

Cornell Law School

From the Selected Works of James Grimmelmann

April 15, 2009

The Google Book Search Settlement: Ends, Means, and the Future of Books

James Grimmelmann, *New York Law School*



Available at: https://works.bepress.com/james_grimmelmann/25/



AMERICAN
CONSTITUTION
SOCIETY FOR
LAW AND POLICY

The Google Book Search Settlement: Ends, Means, and the Future of Books

By James Grimmelman

April 2009

The American Constitution Society takes no position on particular legal or policy initiatives. All expressions of opinion are those of the author or authors. ACS encourages its members to express their views and make their voices heard in order to further a rigorous discussion of important issues.

The Google Book Search Settlement: Ends, Means, and the Future of Books

James Grimmelmann*

I. Introduction

For the past four years, Google has been systematically making digital copies of books in the collections of many major university libraries. It made the digital copies searchable through its web site—you couldn't read the books, but you could at least find out where the phrase you're looking for appears within them. This outraged copyright owners, who filed a class action lawsuit to make Google stop. Then, last fall, the parties to this already large class action announced an even larger settlement: one that would give Google a license not only to scan books, but also to sell them.

It is difficult to overstate the importance of this settlement. The ongoing shift to electronic publishing is arguably the biggest transformation in books since Gutenberg's invention of the printing press. The scale of Google's plans boggles the mind. If the settlement is approved, Google will have the closest thing to a universal library the world has ever seen. We should be enthusiastic about the prospect of creating such a library, and concerned that it may be under the exclusive control of one company. This issue brief will connect this enthusiasm and this concern to the structure of the settlement that gives rise to them both.

Part I of the Issue Brief is an analysis of the lawsuit and its settlement. It discusses Google's original scanning project and the copyright-infringement lawsuit the authors and publishers filed to stop the scanning. It then looks at the settlement Google struck with the authors and publishers to end the lawsuit; the lawsuit has enormous potential both to do good and to cause harm.

Part II then examines one especially important part of the public interest context. Book copyrights are haunted by the specter of "orphan works"—books under copyright but whose copyright owners cannot be found. The orphans are neither making money for their authors and publishers nor effectively available to the public. The settlement authorizes Google to start selling orphan works. That's good for the public to the extent it makes them available again, but potentially bad to the extent it turns Google into a dominant platform with control over a huge catalog of books that no one else has access to.

* Associate Professor, New York Law School. Professor Grimmelmann was a Resident Fellow of the Information Society Project (ISP) at Yale Law School and heads the Public Interest Book Search Initiative (PIBSI) at the Institute for Information Law and Policy (IILP) at New York Law School. PIBSI has received funding from Microsoft, and the ISP and IILP have multiple corporate and foundation sponsors, which are described on their respective web sites. Professor Grimmelmann has worked with Creative Commons and the Electronic Frontier Foundation. Prior to law school, Professor Grimmelmann worked in the computer technology field, serving as a computer programmer for Microsoft. This Issue Brief may be freely reused under the Creative Commons Attribution 3.0 United States license, <http://creativecommons.org/licenses/by/3.0/us/>.

Finally, Part III discusses process. I argue that orphan works issues should be resolved through the legislative process, as the issue substantially affects the interests of large swathes of the public. Indeed, orphan works legislation has been hotly debated in Congress. The settlement tackles the orphan works problem, but through the judicial process. Laundering orphan works legislation through a class action lawsuit is both a brilliant response to legislative inaction and a dangerous use of the judicial power. Many of the public interest safeguards that would have been present in the political arena are attenuated in a seemingly private lawsuit; the lack of such safeguards is evident in the terms of the resulting settlement. The solution is to reinsert these missing public interest protections into the settlement.

I. The Project

“Google's mission is to organize the world's information and make it universally accessible and useful.”¹ Building on its wildly successful web search engine, Google has expanded into searching news stories, blogs, images, videos, and other media. When it started thinking about books, Google worked with publishers through a Partner Program: the publishers supplied it with digital copies of books, and Google made the copies searchable. Type in a search term, and Google would show you what books the term appeared in, along with “snippets”—a few sentences before and after, to give you a little context for seeing how the term appeared in the book. Where the copyright owner agreed, Google also provided previews of a few pages.

In December 2004, however, Google announced that it had also struck partnerships with a number of major research libraries to start scanning their collections. Details of the scanning technology are largely a matter of secrecy and speculation, but the basic outline is well known.² The libraries supply Google with stacks of books from their stacks. Google then uses custom-built machines to hold the books in place and take high-quality photographs of each page. Individual workers stand by the books, turning the pages, and occasionally photographing their hands by mistake.³

Some of the books Google scans are in the public domain, and are thus free for anyone to reuse. Google generally makes them not just searchable but available for download as free PDFs. For books still under copyright protection and whose copyright owners are members of the Partner Program, Google follows their requests in deciding how much of the books to make available. In between, though, is a vast middle ground—perhaps 70% of all books. Google took the position that unless and until it affirmatively heard otherwise from a copyright owner, it would display snippets in search results for books. In this, Google Book Search set itself apart from other similar programs like Amazon's “search inside the book,” which also provides snippets and previews, but which is strictly opt-in for authors, publishers, and other copyright owners.

¹ Google Company Overview, <http://www.google.com/corporate/> (last visited Apr. 10, 2009).

² See Jeffrey Rosen, *Google's Moon Shot*, NEW YORKER, Feb. 5, 2007; Kevin Kelly, *Scan This Book!*, N. Y. TIMES MAG., May 14, 2006, at 42.

