

Green Light for Grey Literature? Orphan Works, Web-Archiving and other Digitization Initiatives—Recent Developments in U.S. Copyright Law and Policy

Tomas A. Lipinski, J.D., LL.M., Ph.D.
Professor
School of Information Studies
University of Wisconsin--Milwaukee
P.O. Box 413
Milwaukee, Wisconsin, USA 53201
414-229-4908, 414-229-6699 (fax)

ABSTRACT

This paper reviews recent legislative and case developments in the area of copyright law affecting the collection, preservation including digitization and dissemination of grey literature. Alternative frameworks for crafting a legislative solution to impediments the copyright present to these uses are discussed. This includes review of pending legislation targeting the problem of so-called “orphan works” offering a limitation on the monetary damages or injunctive relief the copyright owner may be granted and another pending proposal aimed at relaxing the anti-circumvention prohibition of section 1201 that would allow access to compilations consisting primarily of public domain works that are protected by technical protection measures. The recent SECTION 108 STUDY GROUP REPORT also contain recommendations related to preservation (reproduction) and dissemination of both analog and borne-digital works, including a new provision for internet archiving. Finally, recent case law supporting the archiving of various online sub-literatures is reviewed, such as the disputes over caching and archiving by Google and the TurnItIn plagiarism combating service. Short of a legislative solution, the procedural elements affecting copyright enforcement are assessed to determine the legal risk in use of grey literature. These proposals and cases are analyzed and critiqued, with assessment towards solving the copyright issues related to the preservation and use of various grey literatures. Policy failures as well as successes in the United States can assist policy makers in other countries when contemplating copyright issues related to preservation and use of grey literature.

INTRODUCTION

This paper proceeds on the assumption grey literature refers to “any documentary material that is not commercially published and is typically composed of technical reports, working papers, business documents, and conference proceedings”¹ or the “quasi-printed reports, unpublished but circulated papers, unpublished proceedings of conferences, printed programs from conferences, and the other non-unique material which seems to constitute the bulk of our modern manuscript collections.”² In the educational context it could also include recorded lectures and other course content, student papers, thesis’ repositories, etc. The dominant theme of these conceptualizations is the unpublished nature of the literature, but is this true in every case?

A later section of this paper explores the issue of publication status and asks whether in the eyes of the U.S. copyright law these works are indeed unpublished, with the impact of that publication status on use and legal risk discussed. Issues related to the institutional collection and dissemination of grey and other literatures protected by copyright is of increasing interest in the United States, the European Union³ and world-wide.⁴

There are two options pursued in the United States when of crafting legislative or regulatory “solutions” to impediments that the copyright poses to the reproduction (collection, preservation, etc.) and public distribution (circulation, dissemination online, etc.) of protected content. The first is to offer an exemption (or more precisely an affirmative defense) for what would otherwise be an infringing use. Exemptions come in two forms, general (those available to all, such as fair use under section 107) and specific (limited to the particulars of the circumstance, such as the exemption granted to libraries and archives for reproduction and distribution of certain works under section 108). The second option is to offer users some sort of safe harbor or protection from the impact of such infringement. This is typically crafted as a limitation on monetary⁵ and in some cases injunctive remedies⁶ available to copyright owners. In rare instance immunity from any liability whatsoever may be granted.⁷ This paper assess whether the existing and emerging legal climate is amenable to the use of grey literature in the ways that libraries, archives, and other institutional organizations might desire to obtain and make accessible grey literature, through archiving, digitization, etc. The paper explores the current and potential interplay of the two policy options in light of proposals for reform recent case developments and also the dynamics of copyright litigation.

LIBRARY AND ARCHIVE REPRODUCTION AND DISTRIBUTION UNDER SECTION 108

Other than fair use (discussed below) section 108 of the United States copyright law offers qualifying institutions specific reproduction and distribution rights that may be useful in obtaining and distributing collections of grey literature. Section 108 allows for the reproduction and public distribution (circulation for example) of copies or phonorecords⁸ of the collection of a qualifying library and archive for preservation and security of unpublished materials or of published materials in cases of damage, deterioration, loss, or theft, or if the existing format in which the work is stored has become obsolete.

Current Law

In cases of preservation and security under section 108(b), the copy or copies, phonorecord or phonorecords (up to three copies or phonorecords may be made) must be from a work in the current collections of the library or archive and if a digital copy is made, must not be made available to the public in that format outside the premises of the library or archives, i.e., remote access to the material is not allowed afterwards. A copy made under subsection (b) for deposit in another library or archive may be transferred to that library or archive in digital format but the receiving library or archive must not distribute the material in that format or likewise if this institution is the receiving library or archive, i.e., staff cannot make the reproduced material available to patrons in digital form in any manner such as an in-house intranet.⁹ This would allow a qualifying library or archive with a collection of unpublished grey report or proceeding literature of the ABC Association or the XYZ Corporation to make a copy of the collection for preservation or security purposes or even to make a complete copy of the collection for another qualifying library or archive. The library or archive could digitize these collections as well in order to increase searching capabilities of users (staff or patrons) in accessing the content.

However, the digital copies may not be made available outside the premises of the library or archive, but relegated to in-house use alone.

