

---

# The Google Books Settlement: A Private Contract in the Absence of Adequate Copyright Law

Scholarly and Research  
Communication

VOLUME 2 / ISSUE 1 / 2011

Jenna Newman  
*Simon Fraser University*

## Abstract

Internet search giant Google Inc. began digitizing library collections in 2004, confident that scanning and indexing books to display excerpts based on users' search queries were fair uses under U.S. copyright law. Authors and publishers disagreed, and in 2005 representatives filed class action copyright infringement complaints. Rather than litigate, the parties negotiated a settlement that would not only allow Google's original uses, but license Google to use, and sell online, millions of books published before January 5, 2009. This report uses the experience of Canadian scholarly publisher the University of British Columbia Press to illuminate the technical details of the November 13, 2009, proposed amended settlement agreement, and it examines the settlement's economic and cultural costs and benefits and its implications for digital publishing, public access, and copyright law in a rapidly developing digital market. Whatever this settlement's outcome, its proposal underlines the need for meaningful, legislative copyright reform capable of encompassing present technological realities.

**Jenna Newman** is a graduate of the Master of Publishing program at Simon Fraser University, 515 West Hastings Street, Vancouver, BC V6B 5K3 Email: [jenna\\_newman@alumni.sfu.ca](mailto:jenna_newman@alumni.sfu.ca).

## Keywords

Google Books; Google settlement; Library Project; Copyright reform; UBC Press; Partner Program; Orphan works

CCSP Press  
*Scholarly and Research Communication*  
Volume 2, Issue 1, Article ID 010106, 75 pages  
Journal URL: [www.src-online.ca](http://www.src-online.ca)  
Received April 30, 2010, Accepted September 20, 2010, Published February 21, 2011

Newman, Jenna. (2011). The Google Books Settlement: A Private Contract in the Absence of Adequate Copyright Law. *Scholarly and Research Communication*, 2(1): 010106, 75 pp.

© 2011 Jenna Newman. This Open Access article is distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc-nd/2.5/ca>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Table of Contents

### Introduction 3

#### 1. A grand digitization project: Understanding the settlement 4

From scans to settlement: A brief history of the Google Books Library Project 5

Benefits of the settlement 9

Criticisms of the settlement 11

#### 2. Incorporating Google Books in a digital publishing strategy 17

Digital scholarly publishing and the library market 19

Opting in or opting out: Advantages and disadvantages under the settlement 20

A better deal with Google: The Partner Program 22

Combining the settlement terms with the Partner Program 23

#### 3. The practicalities of making a settlement claim: The case of UBC Press 24

Preparing publisher data 25

Technical logistics: A Google database how-to 28

Key settlement concepts for rights-holders 31

The meaning of “in print”: Setting licensed uses and setting rights 32

Next steps for UBC Press 33

#### 4. A steal of a deal: What the settlement offers the parties 35

For Google: More than fair 36

For publishers: As fair as it gets 38

For authors: Unevenly applied 40

#### 5. Copyright reform: Balancing access and control 42

Copyright in a digital world 42

The settlement’s para-legislative implications 44

Alternative digital collections and the public interest 46

Legislative solutions for legislative problems 50

### Conclusion 52

### Acknowledgments 54

### Notes 54

### References 62

### Appendix: Google Books search results, sample screen shots 67

## Introduction

The Google Books settlement has been hailed as an audacious and brilliant move by proponents and critics alike (Lessig, 2010; Samuelson, 2009a). Google's goal of digitizing up to 20 million books drawn from participating libraries has been recast to cut authors and publishers in on the deal. With one comprehensive and complex legal document, Google, the Authors Guild, and the Association of American Publishers have crafted a deal that could transform the digital marketplace for books and could give Google a legal—and exclusive—method to clear rights for some copyrighted works neither it nor anyone else could acquire any other way, excepting changes to U.S. copyright legislation. The following discussion considers the circumstances that led to this settlement and explores its primary components, focusing on the amended class action settlement agreement of November 13, 2009, which in many respects remains similar to the original agreement of October 28, 2008. The settlement makes positive steps in the tricky areas of public access and digital rights, but it remains open to serious legal, economic, and cultural criticisms.

In order to understand the practical effects the settlement may have on class members, it is instructive to consider how and why class members might make a claim. The University of British Columbia Press (UBC Press) is a Canadian publisher with numerous U.S. copyright interests (most of its works are published in Canada and subsequently released in the U.S.) and is a potential member of the settlement's publisher subclass. As such, UBC Press can serve as one example for other publishers and rights-holders, and its experience offers insight into the technical workings of the settlement and the possibilities for rights-holders to use Google Books to their own advantage.

In October 2009, I began to assemble UBC Press' settlement claim. The Press had already decided not to opt out of the settlement, but also to limit Google's uses of UBC Press books if possible. My work on behalf of the Press had two main aspects. First, I studied the settlement agreement to determine how the settlement would directly affect UBC Press and clarify what opportunities and vulnerabilities the settlement presented. A solid understanding of the settlement's terms will assist the Press in the ongoing management of its interests in Google Books. My research in this regard also included many of the responses to and criticisms of the settlement and the amended settlement that potential class members and other interested parties filed with the court, as well as commentary and analysis from academics, legal scholars, authors, and librarians. These varied perspectives supplement this report's focus on a publisher's perspective of the settlement. Second, I identified all UBC Press titles already included in Google Books and submitted an online claim to the settlement administrator on the Press' behalf. In testing this process, I also gathered technical information from the settlement administrator on how Google will recognize claims. Throughout, I considered whether the settlement offers a fair deal to those it affects; it is this report's central question. My primary examination is of the settlement's implications for publishers, particularly those—such as UBC Press—in the niche market of scholarly monographs. Many of these implications apply equally to authors and other rights-holders, but it should be noted that the author and publisher subclasses will have different experiences of the settlement, and it may be most appropriate to consider them separately. This discussion

acknowledges authors' concerns but focuses on a publisher's experience of the settlement. By the same token, the discussion explores the settlement's consequences for public access and the public interest, but it does not, in this space, exhaust these also-significant concerns.

Section 1 presents an overview of the settlement and the legal and technical issues it raises. Section 2 outlines UBC Press' position with respect to the settlement: the Press' strategy, some of its concerns, and the commercial motivations that informed its choices. This section also compares the settlement terms with the Partner Program, Google's other book digitization scheme.

The settlement administrator has created a website that includes settlement information and a database of Google's scanned books; this database will lay the foundation for the Book Rights Registry. All class members are encouraged to register on this site and identify works to which they hold rights. In section 3, I detail the work of assembling the claim and navigating the online claims process as it existed in early 2010, using UBC Press' claim as an example: other rights-holders contemplating the settlement will follow similar procedures. Preparing a claim requires some interpretation and analysis of the settlement's various technical terms, and I have highlighted those most significant in ensuring rights-holders understand how the settlement proposes to use their works.

There are other grounds on which to evaluate the settlement than the primarily legal and technical ones explored in the first section, and in section 4 I turn to questions of fairness, in both economic and practical senses, from the perspectives of Google and the settling publishers and authors. Finally, I undertake a deeper consideration of the settlement's relationship with and potential effects on existing U.S. copyright law. This privately negotiated contract could set precedents outside the law's current scope. Through this lens, I explore the settlement's treatment of the largest party, whose rights are perhaps least explicitly addressed by it: the public. The settlement is neither wholly good nor wholly bad, but perhaps above all else it exposes how wholly inadequate current copyright legislation is at addressing copyright's extension to a digital environment.

### **1. A grand digitization project: Understanding the settlement**

Since its incorporation in 1998, Google has grown to dominate the online search market: over 70% of the world's search queries start at Google's search box (Auletta, 2009). Its search engine's popularity and technical superiority allowed Google to monetize search through the development of now ubiquitous online text advertising. While search and ads remain the core of Google's activities and considerable success, it continues to research, develop, and acquire a number of other online products and services.

One such service is Google Book Search, which applies Google's search algorithms to a growing index of scanned book pages rather than just Web pages. Google Print, as it was first known, was initiated in the early 2000s; it was renamed Google Books in 2005. By 2004, Google had secured the participation of a number of publishers, who agreed to contribute their in-print books to the initiative. This publisher component of the project became what is now called the Partner Program; it grew steadily, if relatively slowly, to include approximately one million in-copyright titles by 2009.

Also in 2004, Google revealed the Library Project, which proposed to scan the entire collections of several leading libraries and add these scans to Google Books. The Library Project was poised to grow significantly larger (and more quickly) than the Partner Program. The libraries' collections encompassed out-of-copyright works as well, and they offered Google access to 15–20 million titles through a handful of library partners, as opposed to the thousands of publishers in the Partner Program, each with perhaps hundreds of titles. However, the Library Project attracted a flurry of controversy that ultimately led to a class action settlement before a New York court.

In order to understand this controversy, it is useful to consider the circumstances from which the Library Project and Google Books emerged, as well as the main features of the proposed amended settlement agreement, especially the changes that attempt to address criticisms of the initial proposal. The settlement offers advantages and disadvantages on several fronts, from its complex mechanics to its potential legal and cultural effects. Although the benefits are compelling, the criticisms remain significant.

#### **FROM SCANS TO SETTLEMENT:**

##### **A BRIEF HISTORY OF THE GOOGLE BOOKS LIBRARY PROJECT**

Google maintains that its development of Google Books was just another aspect of its mission to organize the world's information and thereby make the world a better place (Google Inc., n.d.).<sup>1</sup> Its plan to create a massive digital library would not only bring additional information online but also make Google's search engine more comprehensive. Since this book content would only be available to searches made through Google, the Library Project would, as a side effect, further advance Google's competitive advantage in the search market. Dan Clancy, engineering director of Google Book Search until his move to become engineering director of YouTube in summer 2010, offered another reason: an online collection also purported to bring books to where the readers were, presuming the Internet would supplant libraries or physical collections as students', and future researchers', resource of choice (Clancy quoted in Auletta, 2009). The Library Project was, in some ways, the natural extension of what Google had already done. Google founders Sergey Brin and Larry Page created the search engine that became Google by downloading the entire Web and analyzing it. All the world's books must have seemed like a similarly large, complex, and useful collection of data. It is perhaps not surprising that Google considered scanning and indexing books a fair use of readily available information, not copyright infringement. But unlike the Web—at least the Web as it was in 1996—books have a long history as intellectual property and a number of deeply invested business models and commercial relationships that depend on that property's copyright.

The book scanning initiative's technical challenges were to design a scanner as well as algorithms to de-warp images and render them usable, both to human eyes and indexing computers. Google needed books, so it sought library partners, and in 2004 Google announced that it would scan the collections from the Bodleian Library at Oxford University; the libraries at Harvard, Stanford, and the University of Michigan; and the New York Public Library, representing an estimated 15 million volumes or more.

The Library Project caused authors and publishers concern. Google proposed to scan library books and index their contents, but then only display short excerpts or

“snippets” in response to a user’s search query. Public domain works would be fully accessible and even downloadable, but for copyrighted material Google would only show bibliographic data and short excerpts. Rights-holders could request that Google not scan their works, but many saw the entire scheme as copyright infringement. In the fall of 2005, the Authors Guild of America and several authors filed a class action lawsuit against Google on the basis that the agreement with the libraries “entitles Google to reproduce and retain for its own commercial use a digital copy of the libraries’ archives” and that, in so doing, Google “has infringed, and continues to infringe, the electronic rights of the copyright holders of those works” (*Author’s Guild et al. v. Google Inc.*, 2005, p. 2). One month later, five publishers filed a similar complaint against Google, echoing the Authors Guild’s complaint that Google would scan, store, and display excerpts of books, as well as include those excerpts in its publicly available search results, all of which would “increas[e] the number of visitors to the google.com website and, in turn, Google’s already substantial advertising revenue” (*McGraw-Hill Companies, Inc., et al. v. Google Inc.*, 2005, p. 3); that is, Google would make commercial use of the copies.

For the next three years, authors, publishers, lawyers, legal scholars, and corporations and non-profit organizations with an interest in any of the complaints’ many facets argued over whether and how the case should proceed, including in their debates such issues as fair use, copyright law in digital spaces, the implications of a possible settlement, and competing models. Google continued to scan books, and by the end of 2007 it had 28 libraries participating in the Library Project, along with more than 10,000 publishers contributing to Google Book Search through the Partner Program (Google Inc., n.d.). Two years after that, the Partner Program had more than 20,000 publishers and authors participating (Google Inc., 2009). Google had responded to the two complaints in two submissions to the court in November 2005, but, as Pamela Samuelson points out, in 2010 the case remained “in relatively early stages as a litigation” (2010a, p. 1315 n. 37). Samuelson, a law and information professor at the University of California, Berkeley, suggests that Google continued scanning during settlement negotiations and pending the court ruling because it could reasonably assume the project would ultimately go forward.

It was the Authors Guild that proposed the idea of settlement. As Roy Blount (2008), then-president of the Authors Guild, said, “Our proposal to Google back in May 2006 was simple: while we don’t approve of your unauthorized scanning of our books and displaying snippets for profit, if you’re willing to do something far more ambitious and useful, and you’re willing to cut authors in for their fair share, then it would be our pleasure to work with you.” All parties agreed, and they announced the proposed settlement agreement on October 28, 2008, after two and a half years of negotiations.

Google’s Dan Clancy describes the motivation for seeking a settlement:

We strongly believe in our fair use position, but we didn’t start this project to win a court case on fair use. We started it to provide discovery tools. This settlement is an opportunity to do what, I think, from a user perspective is far better. The snippets we’ve been showing are a far cry from what the user

wants, and really the only solution was a partnership. We assume we would have gone through the courts and won. But once we won, we still would've had snippets. (Clancy quoted in Albanese & Oder, 2009)

The settlement offered the parties an opportunity to expand the terms of the discussion far beyond the initial complaints, and the resulting agreement covered a broad spectrum of uses by Google, libraries, and the public—although neither the libraries nor any entities representing the public interest had an explicit place at the negotiating table. It also called for the creation of an entirely new agency, the Book Rights Registry (BRR). This non-profit organization would act as intermediary between Google, as it applied its sweeping new licence, and the millions of rights-holders who hold a U.S. copyright interest in any work published on or before January 5, 2009.

The court granted preliminary approval to the proposal, and notices were distributed to the many people and entities affected. It was a complex agreement and, due to the U.S. class action mechanism combined with reciprocal rights for foreign nationals under the Berne Convention, it affected authors and publishers around the world. As uncertainty increased among putative class members, the deadlines and next steps were deferred from the spring of 2009 until the fall. The court received some 400 formal objections as to the legitimacy of the settlement, many from non-U.S. citizens and organizations who were shocked that a U.S. case could so broadly encompass their copyrights. Some rights-holders opted out; many more were simply confused, overwhelmed, and deeply uncertain about where their best interests lay as they tried to interpret the agreement. In September 2009, the United States Department of Justice (DOJ) submitted a statement of interest outlining serious concerns, primarily with respect to the settlement's potential failure to satisfy the technical requirements of a valid class action and its seeming inconsistency with antitrust law (United States Department of Justice, 2009). Rather than proceed with the fairness hearing scheduled for October 7, 2009, the parties requested more time to discuss with the DOJ and revise the agreement. An amended settlement was submitted on November 13 and granted preliminary approval on November 19, 2009, initiating another round of notice, a new deadline to opt in or out (January 28, 2010), and a timeline to submit statements to the court in advance of the new fairness hearing on February 18, 2010.<sup>2</sup>

The amended settlement included several key changes, primarily in response to the DOJ's concerns (Amended Settlement Agreement [ASA], 2009).<sup>3</sup> One of the major differences was in the definition of the class. While many books published outside of the U.S. would no longer be covered by the settlement—at least in part due to vehement opposition from France, Germany, and other European Union countries—works originating in Canada, the United Kingdom, and Australia were still included. (The publication cut-off date of January 5, 2009, was unchanged.) Rights-holders from these three countries would now be represented on the board of the Book Rights Registry, with an author director and a publisher director each (Amended Settlement Agreement, 2009).<sup>4</sup> A change to the definition of *commercially available* also addressed the treatment of non-U.S. rights-holders.<sup>5</sup> A book for sale new anywhere in the world (not just within the U.S.) to a buyer in any of the four participating countries would meet the criteria for commercial availability.

A second big change was with respect to the distribution of revenues generated through Google's use of unclaimed or "orphan" works. U.S. copyright extends to 70 years after the author's death; in practice, this means that a large number of works created since 1923 are still in copyright even though they may be long out of print and the rights-holders (or their heirs) are unknown or unlocatable. Because the settlement includes all rights-holders who do not specifically opt out, these orphan works will be swept into the agreement. It can be argued that an opt-out class action would be the most practical way to address licensing for orphan works; it would certainly be strategic and efficient. Even if comprehensive orphan-works legislation would better serve the public interest, legislative efforts in this area so far have stalled.

The initial agreement provided that after five years, the portion of revenues earned through Google's use of works for which no rights-holder had been found would be distributed among all the rights-holders currently registered with the BRR. Many alleged this created a conflict of interest between different class members, since keeping orphan works orphaned would benefit known and registered rights-holders, and therefore the rights-holders of orphan works were not fairly represented as members of the class (United States Department of Justice, 2009). The amendment created an unclaimed works fiduciary to represent unlocated rights-holders. Unclaimed funds would not be used to cover the Registry's own costs or shared among known rights-holders; they would be held for five years, after which the fiduciary could use a portion of the revenues to attempt to locate missing rights-holders. After 10 years, with fiduciary approval, the BRR could distribute the balance of unclaimed funds to charities that promote literacy.<sup>6</sup>

The amended agreement also limited the scope of forward-looking activities addressed by the settlement. Instead of an open-ended clause allowing for non-specified future uses to be developed by Google and the Book Rights Registry, the amended agreement narrowed its scope to three additional possible products or services: books in the collection might be purchased as print-on-demand copies, if they were no longer commercially available; as downloadable files, such as PDF or EPUB, for offline use; or through individual consumer subscription (Amended Settlement Agreement, 2009). The original settlement ensured that Google would always have at least as favourable terms as any the BRR might offer to other third parties. In the amended settlement, this "most favoured nation" clause was removed, allowing the Registry to negotiate new deals with third parties without regard for the terms it offered Google (Amended Settlement Agreement, 2009).<sup>7</sup> Creative Commons licences were explicitly allowed; other changes offered additional options around pricing, revenue splits, and other terms.

The DOJ, in its second statement of interest, acknowledges the amendment's positive changes but still believes the core problem remains: the settlement "is an attempt to use the class action mechanism to implement forward-looking business arrangements that go far beyond the dispute before the Court" (United States Department of Justice, 2010, p. 2). The DOJ is particularly concerned by the notice provided to class members, the settlement's creation of "author-publisher procedures" that attempt to impose a single template for rights and revenue allocations on individualized publishing contracts, and ongoing antitrust risks (such as those around price-fixing or access to orphan works). However, the DOJ seems to prefer working with the parties rather than dismissing the whole enterprise: after all, there is the potential for real public benefit, with Google

footing the bill, in an online collection. (Based on his line of questioning at the fairness hearing, it appears that the presiding judge, Denny Chin, is also interested in practical solutions that could resolve the settlement's possible shortcomings.) In this vein, the DOJ makes several suggestions: revise the settlement to make it opt-in, impose an embargo period before orphan works can be used in Google Books (to give time to seek out rights-holders before their rights are licensed on their behalf), limit the term of the orphan works' licences and evaluate their market at the close of that term, and consider other means for competitors to use orphan and "rights-uncertain" works (United States Department of Justice, 2010).<sup>8</sup> The parties have declined interest in an opt-in settlement or in opening another round of negotiated amendments, but their responses may change once the court has ruled on the settlement's fairness.

### **BENEFITS OF THE SETTLEMENT**

The settlement appears to offer a variety of benefits, both to its direct parties as well as the public. The most obvious—and to many, the most appealing—is the creation of a vast digital collection: the pooled results of many leading libraries, themselves the result in some cases of centuries of care and thoughtful stewardship. Old, rare, or long out-of-print books that are scanned by Google would become available for use to a much broader audience, since access would not be constrained by either the number of physical copies in existence or the number of readers who wish to make use of them. Research would not depend on a scholar's or student's proximity to the physical book's location: in an instant, any qualified user (someone who participates in an institutional subscription or searches from a designated public library terminal) could access a copy of any item from any participating library. Research libraries could share the combined resources of other libraries by purchasing an institutional subscription to Google's collection, which might be of particular benefit to institutions with small collections or those with straitened acquisitions budgets (though it should be noted that the settlement does not specify subscription rates, and these may command a significant portion of acquisitions budgets). In particular, Google Books as envisioned in the settlement would signal a renewed life for orphan works. With new copies unavailable, such books have likely been less accessible than those with a rights-holder to tend them. A reader looking for an orphan or out-of-print work has had to rely on existing accessible physical collections or the second-hand book market; while these limitations are not insurmountable, searches can be time-consuming and inconvenient.

These advantages would not be available only to scholars and students whose home institutions purchase subscriptions to Google Books, although this group of potential users would enjoy flexible and extensive access. Public library terminals would also offer full access to the collection to individual on-site users, and anyone with Internet access could search the collection online and review bibliographic data or read limited preview material. However, these benefits would only be available in the U.S., so for now the "universal library" would be far from truly widely accessible. This is an ironic limitation given how easily digital texts can render geography irrelevant, not to mention that Canadian, U.K., and Australian authors and publishers would contribute their works yet be unable to view or read them.

