

**BEFORE THE UNITED STATES COPYRIGHT OFFICE**

In re: Mass Digitization Pilot Program; Request for  
Comments

Docket No. 2015-3  
Fed. Reg. Vol. 80, No. 110

**COMMENTS OF PUBLIC KNOWLEDGE AND THE ELECTRONIC FRONTIER FOUNDATION**

Attn: Kevin Amer, Senior Counsel for Policy and International Affairs

Public Knowledge and the Electronic Frontier Foundation respectfully submit the following comments in response to the Copyright Office's recent Notice of Inquiry regarding "Mass Digitization Pilot Program; Request for Comments" dated August 10, 2015 (the "Notice"). Public Knowledge is a nonprofit organization dedicated to representing the public interest in digital policy debates. Public Knowledge promotes freedom of expression, an open internet, and access to affordable communications tools and creative works. The Electronic Frontier Foundation is a nonprofit civil liberties organization that has worked for more than 20 years to protect consumer interests, innovation, and free expression in the digital world. Founded in 1990, EFF represents almost 20,000 contributing members. EFF and its members have a strong interest in promoting balanced intellectual property policy that serves both public and private interests. Through litigation, the legislative process, and administrative advocacy, EFF seeks to promote a copyright system that facilitates, and does not impede, "the Progress of Science and useful Arts."

While we believe in expanding access to knowledge through mass digitization, the proposed pilot program does not bring us any closer to achieving that goal. We urge the Copyright Office to reconsider its approach to Extended Collective Licensing.

Earlier this year, the Office released its most recent Report on orphan works and mass digitization, "Orphan Works and Mass Digitization: A Report of the Register of Copyrights" (June 2015) (the "Report"). In the Report, the Office reviewed the landscape of frameworks for mass digitization, concluding that neither the fair use doctrine nor voluntary agreements could lead to true mass digitization - fair use left too much uncertainty, and voluntary agreements left out too many works. From this conclusion, the Office then moved to propose an Extended Collective Licensing (or "ECL") regime as a solution, and published the Notice of Inquiry to which these comments are directed.

Both the Report and the Notice incorrectly merge two related but distinct activities - the act of digitizing, and the act of making full-display copies available. To date, the appellate case law is clear: mass digitization of printed works for the purpose of creating a keyword searchable database and providing accessibility for the visually impaired, are fair uses and therefore beyond the scope of copyright's grant of limited exclusive rights. Indeed, even the display of snippet views of scanned texts has been held a fair use.

The problem then, to the extent one exists, is not about mass digitization. Legal authority to scan, index, increase accessibility, or serve context-providing snippets already exists. The problem is limited to the display and distribution of the entire work, and the transaction costs of securing the rights to do so.

However, the proposed program does little to address these issues. By limiting the scope of the pilot program to nonprofit educational and research uses, which cannot give the user even an indirect commercial benefit, and suggesting that the pilot program should be even further limited by restricting the use of in-print books, and restricting access to students employees and affiliates of the licensed institution, the Office is effectively proposing an ECL for limited e-reserve systems. In addition, we have doubts as to

whether the proposed ECL pilot program will achieve any efficiencies in transaction costs, and concerns about how the Office intends to ameliorate the structural problems necessarily posed by any Collective Management Organization (“CMO”). Finally, we have concerns about the privacy implications of centralized CMO-based licensing system, and the Office’s insistence on imposing Digital-Rights-Management-like requirements on digitization projects.

## **I. The Proposed Pilot Program Is A Solution In Search of a Problem**

As a threshold matter, in reviewing the Office’s proposed contours for an ECL pilot program, it is unclear that it will meet a real demand. While the classes of works identified by the Office are ripe for promoting digitization, the licensees to whom the Office intends to restrict ECLs have expressed little interest, and outright opposition, to such a licensing regime.

The Office’s proposal calls for limiting potential licensees under the pilot program to those making “nonprofit educational and research [uses]... without any purpose of direct or indirect commercial advantage.” In its initial comments, the Library Copyright Alliance, representing the American Library Association, the Association of Research Libraries, and the Association of College & Research Libraries, stated that “any legislative approach that involves licensing, such as extended collective licensing, is completely unacceptable to the library community.”<sup>1</sup> At the roundtables, this view was repeated by other organizations. For example, Debra LaKind of the Boston Museum of Fine Arts noted that from the perspective of a museum, individual licensing was preferred to collective licensing, both as a potential licensee, and licensor.<sup>2</sup> Given the skepticism from two of the largest types of nonprofit educational institutions, libraries and museums, it seems unlikely that the potential targets of the program will actually take advantage of it.