³ See Duncan Riley, *Google Books Adds Hand Scans*, TECHCRUNCH, Dec. 6, 2007, <http://www.techcrunch.com/2007/12/06/google-books-adds-hand-scans/>.

A. The Lawsuit

It didn't take long for copyright owners to raise an outcry over the scanning and the snippets. They saw Google as an intermeddling freeloader, using their books for its own benefit and without paying. They also thought of Google's "scan first and ask questions later" attitude as a dangerous precedent. On September 20, 2005, a group of authors led by the non-profit Authors Guild filed suit against Google.⁴ A month later, a group of publishers coordinated by the Association of American Publishers (AAP) followed with their own lawsuit (which the court consolidated with the authors' suit).⁵ The publishers' suit was a standard lawsuit on behalf of five individually named plaintiffs, all large publishing companies. The authors' suit, however, was styled as a class action. The class—as expanded over the course of the lawsuit—now consists of anyone who owns a United States copyright interest in a book.

The copyright basis of the lawsuits was simplicity itself. United States copyright law gives a copyright owner a number of exclusive rights, including the rights "to reproduce the copyrighted work in copies" (the reproduction right), "to distribute copies . . . of the copyrighted work to the public" (the distribution right) and "to display the copyrighted work publicly" (the display right).⁶ It seems uncontested that Google's scans of copyrighted works constituted "copies" and were therefore *prima facie* infringements of the reproduction right. Similarly, Google's use of snippets in search results could implicate one or more of the reproduction, distribution, or display rights.

The heart of Google's defense was a fair use argument. All claims of copyright infringement in the United States are subject to the fair use defense, a common-law defense now codified in the Copyright Act.⁷ It requires courts to consider whether a particular defendant's uses ought to be excused—even though they would otherwise be infringing—because their social value outweighs any harm to the copyright owner. The copyright owners had, under long-standing print-based fair use principles, a strong argument that Google's uses were not fair. Google was making an unambiguously commercial use, one that required making complete copies of highly expressive books, but that didn't add much original expression of its own.

Google's reply arguments were rooted in a more modern line of decisions that have found fair use in digital cases involving verbatim copies. The common thread connecting these cases was that the copies weren't being passed along to the public; instead they were "intermediate" copies, used as part of the defendant's internal operations. Thus, for example, the Ninth Circuit found fair use where one video game maker reverse engineered a competitor's games to understand how they worked.⁸ Search engine indexing fits this mold; the point is to make a usable overall index, something very different than the indexed works themselves—but which can't be made without copying them along the way.⁹ Looming large in any discussion of fair use

⁴ Authors Guild, Inc. v. Google, Inc., No. 1:2005cv08136 (S.D.N.Y. filed Sept. 20, 2005).

⁵ McGraw-Hill Cos. v. Google, Inc., No. 1:2005cv08881 (S.D.N.Y. filed Oct. 19, 2005).

⁶ 17 U.S.C. § 106(1), (3), (5) (2000).

⁷ 17 U.S.C. § 107 (2000).

⁸ See *Sega Enters. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992).

⁹ See 4 WILLIAM F. PATRY, PATRY ON COPYRIGHT § 10:27.50 (2009).

would also have been Google's argument that trying to license scanning rights for many millions of books would have been so utterly impractical as to make the use impossible.

Many scholars were rooting for Google. They saw the fair use argument as compelling, given the public benefit of the index. As long as Google was merely indexing books, rather than trying to sell their contents, giving copyright owners a right to veto its creation would have run against copyright's purpose of promoting the progress of the useful arts and sciences. What's more, had Google won the fair use claim, it would have created a strong precedent for similar efforts by others to take underused works and to help others figure out what's in them—in effect, helping copyright owners bring them back into a larger, more successful market.

B. The Settlement

This picture changed radically on October 28, 2008, when the parties filed a proposed settlement with the court.¹⁰ According to the settlement's terms, Google will be released from liability, on a class-wide basis, for its past and future scanning and searching.¹¹ In return, Google will pay roughly \$125 million in attorneys' fees and as a one-time compromise payment. It will continue to allow copyright owners to choose whether any content from their books will be shown, and if so, how much.¹² If they allow Google to show excerpts from their books, Google will pay the owners roughly two-thirds of the resulting ad revenue.¹³ A new non-profit Book Rights Registry will collect the payments, hold them for copyright owners, and serve as the interface between copyright owners and Google.¹⁴ Thus, this part of the settlement breaks no new ground; it largely tracks the deal Google already offers to book copyright owners through its Partner Program.

The settlement, however, goes far beyond just scanning and searching. It also authorizes Google to start *selling books*.¹⁵ True, they'll be electronic books (e-books), to be read online through Google's servers. But still, Google is about to have access to a truly gigantic backlist: roughly 10 million titles. That will make it the largest bookstore in the world as measured by catalog size, and could well turn it into the largest bookstore in the world as measured by sales. Copyright owners can set their own prices, or delegate the pricing decisions to Google's all-knowing computers.¹⁶ Again, the Registry will act as the go-between, passing pricing decisions to Google and payments to copyright owners.

¹⁰ Settlement Agreement, *Authors Guild v. Google, Inc.*, No. 1:2005cv08136 (S.D.N.Y. Oct. 28, 2008) available at <http://www.googlebooksettlement.com/intl/en/Settlement-Agreement.pdf> [hereinafter Settlement Agreement].

¹¹ *Id.* at § 10.2.

¹² *Id.* at § 3.5(b).