Despite the limitations on user access outlined in the settlement, Google Books still "has the potential to provide unprecedented public access," and, as such, the settlement

appears to align with libraries' "core mission...: providing patrons with access to information in all forms" (Association of College and Research Libraries, American Library Association, & Association of Research Libraries, 2009, p. 2). In the ideal, Google Books could be a conduit for public access to knowledge that is currently available primarily to specialists and scholars at select institutions. It could be the closest realization to date of the challenge articulated by John Willinsky (2006):

The public's right of access to this [scholarly research] knowledge is not something that people have to earn. It is grounded in a basic right to know. As online technologies appear capable of extending that right to a greater portion of research and scholarship, it falls to the scholarly community to experiment and test just how far such access can be pursued with new publishing models. (p. 125)

However, the settlement's potential realization of public access may be tempered by its shortcomings with respect to other "fundamental library values," including "equity, privacy, and intellectual freedom"—all qualities that could be seen as equally fundamental to the public interest (Association of College and Research Libraries et al., 2009, p. 6). Such shortcomings are discussed with the criticisms of the settlement below.

Through Google Books, publishers and authors would receive a new income stream without making a substantial direct investment of their own resources. The books covered by the settlement are mostly old books; by the time the settlement is implemented, even those published on January 5, 2009, may have drifted into the backlist.<sup>9</sup> The settlement offers ongoing incremental income from several possible sources: revenues from any ads displayed on book information pages in Google Books, fees from institutional subscriptions, and sales revenues from individuals' purchase of online access to specific books (Amended Settlement Agreement, 2009). Google offers a better split than publishers can get from most book retailers, with 63% of the revenues returning to the rights-holders (Amended Settlement Agreement, 2009).<sup>10</sup> Nor do publishers and authors have to worry about providing books to Google or technical resources to the project; Google will take care of it all.

The existence of the Book Rights Registry would create a copyright record of sorts: a list—one that aspires to be comprehensive—of who owns rights to what. Google's licence fees provide an economic incentive to encourage previously unlocatable rights-holders to come forward, and the BRR will be authorized to actively seek them out.<sup>11</sup> Michael Boni, representing the Authors Guild, alludes to the Authors Registry's 85% success rate at locating authors, and an over 90% success rate of comparable collecting societies in the U.K., to argue that the BRR will reduce the number of orphan works (Fairness Hearing, 2010).<sup>12</sup> Although on the surface such assertions sound encouraging, Boni fails to qualify his numbers, and it is unclear if these rates refer to all authors and all works, or some subset (such as known authors who agree to participate after the organization solicits them, or works with accrued fees above some minimum amount). Hadrian Katz, representing the Internet Archive, argues that the fact that Google insists the settlement must be opt-out, "in light of the prohibitive transaction costs of identifying and locating individual rights holders of these largely older out-of-print works," is itself an admission

that the people at Google “know very well they can’t find the rights holders” (Fairness Hearing, 2010, p. 94). There does seem to be an assumption that identifying and locating rights-holders will be easier for the BRR than it is now for Google. More to the point, the settlement allows Google to download these prohibitive transaction costs to the BRR, which is to say to the rights-holders themselves. If the BRR does manage to identify, engage, and register most rights-holders, the resulting record could facilitate the negotiation of other licences. It could even be of broad public use in implementing changes to copyright policy, such as shorter but renewable copyright terms. However, as of yet, there are no provisions to make the BRR database publicly available, nor is there an obvious choice of an entity to receive and manage this database as a public good.

Apart from its generation of a register, the settlement might serve the public good by advancing the purpose of copyright itself: according to the parties, it promotes the creation and distribution of books and other works, with economic incentives as well as access to an appealingly large audience (Google Inc., 2010a). Even as the settlement appears to contradict current international understandings of copyright laws—such as the Berne Convention’s presumption of protection and rejection of a registration process—the parties maintain that the settlement is consistent with the underlying law. Its offer of expanded access to information would seem to support the U.S. constitution’s assertion that copyright should promote the development of public knowledge; it meets the second principle in that it allows creators and rights-holders to earn compensation for their contributions. However, many objectors and critics would disagree with this interpretation, which elides the specific articulations and amendments of U.S. copyright law that followed that framing in the U.S. constitution.

Access to so many books would be a boon to researchers. The collection offers another immensely exciting tool for research simply as a digitized dataset. Alongside the collection available through Google Book Search, Google would host a potentially larger collection, the *research corpus*, which is only available for *non-consumptive research* and as such may contain scans of books that are excluded from *display uses* (Amended Settlement Agreement, 2009). Computational analysis of texts was previously possible on a relatively small scale, but here would be millions of titles spanning centuries of book production, all within easy reach of a computer processor. Any number of linguistic, rhetorical, historical, sociological, or other types of queries could be run against an impressive bulk of information on human culture and thought.

#### CRITICISMS OF THE SETTLEMENT

The settlement hinges on the class action mechanism, and yet its conception of that class may be problematic. Several critics question whether the named plaintiffs are representative of all class members, and even whether the author subclass and publisher subclass are meaningful labels for such a broad, numerous, and variously motivated population. The rights-holders of orphan works continue to be unrepresented. Authors of *inserts* (short fiction, essays, poems, songs) will enjoy more limited rights and benefits than authors of books; some say their interests diverge enough to warrant a separate subclass (Guthrie, Wright, Hyde, & Linden, 2010). Some potential class members may have had no work digitized by Google before the emergence of the settlement, and thus no apparent litigable claim against Google, but the settlement would then give Google permission to digitize their work in future—

at the time the settlement was proposed, Google had not yet scanned all the books published before 2009, but any rights-holders who do not opt out or make specific requests will automatically license Google to digitize, at its pleasure, all their pre-2009 works. “Superstar” authors, who may also have trademark interests in their names or have licensed works for other purposes, are not explicitly represented (Arato, 2009; Guthrie et al., 2010). Foreign rights-holders are treated differently, in that U.S. rights-holders who have not registered with the U.S. copyright office are excluded from the class, but non-U.S. rights-holders who have not registered their work are included. An objection from a group of Canadian authors points out that at least with respect to Canadians, the settlement may violate NAFTA’s requirement for national treatment of investors (which includes publishers and authors, since a book, as intellectual property that can be used for economic benefit, meets NAFTA’s definition of investment) (Bolt, Page, Payerle, Hancock, Levitt, et al., 2010). The notice of the class action must be adequately distributed, and some argue the notice was too confusing, did not reach everyone it needed to, and was not adequately available in translation (Arato, 2009; Raj, 2009). Scott Gant (2009), a lawyer and member of the author subclass, argues that seeking and locating rights-holders—which the settlement proposes that the BRR would do, sometime in the future—is work that should have been done by the parties during the notice period, not after the opt-out deadline.

Proper conception and treatment of the class are not the only requirements to ensure the courts recognize a class action’s validity. The settlement goes well beyond the scope of the original complaint (which concerned only scans and snippets), transforming the class action into a broad commercial transaction addressing full-text display and sales. This too may cast doubt on the legitimacy of the class action (Arato, 2009; Gant, 2009). The settlement provides for a range of future uses. It licenses Google to use the works it has already digitized, but also to scan additional works for similar use; it further gives Google the option to develop certain revenue models for the collection. In doing so, the settlement fails to limit itself to the past wrongs of alleged infringement and could be seen as a strategic route toward separate ends: Google’s desire for a comprehensive index and the plaintiffs’ attempt to earn new revenues while transforming print publications into digital ones.

Perhaps most galling, to some, is that the settlement circumvents existing copyright law and carves out an exclusive private deal for Google in “the formulation of [U.S.] copyright policy,” an area that the U.S. Supreme Court has in numerous prior instances acknowledged “lies solely in the hands of [the U.S.] Congress” (Yahoo! Inc., 2009, p. 2). For rights-holders outside the U.S., the proposal implicates them in a set of rules that diverge significantly from their existing rights and international copyright conventions, and in whose negotiation they had no voice. Pamela Samuelson stresses that class action creates private law, whose application is limited to a particular case and not of broad public use, whereas legislation provides clear guidelines that apply to anyone in similar situations. The implications are not trivial: “Use of a class action settlement to restructure markets and to reallocate intellectual property rights, particularly when it would give one firm a de facto monopoly to commercialize millions of books, is arguably corrosive of fundamental tenets of our democratic society” (2010a, p. 1358). Given the slow progress made to date in the U.S. Congress on orphan works legislation, it is not unreasonable to worry that the settlement may serve as a substitute for public legislation for the near future at least.

By default, the settlement confers membership in the class on authors and publishers unless they have explicitly opted out. Rights-holders who did not know they were affected or who missed the deadline will be bound by the settlement, as will those who intentionally ignored the deadlines as a form of protest. Works with no rights-holder to claim them will automatically be included. The settlement thus offers Google an effective monopoly on digital uses of unclaimed and orphan works; this, coupled with its significant head start on digitizing books at all, means Google will be a difficult—some say impossible—leader with which to compete. Google took a calculated risk and has continued to digitize even while the legalities are debated; the risk for competitors, even if they could also negotiate a licence of their own, is much higher since the existence of Google's settlement could cast doubt on others' claims of fair use. In addition, the BRR will only negotiate licences on behalf of registered rights-holders: the orphan works remain Google's alone until legislation re-writes the rules (Yahoo! Inc., 2009). Not only will there be insurmountable barriers to entry for would-be competitors, but a Google monopoly threatens the long-term viability of the digital collection for users. Academic libraries are painfully familiar with the ever-rising prices of journal subscriptions today, where large publishing conglomerates control key journals and can set aggressive terms of access—subscription prices, bundles—since scholars insist the libraries maintain their subscriptions. Given the likely appeal of Google's digital collection and the foreseeable lack of any comparable products in the market, libraries may find themselves vulnerable to price-gouging in the future, and the settlement offers no explicit protections against this possibility (Samuelson, 2010a).

Google's assurances that it will not seek only to maximize profits and that it has the public interest in mind are not explicit provisions in the settlement, but rely instead on Google's historical institutional culture. Peter Brantley (2009) of the Open Book Alliance is sceptical, but Paul Courant (2009), dean of libraries at the University of Michigan, suggests trust in Google is well placed. The settlement does state that one of its objectives in pricing the institutional subscriptions is "the realization of broad access to the Books by the public, including institutions of higher education" (Amended Settlement Agreement, 2009, s. 4.1[a][i]), but it does not include any review mechanism to determine whether this objective is met. Even if Google resists the temptation to maximize profits, future owners of Google or its collection may have differing views, and Google may assign its rights under the settlement to any successor without seeking the consent of the rights-holders (Amended Settlement Agreement, 2009). Though the BRR will still have the power to approve or negotiate the setting of the pricing strategy (Amended Settlement Agreement, 2009), the settlement does not describe how the BRR will enact or represent the wishes of the rights-holders, and some of the heterogeneous group of publishers and authors may themselves be tempted to set aggressive prices even while others would prefer to maximize access. Rather than leave such risks unaddressed, the settlement might impose certain limits or stipulations, or at least provide for a review mechanism.

Similarly, though Google must provide certain minimum services within five years, there are no provisions for the future of the collection of scans should Google abandon the project for any reason (Amended Settlement Agreement, 2009). Samuelson (2010a) outlines a number of reasons the collection could become unavailable: technical glitches or malicious hackers could disrupt or compromise the collection; subscription

sales may be too low to justify continuing the service (if large numbers of rights-holders exclude their works, or if poor scans and metadata erode scholarly usefulness). The International Federation of Library Associations (2009) suggests Google might discard sections of the collection in order to render the Google Books service more cost effective. Several commentators suggest that Google may simply abandon a project in which it loses interest—perhaps hard to imagine in an undertaking as big and seemingly dear to Brin and Page as this one, but not without precedent.<sup>13</sup> Although Google expended its own private funds to build the digital collection, it also benefited from the publicly funded collections of its library partners in doing so and stands to benefit from the class action's intervention in existing copyright law, if the settlement is approved. Google's collaboration with libraries might suggest that the initiative is not merely a commercial venture in the digital market, but in its scale, scope, and content creeps toward public policy's domain, even if only to underline existing policy gaps around public digital libraries. It would seem a waste if some future decision by Google to not maintain or host the collection consigned usable scans to oblivion, although if in the future government or public entities want to pursue public digitization projects, most scans will also reside with the various participating libraries.

Concerns for the robustness of the collection as well as concerns that Google may not meet the obligations of cultural stewardship could be mitigated if the court proposed the scans be publicly owned and leased to Google. Google would certainly protest—no doubt on the grounds, ironically, that the scans are its own productions and property. In the U.S., more so than in Canada, such state intervention is viewed with suspicion, although antitrust measures have, if not divested U.S. commercial interests, at least limited and dispersed them. Regulatory interventions could effectively support other digital collections, particularly since individual research libraries will each hold portions of Google's collection. Under current copyright laws, participating libraries would be unable to pool their digital collections or make them publicly available. Orphan works legislation, however, might allow more flexible uses of some scans by entities other than Google.

Library associations—the Association of College and Research Libraries, American Library Association, and Association of Research Libraries—support approval of the settlement, but point out that the settlement stipulates “the Court shall retain jurisdiction over the interpretation and implementation of this Amended Settlement Agreement” (Amended Settlement Agreement, 2009, s. 17.23); in other words, “the parties acknowledge this Court's authority to regulate their conduct under the Settlement” (Association of College and Research Libraries et al., 2009, p. 19). The library associations suggest several ways in which users of Google Books, subscribers, researchers, or class members could appeal to the court to review decisions of Google and the BRR or direct those entities to provide additional information (Association of College and Research Libraries et al., 2009).

Some kind of additional regulation might be required if Google is found to be in a position of monopoly, but the settlement as proposed does not impose conditions to prevent monopoly nor stimulate competition, nor does it protect the public interest. To expect it to do so is perhaps unreasonable: it is a contract between private parties. Though its outcomes will be subject to antitrust or competition law, consumer protection law, and privacy law just as any other enterprise, it might be argued that any measures of prevention should

limit themselves to actual violations rather than speculative or potential ones. William Cavanaugh notes that the DOJ antitrust department continues to investigate Google's position in various markets: the new digital book market this settlement develops as well as the search market (Fairness Hearing, 2010). The DOJ, if not the court, will have to determine whether the settlement creates a monopoly. Not only will Google be the only entity with legal access to orphan works, but the settlement gives Google the power to simulate the competitive market and determine appropriate prices for books sold through the consumer purchase revenue model (Amended Settlement Agreement, 2009). In addition, the Registry has the power to negotiate additional licences on behalf of all rights-holders as a group. Both these features could be seen as forms of price-fixing, which is illegal in the U.S. because it suggests that a monopoly, or at least collusion, is in place. The BRR to some extent resembles collecting societies such as Canada's Access Copyright. In the U.S., the Copyright Clearance Center serves similar purposes, although unlike the BRR it requires rights-holders to enrol for its services before it collects revenues for their work; it also negotiates different licences for particular uses rather than one licence for all uses.

Even with the amended settlement's addition of reseller provisions, Google's effective monopoly may endure, since resellers will not control access to the products but will in essence refer customers to Google (Amazon.com, 2010; Amended Settlement Agreement, 2009; Yahoo! Inc., 2009). Law professor James Grimmelmann describes these resellers as "little more than franchisees," since even though they will receive the majority of Google's 37% revenue share, they will have no role in setting prices or other terms of access and "do not create any structural competition" (2010, p. 4). However, telecommunications providers, which are regulated as common carriers, offer a useful parallel here. Third parties are able to resell an identical base product, but might compete on price (or, more accurately, on tolerances for profit margin), additional bundled features, customer service, or ease of use. Resellers of digital books could thus present competition even if they could not alter the fundamental mode of access to Google's collection. Again, this competition would only emerge with some sort of regulatory mechanism to prevent an unchecked monopoly.

The settlement contains no protection for user privacy, even though Google must track all page views of all book content and Google must be able to trace uses back to individuals in order to ensure only licensed uses of the works are made. Even if anonymized, users' search queries are easy to associate with identifiable individuals (Carr, 2008).<sup>14</sup> Many wonder if Google's oversight will lead to a chilling effect on reading (Chabon, Lethem, Ferlinghetti, Romero, Schneier, et al., 2009; Samuelson, 2010a). Marc Rotenberg, representing the Electronic Privacy Information Center, argues that the settlement runs radically against the technical community's efforts to minimize privacy risks, and given that Google "already knows more about internet users than any other company in the world, [and] for its business model relies on the commercial extraction of that information" in order to target ads and search results, the settlement's non-specific and non-binding privacy assurances are completely inadequate (Fairness Hearing, 2010, p. 87). Rotenberg suggests that privacy must be addressed in the design of the technology. In its initial release of Gmail and its recent launch of Buzz, Google has shown that privacy issues may not figure into its design specifications except as afterthoughts.<sup>15</sup> In contrast, libraries are careful not to monitor or disclose personally identifiable patron information. The library associations argue

that the settlement's effective silence on user privacy "stands in stark contrast" to the detail it accords measures to protect the security of digital book files (Association of College and Research Libraries et al., 2009, p. 12). If no competitor emerges to offer service alternatives, Google Books will have little incentive to improve its user protection services (Association of College and Research Libraries et al., 2009).

Some academic authors criticize the settlement for not going far enough in providing public access and claim that the settlement privileges economic interests even if an author's primary interest is public dissemination (Samuelson, Aoki, Armstrong, Auslander, Azzam, et al., 2010; Turk, 2010). Although many publishers (and non-academic authors) might disagree, it is interesting to note that the settlement can fail to satisfy class members on such diverse grounds. The plaintiffs affirm the fundamental commercial nature of the transaction when they attempt to dismiss arguments for open access: "That the reading public may wish to have free access to scientific and other academic works covered by the ASA, or that some academic authors may not wish to exploit their works through the Revenue Models, should not supersede the economic interests of members of the Class" (Authors Guild, Inc., Association of American Publishers, Inc., et al., 2010, p. 94).

Samuelson (2010a) also outlines that readers' fair use rights—for example, copying excerpts or short works for personal or educational use—will be weakened if public access terminals in libraries charge printing fees; in addition, first-sale rights lose ground, since the purchaser of an online Google Book will not be able to lend, gift, or resell it. The settlement allows Google to exclude up to 15% of its scanned books, which could be over a million titles, "for editorial or non-editorial reasons." Editorial reasons are not well defined, though they would be something other than "quality, user experience, legal, or other non-editorial reasons" (Amended Settlement Agreement, s. 3.7[e]). Google must inform the Registry of its omissions, but the settlement does not appear to offer any recourse for appeal or justification. The library associations worry that such omissions will dampen intellectual freedom: "Although public libraries have often contended with demands to eliminate or restrict access to specific books, any collection management decision by a particular librarian affected only that community. Here, by contrast, if Google bends to political pressure to remove a book, it will suppress access to the book throughout the entire country" (Association of College and Research Libraries et al., 2009, p. 15). Intellectual freedom may also be hampered by the requirement that researchers submit their qualifications and research proposals for approval in advance of gaining access to the research corpus for non-consumptive research (Amended Settlement Agreement, 2009; Association of College and Research Libraries et al., p. 16); again, there is no mechanism for appealing Google's or the corpus host's decision. Despite Google's and the plaintiffs' assertion that freedom of expression is important, Google has bowed to government pressure to filter or otherwise control content before. Between the potential for censorship and the absence of privacy protections, along with the lack of comparable alternative digital collections, the settlement offers Google an impressive tool for information control.<sup>16</sup> Even if Google declines to use this power (though its search and ad businesses would benefit from it), others may pursue it; regardless, the settlement again offers weak or no checks to limit this dizzying possibility.

The concerns the settlement arouses are varied. Legal criticisms include whether the settlement is technically valid as a class action and whether it is consistent with, or at

least a reasonable development of, copyright law. Socio-economic criticisms include that the settlement creates a monopoly for Google and gives it unfair access to an otherwise unobtainable market: orphan works. A third broad category of criticisms relates to public risks and public benefits: these are “cultural” or moral criticisms, such as protection of the integrity of the collection, the role or power of libraries, access, user privacy, and intellectual freedom. Many objectors, including the DOJ, focus on the class action validity and antitrust arguments. Since the settlement proposes new rules for rights negotiations that will shape public access in an emerging digital market, it may be entirely reasonable for the court also to consider the settlement’s cultural effects in determining its fairness. Moreover, many class members, especially authors, share these concerns.

On the other hand, the cultural criticisms or appeals to the public interest may seem to impose unreasonable expectations in some respects: after all, Google has spent hundreds of millions of private dollars to create the scans and will spend more to maintain the digitized collection, and the settlement contemplates a private contract. If there is a public policy need for a public digital library initiative, government agencies and legislative bodies can formulate a plan to address it, including allocating public funds to build and maintain it. The existence of a private contract between Google and publishers and authors does not prevent a public digital library initiative in parallel. (Other examples of public digital libraries are discussed briefly in section 5.) The potential advantages of such an extensive digital library are not available to anyone now, and if the benefits of Google Books are not universally or equally available to all users in all circumstances, surely the offer is still better than the current status quo.

Yet in order for the settlement to proceed, the parties are asking for what appear to be certain exceptions to existing laws and identified intellectual property rights (notably, default permission to use copyrighted works unless permission is denied). The court has an opportunity and obligation to review the results of such exceptions or new interpretations, including the far-reaching effects the settlement might have on the sociocultural environment for public inquiry and knowledge. These aspects are in fact raised by the parties themselves as arguments in favour of approval. Google invokes the public benefit of its collection and labels it a library (not a proprietary database or other commercial service): Google claims that “approval of the settlement will open the virtual doors to the greatest library in history” (Google Inc., 2010a, p. 1), and it describes the settlement as “undoubtedly extraordinarily good for all class members and for the general public” (p. 2). The plaintiffs characterize the results of the settlement in a similar fashion: “The benefit to society of enabling anyone, anywhere in America, from the inner city to a rural village, to walk into a public library, sit at a computer terminal and have full Internet access to a multi-million book database is inestimable. No longer would the world’s great literature and vast research resources be kept from the public” (Authors Guild, Inc., Association of American Publishers, Inc., et al., 2010, p. 8). In this light, it is appropriate for the court not only to consider the legal and economic arguments, but to scrutinize just how well, and at what costs, these public benefits might emerge.

