2010

Hurtling Toward the Finish Line: Should the Google Book Settlement Be Approved?

Ivy Anderson

*University of California, Oakland*

Follow this and additional works at: [http://docs.lib.purdue.edu/atg](http://docs.lib.purdue.edu/atg)

Part of the [Library and Information Science Commons](http://docs.lib.purdue.edu/atg)

---

**Recommended Citation**


DOI: [http://dx.doi.org/10.7771/2380-176X.5545](http://dx.doi.org/10.7771/2380-176X.5545)

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
Hurtling Toward the Finish Line: Should the Google Book Settlement Be Approved?

by Ivy Anderson (Director of Collections, California Digital Library, University of California, Oakland, CA 94607)

Recently, Google and the plaintiffs filed their final briefs in defense of the Google Book Amended Settlement Agreement (ASA) that is before the New York Southern Federal District Court. As the rhetoric around the Settlement heats up to white-hot intensity, I’d like to offer a few personal thoughts from my vantage point at the California Digital Library.

The University of California Context

The University of California Libraries are Google’s second-largest library digitization partner; we are also the second-largest library digitization partner of the Internet Archive, thanks to generous funding in the past from Microsoft, Yahoo, the Alfred P. Sloan, and Kahle/Austin Foundations, and other sponsors. In all, UC Libraries have now digitized 2.5M books from their collections through these projects, both in- and out-of-copyright. Within the University of California, some of our closest faculty colleagues are also among the Settlement’s most prominent critics.

In our view, the proposed settlement is hardy perfect; as Google acknowledges in its brief, it’s a compromise among parties with differing agendas and motivations. CDL is a staunch supporter of the underlying aims of the Google Book project to make the knowledge enshrined in the world’s great libraries discoverable and accessible across the globe, and we support the public benefits that will ensue, including the benefits to libraries, if the Settlement is approved. At the same time, public criticism has been good for the Settlement, producing very real improvements in the amended version that is now before the court — improvements that would not have been made without that criticism. Long live democracy!

Digitization Partnerships:
The Promise and the Peril

Like many of the objectors, participating libraries went through their own period of outrage and indignation when details of the Settlement first came to light. What? We would have to buy back access to our own books? Why did Google let us down in abandoning its fair use defense? Why should the parties be allowed to create an artificial revenue model for works that are long out of print, books that would no longer exist at all outside of used bookstores, if the libraries themselves hadn’t purchased and maintained them at great expense over decades and indeed generations? How can they do this without our agreement as to terms, since it is we who have made these books available to them in the first place? Hasn’t our stewardship paid for these books many times over? Isn’t this why copyright law contains unique exceptions for libraries, in recognition of our mission to further the public good? Wasn’t the appropriate use of our own copies in light of fair use principles our decision to make?

The problem with this view, of course, is that libraries did not initiate this enterprise, and we are not its only beneficiaries. The Google project placed two sets of commercial interests as loggerheads, with copyright law in the middle. Admittedly, libraries took a risk in engaging in a partnership so legally entangled.

But let’s be honest: though few seem willing to admit it, revitalizing the world’s heritage of books for a digital age — a task that many considered impossible only a few short years ago — appears within reach today almost entirely due to Google’s enterprising vision. Even the Open Content Alliance, which CDL joined a year before becoming a Google partner, was in some sense a response to GBS (although it had other important antecedents, as well, thanks to Brewster Kahle’s equally inspired vision). When Google’s competitors withdrew their support for that project, no other funders stepped in to fill the breach. The plain fact is that despite the idealistic adjudications of some, the resources required to digitize our cultural book heritage on a grand scale are not likely to be marshaled in the U.S. by libraries and the public sector alone.

At CDL, we’ve done some estimating of what it would take to convert the roughly 15 million unique books in University of California library collections to digital form absent the Google enterprise, using the best alternative technology available today. The answer? Half a billion dollars, and one and a half centuries. And that is just the University of California’s books.

