



OFFICE OF THE PRESIDENT
California Digital Library

415 20th Street
Oakland, California 94612

**[Docket No. 2012-12]
Orphan Works and Mass Digitization; Request for Additional Comments**

**Submitted by Laine Farley
On behalf of
The California Digital Library
University of California
Office of the President**

May 21, 2014
Maria A. Pallante
Register of Copyrights
U.S. Copyright Office
101 Independence Avenue S.E.
Washington, DC 20559-6000

Dear Ms. Pallante:

The California Digital Library is pleased to submit these comments in response to the Copyright Office's Request for Additional Comments on Orphan Works and Mass Digitization which appeared in the Federal Register on February 10, 2014, and the March 10-11, 2014 roundtable discussions. We appreciate the efforts of the Copyright Office to seek broad and continued consultation on these issues.

Introduction and Background

The California Digital Library (CDL), a unit within the University of California Office of the President, is the eleventh University Library in the University of California (UC) system, providing a broad range of digital library services and collaborative leadership to more than 100 libraries on the ten University of California campuses. UC is the largest public research university in the world, spending more than \$4.1 billion annually¹ on direct research, awarding seven percent of the nation's PhDs, and enriching the lives of more than 230,000 students. The University's activities generate a wide range of benefits and services for California's citizens. Directly supporting the university's mission of teaching, research, and public service, its more than one hundred libraries collectively form the largest academic research library in the world. UC libraries house collections that together include upwards of thirty-eight million printed volumes, twenty million pictorial items, two million maps, one million government documents, 1.8 million pamphlets, and 385,000 manuscript and archival holdings, which are accessible to all faculty,

¹ 2013 Accountability Report, Indicator 10.2.5, Direct research expenditures by discipline, Universitywide, 1997-98 to 2011-12 (<http://accountability.universityofcalifornia.edu/index/10.2.5>)

students, and staff at the UC campuses as well as international scholars and community members who visit them. A systemwide, strategic approach to the development of library collections and services has emphasized multi-campus collaboration, the application of new technology, and expanded university-wide sharing of information resources. CDL sits at the nexus of the UC Libraries' collaborative endeavors.

CDL's services include negotiating systemwide content licenses worth \$40 million annually, providing platforms and services for digital archive collections, facilitating shared approaches to the management of print collections, and coordinating the Libraries' participation in largescale book digitization projects. CDL oversees the Libraries' current digitization partnership with Google and organized similar previous partnerships with the Internet Archive and Microsoft. In all, UC Libraries have digitized 3.6 million volumes from their collections through these efforts, three million of which are believed to be still under copyright. The UC Libraries are also founding members of HathiTrust and the second largest contributor to the HathiTrust corpus of eleven million digitized volumes. The issues that are the subject of the current inquiry are thus of vital interest and importance to us.

UC is a prime example of the size and scale at which the nation's largest libraries operate. Collectively, the 4,000 or so academic, research, and large public libraries in the U.S. hold approximately 2.8 billion items in their collections.² Of this enormous legacy, it's been estimated that 70% is still under copyright.³ Many of the published works are esoteric works of scholarship that are long out of print; and although estimates of orphan works are difficult to come by, according to one British Library estimate, well over forty percent of all creative works in existence may be orphaned.⁴ In another analysis, John Wilkin, former Executive Director of HathiTrust, hypothesized that as many as 2.5 million works in the HathiTrust corpus (800,000 of which are US publications) were likely to be orphans.⁵

The size of research library collections is matched by the scale of collective financial investment they represent. The 125 members of the Association of Research Libraries alone have spent \$50 billion in inflation-adjusted dollars on purchased materials since 1923 (that is to say, in the current in-copyright era in the U.S.),⁶ along with an equivalent amount in staffing and operational costs. Since these largest libraries hold fewer than half of the 2.8 billion items held by all academic and research libraries, the overall stewardship investment of America's academic and research libraries since 1923 may be closer to \$200 billion.