The Office frames its choice to limit ECLs to nonprofit educational and research uses as a means to restrict the scope of ECL licensed projects only to those “serving the public interest.”<sup>3</sup> However, nonprofit uses are not the only ones that can serve the public’s interest. The public interest is often promoted not merely by nonprofit organizations dedicated to that purpose, but through policies that shape private parties to act in ways that benefit the public.

One of the largest mass digitization projects ever undertaken, the Google Books Library Project, is an example of a public/private/nonprofit partnership that has tremendously benefitted the public - so much so that the public benefits of the program were specifically highlighted by the district court judge in ruling that Google’s digitization and snippet displays constitute a fair use.

While the Report suggests that “a for-profit entity would not be precluded from undertaking a mass digitization project (such as through a partnership with a nonprofit library or educational institution),” the proposed prohibition against even “indirect commercial advantage” will likely discourage such partnerships from materializing. For example, if a private party partners with a nonprofit to digitize materials, and uses the project to enhance and develop its digitization techniques, which it then deploys in a profit-making context, is that indirect commercial advantage? What if it uses the corpus of digitized works to engage in machine learning or algorithm development that is then integrated into its commercial products?

Similarly, even nonprofit institutions might gain indirect commercial advantage by engaging in mass digitization of their holdings. For example, museums may attract more researchers, more sponsors, or more visitors as a result of digitizing some of their holdings. By pre-determining what kinds of uses can

---

<sup>1</sup> Comments of the Library Copyright Alliance, January 14, 2013 (available at [http://copyright.gov/orphan/comments/noi\\_10222012/Library-Copyright-Alliance.pdf](http://copyright.gov/orphan/comments/noi_10222012/Library-Copyright-Alliance.pdf)).

<sup>2</sup> Transcript of Public Roundtable, March 11, 2014, pg. 227 (available at <http://copyright.gov/orphan/transcript/0311LOC.pdf>).

<sup>3</sup> Orphan Works and Mass Digitization: A Report of the Register of Copyrights (June 2015), pg. 89.

serve the public interest, the Office further diminishes the demand for licenses under its proposed ECL regime.

Some of the specific limitations that the Office contemplates in the Notice only raise further doubts as to the program's value. The Office is considering whether to limit the program to only out-of-commerce books, only books printed before a certain date, only allow access through onsite computers, or only allow access to students, employees, and affiliates of the licensee. Taken together, these limitations suggest that ECLs issued under the pilot program may offer only a limited benefit to institutions over extant resources.

## **II. The Proposed Program Provides Little Additional Efficiency in Licensing**

If ECLs are intended to reduce the transaction costs between potential users of digitized collections and rightsholders, the proposed system does not appear able to achieve this goal. Rather, the transaction costs are merely shifted from the potential licensees to the CMO. Doing this, however, does not reduce location or negotiation costs.

Assuming the CMO is an existing membership organization that can currently license rights to its members' works, it stands in exactly the same shoes that it did before. Any additional efficiencies that might result from the proposed program would therefore come from a CMO's ability to license works to which it currently lack rights.

If efficiencies are reached in this way, it is worth considering the source of that value. An author may not be a member of an existing licensing body for a variety of reasons. Although an author may simply be unaware of the body, it is equally likely that the author has decided not to become a member.<sup>4</sup> Authors may decide that they would like a higher rate for the use of their works than a rights organization may pay through to them; just as easily, they may object to a third party collecting licensing fees for works they would prefer to see more readily distributed.<sup>5</sup> Even authors who have considered, but failed to decisively act, upon an opportunity to join a collecting society are making a rational calculation—the cost of investigating the terms and conditions of that membership, added to whatever rights they would be ceding, may outweigh the estimated benefits from such membership. Efficiencies gained by handing these rights to a CMO would be gained at the expense of such non-member authors.

A CMO funded by deducting operating costs from collected licensing fees will also have certain incentives in how it will expend its resources in locating and distributing royalties. If usage patterns in media and publishing as a whole apply in these licensing programs as they do elsewhere, a small number of works will generate a large proportion of uses. Between popular and unpopular member authors, the location and transaction costs of paying royalties will be nearly identical—which is to say that the cost per dollar distributed is far higher for a less popular work. Furthermore, the costs of finding non-members and searching for orphans will be significantly higher than the costs of finding members.

The Office's proposal to deal with this conflict—segregating search costs for authors to be drawn only from unclaimed funds<sup>6</sup>—actually exacerbates the problem of disincentivizing diligent searches, and effectively acts as a penalty to authors who do not join the collective. The chance for disbursement of leftover unclaimed funds to preferred charities, chosen by members, does not encourage more searching.