¹³ *Id.* at §§ 4.5(a)(ii) (“Google shall pay . . . seventy (70%) of Net Advertising Revenues”), 1.86 (“less ten percent (10%) for Google's operating costs”).

¹⁴ The Registry's governing board will be made up half of author representatives and half of publisher representatives. *Id.* At § 6.2(b). Thus, no other groups will have voting representation at the Registry: not libraries, not readers, not anyone. The parties to the settlement either have not decided or are not yet willing to say who the initial members of the board will be.

¹⁵ *Id.* at § 4.2.

¹⁶ *Id.* at § 4.2(b).

In addition to offering à la carte e-book purchases, Google will also offer an all-you-can-eat subscription service to libraries, companies, colleges, schools, and other institutions.¹⁷ Subscribers will have full online access to most books in Google's scanned collection, and be able to print up to 20 pages at a time. The pricing is very much to-be-determined: Google and the Registry will negotiate a set of prices with the stated twin goals of making money and enabling access.¹⁸ Although it's hard to imagine the full subscription service coming cheap—how much would *you* charge for every book ever?—Google is also allowed to offer a limited, free version of the subscription to colleges and public libraries. Online access will be free at one terminal in each public library building and for every 10,000 students at four-year colleges (every 4,000 at two-year colleges).¹⁹

Even that's not all. The settlement allows the creation of a large research database of every book in the collection, allowing researchers to conduct automated studies that involve computer-based analysis of large numbers of books at once.²⁰ Imagine literary research tracking the occurrence of culturally-loaded terms in novels across the decades, or linguistic studies of pronoun reference in written English, and you have the idea. And we're still not done. The settlement also contemplates that the services Google offers will expand over time. It authorizes the Registry to negotiate the terms of various new business models with Google, including print-on-demand, PDF download, and coursepacks.²¹

C. Public Interest Concerns

The settlement, therefore, initially sounds like a perfect solution all around. The reading public gets access to the enormous all-time backlist of American arts and letters. Authors and publishers open up a substantial new revenue stream, rewarding them for their past creativity and investments—and giving them every incentive to continue producing new literary treasures. Google has the potential to open up yet another gold mine; libraries get a comprehensive new information resource, sometimes for free. What's not to like?

For one thing, there are competition concerns. The ninth circle of antitrust hell is reserved for price-fixers. Section 1 of the Sherman Act declares illegal “[e]very contract . . . in restraint of trade.”²² Legal agreements that cartelize an entire industry for the purpose of coordinating prices are one of the few remaining forms of conduct that antitrust law regards as *per se* illegal. Surely the sophisticated companies with expert lawyers who drafted this settlement agreement wouldn't use it to set up a system of naked price-fixing, right? That's what I thought, too, but consider this: The e-book purchase program²³ creates a system of fixed prices—twelve possible pricing buckets ranging from \$1.99 to \$29.99—and gives every

¹⁷ *Id.* at § 4.1.

¹⁸ *Id.* at § 4.1(a)(i)

¹⁹ *Id.* at § 4.8(a).

²⁰ *Id.* at § 7.2(d)

²¹ *Id.* at § 4.7.

²² 15 U.S.C. § 1 (2000).

²³ The subscription service seems less intrinsically problematic. All-you-can-eat licenses can have procompetitive effects. *See Broadcast Music, Inc. v. Columbia Broadcast. Sys., Inc.*, 441 U.S. 1 (1979) (holding that blanket licenses are not *per se* anticompetitive).

copyright owner the option to let Google pick which price each book will sell for.²⁴ Put another way, the copyright owners may collectively delegate their pricing decisions to a single centralized actor that can examine the market and pick a revenue-maximizing set of prices.

The Registry is also a potentially dangerous monopoly. The Registry's authority to negotiate new business models with Google puts it in a classically dangerous position. The Registry will speak on behalf of an entire industry, more or less. With that concentrated authority, it will have great practical ability to structure these new business models in anticompetitive ways in dealing with Google—or to structure them in ways such that it and Google present a coordinated face to the reading public.

Other concerns relate to intellectual freedom: people's ability to think and to learn, free from outside coercion.²⁵ Obviously, the broad access to books promoted by the settlement promotes intellectual freedom, but other terms potentially threaten it. Take privacy, for example. *The Theory and Practice of Oligarchical Collectivism*, the Index Librorum Prohibitorum, and *The Easy Way to Quit Drinking* remind us that intellectual freedom is closely bound up with reader privacy.²⁶ Google's book access programs, however, being online services, will be capable of tracking millions of users, book by book, page by page. That's a horrifying prospect for librarians who fought the Patriot Act's patron-records provisions. The settlement, however, is largely silent on privacy issues.²⁷

Libraries also fear that Google's power might actually reduce access to books. Libraries are likely to face intense pressure to subscribe to Google's services. Patrons accustomed to ready electronic access through the free public access service will demand that their libraries subscribe to the paid version once the lines at the free terminal start to grow. Outside funders—city councils and university administrators—will ask why the library needs all those expensive books and shelves when everything is available online anyway. Once they are in the system, though, these libraries will be at Google's mercy, no matter how high the prices or restrictive the terms. Once a library has gone digital, going back is almost impossible.²⁸ That is especially worrisome since Google's scans are, in library terminology, "access digitization" rather than "preservation digitization."²⁹ They are meant to be readable today, not to capture every detail that future scholars might care about; moreover, many of Google's digitizations are so marred by misaligned pages, bad metadata, or other scanning mistakes as to be unreadable.³⁰

²⁴ Settlement Agreement § 4.2(b)–(c).

