

Google and the Zombie Army of Orphans

Revised version of a talk given at a symposium on
Google and the Future of Higher Education
Georgetown University Library Scholarly
27 February 2009

James Grimmelmann

New York Law School
57 Worth St.
New York, NY 10013
james.grimmelmann@nyls.edu

This work may be freely reused under the Creative Commons Attribution 3.0 United States license

<http://creativecommons.org/licenses/by/3.0/us/>

I'd like to talk a bit about ends and means in the Google Book Search settlement. I'll try to keep this as simple as possible. Unfortunately, when talking about a settlement agreement with 141 pages and 15 attachments, what's possible isn't always simple.

The lawsuit started out as a challenge to Google's project to digitize and make to make searchable every book it could get its hands on. Google wanted to add books to its search results: type in a term, and Google would tell you what books that term appeared in and where they appeared within the books. Google wouldn't give you the books themselves; instead, Google would tell you what was out there and it would generally be up to you to go track down the actual books themselves if you wanted to read them. It was purely a project about indexing.

I thought this was an unambiguously good thing. Google was creating a general index of everything that had ever been written. An index like that is a remarkable tool for the discovery and transmission of knowledge. Having one greatly accelerates the speed at which people can share creativity and ideas. For researchers, the difference between a good index and a truly comprehensive one is immense.

I also thought the Book Search project ought to be an unambiguous fair use because of the inordinate difficulty of negotiating a deal with all book copyright owner—authors and publishers— to authorize the program. It's not just that some owners would drive hard bargains or hold out for spiteful reasons. It's also that some wouldn't negotiate at all, because no one can find them to negotiate with at all.

A specter is haunting American copyright law—the specter of orphan works. Many, many books under copyright have gone out of print and have copyright owners who can't be found. Perhaps the author didn't leave a will and her heirs don't know that they're now the owners of her copyrights. Perhaps the publisher went out of business, and whoever bought up its assets was thinking of the printing presses, not the copyrights. Problems like these confront a substantial fraction of all books in copyright; they make it essentially impossible to secure permission from everyone affected by Google's scanning. Thus, rejecting Google's fair use defense would have meant vetoing forever its attempts to create a comprehensive catalog. That would have struck me as a great loss.

Although Congress has debated ways of improving access to orphan works, legislation on the issue died in committee in the House last year. The settlement does what Congress hasn't; because the lawsuit is a class action, copyright owners are bound by the settlement's terms unless they speak up. Thus, not only can Google scan books whose owners say, "Yes, please," it can also scan books whose owners don't affirmatively say, "No way!" The class action sweeps in by default

every book whose owner never shows up to say anything at all. Congratulations, orphan works owners, wherever you are: you're about to be part of Google's index.

The settlement, however, goes well beyond just book search. In addition to authorizing Google to scan and index books, the settlement lets it sell full-text access to them, either as individual book purchases or through a subscription to the entire collection. It may well turn into the world's largest bookseller. The deal seems broadly fair for authors and publishers; Google will pass 63% of the revenue from selling books to them. As for absentee orphan works owners, the settlement sets up a Book Rights Registry that will hold on their behalf the money Google collects from selling their books, paying out if and when they show up. In effect, this private settlement through litigation solves a problem that has become stuck in the political process because it's so contentious. The result has great benefits for readers, for copyright owners, and of course, for Google.

The settlement in its current form also has substantial problems. It creates a significant danger that we might give control of the distribution of books and knowledge to one monolithic entity. If it becomes a dominant platform, Google could well become the only game in town for serious online access to many books. The Registry will also have enormous power to establish the terms of access to books and copyrights.

The settlement makes almost no provision for the privacy of readers; Google could, under the settlement's terms, be tracking you as you read page by page. How much Marx; how much Marx Brothers? Google can do disturbing things with that information. Privacy is at the very heart of intellectual freedom.

Similarly, there are few protections in the settlement for consumer rights. If I went to a library and borrowed a book there, or if I bought a paperback at a bookstore, I'd have all sorts of guaranteed rights under copyright law. Google's version of electronic access may well take away many of those freedoms on the ground. I also wonder how consumers will feel if they buy a book only to discover that some of its pages are unreadable due to scanning mistakes—and what recourse they'll have if that happens.