In cases of damage, deterioration, loss, or theft, or if the existing format in which the work is stored has become obsolete under section 108(c), the copy or copies made (up to three copies may be made) are subject to the same limitation on digital distribution, i.e., remote access to the material is not allowed, and the library or archive must first make a reasonable effort to obtain an unused replacement of the published work at a fair price.¹⁰ A “reasonable effort” “will vary according to the circumstances of a particular situation. It will always requires recourse to commonly-known trade sources in the United States, and in the normal situation also to the publisher or other copyright owner (if such owner can be located at the address listed in the copyright registration), or an authorized reproducing service.”¹¹

The legislative history of the digital copying provision of section 108, added by the Digital Millennium Copyright Act,¹² indicates that Congress was concerned with infringement vis-à-vis the proliferation of digital libraries: “Although online interactive digital networks have since given birth to online digital ‘libraries’ and ‘archives’ that exist only in the virtual (rather than physical) sense on Web sites, bulletin boards and home pages across the Internet, it is not the Committee’s intent that section 108 as revised apply to such collections of information...The extension of the application of Section 108 to all such sites is tantamount to creating an exception to the exclusive rights of copyright holders that would permit any person who has an online Web site, bulletin boards, or a home page to freely reproduce and distribute copyrighted works. Such an exemption would swallow the general rule and severely impair the copyright owner’s right and ability to commercially exploit their copyrighted works.”¹³ Thus, an on-premises library or archive use of a section 108(b) or (c) digital copy is the rule. These obvious

limitations of section 108 prompted the recent work of the Section 108 Study Group to consider recommendations for legislative reform.

Proposals for Change: The Section 108 Study Group Report

Various recommendations contained within the Section 108 Study Group Report (Report) would increase the ability of a library or archive (the Report recommends that museums also be added to the list of qualifying institutions) to collect, archive and make other use of grey literature.¹⁴ First, the Report recommends that section 108 be amended to allow third parties through outsourcing arrangements to reproduce the work for later use by a qualifying archive, library or museum. It is often the case the large digitization projects require institutions to engage the services of low-cost often off-shore third parties. Amended section 108 would allow such services if undertaken without subsequent retention or commercial use by the outsourcer and the outsourcer agreed contractually to be subject to effective process, i.e., agree to be sued in United States court.¹⁵

Several recommendations implicate the preservation, digitization and dissemination aspects of collections containing or consisting grey literature and other works. In specific, the Report recommends amendment of section 108 to allow off-site lending of physical-digital content, e.g., a CD-ROM, if that was the original format of the item for preservation or security purposes under subsection (b) and for replacement copying under subsection (c), with a new category “fragile” of work added to subsection (c) “replacement” rights.¹⁶ “[T]he Study Group defines a ‘fragile copy’ as one that is embodied in a physical medium that is at risk of becoming unusable because it is delicate or easily destroyed or broken and cannot be handled without risk.”¹⁷ Depending upon the format, items of grey literature may fall into this category. The

Study Group could not reach consensus on access to virtual-digital, e.g., web-based works, thus the Report made no recommendation regarding this critical level of access in the Report.

Subject to numerous qualifying conditions a section 108 entity would be able to preserve “at-risk published or other publicly disseminated works in its collections.”¹⁸ The most dramatic recommendation relates to the ability of qualifying entities, subject to an opt-out mechanism by owners regarding publicly available online content, to make that content “accessible to users for purposes of private study, scholarship, or research.”¹⁹ However “publicly available online content” is content not protected by password or “requiring an affirmative act by the user to access” and would likely exclude content found on websites that are subject to terms and conditions of use, End User License Agreements (EULA)²⁰ or other control mechanisms that we require click of an “I agree” prompt or other affirmative click of agreement but not websites subject to mere browse wrap agreements.²¹ Qualifying content could be copied or archived by the institution for individual noncommercial use by patrons. Thus sources of grey literature as well as other content located on third party websites could be subject to the archiving provisions of an amended section 108.

SOLVING THE PROBLEM OF ORPHAN WORKS

It may be that archiving and digitization, i.e., reproduction and public distribution of a work of grey literature in its entirety may be impeded by concerns of copyright infringement. Depending on the circumstances as such use may be beyond fair use.

The Problem

It may be that the institutional collectors of grey literature like other users of copyrighted content would be willing to contact the owner and secure permission to use the work, even if

compensation of the owner is involved. However, the owner cannot be identified or located. Given the nature of the provenance of grey literature such content may be particularly susceptible to the problem of orphan works. An “orphan work” is “a term used to describe the situation where the owner of a copyrighted work cannot be identified and located by someone who wishes to make use of the work in a manner that requires permission of the copyright owner.”²² Users that desire to make the use but refuse to do so under any circumstances of legal risk, i.e., that the owner could one day surface and sue for copyright infringement will forego that use. As copyright law is a law of strict liability, these good faith attempts do not impact liability though general efforts of good faith may impact damages.²³ “*Such an outcome is not in the public interest*, particularly where the copyright owner is not locatable because he no longer exists or otherwise does not care to restrain the use of his work.”²⁴

The Solution

Earlier this fall Senate bill S. 2913, the Shawn Bentley Orphan Works Act of 2008²⁵ passed in the Senate is awaiting action in the final days 110th Congress, having been engrossed in the House on September 27. The bill would create new section 514 of the copyright law (title 17 of the United States Code). Proposed section 514 is an example of the second form of policy approach to a copyright problem, i.e., addresses the problem not by creating an exemption but in limiting the so-called bottom line, i.e., damages, the user-defendant faces should litigation by the owner-plaintiff be successful. If the user meets the safe harbor requirements of the provision then the only monetary relief the plaintiff can claim is for reasonable compensation for the infringing use made of the work. Damages (actual or statutory including damage enhancement for willful violations) as well as costs and attorney fees are not available. In some circumstances no monetary relief whatsoever is available. In the instance of derivative uses injunctive relief is also

limited. The question is whether or not limiting monetary liability to reasonable compensation is still too much for some would-be users to afford whereby such user would still forego use of the orphan work. Thus the impact of this solution would not be in the “public interest” to use the language of the Report.