²⁵ See, e.g., AM. LIBRARY ASS'N, LIBRARY BILL OF RIGHTS (adopted 1948, as amended 1980), <http://www.ala.org/ala/aboutala/offices/oif/statementspols/statementsif/librarybillrights.cfm>.

²⁶ See generally Julie Cohen, *A Right to Read Anonymously*, 28 CONN. L. REV. 981 (1996).

²⁷ In one respect, it's worse than silent. The libraries that gave Google books to be scanned received back digital copies. The settlement actually requires that these libraries report all scholarly and classroom uses of their digital copies. Settlement Agreement § 7.2(b)(vii).

²⁸ See generally NICHOLSON BAKER, *DOUBLE FOLD: LIBRARIES AND THE ASSAULT ON PAPER* (2001) (arguing passionately against deaccessioning decisions).

²⁹ See Kalev Leetaru, *Mass Book Digitization: The Deeper Story of Google Books and the Open Content Alliance*, FIRST MONDAY, Oct. 2008, <http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2101/2037>.

³⁰ See Paul Duguid, *Inheritance and Loss? A Brief Survey of Google Books*, FIRST MONDAY, Aug. 2007, <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1972/1847>.

Consumer-protection issues with the settlement's terms are also a concern.³¹ We might worry about the settlement's effects on consumers' rights under copyright law: an e-book cannot be resold or loaned the way a physical book can. We might also worry about the potential impermanence of e-books: when Google shut down its Google Video purchase program, it turned off the servers that made the videos viewable.

The common theme in these concerns is that they all relate to, or are magnified by, centralized power. Price-fixing among two copyright owners is harmless; price-fixing among two million copyright owners is a serious concern. Multiple e-book vendors give libraries and consumers the power to walk away from bad deals; a single monopolist leaves them helpless. Numerous and varied channels of cultural and intellectual exchange lead to a robust public sphere; a single point of control can threaten democracy. Personal information on a hundred readers is a privacy worry; personal information on a hundred million readers is a privacy dystopia. And so on and so forth. The central inescapable issue posed by the settlement is that it will leave Google in an immensely powerful position; it will control access to the single most comprehensive collection of books since the Library of Alexandria.³² It is that dominant position that makes an otherwise sound settlement worrisome. In the next part, I will examine more closely the source of that concentrated power.

II. Context: The Orphan Works Problem

Copyright is designed to increase the supply of creative works available to the public. To do that, it gives creators incentive to create new works by giving them a revenue source; willing buyers pay for copies of the work. Under ordinary circumstances, a user is more than happy to pay a price the owner is more than happy to accept.

An "orphan" work, however, has (or might have) an owner who cannot be found, who may not even know that she is a copyright owner. For instance, think of an author who dies without a will. Her next of kin may have no idea that they are now copyright owners. Or think of a publishing house that gets into financial distress and has to sell itself to a liquidator; the buyer may be thinking of the presses and the office chairs, not the copyrights. Especially with older works that are not currently generating revenue, it becomes all too easy to lose track of ownership records.

The result is a lose-lose situation. On one hand, orphan works owners never see a dime from their copyrights. On the other, it is impossible for a potential user to feel entirely safe when trying to put an orphan work to use, no matter how good her search for the owner. Perhaps it is true that nine out of ten orphan owners will never turn up; the tenth could still file a painful infringement lawsuit. The rational user plays it safe and foregoes the use. For instance, some photo labs will not reproduce old family photographs because of the copyright risk.³³

³¹ See, e.g., James Grimmelmann, *How to Fix the Google Book Search Settlement*, J. INTERNET L., April 2009, at 1.

³² See Robert Darnton, *Google & the Future of Books*, N. Y. REV. BOOKS, Feb. 12, 2009, available at <http://www.nybooks.com/articles/22281>.

³³ U.S. COPYRIGHT OFFICE, REPORT ON ORPHAN WORKS 24–25 (2006), available at <http://www.copyright.gov/orphan/orphan-report.pdf>.

Orphan works thus occupy an unfortunate middle ground between works that are actively being distributed and works in the public domain. If a book is in print and available for purchase at bookstores, the copyright owner is making money off of it; if a book is clearly in the public domain, the rest of the world is able to enjoy it, build on it, and put it to new uses. But orphan works are the worst of both worlds: no one is benefiting from them, financially, creatively or intellectually. This rough tripartite division—actively managed, orphan, and public domain—is critical for understanding the orphan works problem and the structure of the settlement.

Figure 1: Orphan Works Benefit Neither Copyright Owners Nor The Public

		Copyright owner making money?	
		Yes	No
Free for public reuse?	Yes	–	Public domain
	No	Actively managed	Orphan

Copyright’s increasingly long term compounds the problem. Copyright endures for the life of the author plus 70 years.³⁴ That is an immensely long time, and it is an immensely long time to keep careful track of ownership. Under the 1909 Copyright Act (which was in force until 1978), copyright lasted 56 years—but had to be renewed halfway through by filing a form with the Copyright Office. Over 85% of works were not renewed, putting them unambiguously in the public domain; the ones that were renewed had updated ownership records, aiding in the process of finding the owner. For books, orphan-dom is a serious problem. The low renewal rates under the 1909 Copyright Act and the small fraction of books kept in print today suggest that the vast majority of books ever published are not being put to productive use.³⁵ The precise split between orphan works and ones whose owners are rationally allowing them to lie fallow is impossible to ascertain, but no one thinks that the number of orphans is small. Half of all books would be a conservative, lowball estimate.