## **2. Incorporating Google Books in a digital publishing strategy**

Canadian publishers who hold U.S. copyright interests, as many of them do, have the option to participate in the Google Books settlement as members of the publisher subclass. Publishers (and authors) could choose to opt out of the settlement, in which

case they would not receive any compensation through the settlement, they would retain the right to sue Google for infringement, and Google could, but is not legally obligated to, remove their works from the digital collection. Rights-holders who remain in the settlement class have a second set of choices: they may *remove* (a permanent designation) or *exclude* (a reversible one) their works from the collection, either of which would prevent those works from accruing revenues under the settlement; they may also *include* their works in some or all of the settlement's proposed uses and receive a portion of revenues for those uses only. Removal, exclusion, and inclusion—all distinct technical terms in the settlement—are discussed further in section 3.

A publisher that has not opted out may consider various factors in determining its strategies for participating in the settlement. Small publishers and individual authors may believe their strategic choices with Google are limited; large publishers might intentionally restrict their participation in the settlement's schemes and pursue individual parallel agreements with Google directly. In fact, all publishers, and presumably authors too, might obtain a more tailored relationship with Google through Google's own Partner Program. Although the settlement appears to offer one approach for all books and all rights-holders, there remain methods to affect or control Google's licensed uses from within the settlement and in parallel to it. UBC Press has taken advantage of such options.

Canadian rights-holders may have additional distinct concerns with respect to the settlement's treatment of Canadians and Canadian laws. Moral rights in Canada have a more robust tradition than in the U.S. and include protection of a work's integrity. The Canadian Association of University Teachers (CAUT) points out that Google Books proposes to associate authors' work with commercial advertising they cannot choose, which may be inconsistent with their moral rights (Turk, 2010). The argument is stronger for authors, but relevant for publishers. CAUT also points to Canada's existing provisions for unlocatable rights-holders, a model of diligent search and clearance through the Canadian Copyright Board and Access Copyright. The settlement does not refer to this Canadian scheme, but instead develops its own additional layer of bureaucracy—the Book Rights Registry—mostly to the benefit of Google, not necessarily rights-holders. CAUT argues that the Copyright Board “mechanism is flexible, [and] likelier to identify and compensate” Canadian rights-holders (Turk, 2010, p. 6). This may be an overly optimistic assertion, as the transaction costs of the Copyright Board scheme are still high and the Canadian system has not yet been used on the scale of Google's mass digitization project. The settlement's implied conflation of Canadian and U.S. copyright law is problematic, but at the same time it is difficult to separate the commercial circumstances of Canadian and U.S. publishers. Canadian publishers are in many ways already at a disadvantage in reaching the large U.S. market, and the settlement offers a way for U.S. readers to find and purchase access to Canadian books. However, the lack of competition in the U.S.—should Google Books remain an effective monopoly—and the fact that Canadian representatives did not participate in the negotiation of the settlement but Canadian rights-holders may be bound by it could undermine Canadian publishers' ability to develop or exploit other channels to reach the U.S. digital book market. Although concern over the presence of commercial advertising was not a key issue for UBC Press, the economic costs of additional bureaucracies and the settlement's possible effects on emerging market opportunities were factors in determining UBC Press' approach to the settlement.

## DIGITAL SCHOLARLY PUBLISHING AND THE LIBRARY MARKET

The commercial initiatives of the settlement present certain advantages to readers and libraries, and other separate advantages to authors and publishers. In addition, university presses may have different concerns with respect to the proposed settlement than other publishers or rights-holders. Scholarly publishers have often obtained a transfer of all copyright interests from their academic authors. This may change the way these publishers and authors understand who holds the rights for digitized versions of the books, and a scholarly press may have a stronger contractual claim to digital rights than a trade publisher even if digital rights are not specifically mentioned in the contract.<sup>17</sup>

The original library partners were mostly academic libraries, and their collections include many more in-copyright books written by academics and published by scholarly presses than by non-academic authors or published by trade publishers. The original complaints against Google, therefore, directly implicated many scholarly presses and academic authors; however, the settlement was negotiated by more commercially minded elements of the publishing industry and may not represent scholarly interests. The research library market is particularly important for a scholarly press' sales. Though the settlement focuses on older books, the limited commercial markets for scholarly books and the services Google offers the libraries could seriously alter the landscape of this market.

The settlement also allows Google to share scans with those libraries that provided books. Scholarly publishers, and perhaps all publishers, may be interested in limiting the participating libraries' permitted uses of these digital copies. The settlement imposes restrictions on such use, but publishers may see the conversion or supplementation of libraries' physical collections by digital ones as a market opportunity they do not want to share with Google. As copyright law developed, many of its rules never applied to libraries. Libraries preserve and lend the works they hold—activities that could include making and distributing copies—in order to advance public knowledge and learning, and libraries fulfilled this role long before the legal concept of copyright was articulated. Legislation's temporary and exclusive grant of copyright protection to authors and publishers does not wholly govern libraries' use of their works, in part because libraries work in the service of the greater public good and do not commercially exploit their collections. Copyright's typical restrictions are limited so that libraries may freely share and lend the tangible books and records in their collections, but exactly how widely they may do so has itself been limited by practicalities such as the number and location of books. When books are physical objects, library uses remain easily constrained; once books are digitized, lending acquires a different character.

Some might argue that libraries' collaboration with Google to digitize their collections oversteps their historical sharing privileges, since the libraries facilitated Google's reproduction of books for purposes other than merely the libraries' preservation of their own collections. In this interpretation, the settlement could be seen as inconsistent with the underlying copyright laws—specifically, the provisions in Title 17 of the *United States Code* (United States, 2009), section 108, on reproduction by libraries and archives, and section 107, on fair use—which could support some objectors' insistence that the settlement's proposal is properly under the jurisdiction of U.S. Congress, not the court (Amazon.com, 2010; Yahoo! Inc., 2009). Regardless,

scholarly publishers may see the settlement's provisions for libraries as being in direct conflict with their own ambitions to sell e-books or other forms of digital works to libraries or to their natural audience of scholars and students, many of whom will be affiliated with organizations targeted by the institutional subscriptions.

Although some scholarly publishers, more commonly those in the U.S. than in Canada, include trade books in their lists, the core of most scholarly presses' output is scholarly monographs and journals. A significant portion of these titles' sales are to libraries, and as presses have seen overall sales decline over the last few decades, the library market is a key component to the presses' financial survival. Many libraries also face difficult economic circumstances: acquisitions budgets are cut by parent institutions, yet the number and prices of scholarly publications increase. The imperative for both libraries and publishers is to reduce their expenses; unlike libraries, publishers must also generate revenues. Libraries' mandates generally support public access to information, and the settlement seems to advance this cause (although, as discussed above, the settlement's positions on patron privacy and intellectual freedom give some libraries serious reservations about the settlement's offer). Providing patrons with digital versions of libraries' own physical collections seems a natural development of and response to the desires and usage patterns of patrons, and a subscription to Google Books promises a more cost-effective means of "converting" the physical collection than purchasing e-book editions from publishers or wholesalers, especially since Google Books includes items not necessarily in each library's physical collection to begin with. Publishers may lose other backlist sales: the subscriptions will affect how many potential readers are served by each copy of the book (or rather, by each copy's infinite reproducibility). Scholars, researchers, and students may find their needs completely met by the on-call, digital availability of Google Books at their libraries. Individual sales and the previously guaranteed strong sales from a text's course adoption could suffer, from the perspective of the publisher, even if overall readership increases.

Academic authors may be less concerned by this library access than their publishers, as many of these authors primarily want their work to be widely available. Digital publishing promises to maximize reach, and Google Books, poised to become the single largest digital publisher but imposing no exclusivity requirements, may become the natural choice for broad dissemination. Scholarly publishers have been the preferred distribution channel for academic authors, but as online distribution and other options proliferate, this may no longer be the case; for the moment, the imprimatur of the university press is still a critical element. Google Books could threaten publishers' traditional dominance of knowledge distribution, but the settlement creates a distribution platform that offers publishers a way to solidify some role in online distribution. The settlement's economic incentives may draw publishers in even as Google's practices dismay them.

#### **OPTING IN OR OPTING OUT:**

##### **ADVANTAGES AND DISADVANTAGES UNDER THE SETTLEMENT**

UBC Press was never pleased with Google's unauthorized digitization through the Library Project and did not necessarily want to contribute to the Library Project or Google Book Search under the settlement terms. At the same time, walking away from the deal by opting out did not seem to serve the Press' best interests. UBC Press had remained a member

of the publisher subclass under the original settlement in September 2009, and it did not revise that choice in January 2010: the Press did not opt out and will be bound by the settlement, if it is approved. There are several factors that led the Press to this choice.

UBC Press' strategy was to optimize its own position under the settlement while minimizing Google's ability to commercialize intellectual property in which the Press had an interest and to which the Press felt Google had no rights. The settlement offers two commercial incentives to rights-holders: *cash payment*, a minimum U.S.\$60 one-time compensation for any works that were digitized without authorization before May 5, 2009, and ongoing income through three *revenue models*, which include ad revenues associated with display uses, institutional subscription fees, and sales revenues from consumer purchases. The cash payments absolve Google of legal challenge of its previous digitizing acts. The revenue models generate more revenues for Google than for publishers.<sup>18</sup> On principle, UBC Press did not want to allow Google to make display uses of scans it obtained without permission, and the *access uses* (subscriptions and purchases) could compete with the Press' own library and e-book sales. The Press decided to remain in the settlement and collect the cash payments, but remove or exclude its works from the revenue models. This approach prevents Google from making its own uses of the scans and retains for UBC Press the possibility to pursue library and e-book markets without going through Google and the Registry.

In this light, with UBC Press disinclined to pursue the residual income offered under the revenue models, opting out might appear reasonable: if UBC Press would not use the settlement to collect revenues from access uses, then perhaps it should save itself the bother of going through the claims process at all. The cash payments alone are not large sums, with the publisher's share only 50% or less of whatever amount the Registry distributes for each book.<sup>19</sup> However, although the work to assemble and manage a claim is not trivial, opting out does not relieve a publisher of much of that work, as it must still forward to the settlement administrator a complete list of all books that the opt-out covers. There is no economic savings in opting out.

Opting out would allow UBC Press to retain the right to sue Google, perhaps under some future class action, but there is no indication that another collective suit will emerge, and it is not at all feasible for UBC Press to sue on its own.<sup>20</sup> Grace Westcott (2009), a Toronto lawyer specializing in copyright, puts the choice in stark terms: "[F]rankly, it is hard to see the benefit in opting out. An individual rights-holder is unlikely to sue on his own; he forgoes any settlement amounts he's otherwise due, and, given the dominance the Google Settlement will likely have in the digital marketplace, it seems unwise to count oneself out." Google has far more readers than any publisher does, and individual dissenting publishers may ultimately hold little bargaining power.

Finally, the settlement does not impose a legal obligation on Google to remove the works of those who opt out, though "it is Google's current policy to voluntarily honor such requests, if the Books or Inserts are individually specified, are in copyright, and the author or publisher has a valid and unchallenged copyright interest in their Books and Inserts" (Book Rights Registry, 2008b). As Savitha Thampi, associate legal counsel at Access Copyright, pointed out in a December 10, 2009, teleconference on the Google settlement, remaining in the settlement may offer more legal protections than opting

out. For all these reasons, UBC Press did not opt out, though it will restrict Google's use of its works. The Press has deferred the final decision to remove or exclude works and will consider both the court ruling and subsequent developments to the settlement terms, as well as authors' wishes, in making those decisions.

#### A BETTER DEAL WITH GOOGLE: THE PARTNER PROGRAM

For UBC Press, the primary appeal of Google Books is that it makes books discoverable to potential readers where those readers are already searching for information. UBC Press, like many publishers, already has a digitization agreement with Google: in 2005, the Press began submitting books to Google Book Search under the Partner Program, which offers many of the benefits the settlement does, but further limits uses by Google. Each book is included at the publisher's initiation. Publishers, not libraries, submit books to Google, which scans them and makes them discoverable in Google Books and through its search engine. For each book, Google generates an overview page with basic bibliographic data, links to related editions, a brief description, and algorithmically derived representations of the book's content (a tag cloud showing frequently used words, for example). The publisher can control the percentage of content available in a limited, non-continuous preview. Links to buy the book include the publisher's own link first. Information about the book, including the preview settings, can be changed at any time through the publisher's online account. The text ads that Google runs on pages devoted to a single book generate revenues for the publisher, and therefore the author, of that book. Google receives no additional default licensed uses as it does under the settlement. UBC Press has been gradually submitting to the Partner Program all the books to which it holds rights, unless an author has requested her book not be included in Google Books. Ad revenues are minuscule (at least for most scholarly titles), but Google Books does deliver traffic and sales leads, which result in revenues for the Press and its authors.

In terms of a publisher's ability to commercialize its books, the Partner Program is comparable to the settlement and offers the publisher more control more easily. A would-be reader can browse a book, but is directed to purchase the publisher's editions of that book rather than purchase personal access to Google's online version. When the publisher signs in to its Partner account, it can review a range of analytics on user data, page views, and click-throughs. The publisher can generate reports of ad revenues by time period or by book and can drill into this data as it pleases. This level of detail will not be available under the terms of the settlement, since Google will deal with the Book Rights Registry rather than individual rights-holders. The Registry will be able to review and audit Google's information to some extent, but the settlement does not provide for publishers and authors to obtain detailed reports from either the Registry or Google (Amended Settlement Agreement, 2009). Under the Partner Program, publishers receive ad revenues directly from Google; the settlement specifies that Google forward all revenues to the BRR, from which the Registry will first deduct its own operating costs and then remit the balance remaining to rights-holders.

The Partner Program sells neither full-collection access by subscription nor full access to individual books by consumer purchase, but these are two settlement benefits in which UBC Press is not currently interested. The institutional subscription targets an important segment of the market for UBC Press titles—academic libraries and research

institutes—but arguably this is a market UBC Press is already able to reach well. It has relationships with library wholesalers and academic e-book wholesalers, who may negotiate individual sales or bundle purchases with university libraries. A number of UBC Press books were included in the Canadian Research Knowledge Network’s bulk e-book purchase for Canadian libraries; similar deals with other libraries could emerge. In contrast, Google’s institutional subscription may ultimately offer the Press low revenues that cancel out any potential to reach a wider market. In terms of consumer purchase, UBC Press titles tend to be niche books that appeal to a small, focused audience. While reaching that audience online is certainly to UBC Press’ advantage, the Partner Program already ensures titles are visible in Google Books but does not determine e-book pricing or sales strategies. The Press’ specialized titles may command above-average prices, even as e-books; as reference works with long shelf lives, they may flourish differently as electronic or print versions.

Both the Partner Program and the Library Project feed content into Google Books. Users search both collections simultaneously from a single interface.<sup>21</sup> The agreement through which Google received a book governs how Google displays the book’s content, but for most practical purposes a user—in the U.S. at least—will not be aware of how Google obtained the book. Outside the U.S., this will be much clearer. The settlement only deals with, and offers display or access uses in, the U.S. The Partner Program lets Google display book material in any geographic regions for which publishers hold rights, and thus it reaches a broader audience among all Google search users. UBC Press’ audience is not limited to Canada, but it is also not limited to the U.S.; the settlement is.

#### **COMBINING THE SETTLEMENT TERMS WITH THE PARTNER PROGRAM**

UBC Press has participated in the Partner Program since March 2005; over 20,000 other publishers participate as well (Google Inc., 2009). Under the settlement, books covered by separate agreements with Google will be subject to those separate agreements’ terms. For publishers who are also Google Partners, opting in to the settlement may still exclude some books published before January 5, 2009. Any books in the Partner Program will remain in the Partner Program, which will take precedence over the terms of the settlement, unless the rights-holder requests to transfer the books to an alternative agreement.<sup>22</sup> As of 2010, the majority of UBC Press’ titles are included in the Partner Program, which means that even before the Press removes or excludes its list, most of its titles should not be subject to the settlement’s access and display uses anyway.

In 2005, UBC Press agreed with other publishers that the Library Project was a breach of copyright. In a letter to Sheryl Sandberg of Google, Peter Milroy, director of UBC Press, instructed Google not to scan any UBC Press books for the Library Project. Milroy drew a distinction between the books provided under the Partner Program—instances in which UBC Press granted permission for limited preview online only—and the larger pool of titles that Google might draw into the Library Project. An attachment to the letter lists 435 titles already submitted to or then underway in the Partner Program, and Milroy’s letter prohibits Google from including these or any other books with UBC Press’ ISBN prefix in the Library Project (R. Peter Milroy, Director, UBC Press, Vancouver, BC, letter to Sheryl Sandberg, Google Inc., October 27, 2005). It is unclear whether Google respected this request: of the titles Google shows as digitized without authorization, some are also “live” in the Partner Program while some are not.

The Partner Program does not have an institutional subscription component, but although inclusion in the institutional subscription set could increase the public availability of UBC Press book contents to U.S. readers, the commercial advantage this inclusion offers UBC Press as a publisher is less obvious. UBC Press is still developing its own e-book marketing and distribution strategies, and bundled backlist packages targeting academic libraries could be part of them. The Press is loath to allow Google to broker, and profit from, selling access to these books; subscription revenues to individual publishers, while difficult to predict, are likely to be modest.<sup>23</sup>

The Press already derives ad revenues and generates sales leads through the Partner Program; whatever the scope of the Press' participation in the settlement, the Partner Program terms will govern access uses and any advertising that would accompany display uses for most books. These terms include limited preview uses for individuals similar to the settlement's standard preview. However, the other compensation mechanism of the settlement, the U.S.\$60 cash payments that make up for Google's unauthorized digitization, is not affected by the Partner Program agreement. UBC Press is still entitled to cash payments for any eligible books regardless of whether it also submitted a copy of that book at another time to Google for scanning.

It is therefore to UBC Press' advantage to remain in the settlement class *and* participate in the Partner Program. Opting in to the settlement provides UBC Press with a better opportunity to legally restrict Google from making use of its titles under the settlement's blanket licence. While the Press' Partner profile lists a general instruction to exclude titles with an 07748 prefix from the Library Project, claims under the settlement include a specific instruction for the exclusion or removal of each claimed ISBN. UBC Press thus obtains control and flexibility in monetizing digital publishing—Google will not manage its digital sales—but retains all the benefits of discoverability in Google Books.

### **3. The practicalities of making a settlement claim: The case of UBC Press**

Publishers and authors who have not opted out of the settlement have several ways to establish claims to their books with the Book Rights Registry, but completing a claim is the only way to obtain any share of the settlement payments or the revenues Google generates through the proposed licences. As part of the settlement, Google must pay at least U.S.\$45 million in cash payments. Another U.S.\$34.5 million, paid into a settlement fund, will cover administrative costs related to the settlement, including the notice, claims administration, and establishment of the Book Rights Registry (Amended Settlement Agreement, 2009).<sup>24</sup> Rights-holders must register claims by March 31, 2011, in order to receive cash payments. Funds generated through the revenue models can be claimed up to 10 years from the period in which they were earned.

Rust Consulting has been appointed settlement administrator. It manages settlement and claim information at <http://www.googlebooksettlement.com>, which includes a searchable database of all the books Google has or thinks it has as well as other settlement resources.<sup>25</sup> Rights-holders must create an account and can then claim items from this database online. In response to numerous complaints that the online process is too complicated, the Registry also allows rights-holders to simply submit a list of titles with as much bibliographic and other identifying information as is available; the settlement administrator will record these claims in the database on rights-holders'

behalf. However, submitting a claim using the online claims process allows rights-holders better control of their claim as well as more immediate feedback as to the claim's status and subsequent management.

UBC Press' claim highlights the main steps required to register with the settlement administrator, as well as what will likely be challenges common to other rights-holders. To complete a claim, any rights-holder must compare her own list of titles with those identifiable in Google's database, locate and resolve missing titles or other issues, submit the claim, and provide instructions to the BRR regarding display or access uses—unless she is satisfied with Google's default assumptions and permissions. The technical terms and logistics of the settlement are made manifest in the database and the claims process; understanding and evaluating these elements are key to successfully navigating a claim.

#### PREPARING PUBLISHER DATA

To assemble a claim, UBC Press required a complete list of all its titles to ensure it identified all UBC Press books in Google's database. Naturally the Press tracks this information, but over the years, locations and formats have changed to reflect available technology. There was no single source that listed titles from 1971 until the present day. The Press merged data drawn from several sources—the bulk of which came from Press Track, an internal database focused on the management of production data, and C-Press, a second internal database which, along with the University of Toronto Press distribution database, tracks inventory—into an Excel spreadsheet that also included data on titles' digital rights and e-book production status, among other details. This spreadsheet included most (though not all) of the Press' books, and it became the basis for a master list.

Although UBC Press has generally been an early adopter of technology in the publishing process, a digital record of all UBC Press' assigned ISBNs did not exist. The Press had kept a paper register since 1971 that was still being used until the mid-2000s; a digital version was used inconsistently from the late 1990s and regularly from about 2006 onward. The two sources together comprised a complete record of assigned ISBNs and confirmed a number of ISBNs apparently missing from the master list. The master list could then also be used to populate the digital register with titles and authors of older works. Searches retrieved from Google's database in some cases highlighted missing titles (typically older works from the 1970s or 1980s that had been dropped as the Press migrated tracking systems). If Google's results were confirmed as UBC Press titles, they were added to the master list. For conflicting or missing data, cross-references with online third-party ISBN databases, the catalogues of various research libraries, as well as UBC Press' own reference library, which contains physical copies of nearly all its published books, generally supplied answers.

This reference library was a valuable resource for some of the more puzzling inconsistencies. In some cases, Google's database showed more than one title assigned to the same ISBN, or several ISBNs published in the same format and year linked to the same title. Though Google simply had some wrong, these duplicates were in other cases distinct editions (such as an abridged and a complete paperback edition). One book, *Physiological Ecology of Pacific Salmon*, was produced in both hard- and softcover formats under the same ISBN when damaged stock was re-covered. Two others had the wrong ISBN originally printed on their copyright pages and had corrections printed

separately and pasted in; Google listed these under both ISBNs. The reference library also confirmed the format (cloth versus paper), a detail Google's database is frequently missing, and clarified multiple editions' respective ISBNs.

