Before the Copyright Office

Library of Congress

In the Matter of	)	
	)	
Orphan Works and Mass Digitization Request for Additional Comments	)	Docket No. 2012-12
	)	

Comments of the Authors Guild, Inc.

Submitted by Jan Constantine, General Counsel

The Authors Guild respectfully submits these comments in response to the U.S. Copyright Office's Notice of Request for Additional Comments that address topics listed in the Office's February 10, 2014, Notice of Inquiry and that respond to any issues raised during the public roundtables held in Washington, D.C., on March 10–11, 2014.

Crafting legislation to address mass digitization and orphan works will be challenging, but it presents an opportunity: the creation of a National Digital Library. Although stakeholders presented a wide range of views in the Copyright Office's recent roundtable discussions, we believe many would unite around the goal of making vast numbers of out-of-print books conveniently available at desktops, educational institutions and libraries throughout the nation, while ensuring that rights holders are compensated for the value their works bring to such a project.

We urge the Copyright Office to propose collective licensing to address the issues of mass digitization and orphan works.

Background

In 2005, the Authors Guild filed a class action lawsuit against Google for infringing the copyrights of U.S. authors, among others, through its mass digitization of millions of in-copyright books from the shelves of cooperating research libraries. After years at the negotiating table, we reached a proposed settlement with Google, which was rejected in March 2011 by then-District Court Judge Denny Chin on the basis that the proposed settlement would have resolved issues best answered by Congress, not through an agreement among private parties.[1]

With the settlement no longer an option, the Guild pressed on with its suit, but last November, Judge Chin found that Google's large-scale copyright infringement was a fair use of copyrighted material. We have appealed this decision.

Concerning a distinct but related issue, a library consortium called the HathiTrust—formed by many of the same libraries that were beneficiaries of Google's Library program—undertook a mass digitization effort. The Authors Guild was forced to sue the HathiTrust in 2011. The *HathiTrust* lawsuit was triggered by an ill-advised effort designed by the University of Michigan—"The Orphan Works Program"—which purported to identify so-called "orphan works" and, after limited notification to these works' "parents" via the HathiTrust website for a mere 90 days, set out to make full-text copies of these books available for display to the University of Michigan community of upwards of 250,000 people.

After we filed suit against the HathiTrust, we quickly found that many of the so-called "orphan" books had authors or authors' estates that were actually quite easy to find through simple online searches. One "orphan" was the child of an emeritus professor at the University of Maryland whose agent had just negotiated an e-book deal for one of his bestselling novels. A French "orphan" author was living at the time in Paris. The estates of others were represented by major literary agencies or were registered with the Authors

[1] *Authors Guild v. Google*, 770 F. Supp. 2d 666, 677 (S.D.N.Y 2011).

Registry, an affiliated entity of the Guild, or the Authors' Licensing and Collecting Society in the U.K. Pulitzer Prize-winning novelist James Gould Cozzens, another "orphan" author, had left his literary estate to Harvard University, according to Copyright Office records. Rights to one "orphan" were held by our very own Authors League Fund, which provides assistance to authors in dire financial need because of health-related issues or other misfortunes. In light of all this, HathiTrust quickly suspended its Orphan Works Program.

The HathiTrust did not then end the program, however. Instead, it promised to start it anew, after it dealt with the program's flaws. It was not until last month at the Roundtables that we learned from the Director of the Orphan Works Program that it had been officially abandoned.

In October 2012 a district court found HathiTrust's copying within the bounds of fair use. This decision, among other legal developments, emboldened a Creative Commons representative to maintain, at the Roundtables, that print-to-digital copying was merely "format-shifting" and therefore permissible under the fair use doctrine.

It is for Congress, not the courts, to balance the interests of stakeholders in shaping the copyright laws. However, recent court decisions—particularly those discussed above—have improperly tipped that balance, jeopardizing the livelihoods of the creators we depend on for the cultural vitality and open exchange of ideas the Founders recognized as the very lifeblood of a democratic society.

Fair use has been discussed by Congress in recent hearings. Accordingly, these Comments will focus on the issue at hand: the threats and opportunities to book authors posed by mass digitization, and possible legislative solutions to maintaining the proper balance between the rights of authors in their books and the public's interest in making use of those books. We believe that the ultimate solution to both mass digitization and orphan works lies in the establishment of a non-compulsory collective licensing system for a limited set of out-of-print book rights.

