

U.S. House of Representatives
Committee on the Judiciary
Subcommittee on Courts, Intellectual Property, and the Internet

Preservation and Reuse of Copyrighted Works
April 2, 2014, 2:00 p.m.

Statement of

Jan Constantine
General Counsel
The Authors Guild, Inc.
jconstantine@authorsguild.org

COMMITTEE ON THE JUDICIARY
SUBCOMMITTEE ON COURTS, INTELLECTUAL PROPERTY AND THE
INTERNET

Preservation and Reuse of Copyrighted Works
Testimony of Jan Constantine on Behalf of the Authors Guild

April 2, 2014

Mr. Chairman, Ranking Member and all other members of the Committee: My name is Jan Constantine, and I represent the Authors Guild, the largest society of published authors in the country. The Guild and its predecessor organization, the Authors League of America, have been leading advocates for authors' copyright and contractual interests since the League's founding in 1912.

Among our more than 8,500 members are historians, biographers, poets, novelists and freelance journalists of every political persuasion. Authors Guild members create the works that fill our bookstores and libraries: literary landmarks, bestsellers and countless valuable and culturally significant works with more modest sales records. We have counted among our ranks winners of every major literary award, including every U.S. Nobel Prize for Literature¹ honoree, dozens of Pulitzer Prize and National Book Award winners,² as well as U. S. Presidents³ and distinguished members of both Houses of Congress.

We have a 100-year history of contributing to debates before Congress on the proper scope and function of copyright law. It's an honor and privilege to be here today, for the Authors Guild to continue to serve that role before this committee.

¹ Sinclair Lewis (1930); Eugene O'Neill (1936); Pearl S. Buck (1938); William Faulkner (1949); Ernest Hemingway (1954); John Steinbeck (1962); Toni Morrison (1993).

² Among them, John Ashbery, Donald Barthelme, Ray Bradbury, Pearl S. Buck, Robert A. Caro, Rachel Carson, John Cheever, Michael Cunningham, Joan Didion, Annie Dillard, E.L. Doctorow, Jennifer Egan, Ralph Ellison, William Faulkner, Paula Fox, Annette Gordon-Reed, Ernest Hemingway, John Hersey, Oscar Hijuelos, Pauline Kael, Madeleine L'Engle, J. Anthony Lukas, Robert Massie, John McPhee, James A. Michener, Toni Morrison, Joyce Carol Oates, Sharon Olds, Katherine Paterson, Annie Proulx, Philip Roth, Arthur M. Schlesinger, Jr., Upton Sinclair, Isaac Bashevis Singer, Jane Smiley, James B. Stewart, T.J. Stiles, Barbara Tuchman, John Updike, Gore Vidal, Robert Penn Warren, and William Carlos Williams.

³ Theodore Roosevelt, Harry S. Truman, John F. Kennedy.

Mass digitization and orphan works are two issues that I have grappled with for eight years as General Counsel of the Authors Guild. We have been immersed in mass digitization and orphan works issues in *Authors Guild v. Google* and *Authors Guild v. HathiTrust*, in testifying before Congress, in numerous public discussions and presentations, including, just last month, participating in the Copyright Office's roundtable discussions on these very issues.

Today I will focus on the threats and opportunities to our literary culture and the readers, authors, schools, and libraries that depend on it, posed by mass digitization and orphan works, and propose a legislative solution to maintaining the delicate balance between the rights of authors and other creators and the public interest served by providing access to their literary works.

Ad hoc approaches to mass digitization of books and so-called orphan works are rife with problems and seriously endanger our literary culture. There can be no clearer demonstration of the need for Congressional action than the bold and, in our view, blatantly unlawful uses of authors' literary property rights that prompted the Authors Guild to twice file lawsuits against much larger and more powerful adversaries in the past nine years.

I will discuss these "copyshocks," and then look back 50 years to show that there are far better approaches to copyright that we seem to have forgotten. The drafters of the 1976 Copyright Act, it turns out, saw Google coming, anticipating the very fair use problem such a corporation might pose. The drafters clearly thought they had addressed the issue, placing authors firmly in control of their exclusive right to authorize the digitization of their literary works.