Reasonable compensation is defined under proposed section 514(A)(3) as “the amount on which a willing buyer and willing seller in the positions of the infringer and the owner of the infringed copyright would have agreed with respect to the infringing use of the work immediately before the infringement began.” The impact is obvious, users must obtain some evidence or documentation of what that amount might have been, and second keep that evidence or documentation should the orphan owner ever appear one day and the user need to prove qualification under the safe harbor. It is also a requirement of qualification that should the owner appear the user must bargain in good faith, offering to pay reasonable compensation. So again having documentation of what this amount might be is useful especially in cases where the owner appears years after the initial infringement. Considering the duration of copyright in the United States there may be a lengthy period during which this information may be relevant. Assuming the work is still protected by copyright this would be three years after infringing use of the work ceases, as the statute of limitations for commencing a civil action is three years.²⁶ So for a work for which the copyright does not expire until say 2045, where the infringing use commences in 2010, lasts until 2035 when the work is deaccessioned from the library or archive collection, the user would need to keep records of what reasonable compensation would have been in 2010 for 28 years: 25 years of use plus the three years to cover the tolling of the statute of limitations. For situation where the use is continuous, i.e., the work remains a permanent part

of the collection such making the work accessible to the public on a website for example, that would be for as long as the copyright lasts plus three years!

Under proposed section 514(c)(1)(B), a nonprofit educational institution, museum, library, archives, or a public broadcasting entity (or employees of such entity acting within the scope of their employment) can reduce the monetary amount to zero if three conditions are met. First, the infringement was performed without any purpose of direct or indirect commercial advantage (this is different than a situation where the use *results* in a direct or indirect commercial advantage, only the “purpose” must be so). Second, the infringement was *primarily* educational, religious, or charitable in nature (this is not the same “solely” nor does this standard look to entity, the categorizations being those employed more often to describe the nature of the entity rather than as here its conduct, i.e., here the “infringement”). Third, after receiving a notice of claim of infringement and having an opportunity to conduct an expeditious good faith investigation of the claim, i.e., some legal assessment of the merits of the claim of infringement must be undertaken) the infringer promptly ceased the infringement.

The “notice of claim of infringement” does not require that a law suit be filed rather it is more akin to the notice under section 512(c)(3) that triggers an expeditious take-down or restriction of access to content.²⁷ As required under proposed section 514(a)(1) the notice would be made in writing and include the name of the owner and title of the infringed copyright as well as sufficient information regarding the owner or their representative and the location of the infringing content.

Finally in the case of derivative works²⁸ or to be more precise under proposed section 514(c)(2)(B), where the infringer has “prepared or commenced preparation of a new work of

authorship that recasts, transforms, adapts, or integrates the infringed work with a significant amount of original expression,” the court may not enjoin the defendant’s continued use. The concept of integration offers a somewhat broader scope of uses than contemplated by the statutory definition of derivative work. Moreover, the inability to enjoin continued preparation or use in essence creates a statutory license to use the work as long as the “infringer pays reasonable compensation in a reasonably timely manner after the amount of such compensation has been agreed upon with the owner of the infringed copyright or determined by the court.” If the owner refuses to agree during good faith attempts at negotiation, the court may order the owner to accept the reasonable compensation and allow the use to continue. The user must also provide attribution “in a manner that is reasonable under the circumstances to the legal owner of the infringed copyright.” However attribution is only required “if requested by such owner.” It is odd to condition a court ordered attribution upon request by the owner as an initial condition of section 514 qualification is to provide attribution, as discussed below.

Qualifying for the Proposed Section 514 Safe Harbor: The Search

There are several requirements before the section 514 limitations on remedy can apply. First, the user (the proposed language repeatedly uses the word “infringer”) must by a preponderance of the evidence demonstrate that before the use (“infringement”) began he or she “performed and documented a qualifying search, in good faith, to locate and identify the owner of the infringed copyright” and that the search was unsuccessful. This suggests two elements to a search: substantive as to the content of the search or its protocol (“a qualifying search”) and procedural as to how the search is executed (“in good faith”). A user might have access to a list of the proper steps or best practices developed by library and archive professionals but exert little effort to complete or execute those steps practices.

The requirements of a “diligent effort that is reasonable under the circumstances to locate the owner” is detailed in proposed section 514(b)(2). The diligent effort requires “at a minimum” a search of the Internet accessible Copyright Office records (assuming there is sufficient information regarding to the work to construct a search), a search of other authorship and ownership information, the “use of appropriate technology tools, printed publications, and where reasonable, internal or external expert assistance” and “appropriate databases, including databases that are available to the public through the Internet.” These sources might include content made available from third party sources, for example, a web-accessible database of renewal records of published works filed between 1950 and 1992, available at <http://collections.stanford.edu/copyrightsrenewals/bin/page?forward=home> (renewal records). Of course not all of the content will come at little or no cost. In fact a later provision anticipates a diligent effort to include recourse to pay-per-use services (“use of resources for which a charge or subscription is imposed”). A later provision requires review “as appropriate” of Copyright Office records not available to the public through the Internet. This suggests either a trip to Washington, D.C. is in order or at least to the regional federal depository library (as not all partial depositories would have Copyright Office records in the collection). Finally, proposed section 514 anticipates that the Copyright Office (“Recommended Practices” including at least one such statement for each category of work of authorship listed in section 102²⁹) as well as “authors, copyright owners and users” make available best practices to assist users in performing a qualifying search.