Congress Has Considered Mass Digitization and Reserved to Rights Holders the Right to Authorize Digital Distribution of Their Works

The decision to make a work first available in digital form—that is, to *authorize* its digital distribution—is increasingly as significant as an author's decision to first publish her work. The author's ability to control what becomes of a digital work after it leaves the digital store shelf, the security of the digital copy is a concern because of the ease, speed, and low cost of copying digital works. Digital copies, after all, are not subject to the practical limitations of reproducing and distributing paper copies. Ebooks require no paper, no ink; no press, no bindery, no inventory, no packing, no shipping, no warehouse (the digital equivalent is the cloud). If ebooks created through a mass digitization projects are stolen, the market consequences would be devastating. The time it takes to move an ebook from point A to point B, no matter the distance, is measured in microseconds, and each copy is indistinguishable from the original, with none of the wear and tear that print books experience over time.

As early as 1963, in the Congressional hearings that eventually led to the passage of the 1976 Copyright Act, the debate focused on a hypothetical that foretold the threats posed by mass digitization today: could some future, technologically-advanced party construe it to be a fair use to optically scan books for computerized uses?[2]

This hypothetical was suggested by the advent and proliferation of high-speed photocopiers in the 1960s. In order to mitigate the threats to rights holders posed by technological advances, it was proposed in the 1963 hearings that "the rights granted under copyright shall include the right to do *or to authorize*"[3] any of the enumerated rights reserved to the rights holder. Indeed, the final bill, which was enacted into law, included language in Section 106 which enumerates the exclusive rights granted to authors: "Subject to sections 107 through 122, the owner of copyright under this title has

[2] 3 OMNIBUS COPYRIGHT REVISION LEGISLATIVE HISTORY 120-27 (George Grossman ed., 2001).
[3] *Id.* at 123 (statement of Herman Finkelstein, General Counsel, ASCAP) (emphasis added).

the exclusive rights to do *and to authorize* any of the [exclusive rights listed in Section 106]."[4]

This legislation, the 1976 Copyright Act, was carefully calibrated to further the purposes of copyright while taking into account the positions of all stakeholders: authors, publishers, academics, librarians, visual artists and photographers, as well as representatives from the recording and motion picture industries, among others. Yet nothing in that act allows the systematic conversion of entire libraries of books into ebooks. On the contrary, the rights to authorize the reproduction, distribution and display of copyrighted works—all of which are infringed by mass digitization—remain with the copyright owner.

Section 108 and Congressional Authority to Alter the Balance of Rights

The exclusive rights Section 106 grants to the copyright owner are subject to certain limitations. The relevant limitations in the mass digitization context are Section 107's codification of the fair use doctrine, and Section 108's exceptions for copying undertaken by libraries and archives.

The various exceptions contained in Section 108 are narrowly tailored to permit particular types of copying for particular purposes. As such, they are incompatible with the bulk, indiscriminate copying of entire library stacks.

Section 108 permits libraries and archives to reproduce and distribute copyright-protected works for two broad purposes, so long as the reproduction and distribution are not used to gain any kind of commercial advantage: to maintain their existing collections (by replacing damaged, lost or stolen books that are no longer commercially available) and to fulfill the requests of readers and researchers. Section 108(c) permits a library to create at most, and under limited circumstances, three copies of a published work. Clearly Section 108 cannot be used to justify the systematic, wholesale copying that is essential

[4] 17 U.S.C. § 106 (2012) (emphasis added).

to mass digitization schemes, particularly where, as in the Google's digitization project, at least five copies of each book are created, there are no guarantees as to the security of the ebooks, the commercial benefit to the infringer is beyond question, and Google provided the participating libraries with infringing ebooks as compensation for access to their collections.

Section 108 also provides that "[n]othing in this section . . . in any way affects the right of fair use as provided by section 107."[5] Section 108 provides highly specific rules governing library copying. Fair use, on the other hand, developed as a common law doctrine permitting the secondary use of suitable portions of copyrighted works for particular purposes—specifically, when that secondary use was in furtherance of the goals of copyright. The doctrine was not developed for and is not suited to deal with the scope and complexity of mass digitization.