On November 2, 2009, Google filed a list of all the books it had digitized without authorization up to May 5, 2009 (Amended Settlement Agreement, 2009).<sup>26</sup> Google's set of scanned books numbered over 12 million at the time, including books Google had permission to scan; Google continues to scan books now, but these will be authorized scans (assuming the settlement is approved).<sup>27</sup> Yet not all ISBNs for books published up to 2009 are included in Google's database. Three UBC Press titles published before January 5, 2009, are not found at all. Approximately 100 books are listed under either their hardcover or softcover ISBNs, but not both. Google's database includes relatively few e-book ISBNs, but UBC Press had assigned ISBNs for e-book editions of some 800 pre-2009 titles. Google's database includes additional entries that do not use the ISBN, but may use ISSN, LCCN (Library of Congress Control Number), OCLC (the Online Computer Library Center number, or WorldCat number), or nothing at all; while not all of these entries designate actual books, there is no simple way to revise entries to use consistent or multiple identifiers.<sup>28</sup> Various identifiers that appear to refer to the same work may show different statuses for digitization or commercial availability. Though these error rates may be acceptably small to Google given the database's size, they create a more significant effect on individual rights-holders. Since Google continues to acquire or create scans, it may be in UBC Press' best interest to claim online all relevant items currently in Google's database and claim any other assigned ISBNs not included in Google's database by submitting a list. In this way, UBC Press should be able to control the uses of its works in all forms, rather than rely on Google to correctly link multiple formats to a work's single set of instructions.

Google's database allows one entry for each ISBN or other identifier. Although UBC Press uses ISBNs only, not LCCN or OCLC identifiers, the Press' master list has a separate entry for each work rather than each identifier. This allows hardcover and paperback ISBNs (in both 10- and 13-digit formats) and e-book ISBNs to be associated with each other and with a single work. Title, author, publication date, rights status, and other metadata will be consistent across each format of the same book; truly separate editions are tracked as separate entries. This also permits accurate cross-referencing of digitization status and commercial availability. Google will only make one cash payment per *principal work*, so a book that was digitized without authorization in its hard- and softcover versions will only be eligible for a single cash payment (Amended Settlement Agreement, 2009).<sup>29</sup> Conversely, if any format or edition of a work is deemed commercially available, then all formats and editions of that work will be considered commercially available, with Google's default licensed uses set to match (Amended Settlement Agreement, 2009).<sup>30</sup> The settlement administrator has not articulated how Google will reconcile hard- and softcover versions, nor how it will determine whether a revised edition is truly a new principal work. Additional challenges may arise with respect to co-publications or reprint editions, which could involve multiple publishers. The settlement limits Google's obligation to one cash payment per principal work; in practice, the costs of identifying separate works and managing or resolving competing claims will fall to the rights-holders themselves.

The Google database permits broad searches as well as title-specific ones, and users can download search results in spreadsheet format. Three sets of broad search terms located most Press titles: by ISBN prefix, by most common short publisher name (UBC Press), and by full publisher name (University of British Columbia Press). Because of the many possible abbreviations of UBC Press' name and Google's demonstrated inconsistency, there is no reason to believe these searches have successfully retrieved all UBC Press items in the database. Sadly, this abbreviation problem likely affects most scholarly presses, whose names frequently reference both a university and a geographic place name (potentially confusing its works with government publications as well) and whose works may make up the majority of research libraries' in-copyright holdings.

For UBC Press, the ISBN prefix set is most relevant. It contains only a few duplicates and non-UBC Press books under the wrong ISBNs; it also omits about a third of the Press' pre-2009 hardcover and paperback ISBNs. The "UBC Press" set captures additional titles listed by LCCN and OCLC, as well as many titles not actually published by UBC Press. The full-name set includes a wide variety of titles, many of which are published by other units or departments at UBC, but also some legitimate titles not captured previously. Downloading search results from Google's database allowed for close review of the entries, including problem identification (duplicates, ISBN errors, missing/extra editions). It also provided a simple gauge of the number of titles digitized without authorization, as this data is included in Google's database for each entry. In October and November 2009, these search results showed that nearly 54% of titles had been digitized without authorization in at least one format (491, any edition, out of 916 separate works). By January 2010, search results for the same set showed only 41% had been digitized without authorization (379 out of 916).<sup>31</sup> As Google continues to revise its database, these results could continue to change; there is no way that a publisher can verify what Google digitized when. Google appears to be its own auditor in this respect.<sup>32</sup> However, the digitization status does not affect the publisher's claim to the book (merely the publisher's expectation of compensation related to that claim). Since Google continues to update its records in this area, rights-holders should strive to claim all works, not just those that Google believes it scanned without permission.

The next step was to determine to which books UBC Press held rights. Until recently, most contracts did not explicitly address electronic rights; UBC Press considers any cash payments under the settlement to be a form of subsidiary rights and proposes to split these revenues 50/50 with authors. For some books, addenda had been prepared to explicitly address electronic rights in conjunction with the negotiation of the deposit of a large number of e-books in a recent Canadian Research Knowledge Network initiative. A very few contracts have included time-delimited reversion clauses, but in general, UBC Press contracts provide for a book to pass out of print only upon written notification by the publisher or written request from the author. Indeed, with print on demand becoming much easier to manage (both in terms of the Press' current workflow and recent changes with its distributor), UBC Press is contemplating how it could extend the sales life of many of its titles, including older backlist items. Unless rights have reverted or an author has given specific instructions, UBC Press considers any of its titles that exist as active e-books or are available for print on demand to be

“in print.” Additionally, UBC Press maintains relationships with other publishers that lead to co-publication agreements or buy-ins; for these titles, UBC Press may not hold the relevant rights even if the Press is actively selling physical copies.

To confirm the status of each title, information from author files was collated with the Press’ e-book tracking sheet and active sales information. Peter Milroy, the director of UBC Press, also reviewed the lists to confirm titles’ status and to identify titles whose rights belonged to other publishers or to which special circumstances applied.

#### TECHNICAL LOGISTICS: A GOOGLE DATABASE HOW-TO

With a complete list of all UBC Press ISBNs, the Press could then turn to the settlement claim website. Claims for books can be initiated in several ways.<sup>33</sup> “Upload a file of books” uses a spreadsheet listing any number of books; “search for books” offers several potential search fields and accepts whole or partial expressions of the desired search terms. A rights-holder with only a few works to claim might choose the “search for books” option, select the relevant works from the results displayed online, and click the submit button to complete her claim. Open-ended searches (such as by ISBN prefix or publisher name) can be conducted most easily from the “search for books” screen; the results can be downloaded as a spreadsheet for review. To claim by uploading files, a rights-holder would complete a spreadsheet pre-populated directly from the Google database or obtain a blank template and populate it herself. (Open-ended searches, such as by ISBN prefix, via spreadsheet fail because of the large number of matches.)

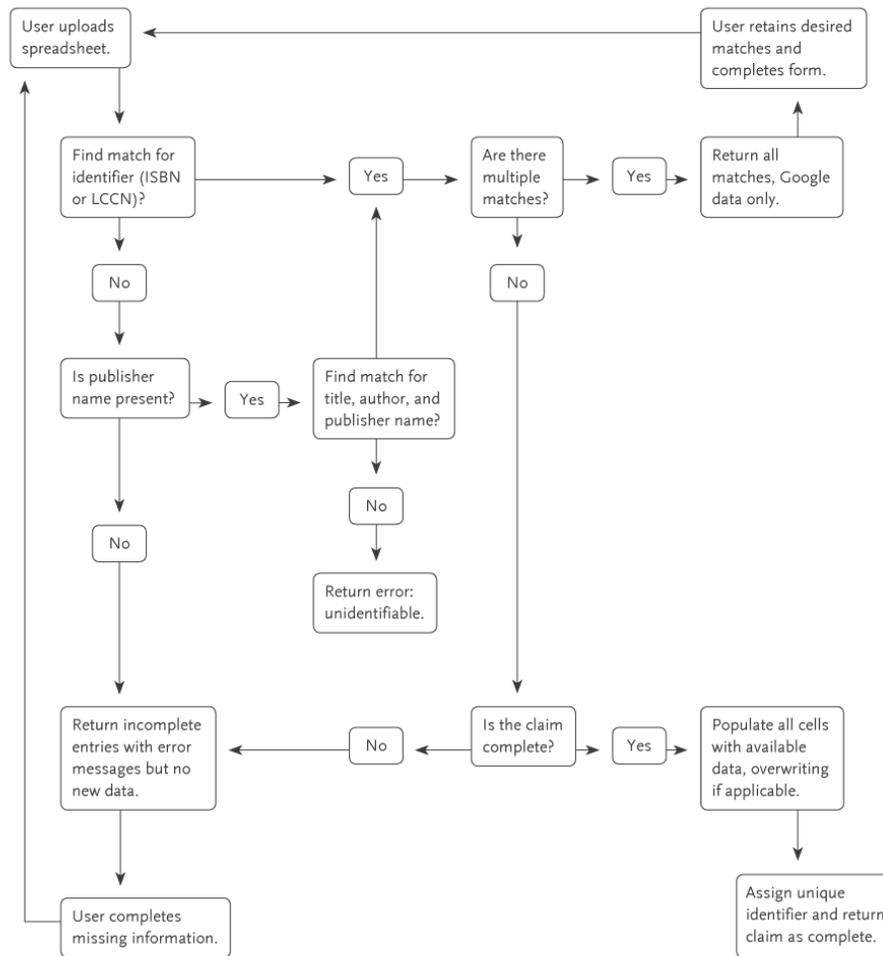
To successfully complete a claim, several criteria must be satisfied. (See Figure 1 for a diagram of the database lookup procedures.) Each book will be described on one line in the spreadsheet, with a number of columns for different data. First, the database considers the identifier, now limited to ISBN or LCCN. If no ISBN is provided, the database next seeks a match using the combination of title, author, and publisher. If this combination finds no match (which could be because too much information is provided, such as including “University of BC Press” when the database lists “UBC Press”), the claim will be returned with the error “unidentifiable.”

If either the ISBN or other bibliographic data produces a match, the database will then check for multiple matches. If more than one entry seems to use the data provided, the claim will be returned with the error “multiple.” Each possible match will be returned populated with Google’s available data on author name, title and subtitle, publisher name, publication year, digitization status, commercial availability, and format (Google uses the terms hardback, paperback, and book, which can denote either of the first two or an e-book). The rights-holder can delete whichever entries are not her books and complete the remaining fields.

If multiple matches are not found, or once multiples have been addressed, the database checks for completeness. A complete claim must include an assertion of rights status. In order to give instructions as to the use of the work, publishers must be *confident* or *highly confident* the rights have not reverted (authors must be confident the rights have reverted).<sup>34</sup> Publishers can confirm commercial availability and challenge Google’s designation if necessary, though this is not required. If some lines remain incomplete, the claim for those items will be denied and will return an error message for each affected

item. Incomplete claims may be the result of an invalid ISBN or LCCN, or a line with only the title and author (no identifier or publisher). Incomplete claims will be returned as submitted, without drawing additional information from Google’s database. The rights-holder can correct or add data and re-submit, at which point the database will reiterate its review. The presence of incomplete items or other errors in the spreadsheet does not block the successful claim of complete items in the same spreadsheet.

Figure 1: Diagram of the Google Books database lookup procedures, as of January 2010



Once the completeness test is passed, the books have been claimed and will appear in the rights-holder’s account online. Fresh downloaded results will show that for each item, Google’s data has populated the fields for author name, title and subtitle, publisher name, publication year, digitization status, commercial availability, and format. Google will also indicate any other claimants (by name) to the book, a Registry identifier (a 103-character alphanumeric string), and the date and time of the most recent update to the claim.<sup>35</sup>

The database records rights-holders’ submissions regarding commercial availability separately from Google’s designation. Certain fields are overwriteable: the assertion of confidence regarding rights status; the option to exclude the book from all display uses (a “no” or blank answer uses the default settings based on commercial availability); the option

to remove the book from Google Books and the library digital copies; and the option to unclaim the book. Exclusion instructions can be more fine-grained, but the spreadsheet allows only for the Google-defined defaults or complete exclusion. To allow certain uses and deny others, a rights-holder must submit instructions on a book-by-book basis, which can be accomplished by logging in to the settlement website, selecting a particular book in the account management screen, and then selecting the specific uses to permit. Instructions can be changed as often as necessary, although any instructions other than Google's defaults or blanket exclusion/inclusion will be cumbersome to manage.

Although the settlement administrator suggests rights-holders can correct their bibliographic data (Book Rights Registry, 2008a), this capability is non-functional as of January 2010. Google may be hesitant to accord millions of rights-holders (not all of them pleased with their lot) read-write powers in its database, but funnelling corrections requests through the settlement administrator seems a painfully slow way to address errors or omissions in the database. Linguist Geoffrey Nunberg (2009a) points out a number of metadata errors in Google's data, including mismatched titles (i.e., clicking a link on one title leads to a different book); errors in differentiating authors from foreword writers, translators, or names in the title; date errors; and bizarre classifications. Nunberg questions "whether Google's engineers should be trusted to make all the decisions about metadata design and implementation for what will probably wind up being the universal library for a long time to come, with no contractual [*sic*] obligation, and only limited commercial incentives, to get it right" (2009a). Paul Duguid (2010) echoes these concerns, underlining in particular that while Google may be making an "irreplaceable, marvellous, fantastic bunch of books," it is a long way from a research corpus of real scholarly use if its metadata does not improve. The collection's utility to researchers is reduced if the metadata is too unreliable or the scans are of poor quality (Nunberg, 2009a, 2009b; Samuelson, 2010a). At the same time, Google's responsibility to "get it right" is amplified by the fact that in taking the role of "the one" digital library, Google has effectively diverted scholarly and library resources away from other, perhaps higher-quality, scanning projects: Duguid (2010) mentions various funded research projects that have withdrawn resources from digitization activities in the wake of Google's proposal to scan everything itself.

Inadequate metadata could have unintended repercussions for the future of the book, especially if books are increasingly virtual rather than physical. Nunberg's argument hints at what researchers and the public might lose in the grand digital library:

Google's great achievement as a Web search engine was to demonstrate how easy it could be to locate useful information without attending to metadata or resorting to Yahoo-like schemes of classification. But books aren't simply vehicles for communicating information, and managing a vast library collection requires different skills, approaches, and data than those that enabled Google to dominate Web searching. (2009b)

Library classification schemes, while not perfect, hold meaning. Different editions of a book can vary significantly in quality, and libraries can evaluate those qualities and present each edition in ways that share with readers this background information (Duguid, 2010). The

contexts of and interrelationships between texts are enriched with creative and semantic analysis that mere indexing may not afford—even an index as all-encompassing as Google’s.

#### **KEY SETTLEMENT CONCEPTS FOR RIGHTS-HOLDERS**

A book’s digitization status (with or without authorization, as of May 5, 2009) determines whether it is eligible for a cash payment. These payments—at least U.S.\$60 per book, and possibly more if relatively few books are claimed and eligible—compensate rights-holders for Google’s failure to obtain permission to copy the works and buy Google the legal right to use the scans.<sup>36</sup> Google alone can affirm whether a book was digitized without authorization. Rights-holders may be able to compare the database information with the digitization provenance that can appear in Google Book Search’s overview pages, but this is prohibitively time-consuming (for publishers, if not authors) and may be inconclusive.

The commercial availability of a book determines the default permitted uses Google is licensed to make (Amended Settlement Agreement, 2009). If the book is commercially available new (used book sales do not count), Google will not make display uses unless the rights-holder grants explicit permission. A book that is not commercially available will be included in display uses, although the rights-holder may explicitly deny some or all display uses. If Google finds no information, it will determine the book is not commercially available. For the convenience of the settlement, commercial availability is used to establish whether a book is in print, but this designation only determines Google’s default uses. Once rights-holders have claimed their work, they can override these defaults and change instructions to Google at any time. If more than one rights-holder has an active interest in a work (such as the author and publisher of an in-print book), Google must respect the more restrictive instructions.

There are several types of display uses, all of which centre on the direct expression of the book’s content. Bibliographic data, a list of frequently occurring words or phrases in the book, or a map of the place names mentioned are not display uses (Amended Settlement Agreement, 2009). Display uses include snippets, front-matter display, limited preview (in various configurations), and access uses (institutional subscription, consumer purchase, and public access). Snippets can be allowed or restricted separately from other types of preview use, as can front-matter display, which includes the title page, copyright page, table of contents, and index. The standard preview is Google’s default preview for most types of books, which allows the user to see up to 20% of the book (but no more than five adjacent pages at a time). Rights-holders may adjust the percentage viewable or may instruct Google to use a different type of preview. Fixed preview includes the front matter up to the table of contents, the index, and a set of pages up to 10% of the total book: the pages displayed will not be selected based on a user’s search and will be the same for all users. Continuous preview allows up to 10% of the book, and does not have the limitations on number of adjacent pages that standard preview does (Amended Settlement Agreement, 2009). No preview is also an option, as is “unlimited” preview, if the rights-holder elects to display 100% of the book.

Access uses, a type of display use, allow users to view an entire book. As with other display uses, Google may by default offer access uses for books that are not commercially available, and rights-holders may override that default. There are three ways for users

to access full texts. With institutional subscription, for which Google sets the price, universities or other educational institutions as well as government or corporate institutions pay an annual or per-semester fee for their members' unlimited access to the collection. Individual users will have the option of consumer purchase, which gives them access to individual full texts online. Users will not download their own copy of the work, but will connect to it through Google. Rights-holders can set the consumer purchase price, or they can allow Google to determine the price through an algorithm that is supposed to maximize sales revenues and distribute prices among Google's designated price tiers (Amended Settlement Agreement, 2009).<sup>37</sup> The third access use is through free terminals at public libraries and at non-profit post-secondary institutions. This public access includes the same collection as the institutional subscription, but some features of the subscription (the ability to copy/paste or annotate text) will be disabled.<sup>38</sup>

#### THE MEANING OF "IN PRINT": SETTING LICENSED USES AND SETTING RIGHTS

The settlement aims to separate in-print from out-of-print books and to permit blanket uses of out-of-print books, while reserving exploitation of in-print books for rights-holders' own products. To do this, the settlement uses a book's commercial availability to create a technical definition of "in print": commercially available books would be designated "in print" and would not be displayed through Google Books. In this way, Google's display and access uses should not cut into existing trade markets, as Google would direct would-be readers of the book to sites from which they can purchase it. The settlement presents itself as creating opportunities to find and read primarily books that are not otherwise commercially accessible.

The settlement's particular definition of "in print" is not entirely in line with the way that term is usually understood within the publishing industry. Generally, a publisher's contract with an author includes some provision for establishing whether a work is out of print and how and when the rights revert to the author. These terms vary. UBC Press contracts, for example, do not set automatic reversion terms, though contracts can be terminated by written notice or request. UBC Press may hold small quantities of stock for a number of years, continuing to promote the books on its own website and through its distribution partners. With the increasing convenience and competitive performance of print on demand, UBC Press anticipates better satisfying long-tail demand for a larger proportion of its backlist. Since scholarly monographs achieve consistently modest sales but hopefully long lives, many authors have been content to let the Press manage and exploit rights indefinitely. As such, many older UBC Press books are still considered in print: contractually, UBC Press retains active rights even if the book is not continuously sold.

Apart from establishing Google's licensed use, the commercial availability/in-print designations also have implications for rights control. The settlement deems rights in an out-of-print book to have reverted to the author, in which case the author is the sole rights-holder. If the rights have not reverted, then both the author and publisher are rights-holders: both can give instructions to Google through the BRR on how to display or sell access to the work, and both share in any revenues from such uses.<sup>39</sup> But if Google fails to find commercial availability, lacking instructions to the contrary, the Registry will assume the book is out of print and rights have reverted to the author, and the publisher may not have any input or receive compensation related to Google's use of the book.

Commercial availability is easier for Google and the BRR to determine than contractual rights status, but the settlement's approach ignores or oversimplifies existing agreements between authors and publishers and is potentially disadvantageous for publishers. A commercially available book is almost certainly in print, but a book may still be officially in print—at least according to UBC Press' contracts—even if it is not commercially available. The settlement acknowledges this possibility and allows rights-holders to override Google's out-of-print designation separately from any challenges to its commercial availability designation. There are two methods to re-classify a book as in print. The first defers to the author-publisher contract terms, and allows that if rights have not reverted to the author according to those terms, then the book is still in print. The second requires the publisher to publicly announce its plan to publish a reprint or revision, and its release must occur within one year (Amended Settlement Agreement, 2009).

The designation of commercial availability is significant to all rights-holders because it determines Google's default use. More important to the publisher, however, is the ability to claim rights at all. In order to complete a claim, publishers must communicate to the Registry their level of confidence that rights have not reverted. This is the procedural technique for informing the BRR if a book is in print according to its author-publisher contract. If the publisher is confident or highly confident that the rights have not reverted, it will have a voice in controlling Google's uses of a work, such as excluding or removing the work. However, the publisher will only receive a share in any payments or revenue if it is highly confident; mere confidence will leave all disbursements to the author alone. Although it can be time-consuming to confirm individual titles' rights status, this is not an unreasonable demand of publishers, which should already have systems to manage rights.

Establishing that a book is in print does not affect Google's default permitted uses—those are still governed by commercial availability. In or out of print establishes whom the settlement recognizes as holding rights. Some critics of the settlement suggest that older contracts never addressed digital rights but that the settlement's author-publisher procedures allow publishers to claim rights to which they are not entitled. Although academic authors often assign full rights to scholarly publishers, a trade publisher's claim may be more ambiguous. The author-publisher procedures also appear to re-write existing contractual arrangements between authors and publishers (in order to set revenue splits paid by the BRR, for example). Authors and publishers may both feel unfairly treated by the settlement's attempt to impose a one-size-fits-all allocation of revenues. Similarly, the settlement may oversimplify rights; authors may have granted different subsets of rights to several publishers or other entities.<sup>40</sup> In this as in other aspects, the settlement privileges Google's administrative convenience over individual rights-holders' justifiable claims.