It is this enormous legacy – three billion items, and \$200 billion in stewardship investment, much of it in public funds – that is at stake in the digital transition. The long-term stewardship role of libraries and the significant financial investment this stewardship represents, along with libraries' public service mission, must be factored into any approach to orphan works.

² American Library Association, *The Nation's Largest Libraries: A Listing By Volumes Held*, <http://www.ala.org/tools/libfactsheets/alalibraryfactsheet22>; National Center for Education Statistics, *Supplemental Academic Libraries Survey (ALS) 2012 Tables to NCES 2014-038*, http://nces.ed.gov/pubs2014/2014038_supp.pdf; Library of Congress, *Fascinating Facts*, <http://www.loc.gov/about/fascinating-facts/>. (viewed May 20th, 2014)

³ Wilkin, John, *Bibliographic Indeterminacy and the Scale of Problems and Opportunities of "Rights" in Digital Collection Building*, Council on Library and Information Resources, 2011.

⁴ British Library, *Intellectual Property: a balance*, n.d. <http://www.bl.uk/listentonature/ip/pdf/ipmanifesto.pdf> (point 5).

⁵ Op cit.

⁶ Unpublished analysis conducted by the Association of Research Libraries, based on annual member statistics.

With this context as a backdrop, we would like to address four of the topics raised in the Copyright Office’s Notice of Inquiry: The need for legislation in light of recent legal and technological developments; Defining a good faith, “reasonably diligent” search standard; Mass digitization, generally; and Extended collective licensing and mass digitization. Our comments also embrace and extend previous comments submitted by the Council of University Librarians of the University of California in February 2013, available at http://www.copyright.gov/orphan/comments/noi_10222012/University-California-Libraries.pdf.

The Need for Legislation in Light of Recent Legal and Technological Developments

Orphan works legislation is not necessary to support library use of orphan works.

The California Digital Library strongly believes that orphan works legislation is not necessary to support library use of digitized copies of orphan works.

The uses that libraries make of digitized orphan works in their collections generally fall into three categories:

- digitization for preservation purposes;
- digitization for transformative purposes such as indexing, full-text search, non-consumptive use, and access for the print-disabled; and
- digitization of archival collections of largely unpublished material for access and use.

In the HathiTrust and Google Books cases, two courts have recently ruled that the first two categories—preservation and transformative use—are fair uses for which copyright permission is not needed, thus the orphan or non-orphan status of any works acted upon in this way is not relevant. While these decisions have been rendered in the context of specific book-based digitization projects, we believe these principles are relevant to library collections of all types. Because these uses enhance but do not substitute for access to the original protected expression (except in the case of print-disabled services), the interests and intent of rights holders, including those whose work was designed for commercial exploitation, are not compromised or harmed.

The third category of use, digitization of archival materials and primary sources, is also regularly undertaken by libraries pursuant to a fair use analysis, consistent with library best practices.

The UC Libraries February 2013 comments cited above contain many examples of our archival collections and the application of best practices; we refer the Office to that document. Another example closer to home is *Calisphere*,⁷ a public gateway developed by CDL to more than 220,000 primary sources—including photographs, documents, newspaper pages, political cartoons, works of art, diaries, transcribed oral histories, advertising, and other unique cultural artifacts — that reveal the diverse history and culture of California and its role in national and world history. *Calisphere* was conceived as a public service of the University of California to support the California K-12 educational

⁷ <http://www.calisphere.universityofcalifornia.edu/>

curriculum and showcase the rich history of our state as a service to California citizens, as well as to support undergraduate scholarship at UC. Content is contributed by the libraries and museums of the UC campuses and from more than 100 cultural heritage organizations across California, including material developed under the auspices of the California State Library's LSTA-funded Local History Digital Resources Project.