Fair use analysis of library activities should not be conducted as though Section 108 did not exist. Congress would not have enacted specific exceptions for libraries and archives in Section 108, with narrowly tailored conditions on any copying, while allowing Section 107 to permit the same activities without any such conditions and limits. To argue that fair use—and not Section 108—governs library copying is to ignore the canonical principle of statutory construction that the specific governs the general.

Nonetheless, two district courts have recently and improperly expanded the doctrine to encompass mass digitization in such a way as to render toothless other provisions in the Copyright Act passed by Congress—especially Section 108. The *Google* and *HathiTrust* courts found that libraries may make digital copies of entire swaths of their collections without the authorization of or compensation to rights holders. In essence, the mass digitization decisions effectively write Section 108 out of the statute, usurping the role of Congress in the process.

[5] 17 U.S.C. § 108(f)(4) (2012).

Given the number and variety of stakeholders affected by mass digitization and orphan works, and the rapid proliferation of the technology that makes these issues so pressing, Congress is the proper branch of government to address these concerns. Indeed, in rejecting our proposed settlement agreement with Google, Judge Chin stated:

> The questions of who should be entrusted with guardianship over orphan books, under what terms, and with what safeguards are matters more appropriately decided by Congress than through an agreement among private, self-interested parties. Indeed, the Supreme Court has held that "it is generally for Congress, not the courts, to decide how best to pursue the Copyright Clause's objectives."[6]

Just two years later, however, during oral arguments in the Google case, Judge Chin questioned the ability of Congress to legislate in the mass digitization arena. When counsel for Authors Guild suggested Congress is the body best suited to alter the balance between rights holders and rights users, Judge Chin replied: "Is anything done in Congress these days? How long would it take reasonably for this to be resolved in Congress? Even the issue of orphan books has been percolating in Congress for years and years."[7]

We believe that Congress is not only the proper body to deal with these issues, it is also the most capable.

Section 108 Does Not Allow for "Preemptive" Preservation

Proponents of mass digitization argue that digitization is a panacea to the preservation problems of libraries and archives. Libraries rightfully worry about the dangers posed to their collections by time and physical decay. This is a valid concern; it is in the interest of all stakeholders, especially authors, to preserve our cultural heritage. Libraries have suggested, however, that the Section 108 exceptions do not reach far enough, because their collections and archives are still vulnerable. Sections 108(b) and (c) are narrowly

[6] *Authors Guild v. Google, Inc.* 770 F. Supp. 2d 666, 677 (S.D.N.Y. 2011), quoting *Eldred v. Ashcroft*, 537 U.S. 186, 212 (2003).

[7] Transcript, Oral Argument, *Authors Guild v. Google*, Sept. 23, 2013 (S.D.N.Y.) (on file with the Authors Guild).

tailored to allow libraries to maintain their collections in a manner that also respects the literary marketplace. However, to extend Section 108—or the fair use doctrine, for that matter—to permit "preemptive preservation" would be incredibly damaging to the market conditions necessary for the production of new works.

Section 108(b) controls library copying for preservation purposes. It permits a library to reproduce and distribute three copies of a work, so long as that work is both unpublished and currently in its own collection. Section 108(c) allows a library to reproduce (but not distribute) up to three copies of a published work in its collection that is "damaged, deteriorating, lost, or stolen" or if the format of the work has become obsolete, so long that the library can't find a replacement copy at a fair price, and that the three copies it does make are not made available to the public outside the library. Taken together, these provisions are tailored to allow libraries to maintain the collections they have already established. However, they also ensure that libraries will continue buying books by providing that (1) preservation of *published* works (that is, works more likely to be available in the marketplace) cannot be carried out by copying, and (2) damaged, deteriorating, lost or stolen works cannot be replaced by copying if a replacement is available at a fair price.

Libraries are a vital part of the literary ecosystem. Not only are they charged with preserving, cataloguing, and making accessible our literary heritage, but the books they purchase constitute a significant percentage of total book sales. Fewer books bought translates—quite directly, as we have seen in recent years—into shrinking incentives for new writers and new works.

Moreover, there is no guarantee that digitization is any more secure or permanent a method of preservation than actual physical preservation, traditionally accomplished by housing multiple physical copies in multiple locations. It would be naïve at this point to think that digital databases are immune from security breaches. Nor can we be sure of the permanent readability of digitized books. The pace of technological change certainly suggests that digital formats will go obsolete more quickly than printing will.