#### **NEXT STEPS FOR UBC PRESS**

UBC Press will claim all its books (including those already in the Partner Program) and share any settlement revenues with its authors. At this time, UBC Press is less concerned with ensuring Google has accurate information on titles' commercial availability. Rather than adjust individual display uses, UBC Press wants to withhold all titles from all settlement display uses, overriding Google's defaults for all books. The Press is more concerned with establishing the legal priority of its own contracts with authors rather than allowing Google's author-publisher procedures to re-write them. UBC Press considers the

identification or negotiation of in or out of print to be a matter between its authors and itself, not something Google or a U.S. class action can determine by decree.<sup>41</sup> The Press is better positioned to identify and locate authors or their heirs than the Registry and can offer authors more efficient disbursements and ongoing claims management.

It is likely that until the March 31, 2011, deadline, Google will continue to revise and expand its database, and the settlement administrator may further refine the claims process. UBC Press can claim books already present in Google's database immediately. To add books, the Press may submit a list to the settlement administrator or may wait in case the online process allows that. The advantage of waiting is that the Press retains control of the data and its accuracy. At the same time, UBC Press has no plans to discontinue its participation in the Partner Program. The clear limits of this program suit UBC Press' interests better than the sweeping licence granted by the settlement agreement.

UBC Press intends to withhold its books from Google's digitized collection under the settlement. There are two types of withholding possible: removal and exclusion. Removal is permanent and means the work will not be available in Google Books at all: its contents would not appear in searches (the indexed scans would have been removed), nor in institutional subscriptions or other access uses. If a removal request is made on or before April 5, 2011, the book will also be removed from the set of library digital copies (LDCs: the scans that Google shares with the libraries that have shared their physical collections with Google) and from the research corpus. Removal requests made after April 5, but before March 9, 2012, will effectively remove a book from Google Book Search and access uses, but individual libraries may keep their digital copies (all books they already hold in their physical collections), with restrictions placed on their use. The research corpus will also retain removed books, though the corpus can only be used for non-consumptive research—none of its books will have their contents displayed.<sup>42</sup> If UBC Press removes its works, it will effectively block Google from extracting any commercial value from those works under the settlement, and it will block libraries from acquiring free digital versions of their own collections, potentially preserving the incentive for libraries to convert or supplement their physical collections with e-books purchased from UBC Press itself or distributors that offer a more favourable or direct relationship to the Press. This incentive may still exist even if the Press does not remove its books from Google Books, since LDCs are not lending copies in the usual sense, though they can be used to replace a damaged or lost book if that book cannot be obtained new at a reasonable price, and excerpts can be used in classrooms or other fair-use educational contexts (Amended Settlement Agreement, 2009).

Exclusion imposes fewer restrictions on a scan's use. A rights-holder can exclude a work at any time and can change from "include" to "exclude" and back again. Works can be excluded from some or all display uses, as selected by the rights-holder, although access uses are subject to a coupling requirement: if consumer purchase is allowed, then the book must also remain in the institutional subscription collection; if institutional subscription is allowed, then the book must also be available through the free public terminals offered to some libraries (Amended Settlement Agreement, 2009; Book Rights Registry, n.d.).<sup>43</sup> If UBC Press excludes its works from all display uses, it retains the ability to experiment with the whole spectrum of proposed uses and evaluate them based on their actual performance. At the same time, Google and participating libraries

retain certain uses of the works that may affect UBC Press' own efforts in digital editions sales. With exclusion, the Press receives no additional revenues, since revenues are only triggered by display uses. However, Google may be able to extract commercial value from the works simply by their ongoing presence in its dataset and in the research corpus; after all, Google's business is founded on its excellence in search, and the larger the pool of data, the better its search algorithms will be.

UBC Press' assertion that virtually all its books are still in print unless authors say otherwise could be cynically interpreted as an opt-out digital rights clearance strategy of its own. However, the Press maintains ongoing communication with its authors, and, unlike Google's, UBC Press' use of authors' work is founded on an initial explicit grant of rights and mutual investments in the work's content.<sup>44</sup> It will keep authors informed of settlement developments, though the Press has deferred its planned author outreach until the settlement's future is clearer following the ruling. The Press' strategy restricts access through the settlement but maintains discoverability through the Partner Program; however, authors can request that the Press remove titles from the Partner Program and can provide their own instructions to the Registry. Pamela Samuelson posits that most academic authors would prefer a more generous interpretation of fair use or even open access, since their interests in books (and intellectual property) emphasize public dialogue and dissemination rather than economic interests (Fairness Hearing, 2010).<sup>45</sup> UBC Press is open to author input on how to include works in Google Books, though it does not at the moment favour open access.

#### **4. A steal of a deal: What the settlement offers the parties**

Technical interpretations of the settlement's fairness—whether it meets the criteria for a valid class action, represents all members' interests, or violates antitrust laws—have been discussed above. However, there are other points to consider in evaluating whether the deal treats all parties fairly. The costs and benefits of the economic and commercial transactions the settlement describes, and their subsequent effects, will not be shared equally among all parties. Google has invested financial resources in developing the Library Project, and if the settlement is approved, Google will incur additional direct expenses in the form of its settlement payments as well as ongoing development and maintenance of the Google Books digital collection. While rights-holders will not make comparable cash outlays, the revenues generated by Google's use of their works may not offer them significant financial returns, either. These revenues will pay for a range of new services from Google (scanning, hosting), but also for the entire edifice of the Book Rights Registry itself.

Google emerged from the settlement negotiations with its cost obligations mostly known, guaranteed revenue rates, and potential additional commercial benefits whose scope will be revealed over time. It is possible that the total cash payments will swell beyond U.S.\$45 million, but even if they do, millions of dollars are surely manageable sums for a company earning billions, and Google also obtains legal access to millions of books for no fees at all, since not all books are eligible for cash payments. Rights-holders, however, are in the trickier position of estimating the potential value of their works in a market just barely emerging. There are no expiry terms in the settlement: rights-holders will only be able to license digitization rights to Google once (although specific display uses can be licensed separately and/or repeatedly, since

rights-holders may switch between including and excluding their works from these uses). Technological changes could mean that in just a few years, the subjects of this settlement might be valued quite differently. Google also had to estimate this potential value, but the stakes for a resource-rich corporation are surely different than for publishers and authors, who have limited resources in comparison and who in most cases do not even know whether or when Google digitized their works. The settlement saddles rights-holders with unknown costs (indirect costs, deducted from their portion of earned revenues, and opportunity costs of selling Google the right to scan their works at what might be undervalued rates) and offers non-guaranteed revenue rates, but it also offers some degree of economic benefit.

#### FOR GOOGLE: MORE THAN FAIR

If the settlement is approved, Google can congratulate itself on a particularly excellent deal. It avoids years of uncertainty, not to mention ongoing legal fees, in litigation. It avoids prohibitive transaction costs by not having to clear rights individually for the works it has scanned already and all the works covered by the settlement and yet unscanned. It will receive a blanket licence to use a broad swath of copyrighted works, and it will enjoy an exclusive position, both as a market leader and with legal peace of mind, in the realm of digital rights: its private licence goes much further than current copyright legislation, particularly with respect to orphan works, for which rights are currently unobtainable in any market. Low transaction costs and legal certainty are key requirements for any mass digitization or digital archiving project (McCausland, 2009). The settlement offers both, to Google and Google alone. It will be years ahead of any potential competitors digitizing print works and may easily end up with an effective monopoly and a leading stake in the emerging markets for digital books. And all this costs Google only U.S.\$125 million—a mere 0.53% of its gross revenue, or 1.92% of its net income, for 2009 alone (Google Inc., 2010b).<sup>46</sup>

Had Google lost its fair use claim in court, it could have faced statutory damages for copyright infringement of U.S.\$750—or as much as U.S.\$150,000—per instance (Hyde, 2009; Samuelson, 2009b). Even if the judge did not fine Google for each of 12 million instances, this liability could have easily totalled far more than the U.S.\$60 or even U.S.\$300 per book that Google pays under the settlement. Google maintains it was confident that scanning and snippets were both permissible fair uses, and it was therefore not afraid of statutory damages. Settling could be viewed as an admission of uncertainty; regardless, settling gives Google far more permissible uses than could ever have been considered “fair.”

Google has suggested its scanned books will not generate vast revenues; after all, the settlement primarily deals with out-of-print books, and the more valuable market is books in print or yet to be published (Clancy, 2009).<sup>47</sup> The revenues Google will share with rights-holders may well be modest, but they represent just a small slice of all the uses to which Google might put the book content as pure data. Adding all books up to 2009 to its dataset could help Google improve its search algorithms and optical character recognition technologies, or develop voice recognition or translation functions. Since Google is a search company with growing media concerns, these are significant benefits—any of which could contribute to increasing revenues or lead to new commercial products or services. It is difficult to value the non-consumptive uses of the

dataset because many of the possible uses are potential uses, in emerging technological territory: no one knows how they will play out, but rights-holders will not participate in the development or monetization of these non-consumptive uses should they emerge.

Dan Clancy, Google Book Search's former engineering director, denies that the books are such a boost for Google's data banks. His argument estimates all the pages of all books Google wants to scan—somewhere between 12 and 20 million books, perhaps an average of 200 pages each, and clearly an impressive dataset (Clancy, 2009). Yet Google has also scanned and indexed some 100 billion Web pages: the book-page dataset is not even on that order of magnitude. In terms of raw data, scanned books thus offer less information than Web pages. However, Clancy's appraisal of "information" seems to consider merely sheer quantity of characters, rather than the varying qualities (syntactical, contextual, semantic, linguistic, cultural) any information-as-text also contains. Given that Google's powerful search function depends on its interpretation of the links between discrete pieces of data, the semantic depth in books surely makes them more valuable than Clancy lets on. Web pages collect information from the last 20 years or so, whereas books offer a sophisticated and robust repository of knowledge that spans centuries. Even if the qualities of individual books offer Google little value—it could be that copious amounts of raw data really are the only valuable element for Google's search algorithms—individual readers will and do find worth in individual books. Google's possession of the books that readers prize will make Google more attractive to search users and further secure its status as online destination of choice.

The existence of the Partner Program—for which services Google receives no fee—suggests that scanning books was already worth something to Google, even without the additional licences the settlement offers. Books offer Google certain advantages that scanning Web pages do not. One of these is stickiness. Google is fundamentally an advertising company. Its text and display ad programs account for 97% of its revenues (Google Inc., 2010b).<sup>48</sup> Its genius is that ads are harnessed to search queries, and the quality of Google's search algorithms means it can target both search and ad results to users' particular interests. In the 1990s, other search engines preferred to be portals and enticed users to stay on their sites; Google from the start was committed to search and willing to "get users out of Google and to their search destination quickly" (Auletta, 2009, p. 52). With Google Books, users will stay in Google's garden, searching within books or clicking from book to book, building rich data trails on user preferences and connections among content. Tracking user behaviour is the key to unlocking and exploiting the considerable power of the "World Wide Computer"; as Nicholas Carr argues, "Every time we write a link, or even click on one, we are feeding our intelligence into Google's system. We are making the machine a little smarter—and Brin, Page, and all of Google's shareholders a little richer" (2008, p. 219). This kind of data has become the currency of any successful online enterprise, which increasingly must be adept at identifying or creating communities and then serving them tailor-made experiences. Data—the more, the exponentially better—is the foundation from which the Web is monetized.<sup>49</sup> While the acquisition of book user data will not make or break Google (it is already much too large for that), it contributes to Google's dominance online.

Google may argue that it has made substantial investments in building its digitized collection, and should be able to derive revenues from it both to offset those costs and

to create a return on investment. But how substantial is this investment?<sup>50</sup> Developing book scanners and de-warping algorithms are fixed costs, even if making scans is not. It is unclear how much Google has put into building the accompanying metadata; much of it was obtained from a wide variety of sources, including libraries, database vendors, and now the rights-holders themselves, who will correct details as they work through claims. The server space to host these scans may be non-trivial, but surely the costs of hosting, maintaining, and extending Google's search functionality to books' content are more than balanced against the other advantages Google may also exploit—not to mention the public relations credit Google accrues for facilitating the largest library to date.

#### FOR PUBLISHERS: AS FAIR AS IT GETS

From many rights-holders' perspectives, the settlement remains a gross distortion of copyright law. First, Google made copies (and continues to make copies) of works without permission. Second, the settlement reverses long-standing copyright practice by putting the onus on rights-holders to request their works not be used, rather than on Google to obtain permission. The fact that Google has also swelled the default uses far beyond even the grey area of fair use seems a further circumvention of copyright law. Much material posted online is intended to be accessible, not protected or explicitly commercialized: in general, the Internet has been a conduit for allowing people to publish freely. Books, and their traditional use of copyright to lever compensation, fit awkwardly in this context. The premise of exclusive exploitation—fundamental to the conception of copyright—is in some ways fundamentally incompatible with online media.

Google's revenues from the deal include 10% off the top to cover Google's administrative, hosting, or other costs, plus another 30% of the remaining nine-tenths of all revenues: 37% in total. While this split is a little better than what publishers have historically received from booksellers (60/40, and now often falling to 55/45), it is unclear how much of Google's cut is required to cover its expenses to sell books. Rather than recoup costs, Google seems to be selling services; the settlement suggests Google's operating costs will be covered by 10% of total revenues, meaning the other 27% is presumably profit or return on investment, if revenue projections are accurate.<sup>51</sup> If those projections prove optimistic, the 27% will cushion Google from taking a loss. The settlement is an expedient way for Google to obtain digital rights, but it also offers a quick fix to publishers and authors struggling with how to commercialize digital rights at all. In exchange for rights plus a cut of the revenue, Google will scan and host the books, sparing rights-holders (and libraries) the trouble, and then manage and secure online material. Google has the technological expertise to block copy/paste or print functions, as well as control preview and blocked pages; a publisher can easily put up a PDF copy of the book, but may have less facility in limiting its access and use, although readily available tools with these functions continue to emerge.<sup>52</sup> Google offers software as a service, but as much as this relieves publishers or authors of the costs of maintaining certain technical expertise, it also locks them into a dependent role in which they are unable to self-direct the use and management of their work. Google may be best positioned now to implement this digitization, but in settling with Google the publishers and authors may contribute to a paucity of alternative models or competitors in the future (perhaps including publishers themselves). Google's head start might encourage others to drop pursuit. If Google is able to successfully commercialize its digital collection, and as the relevant technology

becomes even easier for any publisher or library to obtain, publishers might wish the settlement had been time-limited to allow them to renegotiate once market values—for both the digitization and aggregation rights Google receives as well as the digitization services Google provides—have been established.

Another service rights-holders will pay for is the Book Rights Registry. Before disbursing rights-holders' revenues, the BRR will deduct its own administrative costs (Amended Settlement Agreement, 2009). The settlement sets no specific guidelines or ceiling on how much the Registry might take, but clearly rights-holders will not receive a full 63%. Estimates of deductions range from 10 or 20% up to as much as 50%: New York literary agent Lynn Chu (2009) suggests the settlement creates "a big, costly impersonal bureaucracy, all of whose expenses you are required to pay, to serve as middleman between you and Google. ... The deal negotiated between Google and the plaintiffs—now excitedly looking forward to running BRR themselves—is, as to be expected, largely to their benefit, not yours. Their expenses (of unlimited scope) all come off the top. This will account for 40%–50% at least." Pamela Samuelson is less speculative, but states collecting societies like the BRR "historically have engaged in anticompetitive acts, spent [money] on themselves, not distribute[d] substantial sums to actual rights holders" (2009b, slide 16). Existing collecting societies' practices vary. The Authors Registry withholds a 5% commission but admits this does not cover its operating expenses; its role is also more limited than the BRR's will be, as it only distributes whatever payments it receives and does not represent authors' rights (Authors Registry, 2008). From the royalties it collects, Access Copyright deducts a 20% administrative holdback, to cover its administrative and operating costs, and a 1.5% allocation toward funding its cultural grants program. The remaining funds are distributed to rights-holders. Royalties collected in one fiscal year are distributed in the next; generally, individual payments of less than \$25 may be held until the royalty exceeds \$25, or for up to five years (Access Copyright, 2010b). For these reasons, Access Copyright may in practice retain more, or occasionally less, than its specified 21.5%.<sup>53</sup> The variables affecting the BRR's revenue pool make predictions difficult: no one knows what revenues Google will remit, how many rights-holders the Registry will administer, or what fixed and direct costs the Registry might have. But publishers and authors are paid last: after Google, after the Registry.

The Registry has an open-ended mandate to negotiate ongoing licence developments with Google or with third parties on behalf of rights-holders. The heterogeneity of the class could result in individual rights-holders' sense that the BRR may not work in their best interests. A scholarly press might have very different concerns than a trade publisher for the pricing of the institutional subscriptions, for example, since research libraries—the largest market for scholarly works—will be the primary target of these subscriptions. These challenges are similar to those already facing reproduction rights organizations such as Access Copyright; the potential for conflicting interests remains unless authors and publishers have identical motivations or responses with respect to their creative and economic interests. The sheer scale of the settlement, and its potential to set precedent and create monopoly at such an early stage of the digital book market, adds a new worrisome dynamic.

For Google to create the kind of digital book collection it desires, it needs a register of copyright holders, a simple way of clearing rights, and some method to access the works whose rights-holders are unknown, unlocatable, or disinclined to respond to permissions requests. There is no such register, and existing copyright law does not allow easy mass rights clearance or provide any legally certain way to obtain consent from absent or unresponsive rights-holders. Google must have been impatient with what the government and existing regulations allow and provide, and so naturally welcomed the opportunity to create a para-legislative scheme for itself with the Authors Guild and Association of American Publishers. The settlement defines special exemptions from the laws for Google only: Google becomes government, or at least that government branch that defines and controls digital rights for certain forms of intellectual property. In addition, the settlement's creation of the Book Rights Registry, the regulatory body that facilitates Google's direction of books online, is supported via a sort of tax on rights-holders themselves. Revenues earned by publishers' and authors' creative works will subsidize the services Google needs to run Google Books: the settlement is perhaps not as cost-free as it purports to be.

#### FOR AUTHORS: UNEVENLY APPLIED

Though this report focuses on a publisher's experience of the settlement, it is worth noting that authors bring a distinct set of concerns to the case. The settlement for the most part groups publishers and authors together simply as rights-holders, but as the range of objections and statements of support demonstrate, it is incorrect to assume their perspectives align or that the settlement treats both equally. Authors' assessments differ from publishers' and from other authors'. Some see Google Books as infringement plain and simple; some see an escape from obscurity or a convenient distribution platform; some see collusion between Google and publishers to exploit authors; some worry it will undermine their ability to negotiate and earn revenues with their work even as some welcome any income on work that is otherwise gathering dust. In a further challenge to the representativeness of this subclass, members of the Authors Guild are mostly not academic authors, while most of the scanned books that launched the complaint were written by academics. Though there is much to be said in a detailed consideration of authors' various concerns, it will need to be said elsewhere.

It is worth addressing one line of critique, since scholarly publishers like UBC Press in particular and publishers in general may appear implicated. The settlement administrator's report on initial claims showed that the bulk of books claimed were from publishers. Publishers, comprising only 10% of the claimants in this period, accounted for 71% of the books claimed online (Allen, 2010a).<sup>54</sup> On its own, this is not surprising, since publishers will naturally claim more books than individual authors. Ron Lazebnik, the lawyer representing several large authors' associations, argues that not only were publishers claiming a disproportionate number of books, but Google had designated 56% of these books out of print. Therefore "the majority of books ... are being claimed by publishers who no longer support the hard copy version of these books. This fact pattern demonstrates the exact reason why we believe the settlement is unfair and unreasonable to authors. It is allowing publishers to lay claim to rights and revenues that belong with authors" (Fairness Hearing, 2010, p. 52).<sup>55</sup> It is, however, too early to draw conclusions from these "fact patterns." Google's "out of print" label merely means Google did not find a commercial source for the book, not necessarily

that rights have reverted to the author. The claims deadline is March 31, 2011, and there is no reason to believe the distribution of early claimants is representative of the whole. Rights-holders will likely have better information than Google about the commercial availability of their works, and Google's designations may undergo considerable adjustments. Whether trends develop as Lazebnik suggests, his allegation highlights a key tension between subclasses: objecting publishers are indignant that Google scanned without permission, but many objecting authors feel doubly betrayed by both Google and the publishers themselves.

This fear is not unfounded, but, significantly, it concerns authors and publishers: Google's position is not seriously affected if rights-holders disagree with each other. Authors and publishers may dispute and resolve conflicting claims under the settlement, but this will take time. The settlement makes it relatively easy for publishers to claim, if not keep, active rights. Just as Google might have preferred to acquire millions of books from tens of libraries (the Library Project) rather than from thousands of publishers (the Partner Program), it may be easier for Google or the BRR to deal with publishers under the settlement, which can each speak for hundreds of books, rather than authors, who each speak for fewer books. As noted previously, the settlement primarily serves Google's administrative convenience—from its generous default licences, to shifting the costs of locating rights-holders to the BRR (and therefore the rights-holders themselves), to acquiring orphan works—and this may not align with what is either fair or easily accomplished for any other parties. Perhaps more likely than publishers wresting claims from authors is the possibility that some claimants may take advantage of vulnerable orphan works. The economic incentives proffered by the BRR may encourage anyone with even a slim claim to otherwise unclaimed works to establish themselves as the rights-holder of record. Google need not have accurate rights-holder information in order to apply its licence, and the BRR may have limited incentives or resources to verify data.