I like to compare this to the building of the great Temple of the Sagrada Familia in Barcelona, a city with which my family has an ancestral connection. When my husband’s grandmother left Barcelona as a young girl in the late 19th century, the Sagrada Familia had barely erected its first stone. In 2006, more than 125 years later, her great-granddaughter traveled to Barcelona for the first time, where she was able to observe Gaudi’s monumental edifice, still under construction. At this writing, completion is projected for 2026.

The speed at which Google is converting this content is not without costs of its own. Google’s iterative approach to building large-scale services has drawn criticism from some scholars accustomed to work that is honed and polished before it is released. This is, in part, an argument about means, not ends. Like those progressive JPEG images that start out blurry on the screen and become sharper as the details fill in, Google’s services are improving over time as it continually upgrades and enhances its images and metadata. Over time we will be able to replace those missing or still-blurry pages with better versions. Where the value of the content warrants it, we can selectively invest in more meticulous rendering, textual markup, and other enhancements.

Two cases are illustrative here. CDL has digitized a large number of public domain books with the Internet Archive, some of which have also been digitized in our Google partnership. Although CDL had to suspend its Internet Archive book scanning project earlier this year after Microsoft withdrew its support and additional grant funding proved elusive, we have every expectation that we will take up comparable projects with Internet Archive in future, because its technology is better suited to certain types of uses (better artifactual rendering, for example). The Early English Books Online (EEBO) database...
It's hard for me to see how the Google Settlement is fundamentally about whether Google and rights holders will be allowed to implement a particular set of business models for a certain set of books. I believe the Settlement should be approved, because it will create new and valuable services for libraries as well as consumers. But many of Google's participating libraries have their own plans for these books, plans that do not ultimately depend on the outcome of the Settlement. The greatest risk for libraries if the Settlement is not approved is that further legal setbacks may lead Google to abandon its interest in library digitization altogether. If that were to happen, a unique opportunity would be lost that is not likely to be repeated in our lifetime.

Life Beyond Google Book Search

What of our relationship to the Google Book project itself? Some of the concerns we hear from faculty have nothing to do with the Settlement per se, but rather with the long-term implications of GBS for library collections and services. Let me close with a few words about some of those concerns.

To our scholars who worry that we are about to throw our physical collections overboard in favor of digital surrogates of sometimes uneven quality, I want to say: not to worry. True, libraries everywhere find themselves having to consign more and more of their physical collections to remote storage as campus space grows increasingly scarce and user preferences migrate online. And some libraries — the UC's far less than others — are addressing the space crunch by de-accessioning low-use materials that are widely held with the knowledge that they can borrow these items from another library if need be. (Many cooperative initiatives are now underway to share such information and ensure that enough copies are retained throughout the nation's system of libraries to protect the integrity of the scholarly record.) That train has already left the station, and it's happening independently of large scale digitization. What digitization offers is a valuable complementary mitigation strategy: we can now make those remote collections eminently browsable, saving time and expense both for users and for libraries. As a library user, you can now determine whether that book is really what you're looking for before you request it, not afterward — and in some cases, the digital surrogate may indeed be all that you need. Libraries can promote these "hidden" volumes more effectively to their users, while limiting delivery costs to just those items that are truly wanted. This browsable and/or searchable digital surrogate — which is the quality level that most of the Google mass digitized scans are aimed at — is not a replacement for the original print book, and was never intended to be.

To our scholars who worry that we are outsourcing our library collections and services to Google, again I want to say: please don't worry on this score, either. Far from setbacks might our commission as stewards of the cultural record, we who have opened up our collections to digitization are shouldering this role with vigor. While Google and others are making these books discoverable online to a general audience, the University of California, along with other...
Ivy Anderson was appointed to the California Digital Library effective January 2, 2006. In this role, Anderson coordinates and provides statewide leadership in buying and licensing digital content. Her efforts also extend beyond licensed information, to the acquisition and management of select printed materials and reformatting. She also plays an important part in the university’s strategic goal of addressing the profound problems in the marketplace for scholarly publications. Prior to joining the CDL, Anderson worked at Harvard University Library, where she developed a digital acquisitions program. She is widely known for her work in advancing the goals of libraries and library consortia.