Looking back 50 years is particularly instructive as we consider the challenges of clearing rights to so-called orphan works. When private parties are incentivized to find rights holders, rather than rewarded for failing to do so, they can be remarkably resourceful and successful.

Finally, I will propose a legislative solution, based on well-functioning licensing systems that have been operating around the world for decades. Among other features, these licensing systems have cleared rights to orphan works for years.

Copys shock One: Google & The Universal Digital Library

On December 14, 2004, Google stunned the library, literary, technology, and copyright worlds, announcing it would start scanning millions of copyright-protected books in partnership with the University of Michigan Library. Copyright lawyers re-read Section 108, trying in vain to find anything that might allow Google to engage in so clear a massive copyright infringement. But Google wasn't looking to Section 108. It was gambling that by displaying "snippets" of protected works to its users, all of its copying of entire texts would be deemed a fair use.

Suddenly, the universal digital library, at least the beginnings of one, was rushing toward us.

As details emerged, we were further stunned: Google was swapping authors' ebooks for access. Its agreement with the University of Michigan provided no security guarantees for those ebooks, and Google would be monetizing authors' literary property by running ads alongside its "snippet" displays. Google offered no revenue sharing from those ads. Google had taken it upon itself to put authors at digital risk, with no prospect of digital rewards.

Google's litigation risks were huge; the statutory damages daunting even for a corporation of Google's size. Why roll the dice?

It turns out that Google was suddenly facing a formidable competitor.

Amazon and the A9 Search Engine

On October 23, 2003, Amazon.com launched "Search Inside the Book," which allowed users of Amazon's website to search more than 120,000 books, or 33 million pages, at the time it launched, according to Amazon's press release.

With "Search Inside the Book," not only could Amazon's customers search these 120,000 books, they could preview entire book pages containing the search terms. Customers were required to log in with their Amazon user names and passwords to preview entire book pages, so they could instantly purchase the books they were browsing.

Once customers had clicked on the link to a specific page and signed in with their Amazon.com user name and password, they could preview relevant pages, including the page they selected, and search for other terms of interest within the book.

The announcement was a momentous one in the book world. It was clear to most industry observers that Amazon's ambitions for the program were enormous. Indeed, Amazon had agreements with more than 190 publishers to digitize books for Search Inside the Book, according to its press release, including many of the largest U.S. publishers, such as John Wiley & Sons, Random House, Simon & Schuster, Time Warner, HarperCollins, McGraw-Hill, and Holtzbrinck. The number of participating publishers would rapidly grow to include nearly every book publisher in the U.S.

A critical component of "Search Inside the Book" was that users had to log in with their Amazon user names and passwords before they could view full pages of the books. In that way, those virtually browsing the books would have Amazon's famous one-click purchase button at the ready. By all accounts, Search Inside the Book boosted sales for most of the books in the program, and continues to do so today.

The implications for Search Inside the Book went far beyond the book industry, however. Amazon was offering searchable content that no other entity, not even Google, had offered, and that content was based on contracts negotiated with hundreds of publishers. A November 12, 2003, article in *Wired* magazine made this abundantly clear:

[T]he contents of books may be the only publicly accessible data set with the potential to match Google's Web index both for size and utility. Search Inside the Book makes Amazon the sole guide to tens and ultimately hundreds of millions of pages of information. And while Google's business is vulnerable to any competitor that builds a better search engine, Amazon's book archive is the product of negotiated contracts with hundreds of publishers. Amazon has cornered the market on information that was once hidden away in books. The burden of the physical -

the fact that the database Amazon uses is linked into a complex system involving real things - gives it a stunning, if perhaps temporary, advantage.⁴

Gary Wolf, author of the *Wired* article, wrote about the critical limitations placed on Search Inside the Book:

If you want to read an extensive excerpt, you must turn to the physical volume -- which, of course, you can conveniently purchase from Amazon. Users will be asked to give their credit card number before looking at pages in the archive, and they won't be able to view more than a few thousand pages per month, or more than 20 percent of any single book.⁵