CDL's experience in supporting *Calisphere* and in working with content contributors demonstrates the strength of a best-practices approach to digitization projects that may involve orphan works. The Local History Digital Resources Project guided new curators in copyright evaluation and best practices through workshops and best practice guidelines; many of the LHDRP libraries have subsequently become regular *Calisphere* contributors. Liberal takedown policies complemented by readily accessible contact information on the *Calisphere* website and a clear process for responding to user inquiries provide an avenue for interested parties to engage with us about rights statuses. In nearly a decade since the service has been available, the only known contact from a rights holder was a 2013 example in which a graphic artist asked that certain reproductions of lesser quality be replaced with higher quality versions; the artist worked with the curator of that particular collection at one of our campuses to supply replacement versions, and was also able to obtain high-quality reproductions of some of her other works for her own use from the repository.

Calisphere's community impact illustrates the public benefit of bringing this rich cultural record alive through digitization. Many of our users have personal connections with the materials presented on the website. It is not unusual for CDL to receive feedback from users stating that they have identified family members, friends, or long-gone favorite buildings depicted in the collection.

Sometimes a single image can help inspire a discussion among a community of users. Last year, one particular image became the center of much conversation and personal reflection. The image is a page of *The Aquila*, a 1945 yearbook showing high school students at the Tule Lake Japanese American Relocation center. In late July, we noticed that this image had become very popular, having received more than 7,000 unique visitors on a single day. The spike in usage was a direct result of a link to the image that a user posted to the social bookmarking and networking website Reddit. The user was posting on behalf of her grandmother, who lived in Tule Lake during high school and whose picture appears in the yearbook page. Together, they answered questions from other users about the grandmother's experience living, working, and attending school in the camp. Almost 500 comments and questions were subsequently added to the discussion, with several people stating that this was the first they had heard of internment camps and this period of American history.

Another digital archives example with a more explicit preservation purpose is *California Light and Sound*,⁸ a database of historical moving images and audio from the California Audiovisual Preservation Project developed in partnership with forty libraries, archives and museums that provides "glimpses and whispers" of California's rich audiovisual heritage. Funded by the California State Library, the National Endowment for the Humanities (NEH) and the Institute of Museum and Library Services (IMLS), the California Audiovisual Preservation Project gathers best archival practices for moving image and sound preservation and establishes low-cost, practical standards to help collecting organizations move from the analog age to the digital age. Included in these standards is an extensive best practices guide to

⁸ <https://archive.org/details/californialightandsound>

permissions research,⁹ based on long-standing guidelines developed by the Society of American Archivists.¹⁰ Because of the greater sensitivities that can be involved in audio-visual works and the expense of digitization, item-level searches are conducted for all material in this collection, and permission sought when possible; the resulting copyright determinations, both conclusive and inconclusive, are made available in the public metadata for each work.

That these decisions and practices are sound is evidenced by the many such collections made available by libraries and archives over decades of digitization, and the virtually complete absence of protesting rights holders. When rights holders find these projects, as they not infrequently do, they are typically supportive and happy to see their work valued and made available in a library setting; if they do have objections, privacy rather than copyright concerns are more often what is at issue. We are not aware of any rights holder takedown requests for material in UC archival collections for reasons of commercial interest or economic harm. As the Copyright Office has noted, newer identification and search technologies are beginning to make identifying works and locating rights holders easier than ever, promising to mitigate even further an already rare problem.

For all of these reasons, we believe that the digitization and use of orphan works by libraries is adequately and effectively addressed by fair use, obviating the need for orphan works legislation. However, should the Copyright Office choose to pursue such legislation with the interests of other stakeholders in mind, we strongly urge that the non-commercial uses currently being made by libraries and archives – digitization and other forms of reproduction for preservation purposes, transformative uses such as indexing, full-text search, non-consumptive use, services for the print-disabled, and display of unpublished works in archival collections pursuant to a fair use analysis and following best practices – with limits on remedies and robust takedown mechanisms available to rightsholders – be fully supported in such legislation, either in the form of a fair use savings clause or through more explicit provisions.

Defining a Good Faith “Reasonably Diligent Search” Standard

Fair use is the appropriate strategy for libraries and archives wishing to make non-commercial uses of materials in their collections. For uses that exceed fair use, the Copyright Office should establish basic, minimal search standards that require a search of Copyright Office records coupled with support for a community best practices approach.