If through this search process the owner is located, even though never contacted or once contacted fails to respond to the inquiry the work ceases to be orphan.³⁰ The proposed provision reads in part: “The fact that, in any given situation . . . an owner of the infringed copyright fails

to respond to any inquiry or other communication about the work shall not be deemed sufficient to meet the conditions under paragraph (1)(A)(i)(I)", i.e., "performed and documented a qualifying search." Second, attribution "based on information obtained in performing the qualifying search" must be provided about the owner "in a manner that is reasonable under the circumstances." Attribution will make it easier for owners to identify their works and the unlawful use being made of them. This in turn fulfills the purpose of section 514 (or at least the remission mechanism proposed by the U.S. Copyright Office) which makes it more likely that copyright owners will find users and vice-versa and come to agreement over the use of the work.³¹ Use of the work must also indicate that it is made subject to the proposed section with "the form and manner of which shall be prescribed by the Register of Copyrights." This is likely done to dissuade others from incorrect assumption that because the qualifying user is making use of the work, such as posting of the work on the library or archive website, use of the work by all takers is welcome and free. Rather the provision promotes awareness of uses that are compliant with proposed section 514. In terms of process, the application of the section 514 safe harbor must be asserted in the initial pleading by the defendant. It is assumed this is done to encourage the parties to come to some reasonable agreement or perhaps to assist the court in pretrial motion determinations.

Additional Requirements: When (If) the Owner Later Appears

In addition to these requirements occurring prior to use of the work, other requirements exist should the owner of the orphan work later surface and give "notice of a claim of infringement." The user, after "having an opportunity to conduct an expeditious good faith investigation of the claim" must first undertake a good faith negotiation with the owner over the amount of "reasonable compensation" then "render payment of reasonable compensation in a

reasonably timely manner after reaching an agreement with the owner” (or if ordered by the court to do so). Once it is determined (either by negotiation or by the court) what amount constitutes reasonable compensation that amount must be paid in a “reasonably timely manner.” It is not certain what the requirement of claim investigation is meant to accomplish. It could be to assess the merits of the claim, but this would be odd as the user has likely long since concluded that the use of the work is infringing otherwise there would be no need to have undertaken measures that would qualify for the proposed section 514 safe harbor. This “good faith investigation of the claim” may address procedural aspects of the circumstances, i.e., verifying identification of the owner, the work infringed and the infringing work. Another issue is whether the good faith negotiation obligation must be successful, as the statute appears to anticipate no other option. Notice also, the obligation to negotiate is on the user not the owner. So if the owner wants nothing to with arriving at an arrangement that would compensate for past use and instead desires to sue the user for infringement the owner is free to do so. Of course if the user is able at least to document a good faith attempt to negotiate (as well as the other search and attribution requirements) then the monetary relief available to the owner will be foreclosed for a qualifying nonprofit or limited to reasonable compensation in other circumstances or for derivative uses.

Should S. 2913 or some subsequent variation become law, this discussion provides a basis upon which the user can understand its operation and fulfill the requisite legal obligations or to assess whether the cost of those obligations are not worth the benefit of the safe harbor. In essence, the bill encourages search and documentation of the search for the copyright owner. Oddly, if the search is successful the work is no longer “orphan” and the provision will not apply. Yet in this instance permission may not be forthcoming and so use cannot be made of the work without the threat of legal repercussion anyways. When a work remains orphan and the

copyright owner is not located, the bill offers benefit but at potential high practical (time, record-keeping, etc.) and actual cost (outside exerts and resources) that may be no better in terms of the cost of the associated with a use of the work without application of the safe harbor, the cost of the legal risk of use under present law.

Users of orphan works should also be aware of the potential looming cost of litigation to vindicate a proper attribution and search. Those who move forward with use will be the test cases, carrying the initial cost of this “sorting out,” hopefully to the benefit of subsequent users. It is critical that courts do not interpret the reasonably diligent search, attribution or other requirements too harshly otherwise the litigation-avoidance incentive will not operate properly. Once a precedent has been set to establish some reasonable norms for predicting when the provision would apply, users will have some structure as to what constitutes search and attribution, and owners might be more reluctant to litigate.

A BRIEF COMMENT OF TECHNICAL PROTECTION MEASURES

U.S. law prohibits circumvention of technological protection measures (TPMs) that control access³² as well as the distribution (trafficking) of technologies that control access³³ or distribution (trafficking) of protection technologies that control specific uses of a work, so called “black-box” devices.³⁴ Such devices are those that are primarily designed to circumvent, have limited commercially significant purpose, or are marketed as an anti-circumvention device. It is unknown the extent to which grey literature is disseminated subject to such TPMs. However, the increasing use of TPMs by content owners of “white” literature and in particular formats such as sound recordings such as CDs and audiovisual works such as DVDs suggests that this may be a future problem for grey literature as well. In order for content to be subject to

the prohibition on circumvention the content must be within subject matter and protection of the copyright law, the prohibition does not apply to content not protected by copyright such as that in the public domain.³⁵ Second, the control must be put in place by the copyright owner or with the permission of the copyright owner. If the control is instigated by a third party web site owner of database vendor for example without permission of the copyright owner, the prohibition will not apply.

For qualifying institutions seeking to acquire grey literature that may be subject to such prohibitions there is a statutory exception, at least for lawful circumvention during the acquisitions phase. Section 1201(d) provides a specific exception for qualifying nonprofit libraries, archives, or educational institutions to circumvent an access control in order to make a bona fide determination of whether to purchase an item for its collection or curriculum: “access to a commercially exploited copyrighted work solely in order to make a good faith determination of whether to acquire a copy of that work for the sole purpose of engaging in conduct permitted under this title shall not be in violation of subsection (a)(1)(A).”³⁶ Notice that this exception operates with respect to the section 1201(a)(1)(A) anti-circumvention of access control provision, it does not allow qualifying nonprofit libraries, archives, or educational institution to traffic in either an access or use control.³⁷ Such entities are still prohibited from engaging in conduct that remains a section 1201(a)(2) or section 1201(b) trafficking violation, i.e., sharing the means of the circumvention with another qualifying entity.