Wolf reported that he spoke to Amazon CEO Jeff Bezos about the project, who was emphatic about his purposes:

Bezos is vehement on this point. He has sold publishers on the idea that digitizing hundreds of thousands of copyright books won't undermine the conventional bookselling business. *"It is critical that this be understood as a way to get publishers and authors in contact with customers," he says in an interview at Amazon's Seattle headquarters. "We're perfectly aligned with these folks. Our goal is to sell more books!"*

Bezos has some good evidence to back up his argument. Amazon has consistently added features that have proven to increase book sales. Through its customer reviews, used-book business, and personalized recommendations, the company constantly puts its customers in contact with new titles. *Amazon is a machine that stimulates the acquisitive urge of readers. It appeals to their specialized interests.*

*It makes people buy books.*⁶

⁴ "The Great Library of Amazonia," *Wired Magazine*, November 12, 2003, available at http://www.wired.com/wired/archive/11.12/amazon_pr.html.

⁵ *Id.*

⁶ *Id.* (Emphasis added.)

While the Authors Guild has had its disagreements with Amazon over the years, many of them quite sharp, our concerns are focused on Amazon's behavior, which we view as frequently monopolistic and unfairly undermining the livelihoods of brick-and-mortar bookstores, which many authors depend on for marketing their books. We have never doubted, however, Amazon's mastery of the online bookselling environment. Our conversations with publishers at the time and since then have confirmed that they agreed to participate in Search Inside the Book because they believed it would sell more books, and the evidence Amazon provided them supported that belief. Amazon was unmistakably on its way to making nearly all in-print books available for searching and previewing at its online bookstore.

Google had conquered online content, but Amazon, by negotiating agreements with hundreds of publishers, had gone where Google could not, making tens of millions of pages of carefully written, edited, and published works, all previously off line, available to its shoppers. Amazon, by cornering the market on book content, was suddenly a competitor unlike any Google had ever faced.

The threat to Google's search dominance became even more apparent just five months later, when Amazon unwrapped its beta version of A9, Amazon's search engine competitor to Google.

The A9 search engine offered user content Google could not match: the millions of book pages available (to those who logged in with their user names and passwords) through Search Inside the Book.

A9's launch came at a critical time in the corporate history of Google, six years after the company was incorporated, and, as the *Times* noted,⁷ just one month after Google made its initial public offering of stock. Google had a well-financed, innovative, disruptive competitor in its own back yard.

⁷ "Amazon to Take Searches on Web to a New Depth," New York Times, September 15, 2004, *available at* http://www.nytimes.com/2004/09/15/technology/15search.html?_r=0.

Google Trumps Amazon, by Avoiding Rights Holders and Their Interests

Google didn't stand still. Three months after the formal launch of Amazon's A9 search engine, with its unique access to an expanding list of in-print, copyright-protected books Amazon scanned with the permission of the books' publishers, Google struck back. It announced that it would begin scanning copyright protected books for its search engine.⁸

Google had come up with a way to neatly leapfrog Amazon's scanning efforts: it would scan copyright-protected books without permission, claiming that its scanning and snippet display were covered by fair use (even when it was scanning and displaying in-print books that had been scanned, or were eligible for scanning, by Amazon).

Rather than painstakingly seeking permission from rights holders, who might require a share in advertising revenue, or who might require that users be logged in to one-click purchasing accounts before viewing portions of their books, or who might demand control over the display, storage and security of their books, Google chose instead to negotiate with libraries, swapping authors' and publishers' ebooks in compensation for access to the libraries' repositories.

Google's library project would soon begin displaying snippets of authors' and publishers' in-print books in competition with the authorized displays available at Amazon and BarnesandNoble.com, and other online retailers. To the extent Google succeeds in luring readers to its search engine and away from online bookstores, Google surely harms book sales.