UBC Press does not anticipate such mistrust from its authors. Rather, the Press will inform authors of settlement developments, steward authors' economic interests, forward half of any cash payments and a portion of any other revenues as determined by royalty contracts, and invite authors to diverge from the Press' strategy with Google if they desire. Due to the complexity and ongoing changes of the settlement, UBC Press believes many of its authors will benefit from having the Press as their intermediary.<sup>56</sup>

These are only some areas of potential concern for publishers and authors. However, the Authors Guild and Association of American Publishers are at least as uncertain as Google about their chances of winning the original case in court, and winning that suit would not quell the inevitable drive to digitize. In settling, the plaintiffs spare themselves litigation costs, put a major digital rights question behind them, and receive revenue from putting their work online—and monetizing digital publishing in any form is the holy grail for the industry right now. Unfortunately, rights-holders might need Google more than it needs them: there may be no other recourse to reach readers online as effectively as Google represents. The deal may not be that good, but publishers and authors do not have many other options that include financial compensation. However, a variety of legislative interventions—the DOJ's recommendations or other contingencies—could improve publishers' and authors' lot.

## 5. Copyright reform: Balancing access and control

Copyright laws have a long tradition, and they have been revised on numerous occasions in each of the countries implicated in the settlement agreement. The original spirit of the law, in the U.S. at least, was to support the society-wide benefit of public knowledge development and exchange but also to offer limited market protection so producers could profit from their own creative labours. Some provision to balance owners' and users' rights emerges through the notion of fair use or fair dealing; differing interpretations of fair use were the spark that led to this settlement. The fulcrum at which appropriate balance rests is also a matter of debate. In Canada, the notion of users' rights, and of this balance, are relatively recent concepts to be explicitly recognized in the courts.<sup>57</sup> Although the U.S. appears to offer broader users' rights and better established legislative grounds from which to claim them, twentieth-century copyright reforms in the U.S. have often strengthened owners' rights. Restrictive controls such as long copyright terms and special protections for digital rights management, as well as the presumption of automatic protection regardless of context, can be seen to favour rights-holders over users to the point now that the spirit and the letter of the law have fallen out of phase. Digital communications technologies have placed a particular strain on copyright practice such that the laws appear not only unbalanced but dysfunctional. The Google Books settlement configured the Authors Guild and Association of American Publishers' complaint as an opportunity to address this widening legislative gap, a gap that is a consequence both of the loss of balance and, more pointedly, of technological change.

The settlement is a contract between specific parties and charts one seemingly practical transition from selling print books to selling books online. It does not profess to extend copyright legislation nor provide a usable framework for others to do so, though its approaches and features may ultimately influence the creation of such a framework. The precedents the settlement sets may not be the best models for copyright reform, and it is appropriate for the court to weigh their possible consequences before allowing a private contract to stand alongside, and arguably in exception to, existing copyright law.

Other digitization efforts might better serve the collective interests of rights-holders, libraries, digital service providers, and readers. Various legislative reforms could support alternative digital library projects and address copyright inefficiencies in a general and public forum, rather than reserve such benefits for Google alone. Even if the court declines to refer the class action complaint to U.S. Congress to decide, as some objectors recommend, the settlement certainly underlines the need for and public desirability of congressional action in the shape of meaningful, legislative copyright reform. Canada and other countries might lead or follow such legislative developments, as the problems facing copyright laws are not unique to the U.S.

### COPYRIGHT IN A DIGITAL WORLD

In *Authors Guild et al. v. Google Inc.*, the original complaint lay in a difference of opinion on what constituted fair use, which remains a grey area under copyright law. While Canadian copyright law defines five and only five types of activity that can be considered fair dealing, the U.S. fair use provisions leave room for interpretation. In this case, both sides insisted their interpretations of fair use were correct; both sides were also happy to negotiate a settlement and avoid costly and lengthy litigation;

clearly, both sides were interested in employing the settlement mechanism to address a number of other desirable clarifications on rights and digitization. The uncertainty around fair use may arise because present U.S. copyright legislation offers an inadequate definition of fair use. It could also be argued that retaining ambiguity in fair use provisions is a feature of the legislation, allowing the courts flexibility in order to evolve interpretations and legal understandings organically without having to undertake a complete legislative overhaul for every new situation. Unfortunately, the settlement avoids this beneficial ambiguity as well, since it subsumes any discussion of fair use in a commercial contract.

The copyright laws we have inherited were built for the industrial era, in which making copies presented certain technical barriers (such as access to a printing press) and for the most part required intentional efforts. In this context, “the law of intellectual property placed its triggers at the point where commercial activity by competitors could undercut the exploitation of markets by the rights holder. Copying, performance, distribution—these were things done by other industrial entities,” not by individuals (Boyle, 2008, pp. 50–51). But now, with computers making so much production power so trivially available, “the triggers of copyright—reproduction, distribution—can be activated by individual footsteps” (Boyle, 2008, p. 52). In digital contexts, as Lawrence Lessig elaborates, the architecture of existing copyright legislation “means that the law regulates everything. For every single use of creative work in digital space makes a copy. Thus—the lawyer insists—every single use must in some sense be licensed” (2010, p. 4). Such an approach is impractical and places a heavy burden on users while also failing to acknowledge the structure and possibilities of recent communications technologies. Copyright laws as they stand are ill equipped to address or employ the technological realities now facing publishers, booksellers, and authors or creators. Worse, it is difficult to enforce narrow licensing in a digital environment—where it is all too easy to copy or distribute material—without resorting to draconian measures.

In the absence of sensible digital use guidelines, the public drifts toward the conclusion that poor rules should be broken (Lessig, 2007).<sup>58</sup> It is untenable, in a healthy democracy, to allow the common sense of what is permissible and socially acceptable to diverge so significantly from what the laws actually support or restrict. (Familiar clashes include remixing or reinventing existing work anew; building new tools or knowledge out of existing works; format- or media-shifting, so that a previously purchased work is available in an emergent technological format; sharing content.) Disabling the troubling features of digital technologies would deny many parallel benefits; besides, the technology is everywhere and is not going away. Short of significant market collapse (Nicholas Carr [2008] speculates that the insatiable demand for more data centres could precipitate an IT energy crisis), our media environment will not revert to non-digital. Given the developments in technical possibility and in emerging patterns of use—what people can do and what they want to do have both substantially changed since the last round of copyright reforms in the 1990s—perhaps it is time to consider comparable changes to the regulatory framework that guides such use. The appropriate response is also not necessarily to throw copyright restrictions wide open and bend the law to suit the user who wants immediate, free access to the

creative works of others. Certain types of changes to the law—around copyright term length, or a default of full protection—could allow for flexibility and compromise while still retaining reasonable control for owners and reasonable access for users.

Had this case gone to litigation, the court’s ruling could have clarified the legal status of scanning for indexing, making snippets, and related activities, and it could have offered guidelines to other users in similar circumstances. It could have contributed to the ongoing negotiation of the balance between the two fundamental purposes of copyright, as configured in the U.S. constitution (art. 1, s. 8): first as a system to advance the public good through promoting the “progress of science and useful arts,” and second as a grant of temporary exclusive rights to incentivize creators.<sup>59</sup> However, litigation would require the court to define the limitations on rights-holders’ rights based on the specific evidence of one factual situation: the situation or the interpretation may not scale appropriately to other uses, markets, or “similar” circumstances. A settlement leaves the fair-use status of basic digital acts unresolved. It could as easily spur legislation on—Google’s would-be competitors may become motivated lobbyists in favour of orphan works legislation, for example—as it could offer an interim patch that becomes an excuse to defer the hard legislative work ahead. Neither litigation nor settlement is truly satisfying. A full response to the challenges the settlement attempts to broach requires legislative intervention, both to rediscover the appropriate balance between owners’ and users’ rights and to ensure everyone has reasonable access to the benefits this balance represents.

Of course, neither litigation nor legislation necessarily ensures that the outcome will be this appropriate balance. There are already powerful lobbyists whose commercial interests benefit from further strengthened owners’ rights, regardless of the public costs. In the U.S., the *Copyright Term Extension Act* as well as the *Digital Millennium Copyright Act* (both passed in 1998) increased the scope of copyright protections. Given that the legal system in the U.S. continues to fail to address orphan works legislation, it is perhaps not surprising that the parties in the Google Books case preferred a private settlement. A new round of copyright reform might result in a larger pool of protected and unusable works, licensing schemes that are more restrictive, or a reactionary rejection of technological capabilities. Yet the fact that no one can predict which influences will ultimately sway legislators’ opinions is not reason enough to abandon the legislative process. Even if effective legislation rarely makes it through the thicket of lobbyists and political interests, it does not necessarily justify using the class action mechanism instead, despite its relative expediency; in fact, effecting ad hoc legal change through class action creates other consequences. The settlement may appear to be a safer—and quicker—option than full-scale legislative reform, but the settlement only defers, not resolves, the many thorny issues copyright faces. Whatever shape the next round of copyright reforms might take, it is instructive to consider the possibilities that the proposed settlement, or something like it, models.

#### THE SETTLEMENT’S PARA-LEGISLATIVE IMPLICATIONS

Daralyn Durie, representing Google and defending the settlement as the best solution for book digitization, points out that “the opt-in regime is just the status quo”; that is, the existing system in which those who wish to use (or digitize) a work must seek out and negotiate with individual rights-holders to acquire their permission. “We know it

doesn't work [on a large scale] because if it worked, someone would have done it already" (Fairness Hearing, 2010, p. 146). Mass digitization projects are today technically feasible, but such initiatives, however publicly desirable (and, in the case of orphan works, without apparent harm to the existing market for the work), face prohibitively high transaction costs under the current rules. One solution would be to give up on large-scale digitization projects, though this is unlikely given public appetite and technological capacity: it makes no sense to deny these possibilities because we have yet to identify the model that supports them. Another solution would be to appeal to legislators to address these inefficiencies and create better means of identifying and negotiating with rights-holders. The parties to the settlement have chosen a third approach: to disregard the legislative process and pursue a market solution to the orphan works problem. Part of the question before the court now is whether the ends justify the means.

If approved, the settlement may set precedents that poorly serve the public as well as the authors and publishers who have traditionally brought creative works to the public. One such precedent is that privately negotiated settlements offer more efficient and cost-effective means of addressing legislative challenges than does legislation—a depressing thought for the civic-minded, as it potentially offers corporate interests a way of shaping legal loopholes to order. The DOJ quotes the U.S. Supreme Court in its caution that class actions “cannot carry the large load’ of restructuring legal regimes in the absence of congressional action—however sensible that restructuring might be” (United States Department of Justice, 2010, p. 3); the DOJ seems to acknowledge the existing legislation is lacking but disagrees that a class action settlement is the appropriate response. To some extent the settlement justifies a “take first, ask permission later (if at all)” approach to intellectual property. Unlike most digital pirates, both sides in this case used the perceived wrong as a springboard from which to launch more ambitious negotiations, which also illustrates one effect of concentrations of power: being very large is a good way to get what you want.

The settlement, with its tethered access, also erodes first-sale rights—readers’ ability to do more or less what they will after the book has been purchased. Google Books shifts the point of sale from a good (a discrete instantiation of a creative work) to a service: that of convenient access, provided by Google. Readers pay for ongoing access to a book, not for any one copy of a book. Even as it promises new levels of public availability, Google Books dangles the proposition of perpetual rights, with a new twist: in selling access, it allows for perpetual monetization (and, through ad revenues, even public domain works can be monetized, for Google if not for former rights-holders). Google proposes to share the spoils with creators and publishers, which should encourage them to participate. Yet publishers and authors might find themselves ultimately outsmarted, as history too often shows that clever brokers make more profits than skilful producers.

Lawrence Lessig (2010) sees another significant precedent: the settlement outlines an overcontrolled licensing paradigm in which permission can and must be sought for each and every use. Lessig glides over the fact that permissions are already entrenched in publishing: while brief quotations are allowed, quoting substantial—with “substantial” a matter of case-by-case interpretation—passages requires explicit permission and frequently a fee. Such permissions are required to print books, but not, significantly, to read them: so far, readers’ personal access to and use of books

has been largely uncontrolled. In the architecture of the settlement, however, new depths for licensing and permissions emerge, and not only because it clears rights to so many books in one stroke. Individual users will be identified by geographic location, institutional affiliation, and/or previous online sessions at a minimum (all necessary to ensure only allowed access and uses of the texts occur); it is easy for a computer to track which pages or parts of pages are viewed by whom, and it will likely become easier still to separate specific portions of text or illustrations from surrounding text. As Lessig quips, “The deal constructs a world in which control can be exercised at the level of a page, and maybe even a quote. It is a world in which every bit, every published word, could be licensed. It is the opposite of the old slogan about nuclear power: every bit gets metered, because metering is so cheap” (2010, p. 4). Digital publishing makes two irresistible offers: just as making and distributing digital copies becomes trivially cheap and easy, so too does tracking online access and use—especially for Google. Lessig argues that the appeal of such metering serves corporate interests, while the promise of expanded public access is in fact overshadowed by cementing the presumption that all uses should be licensed. To respect the licences, all uses must also be monitored. This default of control will stop up the flow of knowledge and public access to our own culture; as Lessig puts it, the settlement is more burden than benefit, and it invents a complex legal structure “to make every access to our culture a legally regulated event” (2010, p. 4). Any advantages the settlement seems to offer as a digital library are dwarfed, Lessig concludes, by the settlement’s eagerness to create a digital bookstore, and most of all by its ominous validation of strong copyright as a basic, and appropriate, tool of control. Although the settlement does not invent these ideas—digital reading and online activity are already tracked and regulated at many levels—it does bring them to books, and it does so on a grand scale.

#### ALTERNATIVE DIGITAL COLLECTIONS AND THE PUBLIC INTEREST

The direct parties to the settlement are not the only ones who will be affected by it. U.S. libraries have much to gain under the settlement: outsourcing their own digitization efforts mostly for free, accessing a much broader virtual collection than their own physical acquisitions budgets might allow, expanding patron services. The settlement’s promise of a “universal” (U.S. only), more-convenient-than-ever digital library holds huge appeal for casual and avid readers alike. Digitization opens up new options for users with visual impairments or print disabilities to access books, and the settlement includes provisions to allow these users access to LDCs beyond the regular subscription or other modes of access (Amended Settlement Agreement, 2009). And the public does not directly pay for any of these new services. However, of all parties, it may ultimately be the public that is least well served.

The risk of price-gouging is certainly plausible, but if Google does develop anticompetitive practices, another lawsuit or legislation can address that, as Paul Courant (2009), dean of libraries at the University of Michigan, reminds us. Librarians have long championed public access and freedom of thought and speech over censorship, as well as protected the privacy of patrons’ reading choices. Though many point to the settlement’s lack of concrete user protections in these areas, there is a chance that Google will respect these principles because its historical institutional culture knows them as the right things to do. These are great hopes to pin on such thin promises; explicit and binding guidelines in the settlement would better address such

concerns. The court might also condition its approval on the establishment of review procedures, overseen by the court, so that Google and the BRR are more transparently accountable to users, libraries or subscribers, researchers, and individual class members (Association of College and Research Libraries et al., 2009).

Dan Clancy (quoted in Albanese & Oder, 2009) offers reassurances that libraries (and perhaps through them, the public) will have a strong voice in shaping Google's "[book] product, and the settlement agreement isn't where we define that. This legal document talks about everything we can't do. It doesn't say anything about what we will do." Though Clancy believes this open-endedness will lead to positive developments for both libraries and Google, his remarks underscore that little in the settlement speaks concretely to anyone's long-term obligations to the books themselves and the richness of knowledge they contain.

Happily, Google Books is not the only mass text-digitization project. Even as they participate in Google Book Search, U.S. research libraries have not relinquished their own roles in stewarding cultural knowledge. HathiTrust (<http://www.hathitrust.org>) is an initiative of these libraries to pool and preserve their collections in digital form. While HathiTrust partly relies on Google's scans, its purposes are non-commercial and focused explicitly on supporting the needs of research libraries and scholars. There are other repositories: Gallica (<http://gallica.bnf.fr>), an initiative of France's National Library, digitizes materials from public collections. Europeana (<http://www.europeana.eu>) collects digital items from European libraries, archives, museums, and galleries; it includes published and unpublished texts as well as images, sound, and video. Institutional digital repositories are also becoming increasingly common.

The Internet Archive (<http://www.archive.org>) was established to preserve Web pages and ensure ongoing public access to information and cultural artifacts. Its collection now includes 1.9 million books, as well as music, video, and other digital material. The Open Content Alliance (OCA; <http://www.opencontentalliance.org>) is an initiative of the Internet Archive; it collaborates with libraries and archives to collect digitized text and make it available through the Internet Archive.<sup>60</sup> Heather Morrison (2009), in her analysis of OCA and Google Books, demonstrates OCA's robust network structure: its many collaborators and open nature mean it is flexible and scalable, while its strong sense of purpose ensures consistency among collaborators or "nodes." Google Book Search, she argues, is a poor network: it relies on Google Books as a central control; its partners are not collaborators and have diverging goals; extending its configurations does not scale and will require fresh negotiations. OCA's network is a much better instrument as a long-term public good; in contrast, Google Books, as a single corporate entity, is better suited to its own corporate interests. Of course, OCA's mission is to serve the public good, which may have led it to choose the structure it did; Google Books sometimes pleads public benefits but maintains other goals.

In fact, Google's lawyers have suggested that evaluating the settlement on its merits for the public is inappropriate: despite its benefits for potential users and the public, it is an agreement between the class members and Google, and its treatment of class members' interests is "paramount" (Google Inc., 2010a, p. 3). It is to Google's advantage for the court to consider the fairness of the settlement to its immediate parties; strategically speaking, appeals primarily to the benefits the settlement offers outside the class will not

establish the class action's validity or satisfy its technical requirements. The DOJ claims that the settlement achieves its various public benefits "in spite of and not in furtherance of the basic premises of the Copyright Act" (United States Department of Justice, 2010, pp. 10–11). The primary objective of the settlement is to clarify a commercial relationship between Google and publishers or authors, not to improve legislative interpretation, nor even necessarily to advance the public good or access to information.

These alternative digital collections are generally freely and publicly available; as such, there would be no need to monitor and control which users access which materials, as Google proposes. As limited or collaborative collections, each is unlikely to become the "essential facility" or concentrated control that Google Books might (Association of College and Research Libraries et al., 2009, p. 3); multiple or collaborative collections may encourage the parallel existence of other digital libraries and thus support a robust and diverse digital cultural ecology. Each is also governed by public and/or non-profit entities, which may establish the collections' transparency and accountability to users.

Unlike the other collections mentioned above, only Google offers a revenue stream for rights-holders and explicitly extends the commercial lives of their backlists. Google is already a leader in search, which means lodging material in Google Books should increase a publisher's or author's chance of having that material discovered by potential readers. Other digital libraries claim smaller collections and smaller audiences; the initial advantage with Google's larger collection and audience could easily snowball due to natural network effects (where having either or both more books and more users means Google's search results will be more useful and relevant, which attracts still more users). Other digitizers may also ask the holders of original or physical copies to contribute to scanning or hosting costs, whereas Google Books offers these services without collecting direct fees. Participating in the settlement extends the network-effects advantage to publishers while applying few costs against their bottom lines.

However, Google offers these possibilities just at the point at which mass digitization is getting easier and publishers are, many for the first time, seriously grappling with how to transform their businesses to meet this changing technological climate. Google wants to be everyone's preferred route to monetize their content online, but print on demand, new devices, or other digital collections could significantly alter the landscape for digital books. The settlement's long-term benefits for publishers are uncertain because the market's long-term opportunities are unclear.

Google Books itself could be made to better serve the plaintiffs and the public by adding limits to the settlement or conditions on approval. As Hadrian Katz, representing the Internet Archive, argued at the fairness hearing, one condition that would "realize all the benefits ... [and] solve virtually all the problems" would be to limit the settlement "to those who willingly participate in it": that is, make it opt-in (Fairness Hearing, 2010, pp. 92, 93). The notable problem Katz's recommendation does not solve is that an opt-in settlement would ensure, just as the current copyright regime does, that orphan works remain primarily inaccessible. Alternatively, the DOJ (United States Department of Justice, 2010) suggested limiting the term of Google's licence for unclaimed works to some number of years, after which the market for digital works and access services could be reassessed and the licence either renewed by the Registry

or extended or revised by the court. The settlement appears to grant Google a licence for the term of copyright, after which it could continue to use the texts, now in the public domain. The settlement does not contemplate how Google Books might itself go “out of print” and see its licence revert, although that licence is non-exclusive and does not bar authors and publishers from using other digital collections at the same time. A time-limited licence for all works might reassure rights-holders that they are not locked into a contract that future technological or legislative developments might render less fair. Electronic rights are only just being defined, yet Google obtains a relatively open and generous licence for very small fees. Google pays only for the works it digitized before May 5, 2009, and may pay as little as U.S.\$60—less than the list price of a UBC Press hardcover, and not quite the price of two paperbacks. For works yet unscanned, Google will not pay any fees to digitize and aggregate them. The revenues that Google shares with rights-holders come through Google, but are drawn from libraries and individuals. As digital uses are fleshed out, it may make more sense to divide electronic rights: a right to digitize, a right to aggregate, one to use in non-consumptive research (computational analysis or search tools), one to sell the work in electronic form, and one to sell print products derived from the electronic form (Rowland Lorimer, Director, Canadian Centre for Studies in Publishing, Vancouver, BC, personal communication, March 20, 2010). Although rights-holders can always exclude their works from any form of sales through the settlement, they may not be able to revoke or separate these other electronic rights.