For this reason, then, the prevalence of orphaned books would have played a major role in Google’s fair use defense. Orphan works take the transaction costs argument for fair use—the costs of negotiating a deal would exceed the social benefits³⁶—and ratchet it up to a new level. Orphan owners cannot be found, not at any price. The transaction costs of negotiating with them are effectively infinite. Similarly, truly orphan owners—who cannot be found to be paid—are inherently unable to profit from selling licenses to scan and search their books. Putting these points together yields the powerful argument that rejecting a fair use defense for scanning orphaned books would necessarily mean shutting down a project of great social utility, but without one iota of offsetting benefit to anyone.

³⁴ 17 U.S.C. § 302 (2000).

³⁵ See Christopher Sprigman, *Reform(aliz)ing Copyright*, 57 STAN. L. REV. 485, 497, 519–21 (2004) (collecting sources on renewal rates and in-print status).

³⁶ See Wendy Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and Its Predecessors*, 82 COLUM. L. REV. 1600 (1982).

A. Orphans and the Settlement

If the orphans thus loomed large over the lawsuit, they also cast long shadows across the settlement. Under its terms, by default, Google will be free to make all uses of out-of-print books: scanning, searching, advertising, sale, subscription, and so on. It will set their prices algorithmically, and forward on to the Registry all the money it makes from them. When the copyright owners turn up, they can claim their money and start taking active control of how the books are priced and displayed.

Of course, some owners will *never* turn up. These are the orphan owners. No one knows exactly how many are out there, although to all indications, there are a great many of them. (A quick back-of-the-envelope calculation based on the dollar figures the settlement sets out for one-time payments to copyright owners and the size of the pool set up to pay them suggests that only about 10% of copyright owners are expected to register with the Registry.³⁷) Thus, by the logic of the settlement, the orphan works will be scanned, searchable, and available for sale. The money piling up in their names will sit there, until after five years the Registry will be allowed to apply it to other operating expenses. If any is left over, it will go to charities and to the copyright owners who did turn up.

It's important to recognize the critical role that the settlement's treatment of orphan works plays in its ambitious scope. Because Google is allowed to presume consent of absent copyright owners—precisely the presumption that the plaintiffs objected to when they filed the lawsuit—it sets a default that most of the books in existence in the United States will be part of Google's collection. Every orphaned book is a book whose owner will never reverse the default, will never opt out. Google's book services will be comprehensive on a previously unimaginable scale; the settlement ensures that copyright claims by orphan works owners will not threaten that comprehensiveness.

Indeed, just as Google would have argued the intractability of the orphan works problem in making a fair use case for scanning, that self-same intractability provides the best justification for the comprehensive book-sales programs that the settlement establishes. The orphans are books that are not being made available to the buying public and are not benefitting their copyright owners. The class action gives Google the green light to start making them available again, while collecting money on behalf of copyright owners. That is not an arrangement that would be possible through any kind of voluntary dealing. With orphan works legislation stalled out, the class action settlement may well be the only way of accomplishing this goal.

This basic deal has a lot to recommend it. Society gets access to these orphaned books; Google profits from making them available; money is pushed back at the copyright owners. Some copyright owners will never show up to claim their money; we should shed no tears on

³⁷ The settlement sets aside \$45 million to pay "inclusion fees" to copyright owners whose works were scanned. The minimum inclusion fee is \$60, so the settlement would seem to expect about 750,000 book copyright owners to show up. Compare that figure with the roughly 7 million books Google has scanned, of which about 6 million are in copyright. Jule Sigall, Senior Policy Counsel for Microsoft (and principal drafter of the Copyright Office's study of orphan works) made this calculation in a presentation at a Columbia Law School symposium on the settlement on March 27, 2009. A video recording of his comments, along with a rebuttal from one of the settlement's drafters, is available at <http://kernochancenter.org/Googlebooksettlementrecording.htm>.

their behalf, since there was nothing we could have done for them. Some will show up, at which point they are no longer orphan owners, and everyone is happy about that, too. As for the money that has been piling up in the meantime, if a 63/37 revenue split is fair to present copyright owners, it seems fair to absent ones, as well. This deal is, in its high-level structure, a model for how to do orphan works legislation properly.

The devil, however, is in the details. I have noted a number of problematic details about the settlement: competition, privacy, conflicts of interest, quality, and so on. These devilish details have at least two things in common. First, issues that might not be troubling if they involved a few copyright owners or one of many book platforms become much more worrisome when they involve all copyright owners and a dominant platform. Second, the concentration of power that amplifies all our concerns is *directly traceable* to the orphan works-affecting provisions of the settlement—that is, to its class action nature.

B. A Deal Good for Google Only

Google is not the only player with an interest in scanning books and making them available. Microsoft had a book-scanning program for search purposes, but closed it down in mid-2008. Amazon has the institutional capacity to make books available digitally on a huge scale, and could quite plausibly expand into scanning. The non-profit Open Content Alliance has been busily scanning public-domain works, as well.

In a post-settlement world, however, all of these potential competitors face one insurmountable hurdle: copyright law. Anyone who tries scanning in-copyright books and selling access to them will immediately be sued into oblivion by a mob of angry copyright owners. The only way to avoid that gruesome fate is to get the appropriate licenses, in the time-honored tradition of copyright law: by negotiating advance permission from each individual copyright owner. Here's where the class action nature of the settlement becomes crucial. Precisely because the present lawsuit is a class action, Google can rely on it to force a standardized deal on all book copyright owners at once. Some of them will opt out, to be sure, but the vast majority of them will not. Google only had to negotiate with ten copyright owners to reach this deal; the class action multiplies the effect by a factor of a million.