peer institutions, is creating a robust shared access and provision service for our digitzed books, one that adheres to professional standards, through our partnership in a groundbreaking enterprise called the HathiTrust. If you haven’t heard of HathiTrust yet, you soon will. No UC library user need go to Google to search the full text of our books, or to find accurate bibliographic information, or to view and download those that are in the public domain: s/he can go to http://catalog.hathitrust.org/ and be reassured that those books will be there, in ever-improved versions, for the long term. HathiTrust now numbers 5.4 million volumes from 26 libraries and is growing at a rapid rate, all searchable, all viewable if in the public domain (or otherwise rights-cleared), and all designed to inure to the long-term benefit of the nation’s libraries and their users. The digital library of the future resides not with Google, but with us. And we are building it today.

At the same time, Google, Internet Archive, and others, are providing an invaluable service in bringing the vast holdings of the great research libraries to a worldwide audience and integrating that content with general-purpose internet search services and other content. As one colleague has written, “Who among us has not benefited from a Google search?” In participating in these efforts, we are fulfilling our long-standing public service mission. The Google Settlement, if approved, will further these aims by providing more content, in more ways, to an even wider audience.

But in the end, approval of the Settlement is no make or break event for libraries. Despite the claim that the Google Settlement promises to build “the greatest library in history,” libraries are not leaving the future of information to Google and these other partners alone. Nor need we wait, Godot-like, for fugitive national legislation to begin the work of serving up our cultural heritage in digital form. Through a combination of efforts, including public-private partnerships such as that of libraries with Google, we can go forward in this transformative enterprise together.

This piece was initially published in Anderson’s blog, at: http://www.cdlib.org/cdlibinfo/2010/02/16/hurting-toward-the-finish-line-should-the-google-books-settlement-be-approved.

Column Editor: Donna Jacobs (Research Specialist, Transgenic Mouse Core Facility, MUSC, Charleston, SC 29425) <jacobsdf@musc.edu>

Caldas, his name and the Caldas community are lost as the ones he can hear and see are unable to continue along its course. One by one his shipmates were lost at sea, only lost to the lure of the sea. I will pray for his safe return and that he will never be lost at sea, only lost to the lure of the sea.

Good-byes in his last days in Mobile and boards the ship. His “radar” continues to tell him that something is off as the ship feels funny under his feet. All of the American “gifts” that crew members have purchased for their families back home in Colombia are stowed on the deck and, thus, the seaworthiness of the Caldas is compromised. Luis along with eight other crew members find themselves struggling on the deck during a violently rough sea. A wave washes over the deck, and all are instantly swimming. They fear the boat has sunk, but as Luis realizes a life raft he sees the Caldas crest a wave and continue along its course. One by one his shipmates are lost as the ones he can hear and see are unable to make it to the raft in the boiling sea. For the next 10 days the reader shares his tiny life raft with no water, food, or protection from the grueling sun, all told in simple declarative sentences. García Márquez was the ghost writer for Luis’s recounting, and the story of this phenomenal event was published in 1955 as a series of newspaper articles in the El Espectador. The revelation that contraband was being transported on the military vessel was an embarrassment to the dictatorial government of General Gustavo Rojas Pinilla, which resulted in the closure of the newspaper and led to García Márquez’s “nomadic and somewhat nostalgic exile that in certain ways also resembles a drifting raft.” In 1970 when García Márquez wrote the forward to the book which compiled the series he mused that the publisher was more interested in García Márquez’s notoriety as a Nobel Laureate than the incredible story which he had the opportunity to recount for the world.

So now it is April, and the sea once again sings a siren’s song to the fishermen, sailors, and seafarers. My husband packs for his trip, and I have stashed my copy of The Story of the Shipwrecked Sailor in his bag. I will pray for his safe return and that he will never be lost at sea, only lost to the lure of the sea.

22 Against the Grain / June 2010