The question posed by the Copyright Office in the current Notice of Inquiry is how best to define a “reasonably diligent search standard” and what role, if any, the Copyright Office should have in developing and coordinating the development of standards and best practices for diligent search.

As described above, we believe that fair use covers the most important uses that libraries are making of orphan works today. In the case of digitization for preservation, transformative non-display uses and services for the print disabled, searches are not required.

The question of a diligent search standard thus arises for libraries chiefly in the context of a fair use evaluation of archival collections being digitized for access and display. In these cases, it is typical for the library to evaluate, on a collection-by-collection basis, the need for and likely fruitfulness of a search

⁹ http://calpreservation.org/wp-content/uploads/2014/01/2014CAVPP_Permission_Guidelines.pdf

¹⁰ <http://www.archivists.org/standards/OWBP-V4.pdf>

for copyright holders for particular materials or categories of materials in the collection, taking into account such factors as the age of the material, whether the works in question were created for a commercial purpose or are likely to be of commercial interest to their copyright owners, and related concerns such as privacy and donor restrictions.

Libraries do often conduct searches for these materials, but the cost/benefit of such searches is extremely problematic. Others have spoken and written about the difficulty and expense of these searches, as well as their typical outcomes: an inability to identify rights holders in many cases, and permission or non-response when a rights holder is found. Some of that work is discussed in the previous UC submission and need not be recapped here.¹¹ The recent March Roundtables surfaced fresh instances of these difficulties. For example, Greg Cram (NYPL) described his library's inability to identify rights holders for its collection of 1939 World's Fair photographs, despite having all of the corporate records from the Fair;¹² while numerous rights holders and rights organizations confirmed the tangled rights situation for many works in their own field of expertise and the disinclination of possible rights holders to respond to requests that might impose a burden of research.^{13,14} Moreover, rights holders of work that was not created for commercial purposes in the first place have little incentive to claim an interest in the works they own. Despite (or perhaps because of) these complexities, while non-response rates to libraries and archives are typically high, denial of permission is rare.

At the same time, many individuals in the March roundtables affirmed their support for the appropriate use of orphan works by libraries and archives in fulfilling their educational and non-profit missions, focusing their concerns on orphaned works in other 'real world' contexts.¹⁵ Rights holders and content creators worried that their works intended for commercial use would be misappropriated as orphans without compensation, depriving them of a livelihood; those wishing to make creative uses of orphan works worried that the inability to identify and procure rights was suppressing beneficial uses of older works. Some of these communities want flexible search standards, while others want certainty in order

¹¹ See e.g. Dharma Akmon, "Only with Your Permission: How Rights Holders Respond (or Don't Respond) to Requests to Display Archival Materials Online," *Archival Science* 10 (2010):45-64; Maggie Dickson, "Due Diligence, Futile Effort: Copyright and the Digitization of the Thomas E. Watson Papers," *American Archivist*, 73:2 (2010): 626-636.

¹² Orphan Works and Mass Digitization Roundtables (Hereafter "Roundtables") March 11, 2014, p. 26.

¹³ Roundtables, March 10, 2014. "Because right now ... there are orphan sound recordings from 1980 out there because we can't get the labels to admit that they own it" (Prager, p. 158)."

¹⁴ Roundtables, March 11, 2014.. "...these problems that I run into all the time where no one really knows what to do. And it is not just individual users, but large organizations. National Geographic has a group of images they have no idea whether they own the rights to" (Hill, p. 269); "In music, we have had ... tremendous difficulty over the years keeping track of ownership with mergers and acquisitions of record labels and publishers." (Rae, p.288).