This really is a Google-only deal. Suppose that Yahoo! wants to get into the business. If it starts scanning and gets sued, who's to say that the plaintiffs—who could, by definition, be almost anyone—would file their suit as a class action, be inclined to settle, and be inclined to settle on terms comparable to those offered by Google? Keep in mind that the plaintiff-side attorneys fees payable under the proposed settlement exceed \$45 million. Now that Google is in this line of business, what sensible copyright owner would want to spend tens of millions creating competition—and driving down the price of books? Without a group of “representative” plaintiffs willing to cooperate with the way that the Authors Guild cooperated with Google, Yahoo! would be in the extremely difficult position of trying to file and settle a declaratory judgment action against a gigantic defendant class. Note also that because Google's victory comes from a settlement, rather than a court ruling, it establishes no fair use principles for competitors to draw on in framing their own legal arguments.

Thus, the settlement, by its very class action nature, creates a remarkably effective barrier to entry. No potential competitor could come close to matching the depth of Google’s backlist. As though to make doubly sure, the settlement contains a powerful most-favored-nation clause in Google’s favor; the Registry promises to Google “economic and other terms . . . that, when taken as a whole, do not disfavor or disadvantage Google as compared to any other substantially similar authorizations granted to third parties by the Registry.”³⁸

Thus, Google’s extraordinary market power under the settlement will come from its unique lock on orphan works. These are precisely the works that no one else has legal permission to use, and precisely the works that no one else dares sell *en masse* for fear of the not-actually-orphan copyright owner who turns up with a grudge and a summons. The class action settlement gives Google a green light to make the orphans available; that same light is red for everyone else.

III. Ends and Means

Copyright law is broken, and the orphan works problem illustrates everything wrong with it. Well-intentioned users who can’t honestly acquire the copyright clearances they need? Check. Far-reaching rights that aren’t actually doing much good for creators? Check. Painfully ambiguous fair use standards? Check. Powerful remedies that bear no rational relation to the harms involved? Check. Orphan works are broken for everyone.

A. Legislative and Judicial Solutions

The orphan works problem, in other words, is a textbook example of a *legislative* problem. Something’s wrong with existing law and it affects a lot of people. If we decide to start tampering with the rules, we would like to make sure that the gains are broadly distributed, rather than just accruing to a lucky few. What’s more, the problem affects different constituencies in different, subtly interacting ways. It may be that someone’s ox will need to be gored in order to make progress, but ox-goring is never a decision that should be made lightly. These are the kinds of comprehensive rulemaking and interest-group balancing for which legislation is ideally suited.

Indeed, legislative fixes for the orphan works problem have drawn no shortage of interest. Responding to congressional requests, the Copyright Office conducted its own detailed study of the orphan works problem in 2005 and 2006.³⁹ The resulting study proposed limiting damages to “reasonable compensation” and putting sharp limits on the use of injunctions where “prior to the commencement of the infringement, performed a good faith, reasonably diligent search to locate the owner of the infringed copyright and the infringer did not locate that owner.”⁴⁰

³⁸ Settlement Agreement § 3.8(a).

³⁹ See U.S. COPYRIGHT OFFICE, *supra* note 33.

⁴⁰ *Id.*

Legislation implementing the Copyright Office’s proposal was introduced in the 110th Congress.⁴¹ It passed the Senate but expired in committee in the House. Although the proposal had a number of high-profile supporters in the nonprofit and commercial worlds, it also drew sharp opposition. Some copyright owners, particularly visual artists, thought it went way too far;⁴² Lawrence Lessig thought it “simply wouldn’t do much good.”⁴³ Copyright legislation in the United States tends to follow a consensus model; Congress will not act unless it perceives that most major commercial stakeholders in the copyright system are on board.⁴⁴ At the moment, at least, the political process appears to be logjammed on the orphan works problem; no one has found a way to resolve it that does not make some powerful group scream loudly.

The Google Book Search settlement, then, is a judicial solution to a legislative problem. That makes it appealing. If we care about making orphan works available again, then we ought to be happy that Google, the Authors Guild, and the AAP were able to figure out a way to fit an orphan works solution into a class action settlement. But we should also be concerned that legislative problems demand legislative solutions; in the process of folding and squishing orphan works legislation to make it look like a settlement order, some important pieces have been mutilated.

Think about how the issues I have raised with this settlement would have been dealt with had it been considered as legislation. The competition concerns with designating Google as the unique source for orphan books would have been obvious, and Congress would almost certainly have had to create a generic process that others could use to make those books available. The Registry’s dangerous power would also have been evident and it would have been straightforward to set hard limits on its power to prevent abuses. Consumer and library advocates could have shown up and presented their concerns; Congress would have written privacy guarantees and library freedoms into the legislation.

None of this has happened in the settlement deliberations, at least so far. Google and the plaintiffs have resolutely described it as a private settlement, between private parties, in which the concerns of non-parties simply ought not to register. By judicializing their orphan works solution, the parties to this lawsuit have done an end run around the usual safeguards of the democratic process. The ends may be noble, but the means are not—and the result of choosing the wrong branch of government is that the resulting product is tainted. Freed from the usual checks attendant upon legislation, Google and the plaintiffs have structured the deal to exclude competitors, take money from orphan copyright owners, and potentially oppress the reading public. One should be troubled by this prospect. The public has a right to demand that the concerns that would have been aired in the legislative forum be aired in the judicial one.