In addition there is three year cycle of rule-making, with a de novo review made of requests for regulatory exemption to the circumvention prohibition. The statutory standard for granting the regulatory exemption is whether or not “noninfringing uses by persons who are users of a copyrighted work are, or are likely to be, adversely affected.” In 2006 the standard was

modified somewhat as “the Register has concluded that in certain circumstances, it will also be permissible to refine the description of a class of works by reference to the type of user who may take advantage of the exemption or by reference to the type of use of the work that may be made pursuant to the exemption... must be properly tailored not only to address the harm demonstrated, but also to limit the adverse consequences that may result from the creation of an exempted class.”³⁸ There are six exemptions granted under current law, two of possible relevance to preservation and access: “Computer programs protected by dongles that prevent access due to malfunction or damage and which are obsolete and library preservation of “computer programs and video games distributed in formats that have become obsolete.”³⁹ Likely neither is of much relevance to collections of grey literature but nonetheless indicates that should grey literature be increasingly subject to such controls, short of legislative remedy through amendment of section 1201,⁴⁰ there is an accessible if somewhat cumbersome and limited regulatory process to achieve similar even if not permanent ends.

WEB ARCHIVING AND FAIR USE

Several recent cases in the past two years have suggested that initiatives to engage in systematic archiving of content can be a fair use. In *Perfect 10 v. Amazon.com, Inc.*,⁴¹ the Ninth Circuit concluded that Google’s creation of its thumbnail index of images was fair use, commenting that “the significantly transformative nature of Google’s search engine, particularly in light of its public benefit, outweighs Google’s superseding and commercial uses of the thumbnails in this case.” However, as the index allows users of the Google search engine to be led to infringing sources of the content, Google could be found contributorily liable: “Applying our test, Google could be held contributorily liable if it had knowledge that infringing Perfect 10 images were available using its search engine, could take simple measures to prevent further

damage to Perfect 10's copyrighted works, and failed to take such steps."⁴² A conclusion of fair use was also found in another case involving Google, this time its practice of automatically archiving web sites unless the owner opted out. In *Field v. Google, Inc.*,⁴³ a district court again identified the social good that such preservation projects can achieve: "The fact that the owners of billions of Web pages choose to permit these links to remain is further evidence that they do *not* view Google's cache as a *substitute* for their own pages. Because Google serves *different and socially important purposes* in offering access to copyrighted works through 'Cached' links and does *not* merely *supersede* the objectives of the original creations, the Court concludes that Google's alleged copying and distribution of Field's Web pages containing copyrighted works was *transformative*."⁴⁴ Finally, the impact of the recent settlement by publishers and authors against Google also suggests that such archiving projects will continue to present legal challenge but through decision or settlement will be allowed to continue.⁴⁵ These developments lend support for similar efforts by institutions providing similar social good by preservation of the cultural record. It may be that the same argument could be made in the case of preservation of grey literature when that collection is unique and does not exist elsewhere and the institutions serves as the sole source of the content. A final archive decision not involving Google also stands for the proposition that such initiatives offer a beneficial societal purpose and can likewise be a fair use. In *A.V. v. iParadigms, Ltd.*,⁴⁶ the court observed that as in the Google index, cache and archive cases the "use of Plaintiffs' written works [is] highly transformative. Plaintiffs originally created and produced their works for the purpose of education and creative expression. iParadigms, through Turnitin, uses the papers for an entirely different purpose, namely, to prevent plagiarism and protect the students' written works from plagiarism... makes no use of any work's particular expressive or creative content beyond the limited use of comparison with

other works... provides a substantial public benefit through the network of educational institutions using Turnitin. Thus, in this case, the first factor favors a finding of fair use.”⁴⁷ As a result the use of the student-plaintiff’s papers in the TurnItIn databases was a fair use. In each of the case the use was deemed transformative and even though the entire work was taken in the instance of images in the Google cases or student papers in the *iParadigms* case the complete taking was necessary to accomplish the good purpose. This is in contrast to the recent case involving the Harry Potter Lexicon. The nature of encyclopedias and reference guides being in general transforming, though under the particular circumstances the publisher of The Lexicon: *An Unauthorized Guide to Harry Potter Fiction and Related Material* took more than once necessary to accomplish its good purpose.⁴⁸

FINAL THOUGHTS ON THE USE OF GREY LITERATURE AND THE PARTICULARS OF COPYRIGHT ENFORCEMENT

Other elements of the copyright law may make use of grey literature nonfringing or reduce the likelihood of litigation or the fallout from that litigation should it occur. First, it may be that the content is not protected by copyright. For example, works produced by the federal government are in the public domain.⁴⁹ Other works may have fallen into the public domain due to lapse of protection. The rules can be rather complex. In general works published before 1923 are in the public domain, those works published 1923-1963 with notice and renewal are protected for 95 years from date of publication and those published 1964-1977 with notice (renewal automatic) are also protected for 95 years from the date of publication. Under the 1976 Copyright Act, works created after 1977 are protected for the duration of the author’s life plus 70 years, or if corporate, anonymous, pseudonymous: lesser of 95 years from publication or 120 from creation. If the work is unpublished and created before the first of January 1978 (the effective date of the 1976 Copyright Act) then duration of copyright is 2002 or the author’s life

plus 70 years, whichever is longer. If created before 1978 and published before 2003, then the work is protected for the greater of author's life plus 70 years or until 2047. If the unpublished work was created after 1977, the duration is for the life of the author plus 70 years, or 120 years from creation for corporate, anonymous, pseudonymous authors. When the death date of an author is unknown: a default of 120 years from creation applies. As a result, the period of liability for infringing use of protected content may be lengthy. However, certain particulars of copyright litigation and enforcement may work against litigation and reduce the ultimate legal risk the user of grey literature may face.