¹⁵ Roundtables March 10, 2014. "The most sympathetic case for orphan works is the case of the archive of historical records where works are involved that are very important for researchers, or just important for people to have access to, where it is extremely difficult to determine the provenance of the work -- largely because they were not really created for public distribution in the first place... So I think that it is reasonable to permit access to those kinds of works in a digital archive that, for example, is a library" (Lehman, p. 293); "I wanted to address this notion of educational use and scholarly use. And, certainly, that's already permitted in fair use, as is preservation for archives." "And we keep going back to the discussion of mass digitization of collections of libraries or archives And I think that what most of the cultural nonprofits want to do, in terms of preservation and their collections and out-of-print works and orphan works, is already allowed under fair use." (Shaftel, p. 377 ff)

Roundtables, March 11, 2014. "The digitization issue can be dealt with through a takedown procedure that can be created so that let's say you have a mass digitization of the archives. You put it up. If there was a takedown provision ... to say hey, we don't want that up there because it is ours and you didn't have it, that is one way of dealing with that area." (Prager, p. 351)

to mitigate risk. None of these use cases impinged upon the non-commercial context of libraries and archives, which operate on the basis of a completely different set of concerns. The range of views and the many specific circumstances pertaining to particular communities, uses, and types of works, as well as the many ongoing developments and innovations in search technologies and registries of all sorts all argue for a flexible, rather than a rigidly defined, search standard, as well as for reliance on best practices to guide different communities of users.

These outcomes and reports support the conclusion that *a fair use evaluation is the appropriate strategy for libraries and archives wishing to make non-commercial uses of materials in their collections*. A fair use assessment in the context of digitization and display of archival collections will generally turn on the purpose and character of the use (for non-profit education and research), the nature of the material (generally older unpublished works created without commercial intent), and the effect on the market for the work (generally unpublished, not created for commercial exploitation, and not currently commercially exploited by the copyright holder). Such evaluations recognize the need for searches in appropriate contexts, guiding the library in determining when a rights holder search is appropriate and likely to be fruitful. Crucial to the effectiveness of this approach is that rather than requiring item-by-item searches in every case, libraries are able to rely on professional judgment, experience and expertise in determining when a search is or is not warranted. A legal requirement to conduct a diligent search for every item in an archival collection selected for digitization and display would be prohibitively costly and would disincentivize libraries from making the invaluable cultural heritage materials in their collections available for use by students and scholars.

The Copyright Office should support all of these communities by establishing basic, minimal search guidelines applicable to all types of works, as opposed to legislation, that require a search of Copyright Office records (as those records are being digitized and updated) coupled with affirmation of a community best practices approach to search standards in particular contexts. Such guidelines would provide an incentive to more rights holders wishing to be found to register their work and would encourage additional communities to develop best practices in their area of expertise. Digitization and updating of Copyright Office records, as well as streamlining the registration process, should be an important element of this strategy.

The Copyright Office has asked whether such best practices should be allowed to be developed by user communities themselves or should involve representation from rights holder groups, and what its role should be in the development of such practices. While some communities of rights holders and users may find it productive to collaborate in this area, we strongly agree with Jonathan Band's comments that requiring negotiated practices overseen by the Copyright Office would be unworkable and would slow the evolution of good practices.¹⁶ An attempt to mandate negotiated best practices will repeat the failures of the Conference on Fair Use (CONFU) and result in guidelines that are neither used by nor of use to any stakeholders. By contrast, the successful application of best practices that various communities have adopted to date is a testament to their real-world viability.¹⁷

¹⁶ Roundtables, March 10, 2014, p, 176 ff.

¹⁷ See for example, Society of American Archivists, *Orphan Works: Statement of Best Practices*, 2009, p.3: "As professionals governed by common sense, archivists understand that more effort should be expended to locate the copyright holder of a more recent work because the search has a higher likelihood of success. Similarly, because individuals who have earned their livelihoods through their creative works presumably want to be found, a higher standard should apply to searches for professional authors and artists than for amateurs. Professional creators often do a much better job of making themselves findable through copyright office records, authors and photographers associations, and similar tools and databases, and

Mass Digitization, Generally

Mass digitization should not be conflated with orphan works policy. The term ‘mass digitization’ should be retired from discussions of copyright reform.

We agree with many other commentators that digitization, whether mass or otherwise, should not be conflated with orphan works policy. Digitization in libraries is undertaken for works of all types and in all copyright statuses, for purposes directly related to our long-term stewardship mission, generally under a fair use premise. Orphan works are only incidental to these efforts.