The costs of standing idly by go beyond this particular lawsuit. If the settlement is allowed to stand without question, it will set a terrible precedent for the future. We’re committed

⁴¹ Orphan Works Act of 2008, H.R. 5889 110th Cong. (2008); Shawn Bentley Orphan Works Act of 2008, S. 2913 110th Cong. (2008).

⁴² See ILLUSTRATORS’ P’SHP, GROUPS OPPOSING THE HOUSE & SENATE ORPHAN WORKS BILLS AS WRITTEN (2009), http://www.illustratorpartnership.org/01_topics/article.php?searchterm=00273.

⁴³ Lawrence Lessig, *Little Orphan Artworks*, N.Y. TIMES, May 20, 2008, at A23.

⁴⁴ See generally JESSICA LITMAN, DIGITAL COPYRIGHT (2001).

to a democratic political process, and if we think the process is broken, the solution is to fix it, not to avoid it entirely. Every time we launder a legislative problem through the judicial system with an arguably abusive class action, we erode the rule of law a little more.

B. Substantive Reforms

I have written elsewhere about specific substantive ways the settlement could be improved.⁴⁵ The list is long, but three fairly straightforward changes would suffice to eliminate most of the serious reasons to worry about the settlement. They all relate to issues that would have been on the table if this deal had been hammered out in Congressional hearing rooms, rather than lawyers' offices.

First, the arrangement needs to be genuinely non-exclusive. Right now, the deal is formally non-exclusive in that it does not actively prohibit copyright owners from authorizing Google's competitors to make their books available. But the most-favored nation clause and the class action default will make Google's deal exclusive in practice. That is structurally easy to fix. Anyone else willing to assume the same obligations that Google assumes under the settlement should also be entitled to the same benefits.

If Google isn't the only game in town for the entire back catalog, a remarkable number of other concerns become much less serious. Google cannot hook libraries using drug-dealer pricing (the first time is free) if it has serious competition. Botched scans are not the end of the story if there is another source for the pages. Terms and conditions will be fairer if negotiated in a competitive market. And so on and so forth. Fixing the exclusive licensing regime in the settlement is an absolutely necessary condition for competition to be feasible.

Second, we need to recognize the Registry for what it is: a new collecting society, that is, an institution that issues licenses on behalf of an entire class of copyright owners and divides the money up among them. The best-known collecting societies in the United States are the performing rights organizations—The American Society of Composers, Authors and Publishers (ASCAP), Broadcast Music, Inc. (BMI), and SESAC. They hand out licenses to radio stations, concert halls, large stores and restaurants, and anyone else who wants to perform music publicly without having to negotiate with every individual composer.

Collecting societies have great power, and with it comes great responsibility. Their role—aggregating rights to smooth out licensing transactions—tends to require that we tolerate their accumulation of market power and explicit coordination of pricing. In exchange, however, our copyright and competition authorities must watch them carefully. ASCAP and BMI are subject to antitrust consent decrees, and for very good reason. They have agreed to forego a whole range of dangerous practices, and to live with continuing court supervision to make sure they do not invent new ones.

The basic obligation we need to impose on the Registry is to deal fairly and evenhandedly with all comers. The settlement is already fairly good about this on the copyright owner side; the Registry is prohibited from representing subsets of the settlement class, and thus

⁴⁵ Grimmelmann, *supra* note 31, at 1.

cannot turn copyright owners against each other.⁴⁶ On the Google side, however, the Registry's powers and duties are too vague for comfort. It has vast power to negotiate new business models, but no obligations to be pro-competitive or evenhanded when doing so; its powers to negotiate with Google's competitors are limited and slightly obscure.

Somewhat paradoxically, the best way to make the Registry work well is to universalize it. The settlement, for jurisdictional reasons, limits the class of copyright owners whose works will be controlled by the Registry.⁴⁷ That is another unfortunate artifact of the judicial process; if Congress were to charter the registry, it would not and should not be so limited. The Registry should explicitly be authorized, indeed required, to represent *all* copyright owners with an interest in books: past, present, and future. Thus, it would be able to hand out blanket licenses that have no unfortunate jurisdictional holes (which create licensing transaction-cost problems and create conflicts among copyright owners). If we think the settlement is fair for the settlement class, we have every reason to think that it's fair for *all* book copyright owners.

Third, we need to take the issue of terms and conditions seriously from the very start. The sheer scope of the settlement—especially if it is extended by making it non-exclusive and covering more copyright owners—means that it will profoundly reshape copyright law as it affects books. The fair use, library use, and first sale provisions of the Copyright Act need to be appropriately translated into the digital book access envisioned by the settlement. The settlement currently does a passable job, but there are reasons to be concerned that it may undermine library rights to preserve books and reader rights to share them. Neither of these groups was properly heard from in the negotiations that produced this settlement; their voices would have been a larger part of congressional or agency deliberations.

C. How to Get There

What now? The settlement is still subject to the court's approval. Nominally, the court is considering both whether to certify the plaintiff class and to approve the settlement binding that class. Thus, the court has set a May 5 deadline for class members to file opt-outs and a June 11 fairness hearing. These deadlines follow one of the largest legal-notice programs in history.⁴⁸ During the five-month notice period, Google has spent \$7 million on print ads in 70 languages, in publications as obscure as the *Nauru Bulletin* (circulation 700).