Legal risk is a combination of several factors: the potential for liability, the likelihood of litigation (or threat of litigation) as well as the possibility of settlement and the impact of that litigation (or settlement), i.e., what remedies are available to the copyright owner. What is the potential for liability ("can I be sued?"), how likely is litigation ("will I be sued?") and how likely is it that the infringement will be discovered, and what remedies are available ("what's the bottom line?"). Furthermore, the scope of available damages (and award of costs and attorneys fees) is related to whether the work is published or unpublished, the publication status of the work. Much of the grey literature may be in fact unpublished. Furthermore, there may be an opportunity for damage remission, discussed below, as well.

While the work need not be registered to be protected, it is a prerequisite to litigation.⁵⁰ How many grey literature works are registered? Assuming the work is registered, and this may be a significant assumption in the case of grey literature, the timing of the registration in relation to the infringement and the status (unpublished or published) of the work determines the scope of damages available to the copyright owner. This may impact the decision to sue or not. It is unlikely an owner would undertake the cost of litigation if the monetary award were limited to

actual damages alone or if costs or attorneys fees could not be recovered in addition. Registration must occur before infringement of an unpublished work and within three months of publication for published works in order to obtain statutory damages and attorney's fees.⁵¹

Even if there is potential for a significant award of damages, where the infringement is undertaken by an employee of nonprofit educational institution or library (or the institution itself is liable), and the employee was acting within scope of employment, believed and had reason to believe, that the use was a fair use under section 107, and infringed by reproducing the work, the court must remit the statutory damages awarded to zero.⁵² The possibility of no statutory damages may dissuade a owner from ever suing.

So are works of grey literature published or unpublished? Publication is defined in Section 101 as the "distribution of copies or phonorecords of a work to the public by sale other transfer of ownership, or by rental, lease, or lending." The distribution of copies on a busy street is publication, as is the unrestricted gift of copies constitutes. So too is leaving copies in a public place for anyone to take a publication. However, distributing text at a seminar for use only by the recipients is ordinarily not publication.⁵³ One district court concluded that posting content on the internet is a publication.⁵⁴ "The statutory definition of publication set forth above specifies two rules worth emphasizing: First, publication includes only acts of publication (1) by the copyright owner or (2) authorized by the owner. Unauthorized acts of publication by others do not result in publication. The copyright owner has the sole authority to authorize publication. Second, the phrase 'copies or phonorecords' refers only to plural items. What happens when a single work is distributed? For example, if a single piece of sculpture or book is distributed, rented, or lent, does publication occur? Not necessarily. There must be multiple copies available for distribution, transfer, rent, lease, or lending. Thus, publication occurs only if the single item is one of many

copies available for distribution.”⁵⁵ Thus Internet sources of grey literature are likely published, as are ephemeral reports of organizations that are released to the public, some conference proceedings, etc. Internal organizational documents from a corporation for example that reside in an institutional archive such as a university remain unpublished even if circulated (though this would constitute a public distribution under the copyright law. Once digitized and made accessible on the web, the publication status changes. Yet in either case such documents are likely to be unregistered! If the corporation retained copyright in those documents it could not proceed with litigation until it registered those works. In the proper circumstances the legal risk of using grey literature may be small, the works may be unprotected by copyright or if protected may offer unattractive circumstances for litigation.

CONCLUSION

The expanded collection and dissemination of grey literature (as well as other works protected by copyright) through archiving and digitization is bolstered by recent case law establishing the circumstances under which such initiatives can be a fair use under U.S. copyright law. In addition legislative reform is under way (section 108 and proposed section 514) to increase range of use rights available to institutions regarding protected content including grey literature. Moreover, the particulars of copyright enforcement may also work to minimize the legal risk in remaining circumstances.

¹ Brian Matthews, *Gray Literature: Resources for Locating Unpublished Research*, C&CRL NEWS, March 2004.

² PETER HIRTLE, *BROADSIDES VS. GREY LITERATURE* (1991), as quoted in MOYA K. MASON, *GREY LITERATURE: ITS HISTORY, DEFINITION, ACQUISITION, AND CATALOGUING*, available at <http://www.moyak.com/researcher/resume/papers/var7mkmkw.html>.

³ See, *GREEN PAPER, COPYRIGHT IN THE KNOWLEDGE ECONOMY*, COM(2008) 466/3, available at http://ec.europa.eu/internal_market/copyright/docs/copyright-info/greenpaper_en.pdf, *FINAL REPORT ON DIGITAL PRESERVATION, ORPHAN WORKS AND OUT-OF PRINT WORKS, SELECTED*

IMPLEMENTATION ISSUES (June 4, 2008), available at http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/reports/copyright/copyright_subgroup_final_report_26508-clean171.pdf. See also, Annex: Model agreement for a licence [sic] on digitisation [sic] of out of work prints, available at http://ec.europa.eu/information_society/newsroom/cf/itemdetail.cfm?item_id=3366 (April 18, 2007).

⁴ See, e.g. STANDING COMMITTEE ON COPYRIGHT AND RELATED RIGHTS, STUDY ON COPYRIGHT LIMITATIONS AND EXCEPTIONS FOR LIBRARIES AND ARCHIVES (Seventeenth Session, Geneva, November 3 to 7, 2008) (Prepared by Kenneth Crews), available at http://www.wipo.int/meetings/en/doc_details.jsp?doc_id=109192.

⁵ See, e.g. 17 U.S.C. § 504(c)(2).

⁶ See, e.g., 17 U.S.C. § 512(j).

⁷ See, e.g., 17 U.S.C. § 108(f).