We further recommend that the term ‘mass digitization’ be retired from discussions about copyright reform and libraries because of its potential to confuse and color the debate. The term ‘mass digitization’ arose in the context of specific largescale book digitization partnerships between libraries and third party entities such as Google and the Internet Archive and is most commonly associated with the Google Books Project, but this history and the controversy surrounding some of these projects makes the term unhelpful.¹⁸ From a library perspective, these activities exist on a continuum with other digitization activities, all having the same general purpose and set of goals of improved discovery and findability, preservation, and support for new forms of user interaction and engagement with existing collections. We believe it is more appropriate to refer to these activities simply as library digitization, or more specifically as large-scale or collection-level digitization; they are just one strategy among many to digitize library collections at scale. The manner of digitization should not be the focus of copyright discussions.

That said, the term mass digitization as it has been applied to libraries can also be thought of as a kind of shorthand for the larger goal of converting legacy library collections, both published and archival – those three billion physical items held in U.S. research library collections – to digital form on a systematic scale in order to preserve those collections and make them fit for use in the digital age. This is increasingly recognized as a critical imperative for sustaining the cultural heritage mission with which libraries are uniquely charged, prompting us to digitize more and more collections at whatever scale and pace technology, available resources, and careful stewardship afford.

At the March Roundtables and elsewhere, this has been appropriately referred to as format-shifting,¹⁹ which many assert to be a clear fair use. It is the subsequent uses to which the materials are put that are at issue in a further fair use analysis; and the uses that we have already enumerated – preservation, indexing, full-text search, non-consumptive use, and services for the print-disabled, as well as the display of unpublished works in archival collections subject to specific evaluation by a library or collection curator – we believe also to be fair use. While there are those who dispute this interpretation, we would respectfully point out that the courts have thus far supported it.

orphan works searchers should expend greater effort to pick up the trail that they may have left behind.”

<http://www2.archivists.org/sites/all/files/OrphanWorks-June2009.pdf>

¹⁸ As one roundtable participant noted, “In this country, as we well know, our conversations about mass digitization started in 2004 because of the Google effort. And the term itself really comes out of the Google effort and we are a bit stuck in that model” Roundtables, March 11 (Pilch, p. 101-2).

¹⁹ Roundtables, March 11, 2014. (Carroll, p. 99)

All of this in itself has little to do with orphan works, which are simply one category of work included in library digitization projects. The issue of creating a path for broader uses of orphan works should not be conflated with library digitization. We appreciate that the now-suspended HathiTrust Orphan Works Project may have precipitated this linkage and that in posing the question about the relationship of orphan works and mass digitization, the Copyright Office is seeking authentic solutions to the conundrum of how to enable broader uses of truly orphan works that might satisfy the needs and interests of all stakeholders in similar projects.

In considering this issue, the Copyright Office has also posed a question in the NOI as to whether other potential solutions developed in the context of mass digitization might ameliorate the issue of orphan works. The Office has proposed extended collective licensing as one of these possible solutions. Following our comments about extended collective licensing below, we will put forward a fresh alternative proposal to consider, one that is not limited to orphans and which we believe can make a positive contribution to the discussion.

Extended Collective Licensing and Mass Digitization

An extended collective license (ECL) is not a viable solution for library uses of orphan works.

An extended collective license (ECL) is not an appropriate or viable solution for library uses of orphan works, even if such a license were restricted to published orphan works resulting from so-called ‘mass digitization’ projects.

An extended licensing scheme would be prohibitively costly and would have a significant negative long-term impact on future library purchasing power, purchasing power that is already severely challenged by inflationary pressures and chronic underfunding. The University of California Libraries have thus far digitized some 3.6 million books from our library collections, 3 million of which are believed to be still in copyright. If the British Library’s estimate is correct that 40% of the material in a library’s collection may be orphaned, between 300,000 and over 1 million works in our collections digitized to date might be subject to an orphan works license. And as this represents less than ten percent of our total book holdings, this cost would continue to grow as more works are digitized.