---

<sup>4</sup> Additionally, there are also works whose authors are incapable of consenting, such as orphaned works.

<sup>5</sup> “Why the Authors Alliance Amicus Brief Supports Google's Fair Use Defense,” by Pamela Samuelson, July 10, 2014 (available at <http://www.authorsalliance.org/2014/07/10/google-books-helps-authors-by-making-their-books-discoverable/>).

<sup>6</sup> *Orphan Works and Mass Digitization* (June 2015), pg. 100.

Even with the best intentions, a CMO will necessarily deliver less revenue per use to non-members than to its members. The CMO's ability to gather licensing fees for non-members can thus easily result in those authors losing revenue, as well as control, against their will.

### **III. Essential Definitions and Contours of the Proposal Remain Unresolved**

We recognize that we are asked to comment on an initial proposal for a pilot program, rather than a final proposal, and that many details will be filled in after this initial comment period. Nonetheless, we must note that the current proposal is so vague with respect to key issues that it is difficult to provide helpful input except by making significant assumptions about the programs' possible characteristics. The stated objectives of the program are so broad, and the mechanisms of the program so undefined that, as noted above, it is difficult to identify the specific stakeholders who will benefit from it, and what barriers others may have to finding value in its use.

In particular, the following critical questions remain unanswered, and consensus answers to them seem unlikely to emerge:

#### **CMO representation of non-members**

What would qualify a CMO to represent non-members? This issue was one of the fundamental stumbling blocks in the Google Books settlement—whether the proposed Book Rights Registry could fairly represent the rights of orphan work authors, or other missing or uncontacted rightsholders. The same question pervades this proposal. While Congressional action could conceivably make such involuntary representation legal, the answer of whether Congress or any other mechanism *should* has never been satisfactorily answered.

Apart from the basic question of qualifications, the representation of non-members naturally raises the issue of orphan works. While the Office acknowledges its parallel orphan works proposal in this Notice, it is impossible to discuss any mechanism that represents absent parties without addressing how orphan works will be handled. If a CMO can claim to represent any author who has not affirmatively joined the organization, it will naturally be able to claim representation of works whose authors are not present to object, artificially inflating the scope of its representation in any number of matters, as well as exercising a gatekeeper function for uses of works where no other party would be able to provide licenses or have standing to dispute CMO control of the work.

Whether a CMO can adequately, fairly, and efficiently represent non-members, either at an absolute level or compared to a statutory license, orphan works regime, or other mechanism, is wholly dependent upon how that CMO operates. It remains to be seen whether any CMO would have the incentives to do so under the proposed program; and the success or failure of such a program would have limited value in assessing a program with a broader scope or number of works, a broader scope or number of uses, or a broader scope or number of eligible licensees. For instance, the proposal in the Notice restricts the use of ECLs only to certain uses that provide no commercial advantage. Rather than providing a microcosm of a larger program, this limitation restricts the utility of the proposed licensing system by removing some of the most likely non-fair uses—the uses that actually need licensing—from its scope.

#### **CMO membership**

On what grounds might a CMO assume an author or a rightsholder to be a member? The Notice does not specify its plans for the establishment of CMOs. Could existing organizations of different rightsholders apply for CMO status, or would the CMOs need to be created from scratch? If the former, would existing members of an organization have joined had they been aware of the new ways in which the organization may now exercise control over their works? If the latter, can a proposed CMO acquire a large enough corpus of works through opt-in membership to create a worthwhile license?

Further issues arise with regard to CMO governance. While the Office indicates it will require a CMO to aver that it meets certain standards for consent of its membership, transparency, accountability, and governance, the mechanisms for verification and enforcement of these putative standards are not specified.

For example, the matters of transparency and accountability are key. Near-universal complaints about these very issues dog nearly every aspect of licensing in other areas, even where licensing organizations, their members, and their licensees are under variously stated requirements for transparency and accountability.<sup>7</sup>

Leadership may also be skewed. CMOs would naturally pledge to represent the interests of all of their members collectively. Yet, inevitably, some of those members will be larger, more unified, and more influential than others. Decisions made on behalf of a large set of authors would naturally reflect the interests of the most powerful members of the group, despite potentially large differences between their interests and the interests of small authors, authors whose works are created outside of a strictly commercial context, and others. A CMO substantially influenced by members who themselves control a large catalog of works may find itself steered in directions that prejudice the interests of smaller CMO members, to the benefit of direct deals between licensees and the large members.