It isn't just in-print books that were affected. By December 2004, when Google announced its massive scanning-and-display project, thousands of formerly out-of-print books were being made available by print-on-demand technology, including more than a thousand works by Authors Guild members we made available through our Backinprint.com program. Since then, surely tens of thousands of more out-of-print books are commercially available again,

⁸ "All Booked Up," Google's Official Blog, December 14, 2004, *available at* <http://googleblog.blogspot.com/2004/12/all-booked-up.html>.

brought into print by countless publishers such as university presses, newcomers such as Open Road Media, and publishing programs at Amazon and Barnes & Noble.

Book Excerpts Do Sell Books, at Online Bookstores

The Authors Guild believes that allowing readers to search and view selections of books online can promote the sale of books. But the context for those book displays is crucial. If the book excerpts are displayed in a bookselling environment, then book sales are generally promoted. If book excerpts are displayed in a search engine's advertising-driven environment, then ad sales are generally promoted.

To the extent Google's unauthorized displays of books encourages readers to search at its ad-supported search engine, rather than logging in to Amazon's retail environment, Google is hurting the sales of authors' books. For this reason, and many, many others, authors and other rights holders should have control of when their books are copied in their entirety, and where their books are displayed.

Google, in other words, disrupted the commercial, permission-driven development of book-search-and-display at online bookstores in order to gain a competitive advantage over other search engines. In the process, it distributed millions of ebooks to universities, placing those books beyond the control of authors and publishers, while putting them at plain risk of widespread infringement.

50 Years Ago: The Drafters of the 1976 Copyright Act Saw It Coming

Here's something that the drafters of our 1976 Copyright Act couldn't have imagined – a digital library of all books, accessible to countless users throughout the country, in some manner. Work on that legislation had begun in the 1950s, after all, many of its most important sections complete by 1965, including Section 106, the bundle of authors' exclusive rights, Section 107 fair use, and much of Section 108, governing library copying for preservation, replacement and other limited purposes.

Remarkably, however, much of Google's bold venture had been anticipated, with stunning prescience, and the notion of a national digital library was being bandied about in books, journal articles, and the popular press. At the Seattle World's Fair in 1962, for example, the American Library Association presented its Library of the 21st Century, featuring computer terminals displaying digital books.

Visions of the universal digital library weren't confined to World's Fair exhibits. Books and essays were written about it, including *Libraries of the Future*, by J.C.R. Licklider. Licklider's book is based on a study conducted at the behest of the Council of Library Resources. The book predicts much of what's to come: a digital library accessible online, displayed on high-quality computer screens. The digital library would offer highly refined search techniques, and pages could be printed on high-speed devices.

A debate with Licklider spilled into the pages of the Authors Guild's *Bulletin*: how, exactly, were authors to be paid for their work in this digital library of everything? Mr. Licklider's response: flat payment or by use. He thought flat payment made the most sense.

In a 1964 article republished in the pages of the trade publication *Special Libraries*, artificial intelligence expert Arthur Samuel foresaw that within 20 years from then "one will be able to browse through the fiction section of the central library . . . via one's remote terminal."⁹

In 1965, Curtis G. Benjamin, Chairman of the Board of the McGraw-Hill Book Company, wrote of the "inevitable advent of the automated library system in which documents (including book pages) are exchanged and displayed by photocopy, by microimages, or by more sophisticated electronic-optical devices . . . the library system in which one copy of a printed reference book will serve the present uses of ten or even 20 or more copies."¹⁰ He was "convinced" such a system would become "generally operative in the United States in the foreseeable future."¹¹

Discussions of the digital library also reached Copyright Office and legislative hearings on the new copyright act.

⁹ *Special Libraries*, January 1965.

¹⁰ *Special Libraries*, November 1965.

¹¹ *Id.*

The most remarkable and eerily prescient part of the legislative history, however, was surely a Copyright Office hearing more than 50 years ago, on February 20, 1963. The debate addressed whether it ought to be fair use to optically scan a book for computerized uses. Abe Kaminstein, Copyright Register, dramatically announced that he'd received a telegram from Reed Lawlor, a patent attorney in Los Angeles, with a provocative question for those considering the first draft of the new copyright law.