⁴⁶ Settlement Agreement at § 6.2(b).

⁴⁷ *See id.* at § 1.16 (defining "Book"). Specifically, if a book is a United States work, it is covered by the settlement only if the copyright has been registered with the United States Copyright Office; if the book is a foreign work, it is covered as long as has been made available in the United States. These categories track the jurisdictional categories of copyright law. *See generally* 17 U.S.C. § 101 (2000) (defining "United States work"); 17 U.S.C. § 411 (2000) (requiring registration prior to suit for United States works). The Second Circuit—in a decision currently under review by the Supreme Court—had held that a federal court lacked jurisdiction to approve a class action settlement whose class was drawn more broadly. *In re Literary Works in Elec. Databases Copyright Litig.*, 509 F.3d 136 (2d Cir. 2007), *cert. granted sub nom.* *Reed Elsevier, Inc. v. Muchnick*, 77 U.S.L.W. 3487 (U.S. Mar. 2, 2009 No. 08-103).

⁴⁸ *See* Noam Cohen, *A Google Search of a Distinctly Retro Kind*, N.Y. TIMES, Mar. 3, 2009, at C1, *available at* <http://www.nytimes.com/2009/03/04/books/04google.html>.

This procedural posture complicates the process of raising public interest concerns. In approving or rejecting a class action settlement, a court's typical concern is whether the settlement is "fair, reasonable, and adequate" to class members.⁴⁹ For copyright owners who show up and register with the Registry, the settlement seems excellent. In particular, they can decide at any time to stop allowing their books to be sold and can even opt out of snippet display in search results. Google's new book sales programs offer a new revenue stream, but do not force unwilling authors or publishers to drink from it.

For absent copyright owners—we might call them members of an orphan works owner subclass—the settlement is more questionable. The most directly objectionable terms are the provisions reallocating their revenues after five years. Whether the reallocated revenues are given to copyright owners who have registered or are used to defray Registry expenses that registered copyright owners would otherwise need to pay for themselves, the effect is the same. Money generated from Google's use of orphan works is given not to the orphan work owners, but to other copyright owners. This reallocation illustrates a clear conflict of interest between the non-orphan copyright owners (including, by definition, all of the named plaintiffs) and the orphan members of the plaintiff class. As such, it speaks directly to the adequacy of the representation provided by the class counsel, and should lead the court to scrutinize the settlement closely.

Other public interest concerns, however, cannot be so easily fit into the class action framework. There is no easy way to come in and say that libraries, for example, will be adversely affected by the settlement; that does not really sound like a question of fairness to absent class members. The narrowness of the focus on absent class members, in a case with such profound effects on the public interest in copyright, ought to be a signal that something is seriously wrong with using a class action settlement to bring orphan works legislation into being.

This is not to say that the court is prohibited from considering these urgent larger issues. Settlements may not themselves violate the law,⁵⁰ a rule that comes into play with some of the antitrust issues. Courts also have a duty to ensure that class action settlements are "consistent with the public interest."⁵¹ Objecting copyright owners, intervenors, and *amici curiae* are thus free to call such issues to the court's attention, and some are planning to do so. (My own institute at New York Law School, for example, will be filing an amicus brief.)

Fortunately, there also will be opportunities outside the present lawsuit to raise some of these issues. Google's competitors could file antitrust lawsuits at any time, as could federal and state regulators. Google will have to negotiate the actual terms of its subscription service with libraries and other institutions, which may be able to require fair terms and conditions as part of the negotiations. If consumers themselves do not have the bargaining power to insist on good privacy protections when Google launches its actual book sales programs, it is certainly possible that the Federal Trade Commission and the courts could do it for them.

⁴⁹ FED. R. CIV. PROC. 23(e)(2).

⁵⁰ See, e.g., *Dillard v. City of Foley*, 926 F. Supp. 1053, 1063 (M.D. Ala. 1995)

⁵¹ *Bailey v. Great Lakes Canning, Inc.*, 908 F.2d 38, 42 (6th Cir. 1990).

We should also think about going back to Congress. Last session's orphan works bills stalled, but that does not mean that all orphan works legislation is doomed. The settlement itself could provide a workable model for a book-specific orphan works bill. Just take every mention of "Google" and replace it with a general authorization for anyone to scan and sell books provided they honor the terms of Google's deal with the Registry. Orphan works legislation on this model would be fair to copyright owners; they would get the same deal they do under the settlement. It would be fair to Google, which would remain serenely confident in its ability to take on any competitor so long as the playing field is level. It would solve most of the competition concerns with the settlement to the benefit of libraries, readers, and other book-scanners. That would bring to the table basically every interest group affected by book-specific legislation, making it quite feasible to craft a consensus proposal.

IV. Conclusion

The Google Book Search settlement serves respectable ends through questionable means. The copyright interests in books have been scattered to the four winds over the years, harming both the reading public and copyright owners themselves. True, a class action is a device for gathering together lots of widely scattered interests, but in this case, it's the wrong device. Because this deal was struck through private negotiation among a few parties, it neglects the broader public interest in some critically important ways.

The need for change is real, but at the same time, it's reassuring how eminently *solvable* the problems with the settlement are. The settlement may have emerged from a questionable bargaining process, but the end product bears at least a familial resemblance to an agreement of which we could all feel proud. This settlement does not need to be problematic, and we should not let it be. The court is being asked to place its imprimatur—our imprimatur—on this reshaping of our copyright law and our publishing system. We the people have the right to insist that our interest, the public interest, be reflected in the outcome.