⁸ 17 U.S.C. § 101 defines a phonorecord as “material objects in which sounds, other than those accompanying a motion picture or other audiovisual work, are fixed by any method now known or later developed, and from which the sounds can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device. The term "phonorecords" includes the material object in which the sounds are first fixed.”

⁹ 17 U.S.C. § 108(b).

¹⁰ 17 U.S.C. § 108(c).

¹¹ H. Rpt. No. 94-1476, 94th Cong. 2d Sess. 75-76 (1976), reprinted in 5 United States Code Congressional and Administrative News 5659, 5689 (1976).

¹² Pub. L. No. 105-304, Title IV, sec. 404, 112 Stat. 2860, 2889-2890 (1998) (codified at 17 U.S.C. § 108).

¹³ S. Rpt. No. 105-190, 105th Cong. 2d Sess. 63 (1998).

¹⁴ U.S. COPYRIGHT OFFICE, THE SECTION 108 STUDY GROUP REPORT 31 (2008).

¹⁵ U.S. COPYRIGHT OFFICE, THE SECTION 108 STUDY GROUP REPORT 39 (2008) (“not for any other direct or indirect commercial benefit...contractually prohibited from retaining copies...preserves a meaningful ability...to obtain redress from the contractor for infringement by the contractor”). “The contractor should be contractually required to submit to U.S. jurisdiction and have assets in the United States, or be bonded and insured in this country.” Id at 41.

¹⁶ U.S. COPYRIGHT OFFICE, THE SECTION 108 STUDY GROUP REPORT 52 and 61 (2008).

¹⁷ U.S. COPYRIGHT OFFICE, THE SECTION 108 STUDY GROUP REPORT 54 (2008).

¹⁸ U.S. COPYRIGHT OFFICE, THE SECTION 108 STUDY GROUP REPORT 69 (2008) Qualifying conditions relate security of process (“best practices”), storage capacity, system integrity, identification, retrieval, security of access, and the ability to migrate, audit, afford (in terms of cost), support (in terms of mission) and transfer (in cases of cessation of operations) the preserved content. Allowance should also be made smaller entities to engage in such efforts as well. Id. at 69-70.

¹⁹ U.S. COPYRIGHT OFFICE, THE SECTION 108 STUDY GROUP REPORT 80 (2008).

²⁰ *Ticketmaster L.L.C. v. RMG Technologies, Inc.*, 507 F.Supp.2d 1096, 1108 (C.D. Cal. 2007) “Thus, by the Terms of Use, Plaintiff grants a nonexclusive license to consumers to copy pages from the website in compliance with those Terms. Inasmuch as Defendant used the website, Defendant assented to the terms.” See also, *Druyan v. Jagger*, 508 F.Supp.2d 228, 237 (S.D.N.Y. 2007) (citing *Ticketmaster Corp. v. Tickets.com, Inc.*, 2003 WL 21406289 (C.D. Cal. 2003); and *Register.com v. Verio*, 126 F. Supp. 2d 238 (S.D.N.Y. 2000)): “First, courts have consistently held that the use of a website for such purposes as purchasing a ticket manifests the user's assent to the Terms of Use, and that such terms constitute a binding contract as long as the terms are sufficiently conspicuous.”

²¹ *Specht v. Netscape, Inc.*, 306 F.3d 17, 32 (2d Cir. 2002) (footnote omitted): “Internet users may have, as defendants put it, ‘as much time as they need[]’ to scroll through multiple screens on a webpage, but there is no reason to assume that viewers will scroll down to subsequent screens simply because screens are there. When products are ‘free’ and users are invited to download them in the absence of reasonably conspicuous notice that they are about to bind themselves to contract terms, the transactional circumstances cannot be fully analogized to those in the paper world of arm’s-length bargaining.” See also, Blaze D. Waleski, *Enforceability of Online Contracts: Clickwrap vs. Browse Wrap*, *e-Commerce Law & Strategy*, November, 2002, Vol 19, no. 7 (no pagination in Westlaw).

²² U.S. COPYRIGHT OFFICE, REPORT ON ORPHAN WORKS 15 (2006).

²³ See, e.g., *Lowry’s Reports, Inc. v. Legg Mason, Inc.*, 271 F. Supp. 2d 737, 746 (D. Md. 2003) (“The fact that Legg Mason’s employees infringed Lowry’s copyrights in contravention of policy or order bears not on Legg Mason’s liability, but rather on the *amount of statutory and punitive damages and the award of attorneys’ fees.*” (emphasis added).)

²⁴ U.S. COPYRIGHT OFFICE, REPORT ON ORPHAN WORKS 15 (2006) (emphasis added)

²⁵ S. 2913, 110th Congress, 2d Session (April 24, 2008) (Shawn Bentley Orphan Works Act of 2008).

²⁶ 17 U.S.C. § 507.

²⁷ 17 U.S.C. § 512(c)(1)(A) (“upon notification of claimed infringement as described in paragraph (3), responds expeditiously to remove, or disable access to, the material that is claimed to be infringing or to be the subject of infringing activity”).

²⁸ 17 U.S.C. § 101 defines a derivative work as “a work based upon one or more preexisting works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which a work may be recast, transformed, or adapted.”

²⁹ 17 U.S.C. § 102 provides that “[w]orks of authorship include the following categories: (1) literary works; (2) musical works, including any accompanying words; (3) dramatic works, including any accompanying music; (4) pantomimes and choreographic works; (5) pictorial, graphic, and sculptural works; (6) motion pictures and other audiovisual works; (7) sound recordings; and (8) architectural works.”

³⁰ See also, U.S. COPYRIGHT OFFICE, REPORT ON ORPHAN WORKS 97 (2006) (“[O]nce an owner is located, the orphan works provision becomes inapplicable.”).