KAMINSTEIN: I was going to hold this for later on, but I have a telegram from Reed Lawlor, who says, "*I suggest you consider adding the following section 6: 'In any event reproduction of a copyrighted work in machine readable form for use in the analysis, citation and reasonable quotation of the work by means of an information storage and retrieval system shall be considered a fair use.'*". We were going to hold this for the discussion of fair use, but I certainly have no objection to opening up the subject here. Did you want to comment on it?

BROWN: If, as Mr. Lawlor suggested, you have a machine which simply absorbs information for the sake of giving you citations later, and which does not have the capacity to print it out again or to make copies, then it seems to me that that machine might be considered as simply an adjunct to note taking-in the sense that one can absorb a copyrighted work to make proper use of it, which may or may not be considered "fair use."

SCHIFFER: *With the way these things seem to be going, there's a good possibility that, within the lifetime of this statute, they're going to eliminate printed books for most purposes or for many purposes. In other words, if we take Professor Brown's view that you can put material into these machines as a matter of note taking, you may find that, for practical purposes you have eliminated the market for the book entirely.*

I think that the only way to handle these things is to make machine uses, in all forms, subject to exclusive control of the author, except to the extent that the use actually made by the machine is not a use within the concept of substantial taking in the ordinary sense. But the idea that you can feed a book into the machine in its entirety and then make it available to the world at large (as will undoubtedly happen; there are many library-types of computers under study now) will inevitably hurt the copyright proprietor to an extent which cannot be intended here.

FINKELSTEIN: Well, may I then concentrate on one implication of what Fulton says, because his bringing that up brings this to mind. A previous bill - the Vestal Bill, as I recall it - would have given the right "To make any form of record in which the thought of an author may be recorded and from which it may be read, reproduced, performed, exhibited, represented,

delivered, transmitted or communicated" - language along those lines. Is there any reason for not having the grant to the copyright owner that broad? And might not a grant that broad include these information retrieval systems?

And, when we talk about information retrieval systems at another point, why do we say "information retrieval?" Suppose it merely retrieves entertainment? If the right that you are giving for this kind of a system relates only to information, you may be excluding the kind of thing that comes out of one of these big tape machines, something we don't think of now: a complete form of entertainment.

May I just add one thing while I have the floor? *I wonder if at the beginning, right in the introductory sentence-this is a matter of drafting-we couldn't say "... the rights granted under copyright shall include the right to do or authorize any of the following with respect to the copyrighted work."* The reason for suggesting "authorize" here is this. Suppose ASCAP, or BMI, or any of the other licensing organizations authorizes its licensees to perform a certain work. I doubt whether that would be an act of contributory infringement, but I think there should be liability there. *It would seem to me that the mere authorization to make the use of the copyrighted work, that particular work, ought to subject the person making the authorization to liability even though he may not be a contributory infringer.*

SKIPPER: I am James Skipper, representing the Association of Research Libraries, and I'd like to speak for just a moment on this problem of information storage and retrieval. The point was made that, if a text goes into a computer and the entire text is printed out, this is a violation of copyright, I would be inclined to agree with this principle. However, I would hate to see anything written into the law that would inadvertently inhibit research.

Let me cite an example. The English Department at Cornell University now has a computer concordance program going. They are using the computer to analyze the works of literary authors. And to do this they have to feed the whole text of what the author wrote into the computer. The printout is an analysis; it is not the complete text, but the text has to be in the machine before the analysis or concordance can be obtained.

With the potential of optical scanners, with the potential of indexing in depth for information retrieval, it is becoming increasingly necessary to feed the whole text in to the computer apparatus. But what you're getting out is an analysis. You're indexing literature; you're not printing out the whole text.

I like very much the suggestion, contained in the telegram read by the Chairman, that some consideration of fair use be given, especially with respect to the printout phase of this information retrieval problem.

ROTHENBERG: The information storage systems being defined now will not be confined to use in a library for literary analysis. For example, law offices will have sending and receiving sets to obtain information from a storage system at the nearest large law library. This may reduce substantially the need by law offices for many textbooks. Yet the information obtained from the machine, at any one time, might constitute fair use within the traditional sense.