Second, given the difficulty in locating orphan rights holders, a present rights holder would not be available to receive compensation in the vast majority of cases. A great deal of expense would be consumed by administration of the organization itself as well as in fruitless or inconclusive searches, with surplus funds ultimately going to beneficiaries unrelated to the copyright owner. With few rightful claimants, the revenue siphoned from library budgets to sustain an ECL would serve no valid purpose of copyright.

Third, it is not at all established that absent rights holders have a continued economic interest in their work. Without an organized outreach effort of any sort, HathiTrust has opened up access to 14,000 works at the request of individual rights holders who have discovered their works in the HathiTrust corpus and requested that they be made fully available for use, without remuneration. A new authors group, the Authors Alliance, has recently been formed to speak for the interests of academic and other authors who would like to see broader access to their works. Many long-forgotten authors and rights holders of works in academic and research library collections are surely in this category.

Fourth, library investment in the stewardship of collections is already very great. As we described earlier, academic and research libraries in the U.S. have invested some \$200 billion in the acquisition and long-term stewardship of their collections over the past century alone. Indeed, publishers and rights holders frequently turn to libraries for out-of-print copies of their published work when needed for digitization and other projects, because it is libraries that have invested in their care. A collective license would subject libraries to a form of double-jeopardy for having preserved these older works and kept them in circulation and available to the public long after their copyright owners have abandoned them.

Rather than placing a backward-looking tax on old books that aren't economically viable, scarce library budgets are better spent on new content purchases that directly benefit and amplify the reach of today's authors and stimulate new creative work, advancing the fundamental purpose of copyright.

Lastly, since libraries are making fair uses of digitized works in their collections now, and have no interest in an extended collective license for broader use, it is unclear what purposes or markets an extended collective license might serve. An ECL is intended to address a market failure, but a market failure requires a market clamoring to be served. The March Roundtable discussions revealed little interest in extended collective licenses among other stakeholders, although support for voluntary licensing for commercial uses was strong.

Libraries themselves are part of that vibrant voluntary licensing market today. Libraries are major contributors to the economic well-being of the content industries, spending some \$4-5 billion dollars in content purchases every year.²⁰ The market for both current and retrospective digital content is strong. We would go so far as to posit that at least as far as libraries are concerned, the tensions we are experiencing around the digitization of legacy collections are a bounded problem, a bubble of sorts between the historical public domain on the one hand and the new realities of online content provision on the other. The services that libraries have adopted to bring their existing legacy collections into the digital age already coexist with a robust digital marketplace. They should be recognized as a complement rather than a threat to that marketplace, and supported accordingly; precisely the balance between private and public interest that copyright aims for.

A Better Solution for Libraries

A better solution for libraries would be to allow the display of digitized versions of in-copyright published works in their collections under limited conditions which mimic access to the originals, designed with appropriate safeguards for works in which there is or may be a commercial interest.

At the March Roundtables, the Copyright Office raised the question of how to provide for library display of digitized in-copyright works. An approach that permitted the controlled display of digitized in-copyright works contained in a library's collection would be a viable solution to this problem.

²⁰ Data sources: ALA, NCES, and STM – see <http://www.ala.org/tools/libfactsheets/alalibraryfactsheet04>, <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2014038>, and http://www.stm-assoc.org/2012_12_11_STM_Report_2012.pdf

Such an approach, whether achieved through a formal library exception or some other framework, could be carefully crafted to include adequate safeguards against redistribution and other potential market harms for rights holders.

Measures that could be considered might include:

- Restricting display to only the users of the library that owns or owned an original copy of the work, using well-established authentication mechanisms
- Restricting the number of simultaneous users to the number of physical copies owned by the library (one copy = one user)
- Employing controlled access mechanisms such as streaming or other technological protection measures, through which users are prevented from downloading or retaining a permanent copy of the item
- A carve-out for materials of current or potentially current commercial interest, such as a restriction on display of materials published within the last 20 or 25 years
- A robust takedown or opt-out mechanism

A carefully crafted approach along these lines would solve the ‘last mile’ problem of collection-level digitization of published works—the inability to display the full text. Legitimizing this through an agreed-upon framework with carefully designed market safeguards would solve the problem of library digitization of older published works for both libraries and rights holders.