How can one raise these issues in a way that makes sense in the context of the settlement? The public-interest principles I've mentioned—competition, privacy, and consumer rights—don't fit naturally into a purely private agreement worked out between the litigants. Shouldn't authors and publishers just be concerned with getting the most money out of the system? What right does anyone have to complain that they struck an amicable deal with Google?

No, the real issue here is the use of that class action device. The lawsuit didn't need to be a class action to answer the question of whether Google is legally permitted to scan books and make them searchable. Authors and publishers who objected to what Google was doing could simply have sued it for infringing their particular copyrights. They might have won; they might have lost. Either way, the ruling would have resolved the central legal issue.

But that's not how it was done. Instead, a few copyright owners filed the lawsuit in the form of a class action, dragooning everyone who owns a copyright into being one of the plaintiffs along with them. That includes a lot of people in this room. It may well include you. Until just now, did you realize that you were a plaintiff? I'm still not sure whether I'm a class member or not.

Out of that huge plaintiff class, there's one very large subclass unlikely to benefit from this agreement. I'm thinking of the very orphan works owners whose existence made this situation so problematic for Google in the first place. The orphan works owners are by definition the ones who can't be found. They couldn't be found to negotiate with Google originally, they can't be found to show up and claim their money under the settlement, and most importantly, they can't be found to show up and object that maybe the terms of this deal aren't what they would have wanted for themselves or for society.

Of course, this class-action bargain is necessary to make the deal work. It gives us, the reading public, access to all these books going forward. But as part of that deal, the small number of people who are actually in court are lining their own pockets, and not always in ways that are in society's best interests.

Thus, Google gets a legally backed head start on digitizing books and making them available. None of its competitors can go and start selling the whole corpus of books without getting individual permissions from every copyright owner. But of course they can't do that for the same reasons that Google couldn't initially do it. They'll have to file their own class action lawsuit, and there's no guarantee that they could find somebody to settle with them. Google has this market locked up: no one else can get the same kind of legal permissions it can.

I've heard the argument that we have class actions precisely so that courts can reach this sort of result. Isn't the whole purpose of a class action to resolve issues all at once like this? Yes and no. A typical class action involves someone who's made a dangerous drug or a product that doesn't work. Everyone who's bought one or is affected by one sues, the company pays out a pile of money, and it gets split up among everyone. It's compensation for wrongs done in the past.

This class action, though, this one is special. It's not just Google ponying up for past wrongs. Instead, this is a structural settlement; it reshapes the entire book industry by giving Google and Google alone access to this comprehensive out-of-print backlist. To make that happen, the settlement takes away the rights of people who aren't before the court. Indeed, knowing what we do about the orphan works problem in copyright law, we know that these absent class members are highly unlikely to be able to do anything about this massive giveaway to Google taking place supposedly in their name.

It's a version of Russell's paradox, applied to class action litigation. There's a class here that consists of all people who don't realize they're part of it. Under the guise of this class action, the named plaintiffs have been able to use the huge collection of orphan works copyrights as a bargaining chip. The named plaintiffs negotiated away everyone else's rights, lining up all those millions of books for Google's benefit. The orphans have become zombies, raised from the dead by the dark magic of a class action, turned into a shambling army under Google's sole control.

This, I submit to you, is not the way things ought to be done a democracy. We have political processes for resolving major social issues. We have a Congress; it holds hearings and passes bills. We have administrative agencies that can take expert advice and make reasoned decisions. The courtroom isn't supposed to be the place where we resolve huge issues that involve the carefully regulated copyrights of multi-million-member classes. Litigation is structured to sort out individual adversarial you-versus-me disputes. It's a uniquely bad way to sort out complex, sweeping questions—such as how we get at all of our information and all of our books.

The settlement is still a net positive for society: good things can come of corrupt practices. But we should be concerned that this isn't how it ought to be done. The parties here have reached a result that's different from what society has a right to expect. Perhaps the regular political processes are too jammed-up to be able to come up with something good for book search and orphan works on their own. But we shouldn't let that fact stop us from demanding something better than, "It's better than nothing."