Accordingly, the copyright owner must control the right at the very outset when his book is being placed into the machine, because it is the cumulative effect of the multiple fair uses which will effectively destroy the value of his copyright.

GOLDMAN: *Are you suggesting that the need is to control putting the work into the machine?*

ROTHENBERG: Yes.

GOLDMAN: *And then you don't have to worry about the use by taking it out of the machine?*

ROTHENBERG: Then it will be merely by contract; whatever arrangement the copyright owner wishes to make with Remington Rand or whoever the company is, or the program.

IRWIN KARP (Authors League/Authors Guild): As Mr. Rothenberg points out, you've got another separate problem when it comes to using the machine in lieu of the book to begin with in the type of operation where you are actually wiping out the markets for multiple copies. *There I don't think you can solve it in any other way than by controlling the right to put it in. I think you have to control both the right to put it in and the right to take it out.*¹²

And so it came to pass. The next version of the copyright bill, the 1964 draft, provided: “the owner of the copyright under this title *has the exclusive rights to do or to authorize* and of the following” (emphasis added). This language filled a frightening loss-of-control void in authors’ rights.

¹² 3 OMNIBUS COPYRIGHT REVISION LEGISLATIVE HISTORY at 120-27 (George Grossman ed., 2001) (remarks of Abraham Kaminstein, Register of Copyrights; Reed Lawlor; Harry Rosenfeld, Ad Hoc Committee of Educational Organizations and Institutions; George Schiffer, Schiffer & Cohen; James Skipper, Association of Research Libraries; Abe Goldman; Stanley Rothenberg; Irwin Karp, Authors League of America) (emphasis added).

Indeed, the next draft of the copyright bill included language very close to this, which now reads 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 the exclusive rights to do *and to authorize* any of the [exclusive rights listed in Section 106].”¹³

In a point that is essential to the issue at hand, underscoring the fact that mass copying was an explicitly consideration in balancing the rights of creators and users under the 1976 act, a representative of the Association of Research Libraries, no less, conceded that “if a text goes into a computer and the entire text is printed out, this is a violation of copyright.”¹⁴

This legislation that was passed in 1976 was carefully calibrated to further the purposes of copyright while taking into account the positions of the various stakeholders. Indeed, in the years leading up to the passage of the 1976 act, many stakeholders had seats at the table, including authors, publishers, academics, librarians, visual artists and photographers, as well as representatives from the recording and motion picture industries. But nothing in the 1976 act allows the systematic digitization of entire libraries of books. On the contrary, it is clear that 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.

Problem solved, we must have thought.

Copys shock Two: Hathitrust & “Orphan Works”

In 2011, a library consortium—known as HathiTrust—formed by many of the same libraries who were beneficiaries of Google’s Library program by receiving Google’s ebook versions of their physical library books from Google in exchange for providing access to their vast collections of library books, announced it was embarking on its own “Orphan Works Program.” The program, designed by the University of Michigan, purported to identify so-called “orphan works” and, after limited notification to these works’ “parents” through the HathiTrust website for a mere 90 days, set out to make full-text ebooks available for download and display to the University of Michigan community of upwards of 250,000

¹³ 17 U.S.C. § 106 (2012) (emphasis added).

¹⁴ 3 OMNIBUS COPYRIGHT REVISION LEGISLATIVE HISTORY 123 (remarks of James Skipper, Association of Research Libraries).

people. Once again, the Authors Guild was compelled to bring litigation in order to protect the authors' rights.

After we filed suit, 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 Registry 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. It 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 Copyright Office's roundtable discussions did we learn from the Director of the Orphan Works Program that it was officially abandoned.)

50 Years Ago: We Knew How to Find Rights Holders. "Going to Timbuktu"

There had even been a forerunner to Amazon's race with Google to obtain permission to vast databases of books, as print-on-demand technology went mainstream in the early 1960s, a few years after Xerox acquired the Copyflo, a machine capable of printing images from microfilm onto paper. Even as stakeholders gathered in Washington to debate the balance of rights to be struck in the next Copyright Act, a Xerox subsidiary called UMI (University Microfilms, Inc.) of Ann Arbor, Michigan, was playing the role of Amazon, stealing a march on its competitors Bell & Howell and 3M.