To be eligible for such treatment, the works under discussion would have to be held and have been legally acquired by the library in question. The display use thus enabled would be a clear instance of format-shifting, enabling the library’s users to access the digital surrogates as if they were the originals.

Some rights holders will object that this proposal entails providing two copies – one digital, one physical – where formerly there was only one. However, there is important additional context for this proposal, arising from the physical reconfiguration that is taking place in libraries throughout the country, which is in fact causing a reduction in the number of physical copies of published works in library collections. Academic libraries in the U.S. collectively house more than 1 billion books and journals; many of these libraries are running out of space. Like many of our peers throughout North America, the University of California Libraries are currently involved in a long-term space planning effort to free up library space for newly acquired collections as well as for transformative uses of library buildings such as group study and digital commons areas. At the same time, the circulation of books in libraries is declining sharply²¹ as research moves online. Consequently, an enormous amount of work is going on in libraries today to reduce the physical footprint of library book collections, with many libraries consolidating their collections in lower cost shared storage repositories located at a significant physical remove from the user. The University of California maintains two such shared storage facilities in the northern and southern portions of the state, in which are housed some twelve million of the 38 million volumes in our

²¹ See for example Anderson, Rick, *Print on the Margins: Circulation Trends in Major Research Libraries*, *Library Journal*, June 2, 2011. <http://lj.libraryjournal.com/2011/06/academic-libraries/print-on-the-margins-circulation-trends-in-major-research-libraries/#>

collections. These facilities themselves are also running out of space, necessitating even newer approaches to shared collection storage and, in some cases, de-accessioning of physical collections.

These developments are relevant to the conversation about copyright and full display of digitized works, because the challenge of library space is one of the major drivers for large-scale digitization of our collections; even as we sequester a dwindling number of physical copies in off-site storage, our books are more findable now thanks to full-text indexing of the digitized copies. This is good news for everyone: consigning these books to obscurity isn't in the interest of authors any more than it is in the interests of readers. But while indexing and full-text search are doing much to compensate for the inconvenience of remote storage, common sense says that as the habits of students and scholars move online and the physical copies are deprecated in importance, it should be possible to let our users read the digital surrogates as substitutes for the originals. In fact in a very real, if not literal, sense, the digitized copies *are* the original library copies, with all of their markings, marginalia, and other venerable badges of use. And as a result of this vast thinning of the library forest that is slowly but surely taking place, many of these digital copies are arguably in their effect, replacement copies – where once there may have been 50 physical copies housed in as many libraries, there will soon be 40, or 30, or 20, with a digitized surrogate standing in for these newly de-accessioned copies. While some of the books being physically de-accessioned in libraries may be relatively recent works for which a digital version is commercially available and which libraries will acquire in that new format, many others are old, esoteric, and long out of print – legally acquired copies of mostly very old books that libraries have collectively spent billions of dollars to purchase and additional billions to house and maintain for nearly a century.

A new copyright framework, which mimics the access conditions of the physical books and designed with appropriate safeguards for works in which there is a commercial interest – something that we fully believe is achievable - would be a far better solution to the display of older in-copyright works and would better serve the underlying purpose of copyright.

Thoughtful discussion of such an approach deserves equal consideration to other options that are being put forward to address the more limited problem of orphan works and should be part of the conversation around copyright reform.

—

We thank the Copyright Office for soliciting these comments and devoting such thoughtful attention to a viable way forward for orphan works, library digitization, and the many additional issues posed by copyright in the digital era. The California Digital Library is pleased to be part of this dialogue and looks forward to further opportunities to participate in these consultations.

Sincerely,



Laine Farley
Executive Director
California Digital Library

CC: University of California Council of University Librarians