The Notice asks about dispute resolution, but only with regard to disputes between a CMO and a licensee; the structure and operations of the CMO itself are also sources of substantial dispute. Would CMO failure to properly represent its members, disclose its dealings transparently, or abide by its agreements subject it to effective legal liability? Could failure to meet Office standards result in suspension or rescission of a CMO's licensing authority, and can such enforcement be applied fairly and effectively?

### **CMO designation**

The Office suggests that it be granted the ability to designate entities as CMOs, and suggests broad qualifications for eligibility. However, the characteristics, achievable objectives, and desirability of a licensing program depend upon exactly how a proposed CMO might meet these criteria. If the Office is to be the gatekeeper to collective management, it can lend additional clarity to this proceeding by stating with more specificity how it imagines a responsible CMO would ensure transparency, accountability, and good governance. However, the Office has not provided any more detail in its Notice beyond the same general statements present in the Report.

The Report's focus upon a proposed CMO demonstrating experience in administering collective licenses highlights a number of potential issues. If, as the Office proposes, the CMO is to represent non-members, experience with voluntary collective management may not be as applicable. More generally, the mere fact that an organization has issued collective licenses in the past does not in itself indicate that it has done so well.

Another foundational question is whether multiple CMOs might be contemplated. In the music licensing space, for instance, the presence of multiple PROs is often cited as an advantage in providing competition between them; it has also been criticized for reducing effective transparency with regard to catalog scope.<sup>8</sup> Particularly if CMOs are allowed to represent authors who have not actively signed up for membership, they would have conflicting interests over those authors' rights. On the other hand, a system

---

<sup>7</sup> See, e.g. "Fair Music: Transparency And Payment Flows In The Music Industry", by Rethink Music, a Project of Berklee Institute of Creative Entrepreneurship (available at [https://www.berklee.edu/news/fair\\_music\\_report](https://www.berklee.edu/news/fair_music_report)); "Cautionary Tales About Collective Rights Organizations", by Jonathan Band (available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2149036](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2149036)).

<sup>8</sup> Comments of Digital Media Association In the Matter of Review of ASCAP and BMI Consent Decrees, pgs. 13-19 (available at <http://www.justice.gov/sites/default/files/atr/legacy/2014/08/13/307972.pdf>).

designed with a single CMO in mind might require a heavier regulatory oversight burden to ensure that its monopoly status is not abused.

### **Scope of included works**

The Notice proposes “literary works” as one category of included works. Presumably, as with the Report, this is intended to cover books, specifically. It is worth noting that “literary works” would encompass a much larger category of works than books or even published print matter, to include electronic works such as computer programs, as well as matter created in and for radically different contexts, such as web pages, social media postings, and more.

### **User Privacy**

Licensing schemes necessarily would require some information about who is accessing works. To the extent that any system would collect user information, protecting user privacy would be paramount. This can certainly be a concern with images, but is heightened even more with books and other literary works.

Users should be able to access and use content without worrying that the government or third parties may be reading over their shoulders, accessing either their personal information or information about their reading habits. To ensure that any information it stores linking users to the content they view or purchase is not freely disclosed to the government or third parties, CMO authorization should be conditioned on commitments that the collecting agency and its licensees (1) will not disclose information about users and their identifiable uses or purchases to government entities or third parties absent a warrant or court order (unless they are barred from doing so by law); and (2) will notify the user prior to complying with any government or third party request for her or his information, unless forbidden to do so by law or court order.

In addition, just as readers may anonymously browse books in a library or bookstore, readers should be able to search, browse, and preview content without being forced to identify themselves. Thus, a CMO or licensee should ensure that searching and previewing content does not require user registration or the affirmative disclosure of any personal information. CMOs and licensees must commit that they will not connect any information they collect from individuals with those same individuals’ use of other services without their specific, informed consent; purge all logging or other information related to individual uses of no later than 30 days after the use to ensure that this information cannot be used to connect particular books viewed to particular computers or users; and allow users of anonymity providers, such as Tor, proxy servers, and anonymous VPN providers, to access and use the system.

In the interest of transparency and enforceability in the protection of reader privacy, at a minimum, each CMO and licensee should also provide a robust, easy-to-read, and easy-to-access notice of its privacy provisions; ensure that any commitment it makes to protecting privacy is legally enforceable and that all data it collects about its users is stored such that it is subject to U.S. legal protections; and annually publish online, in a conspicuous and easily accessible area of its website, the type and number of requests it receives for information about its users from government entities or third parties.

A CMO or licensee that cannot meet these standards should not be entrusted with user data.