A rights race was developing: a race to *clear* rights, that is. Bell & Howell and 3M—both technology juggernauts in their time—were competing with UMI to find the most rights holders and corner the market in print-on-demand books.

In ads in the *Special Libraries Journal*, UMI bragged about its prowess at finding authors: “If we don’t have the book you’re looking for, we’ll find . . . it, clear copyright, pay royalties and send it to you. Whether we find the book in Timbuktu or in our collection of 50,000 old and new titles; whether the original is \$10 or \$10,000, the cost is the same.”¹⁵

These companies’ stance to rights was in sharp contrast to what we have come to expect from the practices of a company like Google today: they were competing against each other publicly to see *who could locate the greatest number of rights holders*.

In 1960 UMI had announced its “O-P Books” Program. After obtaining the rights for an out-of-print book, technicians would microfilm the material and then print it on book-quality paper. UMI was also able to offer specially-designed collections of works whose rights they had obtained to university libraries—as either microfilm collections or as individual, bound reprints of the original books. For example, UMI was able to offer at a reasonable price a collection called “Russian Language Works,” which contained more works than the Library of Congress’s Russian collection. “Books of this sort are virtually unobtainable on the open market,” UMI explained, “and when they can be obtained it is often at an unreasonably high price.”¹⁶

By March 1965, the UMI Russian library consisted of 2,000 out-of-print books. The total number of titles it offered was over 15,000. “In fact,” boasted UMI, “we can supply almost any book that has ever been printed in any language. For as little as 4 cents a page.”¹⁷

Of course, UMI’s success only motivated its competitors to clear more rights themselves. A Bell & Howell advertisement in the April 1965 issue of *Special Libraries* magazine lauds a

¹⁵ *Special Libraries*, April 1968.

¹⁶ University Microfilms, Inc.: Ann Arbor, Michigan. Ann Arbor, Mich.: University Microfilms, 1961.

¹⁷ *Special Libraries*, March 1965.

5,000-title-strong catalog of out-of-print titles available.¹⁸ But in the end, UMI won the rights race. By August 1965 UMI offered over 57,000 out-of-print titles.¹⁹

This race to clear the rights to out-of-print books shows the market functioning as it should—companies vigorously competing against each other, driving down prices, and doing so in the proper, responsible way. Rights holders were compensated. Access to books was increased. And in doing so, the benefit to the public is clear: increased access to our literary and cultural heritage at prices libraries and even individuals could afford.

After HathiTrust released the list of candidates for its Orphan Works Program, we had a look at the 1977 UMI catalogue. The very first book listed on the HathiTrust list of “orphans”—*Preachers Present Arms*, by Ray Hamilton Abrams—was listed in that catalog, along with seven other “orphans”: *Between Two Wars* by the Princeton psychologist James Mark Baldwin; Richard Allen Foster’s critical study *The School in American Literature*; a University of Michigan report, *Group Influence in Marketing and Public Relations*; a 1953 biography of the composer Stephen Foster by the music historian John Tasker Howard; Claude Searcy McIver’s literary study of the novels of W.S. Maugham; *Robert Gould: Seventeenth Century Satirist* by Eugene Hulse Sloane; and Henry Justin Smith’s *It’s the Way It’s Written*.

If the rights to so many works could be cleared using mailmen and pre-Google research tools, what excuse remains for the tech giants of today—with greater resources and less standing in their way—to build their empires on the backs of uncompensated creators?

Solution: Non-Compulsory Collective Licensing of a Limited Set Of Rights

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 answer is to allow for non-compulsory, collective licensing system of a limited set of out-of-print book rights, which would the way for a *real* digital library, not the mere

¹⁸ *Special Libraries*, April 1965.

¹⁹ *Special Libraries*, July-August 1965.

excerpts and snippets currently offered by Google Books. And, critically, the books subject to the license would be out of print, to avoid disrupting commercial markets.