The Google Books Settlement as a Model for a Collective Licensing Legislation

In 2009, the Authors Guild and Google proposed a settlement agreement to our class action lawsuit. If nothing else, this shows that rights holders and rights users are capable of coming to the table and arrive at a solution which serves the interests of all stakeholders and also promotes the goals of copyright law.

A proper solution would ensure that rights holders are compensated for the value their works bring to such projects, and that the proper security measures are in place to protect those works, while at the same time allowing the public to benefit from mass digitization projects. The solution is collective licensing of out of print books, following the highly successful models in use overseas, which provided the model for our proposed Google Books Settlement. This paves the way for a *real* National Digital Library, not the mere snippets offered by Google Books.

Negotiating individual licenses render is impracticable, considering that millions of copyright-protected out of print books are to be digitized. The transaction costs are far too high. This situation is ripe for a collective licensing solution.

Congress has already enabled collective licensing in the copyright context. Performing rights organizations are identified in Section 101 of the Copyright Act. Section 115 provides for a compulsory license for the making and distribution of phonorecords. And the Copyright Clearance Center—a not-for-profit organization—is a long-established American example of a successful licensing service.

Collective licensing has met with great success in foreign nations. Nordic countries, for example, have been using collective licensing for the better part of 50 years, with near-universal approval. And in 2011, the European Union issued a memorandum to Member States urging them to solve the problem of "orphan works" in the mass digitization context by establishing collective licensing societies. France passed collective licensing legislation in 2012, and Germany followed a year later. The United Kingdom has

recently announced a commitment to introduce a collective licensing scheme to license orphan works. In short, collective licensing can enable uses that are unauthorized but beneficial by making these uses subject to a payment to the rights holder.

The ingredients of a successful collective licensing system are straightforward. First, Congress would enable the creation of a collective management organization (or collective management organizations: visual artists and others may want their own). Second, Congress would set limits on the types of licensing the collective licensing organization could manage. For example, it would be limited to out-of-print books only, and downloading or printing of the ebooks would be limited to excerpts of a certain length, to help assure that unauthorized ebooks are not available to the public, usurping the author's right to choose when to publish an out of print book. Third, Congress would create a third-party regulatory body to ensure that the collective management organization does not abuse its monopoly power in negotiations with universities, colleges, libraries and others. Fourth, licensing would not be compulsory. Rather, authors, publishers and other rights holders would be empowered to remove all their works from the database, or to exclude works from any and all uses. Fifth, the collective licensing organization should not be able to collect its administrative fee for a work until it locates and pays the rights holder. This assures diligent, ongoing searches for rights holders.

Collective licensing is in the interest of the Authors Guild membership and authors in general. For over ten years the Authors Registry, which the Authors Guild helped found and supports financially, has acted as a payment agent for foreign collecting societies who send revenues from secondary uses (such as photocopying and library lending) of books to be paid out to U.S. authors. To date, the Authors Registry has paid out more than $22 million to more than 10,000 authors. Last year alone it distributed over $2.8 million.

The Registry has been quite successful at finding rights holders to out-of-print books. A few years ago, he Registry conducted a sampling of out-of-print books for which the

Registry had received royalties. With one full time and one part time employee, the Registry had found more than 80% of rights holders of out-of-print books. Longer-established collecting societies—such as the ALCS in Britain—claim success rates of 90%. (Incidentally, these success rates demonstrate that the "orphan works" problem, at least for published books, is vastly overblown.) Such efforts shrink the orphan works problem for books, and locating rights holders should be a primary goal of any legislative solution.

Collective licensing is also in the interests of libraries, universities, colleges, and other institutions that would establish mass digitization projects. In the proposed Google Books Settlement, collective licensing presented us with a way to resolve the Google dispute which the parties to the settlement (including university libraries as third-party beneficiaries) all believed was a win-win-win. Authors, publishers and rights holders gained control over and compensation for the uses of their works, digitization of their works by Google (for those who wanted it), and out-of-print works would be able to find new markets. Libraries benefitted by the digitization of their collections, allowing for the preservation of copies. Google benefitted by the increased revenue enabled by the traffic the project would bring to its search engine, the releases it received for past infringements, and the lawful future uses it could make of its digitized copies.

A National Digital Library that solves the mass digitization and orphan works problems for books is within reach. We urge the Copyright Office to do what it can to make it real.