The court might also impose caps on profitability to prevent price-gouging and protect users should Google one day sell its assets to another entity. Regulations to prevent monopoly could define terms by which Google must license its scans to third-party resellers. A Google monopoly could play out another way: if Google commands enough of the market, it could drive prices down, even below cost, to dissuade competition or non-compliant partners—its pre-set price tiers start at U.S.\$1.99, and the settlement’s vision for collection-wide pricing proposes 41% of all books priced under U.S.\$5 and only 19% priced over U.S.\$10 (Amended Settlement Agreement, 2009).<sup>61</sup> These prices are much lower than those the existing publishing industry requires and may fail to sustain many creators and producers. Although rights-holders may set their own prices rather than rely on Google’s algorithmic defaults, in practice it would be cumbersome for rights-holders to manage independent pricing (in 2010, non-default price data could not be uploaded to the Book Rights Registry but had to be entered manually, book by book, although this mechanism may be developed in future). If a critical mass of rights-holders forgoes the extra effort of adjusting their prices, it would be difficult for those who do not want to use the algorithm to resist the emerging “market rate” for a digital book. The settlement offers Google, and the BRR, considerable power to shape the market. Publishers, authors, and in the end readers may be unable to withstand the market that suits Google. Even if low prices seem to lower barriers to public access, once publishers and authors cannot sustain themselves at those levels, their possible disappearance will erode the range of resources to which the public has access.

While any of these possible additional settlement stipulations could improve the digital landscape for publishers and authors, more meaningful solutions—and ones likely to also tend the public interest in the face of copyright’s controls, as well as spark international dialogues—will be found in copyright reforms via legislative channels.

## LEGISLATIVE SOLUTIONS FOR LEGISLATIVE PROBLEMS

The complaint against Google involves important issues in copyright law, fair use or fair dealing, and digital publishing, but a class action settlement does not further the public, legal resolution of any of them. Instead, Google and the plaintiffs create a private contract: unsettling and “arguably corrosive” to democracy (Samuelson, 2010a, p. 1358; Samuelson, 2010b). Specifically, rights-holders who do nothing or who remain unaware of the settlement’s effects may unwittingly give up rights in a way that, according to all the laws other than the settlement agreement, cannot occur. It is not surprising that the settlement, which covers controversial legislative territory, should excite such a broad spectrum of disagreement among its many potential participants. There is a range of opinions from publishers, authors, librarians, and legal scholars: some copyright progressives object because the settlement provides too-limited access, some because it provides too much.

The case proceeded to settlement not only because the broad commercial ambitions of the parties stretched copyright, but because copyright legislation cannot now address the questions this digitization project raises. The current copyright regime has accorded owners generous term lengths, but in removing the formality of registering copyright, it leaves would-be users with no practical recourse to find those owners. Crippling transaction costs, for orphan works if nothing else, shut off public access. One reform might be shorter copyright terms: these would still support innovation, but would not persist past their useful lives to clutter the paths of those who would build on that innovation. Shorter copyright terms would also render this settlement much less contentious simply by transferring more works to the public domain sooner and removing the rights-holders from the negotiation, but it would be difficult for some rights-holders to adjust their expectations and their business models’ dependence on long terms of exclusive exploitation and withstand any temporary financial hardship that could impose. For some, such hardship would leave permanent effects.

If term reduction is too dramatic a change, at the very least progressive orphan works legislation would stimulate competition to Google and open the field to public, non-profit collections. It is hard to defend the value of copyright protection that, by unintentional default, locks so many works away from any uses. Copyright extension, as James Boyle (2008) outlines, imposes a reduction in social welfare or public value while providing little additional real value to the creator. Some period of protection will encourage the labours of invention, but it seems unlikely that the slight potential value of 50- or 70-year terms as opposed to, say, 30-year terms will affect the creator very often. Many works’ commercial lives expire well before their thirtieth year, let alone 50 years after the author’s death. Those additional years will always take something away from the public.

Term reduction might also be mitigated and complemented by requiring creators and producers to register their copyright but allowing them to renew its term. Those who want and are able to extract commercial value would enjoy the protections of copyright, but works whose commercial value is spent could easily be used and appreciated for their cultural, historical, and educational values. The orphan works problem, and the need for orphan works legislation, would disappear. Such a system might require a clearer articulation of moral rights than U.S. laws currently include. It would certainly bring with it new challenges: the lack of automatic protection would

impose a burden on rights-holders, and some rights-holders might be taken advantage of and find their labours exploited by others. These spectres should not overshadow the failings of present copyright legislation, which is not immune to similar problems of its own. A default of improved access with the option of control, rather than a default of control with the option of public or fair use, may better suit the technological climate that intellectual property works now inhabit. Once practices and expectations are realigned, such a system could be as effective as the current one, though subject to different particular vulnerabilities. Copyright reform along these lines seems unlikely, since automatic protection with no formality (that is, no requirement to register) is one of the fundamental principles of the Berne Convention, now “the foundation of modern copyright law” in most countries (McCausland, 2009, p. 379).

Litigation in this case could improve interpretations of fair use and perhaps resolve some of the ambiguity of those provisions, but if the parties implement a settlement, everyone—libraries, publishers, authors, competitors to Google, the public—is denied this resolution. It is telling that the Authors Guild and Association of American Publishers were keen to expand Google’s licence and overlook the purported infringement. Samuelson (2010a) claims that she and many other legal scholars have examined the situation and believe the fair use argument is sound; Lessig concurs, though he cautions that litigation was no sure thing, since “courts have been known to reach the wrong conclusion in copyright cases” (2010, p. 3)—although in copyright cases, at least one side is likely to believe the courts arrived at the wrong decision regardless of the outcome. Even if the court concluded that Google’s actions were not fair use, it is hard to imagine that this interpretation would stem the digital threats to rights-holders’ traditional commercial interests. A settlement, however, offered Google and the plaintiffs much more control in crafting a digital update to copyright protections than they could have expected in the legislative process. For the parties, if not for the public, the settlement is the safe choice.

The settlement represents, as any settlement must, a number of compromises for publishers, authors, libraries, and readers. But it also staves off the more frightening compromise that publishers and authors may soon find themselves forced to make. The settlement carves a recognizable model of rights, permissions, and licences into the wildly unknown territories of digital publishing. Google, behemoth of the online world, with a vast network of incremental ad revenues threaded through it, may be the best bet publishers have for monetizing their works online. Rights-holder revenues will likely be modest for most, but digital publishing has so far rarely posted high returns (except for those who provide the access, the platform, or the devices to reach it); any publishing, for most who pursue it now, rarely leads to high returns. In the short term, the settlement spares publishers from the need for, and deprives them of, a useful blueprint to develop their own business models or holistically address the underlying rights-sharing issues digital possibilities raise. But it also lets publishers run a large-scale experiment using primarily backlist or out-of-print titles, with the option to add current, in-print titles to the mix when they like. Perhaps out of new possibilities and demonstrable usage patterns, the future of publishing will emerge.

Legislation can and should be adjusted over time, but it should also take a long-range view of its subject: law-making should not be done lightly. Legislation sets regulatory

standards that endure, though it tends to be clearer after the fact whether those standards were well or ill chosen. Whatever the outcome of the settlement, it does not absolve legislators of continuing to update copyright laws to reflect the needs and practices of both those who create and produce expressions of ideas as well as the public interest that attends, and builds on, these ideas. Sally McCausland, in her discussion of the settlement, orphan works, and mass digitization, observes, “The Google scheme inherently relies on regulatory intervention—in this case via the class litigation rules—to underpin it. The copyright system itself is a legal intervention in the market, which is in a constant feedback loop responding to new technologies and the uses they permit” (2009, p. 389). Copyright legislation is not neutral, natural, or static: it is a device crafted by governments to both protect and serve all the individuals who make up the public. When it ceases to function appropriately, we must demand—and allow—that it change.

### **Conclusion**

The Google Books settlement proposes to accomplish certain things in a relatively efficient and practical manner, namely the mass digitization of books and the rescue of orphan works from their current invisibility. However, criticisms of the settlement abound: on legal grounds, from the technicalities of class action validity to the settlement’s revision of the fundamental legal basis for copyright; on economic grounds, such as whether it creates a monopoly or compensates all parties fairly; and on cultural or moral grounds, including its weak requirements of Google as cultural steward and the precedents the settlement could set for the ongoing development of copyright law and modes of public access. Rights-holders contemplating the settlement should recall that the Authors Guild and the Association of American Publishers went from articulating a specific infringement complaint to appointing themselves agents—for their own members, but for many more individuals and entities who are not members of these organizations—in negotiating a huge parcel of digital rights. The reception of the settlement thus far makes it clear that not all rights-holders feel their interests are well represented in the settlement agreement.

Authors, especially those whose books are out of print, have expressed various responses to the settlement and must appraise for themselves whether its benefits outweigh its risks, costs, or perceived wrongs. Economic interests will play a role, but so will authors’ desires for audience reach or public reputation; the priority of these possibly conflicting interests will vary from individual to individual. Publishers are not solely motivated by economic interests either. Many build their lists with goals other than profit: they want to effect cultural change, advance public knowledge, or support and contribute to the creation of beautiful, powerful works. But a publishing house is also a business, with rights to perhaps hundreds of titles; in this context, many publishers may be reluctant to dismiss the settlement’s financial and other incentives. The settlement retains a role for publishers in creating an online distribution model. It also offers certain commercial opportunities that may make financial sense from a pure business perspective—although time will tell if whatever revenues a rights-holder earns justify the administrative burden of maintaining claim information.

If they have not opted out, publishers (and authors) should register their claims before the March 31, 2011, deadline so as to maximize their economic benefits and receive the one-time cash payment for each book Google digitized without authorization.

Publishers should then evaluate their strategies for selling books online: once they opt in, rights-holders must still decide how they want to participate in the settlement or collaborate with Google. The settlement offers a convenient platform for selling into every library in the U.S. and for reaching individual U.S. consumers, but the revenues from these sales may be relatively modest. Publishers must consider their other options for promoting or selling backlist titles in particular. The settlement does not affect any book published after January 5, 2009, and while publishers may have their own plans to create digital formats for recent and upcoming releases, limited resources might mean that older titles do not warrant special efforts. At the same time, highly specialized titles may command high sales prices so long as they are simply discoverable online, in which case the publisher may prefer to sell e-books or print copies itself, using Google to generate traffic but not to broker sales. Other titles may show better revenue potential from incremental usage fees through institutional subscriptions.

If a publisher decides that the incremental income of the settlement's three revenue models—online text ads, institutional subscription fees, and consumer purchases—is not appealing, that publisher could remove its books completely, exclude them and retain the ability to experiment with the revenue models at a later date, or even consider alternative agreements with Google. Large publishers may be able to negotiate individual agreements, but for the many small and medium-sized enterprises contemplating the settlement, the Partner Program is a viable alternative. The Partner Program offers a limited revenue model—only ad revenues—but also does not require as vigilant management since Google can only use the scanned book in the ways the publisher instructs: no default licences apply based on commercial availability or other shifting criteria. Although the Partner Program avoids the risk of Google unexpectedly applying a default licence, a publisher that collaborates with Google through the Partner Program is barred from the institutional subscription or consumer purchase options.<sup>62</sup> However, another economic advantage of the Partner Program is that any revenues earned are only divided between Google and the rights-holder; they do not support the operations of the Book Rights Registry.

For maximum flexibility and control, a publisher might opt in to the settlement, exclude books (and include them later if Google Books proves financially successful), and maintain an agreement under the Partner Program, which would allow the publisher to earn some revenues even while it waits to see how the settlement's other revenue models work in practice. Partner Program revenues—all of which are ad revenues—are derived from indexing a book to determine the presence of certain words or phrases and do not depend on selling access to the book's content.

A publisher can have a Partner agreement with Google and be a member of the settlement class. Individual books, however, can only be subject to one agreement at a time. Cash compensation is available, for a limited time, regardless of each book's governing agreement. To enjoy any other terms the settlement proposes, rights-holders can join the class at any time by registering their claims: there is no time limit to opt in. The settlement's terms and revenue models will apply to all books not covered by a separate agreement with Google, which means rights-holders should be free to make other arrangements with Google (like the Partner Program) or to terminate those other arrangements and resort to the settlement agreement at any time.<sup>63</sup>

The settlement may be an efficient addendum to existing copyright practices and laws, and it may offer concrete economic benefits to individual rights-holders, but these practical advantages may not be sufficient to establish the settlement as a fair and appropriate solution to copyright's shortcomings or the challenges of digitization. The settlement may serve publishers' and authors' individual or immediate interests even as it erodes their collective and long-term ones. The public, too, has a significant vested interest in the subjects of the settlement—the books themselves, repositories to centuries of knowledge and creativity—as well as the legal and cultural environment the settlement endorses. A detailed account of the settlement's economic and cultural costs and benefits is instructive, but more importantly the settlement highlights the structural and technological deficiencies of existing copyright law. Long copyright terms and the presumption of total rights protection have created a copyright regime that privileges the potential for commercial exploitation regardless of whether that exploitation is feasible or even desired by the creators themselves. This regime is also particularly ill equipped to recognize digital possibilities. Whatever happens to this settlement, such tensions continue to strain copyright's rules.

A number of conditions on approval could address criticisms of the settlement, but perhaps the best way to ensure Google, publishers, and authors are all treated fairly is to pursue copyright reform, not private contracts, to address the legislative problems that the settlement tries to engage. Legislative changes with respect to intellectual property rights have been slow to reflect everyday technological realities. The existence of the settlement, and much of its reception, demonstrates that private interests and public appetites are eager to move beyond the limits of the current regulations. Copyright reform will be fraught with challenges of its own, but the existing legal framework—in Canada as in the U.S.—is increasingly inadequate for accommodating common and emerging practices and capabilities: copyright law has swung out of balance. The settlement may serve as an early test bed for certain possibilities, including digital distribution and access, and the imposition of limited formalities on rights-holders. However, as a private contract, it is an insufficient guide for legislative development. The trouble with copyright does not affect Google alone. The public interest demands more broadly applicable solutions, and these will be achieved—eventually, and possibly with great difficulty—through copyright legislation. We may get copyright reform wrong, as arguably we have done in the past, but that fear should be allayed if we also recall that we have the power to revise our legislative interventions until we get them right.

### **Acknowledgments**

Rowland Lorimer and John W. Maxwell of Simon Fraser University read an earlier version of this report; their queries and challenges were much appreciated. I am also grateful to Peter Milroy at UBC Press for inviting me to assemble the Press' settlement claim. Finally, I would like to thank the two anonymous reviewers for their thoughtful comments.

### **Notes**

1. Sergey Brin and Larry Page's core search technology sprang from a graduate school project focused on digital libraries, in which indexing and analysis of data interrelationships could help users find and evaluate the relevance of books.
2. Letters of objection or of support from any class members who did not opt out were accepted until January 28, 2010. The DOJ had until February 4, 2010, to submit its

- statement on the amended settlement to the court. The parties then also filed briefs on February 11, 2010.
3. See Amended Settlement Agreement, 2009, “Attachment N: Supplemental Notice,” for a summary of amendments.
  4. The board will include an equal number of author subclass representatives and publisher subclass representatives, and each country will be represented by at least one director from each subclass.
  5. Technical terms are indicated with italics at first appearance. The settlement’s definitions of these terms can be found in Amended Settlement Agreement, 2009, article 1.
  6. Procedures for unclaimed funds are described in full in Amended Settlement Agreement, 2009, section 6.3.
  7. Amended Settlement Agreement section 3.8 eliminated the “Effect of Other Agreements” clause and now only includes “Effect of Changes in Law,” which allows Google to take advantage of future legislative changes that may provide better terms, such as with respect to orphan works.
  8. The DOJ makes eight specific recommendations, some of which follow directly from these four.
  9. Since rights-holders have until March 31, 2011, to claim their works, Google Book Search would not be fully available until at least some time in 2011, assuming no further legal or technical barriers. Realistically, legal and technical barriers are both likely.
  10. As discussed below, in practice rights-holders will not receive this full share.
  11. Samuelson (2010a) describes the U.S.\$60 compensation for books that were digitized without authorization—only available to those who register a claim before March 31, 2011—as a reward for or encouragement of early registration.
  12. The Authors Registry in New York collects and remits fees for use, but does not negotiate these fees or represent authors.
  13. Kevin Poulsen (2009b) describes the “last time Google assembled a giant library that promised to rescue orphaned content for future generations. And the tattered remnants of that online archive are a cautionary tale in what happens when Google simply loses interest.” He is referring to Usenet, an archive of online articles from the 1980s and 1990s; despite more than a year of inaction on these known problems prior to Poulsen’s article in *Wired*, Google responded immediately to address his charges (Poulsen, 2009a).
  14. Carr details the ease with which two reporters used anonymized AOL searches to positively identify the searcher (along with her address and phone number) and other examples of how easily the Internet reveals the personal data of its users.
  15. Gmail “reads” users’ personal email messages in order to serve relevant ads, and it met with some resistance when it was launched for this reason and for its initial omission of a delete-message function. The social networking tool Buzz launched with a misstep when it automatically populated user profiles (essentially publishing users’ address books) and offered no way for Gmail users to decline to participate, though Google adjusted the service in response to criticisms. Tim Wu (quoted in Auletta, 2009), a law professor at Columbia University, believes Google has a “total deaf ear to certain types of issues. One of them is privacy”; the people at Google “just love that data because they can do neat things with it” (p. 198).
  16. Nicholas Carr describes the oscillations between freedom and control in the development and application of information technologies, concluding that history demonstrates “that the most powerful tools for managing the processing and flow

- of information will be placed in the hands not of ordinary citizens but of businesses and governments,” and that their interest is control (2008, pp. 208–209).
17. However, authors may still argue that digital rights not contemplated in the original contract remain outside the contract. *Random House, Inc., v. Rosetta Books LLC* (2001) suggested that the right to publish a work in “book form” does not necessarily extend to e-books or digital rights; the idea that exclusive licences do not necessarily include future new uses was upheld in the appeal (*Random House, Inc., v. Rosetta Books LLC*, 2002).
  18. Google is guaranteed 37% of these revenues. Publishers and authors together may receive more than 37%, and authors of out-of-print works, who need not share with publishers, may receive a larger portion. Depending on the royalties or splits negotiated in their contracts, publishers’ portions of the proceeds from digital rights activities may or may not be more than half of the 63% they share with authors. It is also unclear how much of the 63% reserved for rights-holders will trickle down to rights-holders after the Registry subtracts its own undefined expenses.
  19. See Amended Settlement Agreement, 2009, “Attachment A: Procedures Governing Authors and Publishers,” sections 6.2 and 5.5. For out-of-print works published before 1987, 35% of any revenue will go to the publisher and 65% to the author; if published in 1987 or later, the split will be 50/50. In these cases, the revenues will be distributed by the Registry directly to each rights-holder. For in-print works, the Registry will remit revenues to the publisher only, who will accord the author whatever splits her contract specifies. UBC Press considers the settlement’s cash payments to be subsidiary rights revenue and will pay its authors 50%.
  20. The original settlement required parties to give up their rights to sue should any disputes arise and instead submit to binding arbitration. This has been relaxed: parties may agree to resolve disputes other than through the suggested arbitration process (Amended Settlement Agreement, 2009).
  21. Although both Partner Program and Library Project scans end up searchable in Google Books, we should not conclude that the two scanning programs are equivalent. Library Project scans result in the creation of *library digital copies* (LDCs) such that Google provides the libraries with electronic versions of their physical collections (see Amended Settlement Agreement, 2009; Google, 2010c). Scholarly publishers see research libraries as the natural market for purchasing their e-books. UBC Press does not bundle its e-books with print purchases, but in fact charges the same rate for an e-book as for a hardcover book. While e-books are still not widespread, some libraries are undertaking bulk purchases of e-book versions of current and recent books (e.g., the Canadian Research Knowledge Network project). Google’s provision of an LDC could undercut the library market for purchasing e-books. The amended settlement agreement (2009) specifies that libraries can only use their LDCs to replace a physical copy if a replacement copy cannot be obtained new for a reasonable price. Display and access uses are allowed for users with print disabilities or as excerpts for educational or classroom use. The LDCs could be valuable as finding tools or in non-consumptive research. For books out of copyright, the LDCs would add significant flexibility and richness to the collection, since they could be used much more expansively.
  22. The amended settlement agreement states that “the payments required to be paid to the Registry set forth in Article IV (Economic Terms for Google’s Use of Books) shall not apply to those Books and the payments set forth in the applicable direct agreement

- will apply” (2009, s. 17.9). Article 4 governs the revenue models, and the cash payments are covered separately in article 5; thus even books also included in a separate direct agreement are eligible for cash payments if all other criteria have been met.
23. Section 1.2 of “Attachment C: Plan of Allocation” of the amended settlement agreement (2009) specifies a target one-time *inclusion fee* of U.S.\$200 per book; books may also accrue *usage fees* (s. 1.1). The Book Rights Registry will hold collected subscription revenues for the first 10 years, with 25% apportioned to the inclusion fee subfund (this amount is further divided, with 80% earmarked for books and 20% for inserts), and 75% apportioned to the usage fee subfund. If the 10-year target for a book’s inclusion fee is U.S.\$200, extrapolating these figures suggests around U.S.\$750 in usage fees (on average, for a 10-year period; particular books could vary from nothing to much more). The BRR will also take its cut before these revenues are shared between author and publisher. In terms of income earned, UBC Press may well see higher returns selling one or two copies of each backlist title per year than on a title’s inclusion in the institutional subscription.
  24. This accounts for U.S.\$79.5 million of the U.S.\$125 million settlement. The other U.S.\$45.5 million will go to the plaintiffs’ lawyers. Pamela Samuelson (2009b) questions whether these legal fees are excessive. Only two law firms are involved, yet the legal fees are more than the amount reserved for all the authors and publishers combined. This could be interpreted as an abuse of the class action process.
  25. While the site’s general information is available to all, the searchable database is only available to signed-in account holders. Creating an account is simple and has no fee or other barrier.
  26. The list is not publicly available, though presumably the database reflects it.
  27. Dan Clancy (2009) confirmed the collection had topped 12 million books in his remarks at New York Law School’s D is for Digitize conference. This may not be 12 million unique works, as Google may scan more than one copy (Samuelson, 2010a, p. 1308 n. 2). The results for UBC Press suggest this is definitely the case; in addition, Google may count items that are not books at all, such as ISSN or the ISBN for a multi-volume set, or erroneously group separate books as variants of the same work, as in its treatment of UBC Press’ annual volumes of *Canadian Yearbook of International Law*.
  28. Google’s database can have multiple identifiers—ISBN, LCCN in several versions, and OCLC in several versions—in the identifier field for a single entry. The reports that the database generates use only one identifier and identifier value, and as such do not offer any clear way for a rights-holder to confirm whether identifiers that represent the same work or edition are linked, or to link database entries with no identifiers to their correlates. In an email message to the author on January 22, 2010, Kory Kennelly, a project manager with Rust Consulting, confirmed that Google is working on ways to link editions and translations (Kory Kennelly, Project Manager, Rust Consulting, Minneapolis, MN, personal communication, January 22, 2010). As of January 2010, individual rights-holders are not able themselves to add or correct bibliographic data (ISBN, title, author, publisher name, publication year), although Google invites rights-holders to submit lists of errors that it will correct itself in due time.
  29. Two books whose content is identical, except for ISBN or format, are treated as one work with only one cash payment. Two books that contain the same principal work but also contain other variant materials (such as different forewords or annotations) could be eligible for additional cash payments, though possibly only once as a book, with other portions treated as inserts.