This limited set of rights would include display rights, so that colleges, universities, school libraries, public libraries and other institutions would have ready access to millions of copyright-protected works. Print and ebook rights would not be part of the package—only the author or other rights holder could authorize such uses.

Congress has already acted to enable 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.

Moreover, collective licensing solutions have met with great success around the world. Nordic countries, for example, have been using collective licensing for the better part of 50 years, with near-universal approval.²⁰ In short, collective licensing can enable uses that are unauthorized but also beneficial by making these uses subject to a payment to the rights holder.

First, a collective management organization would have to be established by Congress: a third-party regulatory body with proper oversight to ensure it does not abuse its monopoly power, and empowered by Congress to negotiate with all stakeholders—libraries, authors, publishers, end-users—to determine the scope of the rights granted and the appropriate licensing fees. This collective would be authorized to license a limited set of rights in certain types of works for the negotiated fee—unless, that is, the rights holder opts out of the license. Licensing would not be compulsory. Rather, authors, publishers and other rights holders would be empowered to remove all their works from the database, or exclude works from any or all uses. In order not to disturb existing commercial markets, in-print works should not be displayed without an opt-in by the rights holder.

²⁰ 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.

Collective licensing is also in the interests of libraries and the companies that seek to establish mass digitization projects. Authors, publishers and rights holders gain proper control over and compensation for the uses of their works. And best of all, colleges, universities, schools and public libraries, and all of the communities they serve, benefit by gaining access to vast collections of books. Our research libraries would become available to all, a great level-the-playing-field leap in access to our literary, scientific, and cultural heritage.

The “Orphan Works” Problem: A Way Forward

For more than ten years the Authors Registry, an affiliate of the Authors Guild, has acted as a payment agent for foreign collecting societies who send revenues from secondary uses (such as photocopying) of books to be paid to U.S. authors. To date, the Authors Registry has paid more than \$20 million to U.S. authors, distributing \$2.8 million last year. These payments are for both in-print and out-of-print books. With one-and-a-half employees on the project we have been able to find a substantial percentage of recipients. The Authors Registry sampled its success at finding rights holders of out-of-print works: it found more than 80% of such rights holders. Longer-established collecting societies—such as the ALCS in Britain—claim success rates of 90%. The “orphan works” problem, at least for books, is vastly overblown.

Some basic principles have emerged regarding so-called “orphan works:”

- 1. Diligent searches are not the answer to the orphan works problem.* This approach appeared promising to many, including us, but it simply doesn’t work in practice. The incentives are all wrong, rewarding failed searches with uncompensated use of copyright-protected materials. Diligent searches may prove to be part of the solution, but such searches must be coupled with the payment of a reasonable license fee for the proposed use.

- 2. The orphan works problem, at least for rightsholders in books, appears to be vastly overstated.* Those holding rights to in-print books can readily be found, of course. The difficulties in finding authors of out-of-print works are not nearly as daunting as some have suggested. In 2009, for example, out of a sample of 1,000 authors of out-of-print books for which the Authors Registry had collected overseas photocopy

royalties, the Registry located and paid more than 87%. Similarly, the Guild's 2005 survey of members found that 85% had "never" or "rarely" been unable to reach a copyright owner to request permission to use a copyrighted work.

3. Foreign licensing and collecting organizations have been efficiently licensing orphan works for decades. We should learn from their examples. Collective licensing for a well defined, limited set of uses may be the only means of addressing the complex compensation, control, and security issues raised by the mass digitization of books.

4. Uses permitted under any orphan works regime should be carefully circumscribed, to avoid damaging literary markets here and abroad. Any potential uses should be carefully weighed, with a strong preference given to uses that help authors, artists, filmmakers and others make new creative works. Rote copying of entire works, permitting a user to take on the publishing function, should be avoided. Care should especially be taken to avoid disrupting existing, well-functioning permissions markets served by literary agents, publishers, and authors.

Although arriving at solutions to the problems posed by orphan works appears to be more challenging than ever, we still believe that solutions that respect authors' rights and strengthen the literary markets that copyright is intended to foster are achievable.