (the appropriate amount) looks to whether the amount of material used was appropriate in light of the user's transformative purpose. Some say using less favors fair use, but courts disagree. Courts have said that using the entire work can be appropriate, while even seemingly small

portions can be excessive if the use is not transformative. In software preservation, the appropriate amount often is the whole work.

And what about the effect on the market, the fourth factor? Since fair use means that the user will not be paying a license fee, won't there always be an adverse market effect, since the rightsholder will be getting less money? Not when a transformative purpose takes the use out of the primary and secondary markets for the work. Courts have recognized, repeatedly, that copyright owners are not entitled to control the "transformative markets" for their works. For example, search engines copy the entire contents of millions of copyright-protected websites every day in support of an extremely lucrative commercial business, but courts have ruled again and again that they have no negative effect on the traditional, reasonable, or likely-to-be-developed markets for websites. The software preservationists who came together to make this Code, many of whom had strong personal ties to the software industry, were acutely aware of the markets for software, and showed great concern—reflected in the document—for them.

A few additional points to consider:

THE PRESERVATION NARRATIVE MATTERS.

The importance of transformativeness in fair use reasoning makes it useful for those exercising the right to understand their own reasons for doing so. They need to know what the new function, purpose, or context of their use is, and why they are using the amount they are. This can be done formally, for instance by keeping notes, or informally. The ability of users to explain clearly what they were doing and why has been decisive in many fair use cases. In the unlikely event that preservationists receive a request to "cease and desist," their ability to explain their own fair use rationale is an extremely helpful deterrent to litigation. After all, lawsuits are expensive and risky for copyright holders, too; they must pay their own attorneys' fees and, if they lose, they may be ordered to pay the user's fees, as well. There has never been a copyright lawsuit directed against institutional software preservation practices, as of 2018.

PEER CONSENSUS ABOUT PRACTICE MATTERS.

Especially where there is no case law specifically addressing a category of uses, such as in software preservation, lawyers and judges consider collective expectations and values —that is, whether the user acted reasonably in light of standards of accepted practice in a particular field. That is why this Code, which articulates a consensus in software preservation around best practices, is valuable. It is valuable to potential fair users ("What do my peers regard as the right thing to do?"). It is valuable to potential challengers ("Am I looking at outlier behavior or something endorsed by the field?"). And finally, it is valuable to judges ("What do experts in this community regard as good practice?").

GOOD FAITH MATTERS.

While it does not appear in the text of the statute governing fair use, courts, lawyers, and potential litigants often take overall good faith into account. As this Code makes clear, librarians, archivists, and curators who preserve software agreed that good faith is shown in a number of ways, among them providing robust attribution and other information about software in their collections, by making clear a mission-centered reason for collecting and reusing software, and by taking reasonable steps against substitutional use of collection material.

APPENDIX TWO: OTHER COPYRIGHT-RELATED ISSUES

In our discussions, many preservation professionals expressed concern that even if their copyright fair use rights were clarified, they would be prevented from exercising them by other legal restrictions. They feared that essential preservation activities might violate state contract law, in the form of software license agreements. They also worried about the anti-circumvention provisions included in the 1998 federal Digital Millennium Copyright Act. Both of these concerns are routinely overstated in the software preservation community as grounds for pessimism about the potential of fair use.

LICENSING.

Since the mid-1980's, vendors of enterprise and consumer software have taken the position that while physical media may be sold, program content is merely licensed, on terms linked to each copy of the program. End-user license agreements (EULA's) once accompanied the retail copies of programs found in software collections. The exact terms of these agreements are frequently undocumented, but curators fear that licenses may override fair use. Not every "shrink-wrap" or "click-wrap" contract necessarily is enforceable. But courts have held that—in a narrow range of specific circumstances—license terms can be enforced even when they prohibit uses that copyright law itself permits.

Nonetheless, there are four reasons why legacy vendor licenses should not typically be a problem for software preservationists.

- Prohibitions that bar the fair use activities essential to preservation, teaching, and research, as described in this Code, aren't likely to be found in ordinary license agreements, new or old. When they crop up today, experienced negotiators are increasingly vigilant in negotiating their removal, while existing agreements that ban reverse engineering

or commercial redistribution generally don't bear on preservation activities. While most software licenses affirmatively authorize only a few activities (typically, running the software for its intended purpose), they don't expressly rule out others. This means that users must look outside the agreement, to fair use, for instance, for authorization.

- The Supreme Court has recognized the constitutional dimensions of fair use, and general waivers of personal liberties are generally disfavored as being "against public policy." It's likely, then, that even a broadly worded prohibition would be narrowly interpreted, if honored at all, in order to avoid such conflicts.

- Although it is under some pressure, the ancient common law doctrine known as "privity of contract" still applies to software licensing. Simply put, someone who receives a resold or gifted object doesn't automatically take on contractual obligations that originally came with it. Thus, the terms of commercial license agreements generally cannot be enforced against software collections if they didn't themselves agree to be bound when they acquired "second hand" program copies.