³¹ U.S. COPYRIGHT OFFICE, REPORT ON ORPHAN WORKS 97 (2006) (“The primary goal of this study is to prompt owners and users to find each other and commence negotiation—it is *not* intended to allow use of works in disregard of the owner’s wishes after that owner has been found.”)

³² 17 U.S.C. § 1201(a)(1).

³³ 17 U.S.C. § 1201(a)(2).

³⁴ 17 U.S.C. § 1201(b).

³⁵ See, *Online Policy Group v. Deibold, Inc.*, 337 F. Supp. 2d 1195 (N.D. Calif. 2004); and *Lexmark International, Inc. v. Static Control Components, Inc.*, 387 F. 3d 522 (6th Cir. 2004).

³⁶ 17 U.S.C. § 1201(d)(1).

³⁷ See, 17 U.S.C. § 1201(d)(4) (“This subsection may not be used as a defense to a claim under subsection (a)(2) or (b), nor may this subsection permit a nonprofit library, archives, or educational institution to manufacture, import, offer to the public, provide, or otherwise traffic in any technology, product, service, component, or part thereof, which circumvents a technological measure.”).

³⁸ 71 Fed. Reg. 68472, 68473 (November 27, 2006).

³⁹ 37 C.F.R. §201.40 (2006).

⁴⁰ See, e.g. H.R. 1201, the Freedom and Innovation Revitalizing U.S. Entrepreneurship Act of 2007 (FAIR USE Act of 2007), 110th CONGRESS, 1st Session (February 27, 2007). The FAIR USE Act would amend section 1201 to allow for circumvention of an access control protecting compilations consisting primarily of public domain works, a work of “substantial public interest” for purposes of “criticism, comment, news reporting, scholarship, or research” and would allow circumvention in order to exercise rights under section 108(c) privileges, but excluding works in “obsolete” formats.

⁴¹ *Perfect 10 v. Amazon.com, Inc.*, 487 F.3d 701, *13 (9th Cir. 2007).

⁴² *Perfect 10 v. Amazon.com, Inc.*, 487 F.3d 701, *19 (9th Cir. 2007).

⁴³ *Field v. Google, Inc.*, 412 F.Supp.2d 1106 (D. Nev. 2006).

⁴⁴ *Field v. Google, Inc.*, 412 F.Supp.2d 1106, 1119 (D. Nev. 2006) (all emphasis added).

⁴⁵ *The McGraw-Hill Cos. Inc. v. Google Inc.*, No. 05 CV 8881 (S.D.N.Y. filed Oct. 19, 2005); and *Authors Guild v. Google Inc.*, No. 05 CV 8136 (S.D.N.Y. filed Sept. 20, 2005). See, MOTION to Approve /Notice of Motion for Preliminary Settlement Approval (October 28, 2008); and STIPULATION AND ORDER FOR AMENDMENT OF PLEADINGS (October 30, 2008) available at <http://news.justia.com/cases/featured/new-york/nysdce/1:2005cv08136/273913/>.

⁴⁶ *A.V. v. iParadigms, Ltd.*, 2008 WL 728389 (E.D. Va. 2008).

⁴⁷ *A.V. v. iParadigms, Ltd.*, 2008 WL 728389. *6 (E.D. Va. 2008).

⁴⁸ *Warner Brothers Entertainment, Inc. v. RDR Books*, 575 F.Supp.2d 513 (S.D.N.Y. 2008). Regarding the books in the Harry Potter series: “Other times, however, the Lexicon disturbs the balance and takes more than is reasonably

necessary to create a reference guide. In these instances, the Lexicon appears to retell parts of the storyline rather than report fictional facts and where to find them.” Id. 548. Regarding the companion books to the series the use is less transformative: “The Lexicon’s use of copyrighted expression from Rowling’s two companion books presents an easier determination. The Lexicon takes wholesale from these short books. Depending on the purpose, using a substantial portion of a work, or even the whole thing, may be permissible... In this case, however, the Lexicon’s purpose is only slightly transformative of the companion books’ original purpose. As a result, the amount and substantiality of the portion copied from the companion books weighs more heavily against a finding of fair use.” Id. at 548-549.

⁴⁹ U.S.C. § 105 (“Copyright protection under this title is not available for any work of the United States Government.”).

⁵⁰ 17 U.S.C. § 411: “[N]o action for infringement ... shall be instituted until registration of the copyright claim has been made...”

⁵¹ 17 U.S.C. § 412 (“[N]o award of statutory damages or of attorney’s fees ... shall be made ... [for infringement of] an unpublished work commenced before the effective date of its registration; or ... commenced after first publication of the work and before the effective date of its registration, unless such registration is made within three months after the first publication of the work.”).

⁵² 17 U.S.C. § 504(c)(2).

⁵³ Compendium II, Copyright Office Practices § 905.02 (1984). See also, *Estate of Martin Luther King, Jr. v. CBS, Inc.*, 194 F.3d 1121 (11th 1999) (“I Have a Dream” speech of Dr. Martin Luther King heard by thousands and broadcast to thousands more was not a publication.)

⁵⁴ *Getaped.com v. Cangemi*, 188 F.Supp.2d 398 (S.D.N.Y. 2002). Events: website revised in June, infringement in July, registration in August, litigation follows. “Consequently, when a website goes live, the creator loses the ability to control either duplication or further distribution of his or her work. A webpage in this respect is indistinguishable from photographs, music files or software posted on the web—all can be freely copied. Thus, when a webpage goes live on the Internet, it is distributed and ‘published’ in the same way the music files in *Napster* or the photographs in the various *Playboy* decisions were distributed and ‘published.’” Id. at 402.

⁵⁵ JOHN W. HAZARD, JR., COPYRIGHT LAW IN BUSINESS AND PRACTICE § 1:5 (2008).