### **Technical Protection Measures (“TPMs”) or Digital Rights Management (“DRM”)**

The Office recommends requirements for digital security measures to control access to collections and asks for specific measures. We do not believe that the Office should be involved in determining security standards, and it certainly should not require CMOs to implement TPMs or DRMs in order to operate. Nor should the Office involve itself in determining the technical standards under which licensees should be required to implement TPMs or DRM. DRM technologies frequently harm consumers, undermine competition and innovation, and unnecessarily preempt users’ fair uses of copyrighted content – all while making no appreciable dent in “digital piracy.”

First, DRM helps industry leaders dominate digital media markets and impede innovation. DRM, particularly in online-accessible and downloadable media, has been used to engender both horizontal market share and vertical integration of online media systems. Notably, this has led to firms exerting market dominance in downloadable music and ebooks, for instance. A CMO insisting upon DRM systems is in a position to force those systems' adoption by its licensees, granting an anticompetitive windfall to that systems' vendors as licensees are left to decide between systems for their other collections.

Second, DRM endangers consumers by rendering their computers insecure and violating consumers' reasonable expectations of privacy. In addition to several instances where DRM created serious security vulnerabilities<sup>9</sup>, many other DRM systems function by reporting back to third parties user identities and behavior, undermining the core value of reader privacy.<sup>10</sup>

Third, DRM harms consumers by degrading products and restricting consumers' ability to make otherwise lawful uses of their personal property, upsetting the traditional balance between the interests of copyright owners and the interests of the public. Many implementations not only restrict use of the copyrighted works they are intended to cover—frequently preventing uncontroversial fair uses like quoting, excerpting, and research—but also hamper the normal functioning of their computers and other devices.

Additional problems would arise from the need to determine what constitutes adequate security. In many instances, different parties would define that “security” differently. Users would be far more concerned with the protection of their personal information and reading privacy; licensees would want to limit data breaches of personal information, denial-of-service attacks, abuse of their computing resources for malicious purposes like botnets, and attempts to tamper with the integrity of their collections; CMOs and their members would want to prevent infringement of their works. Many of these separate objectives can be in tension, as noted above; balancing those interests can be a highly technical and context-specific endeavor.

Furthermore, preventing inefficiencies and abuses in the contracting process would require careful assessment of different vendors' systems; failure to properly assess the costs, benefits, and liabilities of each could lead to the creation of a standard that fails to protect the security of users, licensees, or CMOs. Meanwhile, designating characteristics of such systems may result in particular CMOs or licensees being unable to adopt more effective means for protecting security, or preventing licensing projects from moving forward by requiring entities unforeseen by the Office or the CMO to adopt inapposite technical requirements.

---

<sup>9</sup> Deirdre K. Mulligan and Aaron Perzanowski, “The Magnificence of the Disaster: Reconstructing the Sony BMG Rootkit Incident,” *Berkeley Technology Law Journal*, Vol. 22, p. 1157 (2007) (available at <http://www.ischool.berkeley.edu/research/publications/mulligan/2007/the>); Luigi Auriemma and Donato Ferrante, “EA Origin Insecurity (When Local Bugs Go Remote.. Again)” (available at [http://www.revuln.com/files/ReVuln\\_EA\\_Origin\\_Insecurity.pdf](http://www.revuln.com/files/ReVuln_EA_Origin_Insecurity.pdf)).

<sup>10</sup> See e.g., *Adobe Spyware Reveals (Again) the Price of DRM: Your Privacy and Security*, by Corynne McSherry, October 7, 2014. (available at <https://www.eff.org/deeplinks/2014/10/adobe-spyware-reveals-again-price-drm-your-privacy-and-security>).

#### IV. Conclusion

We believe that the mass digitization of expressive works has been, and will be incredibly beneficial to the public. Whether by increasing accessibility for the disabled, or increasing discoverability of long forgotten works, or helping advance machine learning and other technologies, digitization of our culture and knowledge is a commendable goal. Nevertheless, we believe that the Copyright Office's vision of an Extended Collective Licensing system falls well short of that goal. We hope the Office will reconsider pursuing legislation to that effect.

Respectfully submitted,

Sherwin Siy  
Vice President of Legal Affairs  
Public Knowledge  
1818 N St. NW, Ste. 410  
Washington, DC 20036  
(202) 861-0020  
ssiy@publicknowledge.org

Raza Panjwani  
Policy Counsel  
Public Knowledge  
1818 N St. NW, Ste. 410  
Washington, DC 20036  
(202) 861-0020  
raza@publicknowledge.org

Corynne McSherry  
Legal Director  
Electronic Frontier Foundation  
815 Eddy St.  
San Francisco, CA 94109  
(415) 436-9333  
corynne@eff.org