- Perhaps most important of all, in the unlikely event of a license-based legal challenge, a preservation program's financial exposure would be trivial or non-existent. Whereas copyright law (which wouldn't apply) allows for court-ordered "statutory damages," remedies in breach-of-contract cases depend on proof of so-called "actual damages." Actual damages are difficult even to conceptualize where preservation and support for teaching and scholarship are concerned. This may well be why there has never been a breach-of-software license case arising out of software preservation efforts undertaken by a memory institution.

ANTI-CIRCUMVENTION.

The preservation workflow may involve some workarounds to technological protection measures (TPMs) on legacy software, which range from encryption to passwords. The ban on breaking or avoiding these digital content locks in Section 1201 of the 1998 Digital Millennium Copyright Act

has caused some preservationists to avoid preserving software, leaving gaps in the archival record. In the future, however, concerns about this category of legal liability should recede.

Fair use itself is not a defense under these "paracopyright" prohibitions and related penalties. Because the penalties are not part of the Copyright Act, they are not subject to copyright's ordinary flexibilities, including the fair use doctrine. And although Section 1201 incorporates certain exceptions of its own, they do not apply to software preservation. As of 2015, the triennial process for granting new exemptions under the authority of the Librarian of Congress has yielded only one narrow rule affecting software preservation, for "video games for which outside server support has been discontinued, to allow individual play by gamers and preservation of games by libraries, archives and museum[s] (as well as necessary jailbreaking of console computer code for preservation uses only)."

This may change soon. In the rulemaking whose results will be announced late in 2018, the Software Preservation Network and others petitioned for an exemption applicable to software preservation in general. One hopeful sign is that most representatives of the commercial software industry did not oppose this proposal during public hearings. Winning an exemption depends on broad agreement that the activities it would enable are themselves otherwise lawful under basic copyright law. More general understanding of the many ways in which fair use applies to software preservation will help, and this Code is part of that effort. If the exemption is not granted, there will be an opportunity to seek an exemption again in the next round.

But what if the highly unlikely happens, and in the meantime, there is a circumvention-related lawsuit against a software collection? Some courts have suggested that in cases of overreach, they may be prepared to step in and define a judge-made, policy-based defense to liability for unauthorized circumvention. Indeed, this is how—back in 1841—fair use first was recognized as a defense to copyright infringement.

DONOR AGREEMENTS.

Of course, collections need to honor the terms of donor agreements, which can override the right to employ fair use. This is why negotiators involved in defining the terms of donor agreements should ensure the appropriate flexibility to permit software preservationists to do their jobs well. The donor may need to understand why putting as few restrictive terms as possible will allow for the most public benefit from their generous donation.

LEGAL ADVISORY BOARD:

Sharon Farb, Associate University Librarian for Collection Management and Scholarly Communication, UCLA

Joseph C. Gratz, Partner, Durie Tangri, San Francisco

Michael J. Madison, Professor, University of Pittsburgh School of Law

Matthew Sag, Professor, Loyola University Chicago School of Law

Pamela Samuelson, Richard M. Sherman Distinguished Professor, UC Berkeley School of Law

COORDINATING ORGANIZATIONS:

The Association of Research Libraries (ARL)is a nonprofit organization of 125 research libraries in Canada and the US whose mission is to advance research, learning, and scholarly communication. The Association fosters the open exchange of ideas and expertise, promotes equity and diversity, and pursues advocacy and public policy efforts that reflect the values of the library, scholarly, and higher education communities. ARL forges partnerships and catalyzes the collective efforts of research libraries to enable knowledge creation and to achieve enduring and barrier-free access to information. ARL is on the web at ARL.org.

Center for Media & Social Impact (CMSI), based at American University's School of Communication in Washington, D.C., is an innovation lab and research center that creates, studies and showcases media for social impact. Focusing on independent, documentary, entertainment, and public media, CMSI bridges boundaries between scholars, producers, and communication practitioners who work across media production, media impact, public policy, and audience engagement. Directed by American University Professor Caty Borum Chattoo, CMSI was founded (as the Center for Social Media) in 2001 by American University Professor Patricia Aufderheide. www.cmsimpact.org

The Program on Information Justice and Intellectual Property (PIJIP), co-founded by Prof. Peter Jaszi, promotes social justice in law governing information dissemination and intellectual property through research, scholarship, public events, advocacy, and provision of legal and consulting services. The program is a project of the Washington College of Law at American University in Washington, DC.

University of Virginia Library has been the center of the University since its founding by Thomas Jefferson in 1819. UVA is known not only for its extensive library system but also for the work it does in preserving works of scholarship both physical and digital. The UVA Library is also a leader in the creation of digital archives that ensure the safety of scholarship for generations to come.

More information
arl.org/fairuse
cmsimpact.org/software
wcl.american.edu/bestpractices

ARL.org
CMSimpact.org
pijip.org

ISBN 978-1-948964-02-9
9 781948 964029
US$2.00
50200