The Orphan Wars

"Orphan books"—books that are in copyright but whose copyright owners can't be found—have been in the news lately, thanks to lawsuits over Google's plan to scan a copy of every book ever published. What started as a project to make a better search engine has gradually become a focal point for debate over whether the legal system can find a way to rescue the orphans from copyright limbo. Some of the libraries working with Google have announced plans to make available to their patrons digital versions of the books they think are orphans; an authors' group has sued to stop them. In this column, I'll review the convoluted history of the Google Books lawsuits, with an eye toward what they might mean for orphan books.

In 2004, Google started working with research libraries to scan books. Google used its scanned copies to launch a search engine; it gave the libraries digital copies of their books in return. Many of the libraries, in turn, kept their copies with HathiTrust, which works like a digital version of a shared off-site storage warehouse. HathiTrust makes multiple copies of each file, storing versions on hard drives and tape backups at the University of Michigan and Indiana University Libraries. It offers bibliographic information about the books to the public and provides a full-text search engine.

Google miscalculated how much the scanning project would upset copyright owners. In 2005, the Authors Guild (a membership organization of about 8,000 authors), claiming to represent all copyright owners, filed a class-action lawsuit against Google. Five major publishers soon filed their own suit. Neither group sued the libraries. At the time, observers thought the suits would turn on fair use: could the complete scans be excused on the grounds that Google would show only short "snippets" to search users?

Instead of litigating that question, however, Google and the authors and publishers spent years negotiating and defending the infamous "Google Books settlement," which was unveiled in 2008. Under the settlement (which didn't work out), Google could have commercialized the scanned books, sold them online, and shared the money with copyright owners. Libraries and universities would have been able to purchase an "Institutional Subscription" providing unlimited access to millions of books in the collection. Google's library partners would also have received immunity for their part in the scanning, provided they agreed to strict limits on what they could do with their own copies of the scans.

Opposition to the settlement was worldwide and vociferous, and the federal court rejected it in March 2011, on grounds too numerous to detail here. It now seems likely that the publishers will settle on much narrower terms and walk away. The Authors Guild, on the other hand, seems poised to litigate against its erstwhile ally with a newfound ferocity.

The Google Books settlement both was and wasn't about the orphans. It was about orphan books in the sense that it would have made them broadly available again. The trick was that by requiring class members to opt out if they didn't like the terms, the settlement all but guaranteed that orphan owners (who by definition can't be found) wouldn't take their books out. But it had no special category for "orphan books"; they were just swept up in its larger scheme. And now that the settlement is off the table, the lawsuit against Google is unlikely to have anything definitive to say about what can and can't be done with orphans.

The orphans took center stage in the spring and summer of 2011, as HathiTrust members created the Orphan Works Project. This project was intended to investigate the available author and publisher information about potential orphan book. If neither an author nor a publisher could be located and contacted for a book and the book was out of print, it would be flagged as a possible orphan. If at any time a copyright owner was identified and located, the book would be removed from the list of orphan candidates.

Then, in June, the University of Michigan Library announced that it would take these identified orphans and make them available for full view to its community. Other universities announced their participation later in the summer, planning to make books flagged as orphans in their libraries available to their own affiliates. The first batch of books was scheduled to go live on October 13, 2011.

On September 12, the other shoe dropped. The Authors Guild filed a lawsuit against HathiTrust and five of its member libraries, including the University of Michigan, Indiana University, and Cornell University. The lawsuit strongly resembles the one filed against Google—they are both based on the theory that scanning, by itself, is an infringement, even if no one sees the book—but many of the details are strikingly different. The suit against Google sought damages, which could have reached into the trillions of dollars. The suit against HathiTrust and the libraries seeks

> Building consensus will be a difficult matter while the suit is under way: hope, fear, and anger will tug at stakeholders in subtle and complicated ways.

By JAMES GRIMMELMANN

only a declaration that the libraries are infringing, an injunction against further scanning, and the right to "impound" the digital copies so that the libraries can't get to them.

Unlike the suit against Google, this new one isn't a class action. Instead, the Authors Guild is suing as an "associational plaintiff" on behalf of its members. It's joined by a few foreign authors' groups and a handful of individual authors (most of whom are officers or board members of one of the authors' groups).

The law here is unclear. Section 108 of the Copyright Act lets libraries make certain kinds of copies for preservation and research use, but only in narrowly defined circumstances. That leaves fair use. In the case of the storage of the scanned copies, the libraries' fair use case is even stronger than Google's: they're using the copies for preservation, and unlike Google, they don't show even snippets. The Orphan Works Project, however, is legal *terra incognita*.

A few days after filing the lawsuit, the Authors Guild started digging into the initial list of candidate orphan books posted by HathiTrust. With a few phone calls, it was able to find one book's author, J. R. Salamanca. The guild invited its members to dig through the rest of the list and was quickly able to find copyright owners or literary agents for a number of other books. (The most eyebrow-raising was Walter Lippmann's *The Communist World and Ours*.) All of these books had gone through HathiTrust's workflow and had been flagged as potential orphans: if another month had gone by with no one speaking up, they would have been made available online to the University of Michigan's library patrons.

Although some commenters dismissed these findings as anecdotal, they were still embarrassing to HathiTrust. These were, after all, among the books it had chosen as being among the most likely to be orphans. HathiTrust quickly suspended the Orphan Works Project in order to review its procedures. After some deliberation, it announced that it remained fully committed to the project,

which it would relaunch with a more detailed workflow.

What might this new lawsuit mean? "Nothing" is a distinct possibility: the suit faces some substantial procedural hurdles. Perhaps the most significant is that the Authors Guild may not have standing to object to the Orphan Works Project. It could be that none of the commercially successful plaintiffs are actually at risk of having their books included in the Orphan Works Project, which focuses on out-of-print books.

But if the suit does reach the orphan issues, it may have an enormous effect. A ruling for the libraries could give them a broad privilege to start making older out-of-print books available digitally when it seems unlikely that a copyright owner can be found. A ruling for the authors, on the other hand, could significantly limit the scanning of books, pamphlets, letters, and other works, even for archival and preservation purposes.

This suit also upends many of the conversations taking place around the Digital Public Library of America (DPLA), since whatever happens here will substantially shape the legal environment within which the DPLA will operate. The U.S. Copyright Office, which has tried to promote orphan works legislation, is launching a study of mass digitization. But one wonders how much progress it can make. Building consensus will be a difficult matter while the suit is under way: hope, fear, and anger will tug at stakeholders in subtle and complicated ways. And the suit gives Congress yet another excuse to keep well clear of intervening with orphan works legislation for the time being.

The Orphan Wars are upon us, I fear. We might have hoped that they would be the Orphan Discussions, or perhaps the Orphan Debates, but no. The Orphan Wars it will be. ■

James Grimmelmann (james.grimmelmann@nyls.edu) is Associate Professor at New York Law School.