30. If a revised edition is commercially available, previous editions, even if out of print, will be treated as commercially available as well in setting default display uses.
31. In an email March 10, 2010, Kory Kennelly explained this dramatic shift: in November 2009 Google updated the database “to reflect books that were scheduled for digitization, but ended up not being digitized. In this update, Google reviewed the digitization status of books to further ensure the accuracy of the designation, and subsequently made further updates. Google has confirmed that the current digitization status of the books UBC [Press] is inquiring about are accurately listed in the books database” (Kory Kennelly, Project Manager, Rust Consulting, Minneapolis, MN, personal communication, March 10, 2010). The Press has no way to verify this claim.
32. There is no reason to assume a Partner book would only be scanned through the Partner Program. The Google Books overview page may include digitization source and date for Library Project scans, as seen with the book *Killer Whales: The Natural History and Genealogy of Orcinus orca in British Columbia and Washington State*. UBC Press initially provided the 2000 edition through the Partner Program with limited preview and later requested no preview. Google also has the 1994 and 2000 editions available in snippets only; these two include, under the “More book information” section, two additional metadata fields: “Original from” and “Digitized.” Both books were digitized from the University of California on August 7, 2009. See the appendix for screen shots that illustrate these variations.
33. The claims process and settlement rules for inserts are not necessarily the same as those for books. Since UBC Press does not have any inserts that it will not already claim as books, this report deals only with books.
34. “Highly confident” means the publisher bases its claim “on the individual Book or contract for the Book”; merely “confident” means the publisher bases it “on the type of Book or type of contract for the Book”—that is, relying on generic data rather than specific and confirmed data (Amended Settlement Agreement, 2009, s. 13.1[c][ii]2).
35. Although the Registry identifier appears to be a convenient stable marker, allowing revisions as necessary to any of the ISBN, title, or author fields, this may not be the case. In an email on January 29, 2010, Kory Kennelly confirmed that the Registry identifier may not be static and is primarily for Google’s own use in maintaining the database (Kory Kennelly, Project Manager, Rust Consulting, Minneapolis, MN, personal communication, January 29, 2010). However, results with UBC Press claims indicate that a claimed book always uses 103 characters, while an unclaimed book’s identifier varies in length and is typically 200 to 300 or more characters.
36. The settlement requires that Google pay at least U.S.\$45 million in cash payments, with minimums of U.S.\$60 per principal work, U.S.\$15 per insert, and U.S.\$5 per partial insert. (An insert could be an essay, chapter, short story, foreword, quotation, excerpt, or other non–book length work contained in a collection or within or alongside the book’s principal work [Amended Settlement Agreement, 2009].) If a large number of rights-holders claim works, Google may have to pay out more than U.S.\$45 million in order to award the minimums required. If few rights-holders claim works, Google will increase the individual payments up to maximums of U.S.\$300, U.S.\$75, and U.S.\$25 until the full U.S.\$45 million is disbursed (Amended Settlement Agreement, 2009). It seems unlikely that rights-holders would receive substantially more than the minimums, since U.S.\$45 million would cover only 750,000 books (or perhaps 500,000 books and 1,000,000 inserts). As of May 5, 2009, the digitization cut-off date for cash payment eligibility, Google

had scanned between seven and twelve million books; even discounting public domain works, Partner Program works, and duplicates, claims could easily surpass U.S.\$45 million. As of February 8, 2010, 1,125,339 books and 21,829 inserts had been claimed, though not all of these are necessarily eligible for cash payments. These interim figures on claimants and claims are drawn from the settlement administrator's declaration (see Allen, 2010a).

37. The pre-set tiers—specific percentages of all books at 12 set price-points—suggests price-fixing to some critics. James Grimmelman points out that placing a predetermined portion of books in each pricing bin is “mathematically incompatible” with an algorithm that also sets the price for each book independently in a simulation of the competitive market (2010, p. 5).
38. These three modes of access are the only access uses proposed for immediate implementation. The settlement also proposes three additional access uses as possible new revenue models. These may or may not be established in future: print on demand, to create physical copies of books not otherwise commercially available; file download, such as PDF or EPUB files for personal use; and consumer subscription, to offer individuals access to the full collection or a subset of the collection, similar to the institutional subscription (Amended Settlement Agreement, 2009).
39. Work for hire may be controlled solely by the publisher or other rights-holder rather than the author.
40. These issues also comprise the main complaints of the DOJ (United States Department of Justice, 2010) with respect to the author-publisher procedures, which it contends cast doubt on the validity of the class action: specifically, the author-publisher procedures overturn distinct individual contracts, which could create conflicts of interest between class members and undermine the conception of the class. Many objecting authors echo these concerns.
41. Google can challenge a rights-holder's assertion of commercial availability, but it seems that Google will not challenge the author-publisher procedure tests. (Google does not have access to author-publisher contracts, so it would have no grounds.) To re-classify a book based on its contract terms rather than its commercial availability, a publisher could be required to produce the contracts for the Registry, especially if there are disputes between publishers and authors over in-print status (Amended Settlement Agreement, 2009).
42. Google's obligations to remove books are specified in ASA s. 1.126. If a book has not yet been digitized when a removal request is made (at any time), Google will use reasonable efforts not to digitize the book at all. Removal requests made after March 9, 2012, will only be honoured if the book has not already been digitized; otherwise, the rights-holder will only be able to exclude a book (Amended Settlement Agreement, 2009).
43. The coupling requirement works to Google's advantage. Publishers already have ways to negotiate bulk deals with libraries, but have fewer channels through which to reach individual consumers online. If publishers want to maximize their use of Google to reach individuals, they must give Google what it wants, which is to build a large collection to support subscription sales.
44. As do many scholarly publishers, the Press invests significant editorial work (as well as production and promotion) in its books and may derive just enough revenue to cover its direct and often substantial costs; academic authors may expect little

- commercial reward from book sales, but still enjoy tangible benefits from its publication through tenure or promotion regardless of its sales levels.
45. All the academics with whom Samuelson has spoken (at a minimum, presumably the 150 co-signatories to her objection letters) have “to a person” agreed that they would prefer their out-of-print books to be open access (Fairness Hearing, 2010, p. 55).
  46. The unaudited 2009 figures list total revenues of U.S.\$23.65 billion and net income of U.S.\$6.52 billion.
  47. Clancy claims out-of-print books represent only 3% of the book market.
  48. This figure was stable throughout 2008 and 2009.
  49. See also Tim O’Reilly on Web 2.0 business models: in the section “Data is the Next Intel Inside,” he characterizes data as “a sole source component in systems whose software infrastructure is largely open source or otherwise commodified,” while “control over the database has led [or can lead] to market control and outsized financial returns” (2005, p. 3).
  50. Dan Clancy claims “Google has spent hundreds of millions of dollars researching, developing, patenting and implementing cutting edge digital scanning technology” since 2004, but he does not further specify these expenses (2010, p. 1). It seems likely that Google will be able to apply much of its research, patents, and technologies to other applications beyond rights-holders’ works within the settlement.
  51. The breakdown of Google’s 37% share to include 10% for Google’s operating costs is described in ASA, s. 2.1(a) (Amended Settlement Agreement, 2009).
  52. Examples include applications such as Adobe Digital Editions and Texterity. Arguably, even Google’s security measures to control display will offer limited resistance against hackers and ongoing technological developments, including applications that convert any screen capture into a PDF file and permit users to “rip” copies of DRM-protected e-books or books online, as detailed by Lazebnik (2010). A business model that relies on DRM will constantly be fighting to stay a step ahead of users’ ingenuity at breaking locks.
  53. In 2009, Access Copyright’s distributions totalled 76.01% of its revenues. In 2008, distributions rose to 81.8% of revenues, but that year included the distribution of several years’ accumulated unclaimed funds. In 2007, distributions represented only 62.64% of revenues; in 2006, 64.94%. Total revenues in this period were as high as \$37.285 million in 2007 and as low as \$34.194 million in 2006. Access Copyright considers only some of the revenues not distributed as operating expenses or other expenses; these expenses account for between 18% (2008) and 24% (2006 and 2009) of total revenues (Access Copyright, 2008; 2009; 2010a). The variations between revenues not distributed and the target 21.5% may be due to undistributed royalties, the time-lag between collection and distribution, or other circumstances. The discrepancy between revenues not distributed and annual expenses has been relatively high at times, but in 2008 and 2009, it seems to have disappeared: the revenues were equivalent to total expenses plus distributed royalties of the same year.
  54. As of February 8, 2010, 4,392 of the 44,450 claimants were publishers and a total of 1,125,339 books were claimed. These figures include both online claims and hard-copy claims. Allen offered additional details on online claims. There were 1,107,620 books claimed online, of which Google had designated only 488,089 as commercially available and therefore in print. Of the books claimed online, publishers claimed 787,942. Exhibit D (Allen, 2010b) shows summary data on numbers of claims and claimants, but note that the first table contains an error and

- transposes the number of online claiming accounts of agents and publishers, which also affects the averages cited in this table.
55. Lazebnik represents the Science Fiction and Fantasy Writers of America, the American Society of Journalists and Authors, and the National Writers Union. Again, the concerns of these primarily non-academic authors and the relationships they have with their publishers could be quite different from those of the many academic authors whose works fill the participating libraries' shelves.
  56. For example, UBC Press will be aware of which authors have not claimed their works and could alert or assist those authors. Revenues for books deemed out of print are paid separately to authors and publishers, but revenues for books deemed in print are paid to publishers and flow through to authors, as with royalties. UBC Press may thus pass along revenues from the settlement even if individual authors do not want to invest time in managing their claims themselves, although the settlement does not explicitly describe what happens if the publisher claims an in-print book and the author is unlocated by the BRR. In this case, the publisher should receive only its own share until the author registers with the BRR (Amended Settlement Agreement, 2009).
  57. The Supreme Court raised these points in its decision in *Théberge v. Galerie d'Art du Petit Champlain inc.* in 2002, and in *CCH Canadian Ltd. v. Law Society of Upper Canada* in 2004. Previously, fair dealing was frequently conceived "as nothing more than a limited defence to infringement" (Murray & Trosow, 2007, p. 76).
  58. Lessig describes the lack of common sense-compatible rules as a contributing factor to the growing extremism from rights-holders (or, more accurately, their corporate designates) who seek to restrict completely, totally, and automatically all uses of works, and users who balk at these restrictions and penalties and reject copyright entirely. What unfolds, Lessig concludes, is that "ordinary people live life against the law ... knowing they live it against the law. That realization is extraordinarily corrosive" to both democracy and business: ultimately, neither the public nor private interests will be served by this approach. And yet this is the approach the U.S. has adopted so far.
  59. Canadian copyright law does not include a similar visionary purpose to contextualize its form and rules: the "Canadian Constitution merely lists copyright as an enumerated power of the federal government, with no rationale or guidance provided" (Murray & Trosow, 2007, p. 209 n.12).
  60. The Internet Archive was founded in 1996, and OCA was initiated by the Internet Archive and Yahoo! in 2005.
  61. Rights-holders may set their own prices, but absent specific instructions, Google's defaults will apply (Amended Settlement Agreement, 2009).
  62. This barrier may not be permanent: other forthcoming Google services or new agreements, such as Google Editions, may include versions of consumer purchase and even subscription access of some kind.
  63. If a book was removed from the settlement—the permanent designation, stronger than exclusion—Google might decline to digitize it again under the same terms. However, the settlement explicitly provides for unlimited changes between include and exclude; this flexibility is the main advantage exclusion has over removal. Partner Program agreements can be terminated by either the publisher or Google, and the settlement by default encompasses all books whose rights-holders did not opt out.

## References

- Access Copyright. (2008). *Annual report 2007*. Toronto, ON: The Canadian Copyright Licensing Agency. URL: <http://www.accesscopyright.ca/docs/AnnualReports/2007AnnualReport.pdf> [March 27, 2010].
- Access Copyright. (2009). *Annual report 2008*. Toronto, ON: The Canadian Copyright Licensing Agency. URL: <http://www.accesscopyright.ca/docs/Annual%20Report-final.pdf> [March 27, 2010].
- Access Copyright. (2010a). *Annual report 2009*. Toronto, ON: The Canadian Copyright Licensing Agency. URL: [http://www.accesscopyright.ca/docs/AnnualReports/Annual\\_Report\\_2009-Final.pdf](http://www.accesscopyright.ca/docs/AnnualReports/Annual_Report_2009-Final.pdf) [April 12, 2010].
- Access Copyright. (2010b). *Schedule of distribution payments to rightsholders*. URL: <http://www.accesscopyright.ca/Default.aspx?id=255> [July 2, 2010].
- Albanese, Andrew Richard, & Oder, Norman. (2009, May 1). Corner office: Google's Dan Clancy. *Library Journal*. URL: <http://www.libraryjournal.com/article/CA6652445.html> [January 15, 2010].
- Allen, Tiffany. (2010a, February 11). Declaration of Tiffany Allen. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). Doc. 953. URL: [http://thepublicindex.org/docs/amended\\_settlement/Allen\\_declaration.pdf](http://thepublicindex.org/docs/amended_settlement/Allen_declaration.pdf) [February 15, 2010].
- Allen, Tiffany. (2010b, February 11). Exhibit D. Declaration of Tiffany Allen. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). Doc. 953-3. URL: [http://thepublicindex.org/docs/amended\\_settlement/Allen\\_declaration\\_exhibit\\_D\\_color.pdf](http://thepublicindex.org/docs/amended_settlement/Allen_declaration_exhibit_D_color.pdf) [February 26, 2010].
- Amazon.com. (2010, January 27). Objection of Amazon.com, Inc., to proposed amended settlement. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). Doc. 823. URL: [http://thepublicindex.org/docs/amended\\_settlement/amazon.pdf](http://thepublicindex.org/docs/amended_settlement/amazon.pdf) [February 1, 2010].
- Amended Settlement Agreement (ASA). (2009, November 13). *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). URL: [http://www.googlebooksettlement.com/r/view\\_settlement\\_agreement](http://www.googlebooksettlement.com/r/view_settlement_agreement) [November 16, 2009].
- Arato, Cynthia. (2009, October 9). *L is for lawsuit*. Panel discussion at D is for Digitize, New York Law School, New York, NY. URL: [http://www.nyls.edu/centers/harlan\\_scholar\\_centers/institute\\_for\\_information\\_law\\_and\\_policy/events/d\\_is\\_for\\_digitize/program](http://www.nyls.edu/centers/harlan_scholar_centers/institute_for_information_law_and_policy/events/d_is_for_digitize/program) [January 26, 2010].
- Association of College and Research Libraries, American Library Association, & Association of Research Libraries. (2009, May 4). Library association comments on the proposed settlement. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). Doc. 100. URL: <http://wo.ala.org/gbs/wp-content/uploads/2009/05/Library-Association-Comments-on-the-Proposed-Settlement.pdf> [June 25, 2010].
- Auletta, Ken. (2009). *Googled: The end of the world as we know it*. New York, NY: Penguin.
- Authors Guild, Inc., Association of American Publishers, Inc., et al. (2010, February 11). Plaintiffs' supplemental memorandum responding to specific objections. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). Doc. 955. URL: [http://thepublicindex.org/docs/amended\\_settlement/Supplemental\\_memorandum\\_of\\_law.pdf](http://thepublicindex.org/docs/amended_settlement/Supplemental_memorandum_of_law.pdf) [March 25, 2010].
- Author's [sic] Guild et al. v. Google Inc.* (2005, September 20). Case No. 05 CV 8136 (S.D.N.Y.). URL: <http://www.authorsguild.org/advocacy/articles/settlement-resources.attachment/authors-guild-v-google/Authors%20Guild%20v%20Google%2009202005.pdf> [January 28, 2010].
- Authors Registry, The. (2008). FAQ. URL: <http://www.authorsregistry.org/disc.htm> [March 15, 2010].
- Blount, Roy, Jr. (2008, October 28). *\$125 million settlement in Authors Guild v. Google*. Authors Guild member alert. URL: <http://www.authorsguild.org/advocacy/articles/member-alert-google.html> [January 25, 2010].

- Bolt, David, Page, Kathy, Payerle, George, Hancock, Lyn, Levitt, Peter, et al. [177 Canadian authors]. (2010, January 28). Letter to the court. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). URL: [http://thepublicindex.org/docs/amended\\_settlement/canadian\\_authors.pdf](http://thepublicindex.org/docs/amended_settlement/canadian_authors.pdf) [March 25, 2010].
- Book Rights Registry. (n.d.). Claim form for members of the publisher sub-class. URL: <http://www.googlebooksettlement.com/intl/en/claim-forms.zip> [October 22, 2009].
- Book Rights Registry. (2008a). *Instructions regarding claiming template spreadsheet: How to use the spreadsheet to manage claimed books*. URL: <http://www.googlebooksettlement.com/help/bin/answer.py?answer=128054&hl=en> [January 6, 2010].
- Book Rights Registry. (2008b). What happens to the books of authors and publishers who opt out of the amended settlement? *Google Book Settlement FAQs* (Question 21). URL: <http://www.googlebooksettlement.com/help/bin/answer.py?answer=118704&hl=en#q18a> [January 25, 2010].
- Boyle, James. (2008). *The public domain: Enclosing the commons of the mind*. New Haven, CT: Yale University Press.
- Brantley, Peter. (2009, September 3). *Is "trust us" good enough?* Open Book Alliance. URL: <http://www.openbookalliance.org/2009/09/is-trust-us-good-enough/> [January 25, 2010].
- Carr, Nicholas. (2008). *The big switch: Rewiring the world, from Edison to Google*. New York, NY: W.W. Norton & Company.
- Chabon, Michael, Lethem, Jonathan, Ferlinghetti, Lawrence, Romero, Anthony D., Schneier, Bruce, et al. [28 authors and publishers]. (2009, September 8). Privacy authors and publishers' objection to proposed settlement. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). Doc. 281. URL: [http://thepublicindex.org/docs/letters/privacy\\_authors.pdf](http://thepublicindex.org/docs/letters/privacy_authors.pdf) [January 25, 2010].
- Chu, Lynn. (2009). *On the Google book settlement: Part 2*. Writers' Representatives blog. URL: <http://www.writersreps.com/feature.aspx?FeatureID=157> [January 25, 2010]. (revised June 12, 2009)
- Clancy, Dan. (2009, October 9). *I is for industry*. Panel discussion at D is for Digitize, New York Law School, New York, NY. URL: [http://www.nyls.edu/centers/harlan\\_scholar\\_centers/institute\\_for\\_information\\_law\\_and\\_policy/events/d\\_is\\_for\\_digitize/program](http://www.nyls.edu/centers/harlan_scholar_centers/institute_for_information_law_and_policy/events/d_is_for_digitize/program) [January 26, 2010].
- Clancy, Dan. (2010, February 11). Declaration of Daniel Clancy in support of motion for final approval of amended settlement agreement. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). Doc. 946. URL: [http://thepublicindex.org/docs/amended\\_settlement/dan\\_clancy\\_declaration.pdf](http://thepublicindex.org/docs/amended_settlement/dan_clancy_declaration.pdf) [February 15, 2010].
- Courant, Paul. (2009, October 9). *K is for keynote*. Presentation at D is for Digitize, New York Law School, New York, NY. URL: [http://www.nyls.edu/centers/harlan\\_scholar\\_centers/institute\\_for\\_information\\_law\\_and\\_policy/events/d\\_is\\_for\\_digitize/program](http://www.nyls.edu/centers/harlan_scholar_centers/institute_for_information_law_and_policy/events/d_is_for_digitize/program) [January 26, 2010].
- Duguid, Paul. (2010, January 7). *Is Google good for history?* Panel discussion at annual meeting of the American Historical Association, San Diego, CA. URLs: Part 1: <http://www.youtube.com/user/hnneditor#p/u/40/wfDkdhE6m80>; Part 2: <http://www.youtube.com/user/hnneditor#p/u/39/E38JOMJ-S7g>; Part 3, <http://www.youtube.com/user/hnneditor#p/u/38/xoDwiwoyaT4> [July 2, 2010].
- Fairness Hearing. (2010, February 18). *Authors Guild, [Inc., Association of American Publishers, Inc.,] et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.) [Transcript]. URL: [http://thepublicindex.org/docs/case\\_order/fairness-hearing-transcript.pdf](http://thepublicindex.org/docs/case_order/fairness-hearing-transcript.pdf) [March 7, 2010].
- Gant, Scott E. (2009, August 19). Objection of Scott E. Gant to proposed settlement, and to certification of the proposed settlement class and sub-classes. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). URL: <http://thepublicindex.org/docs/objections/gant.pdf> [January 31, 2010].

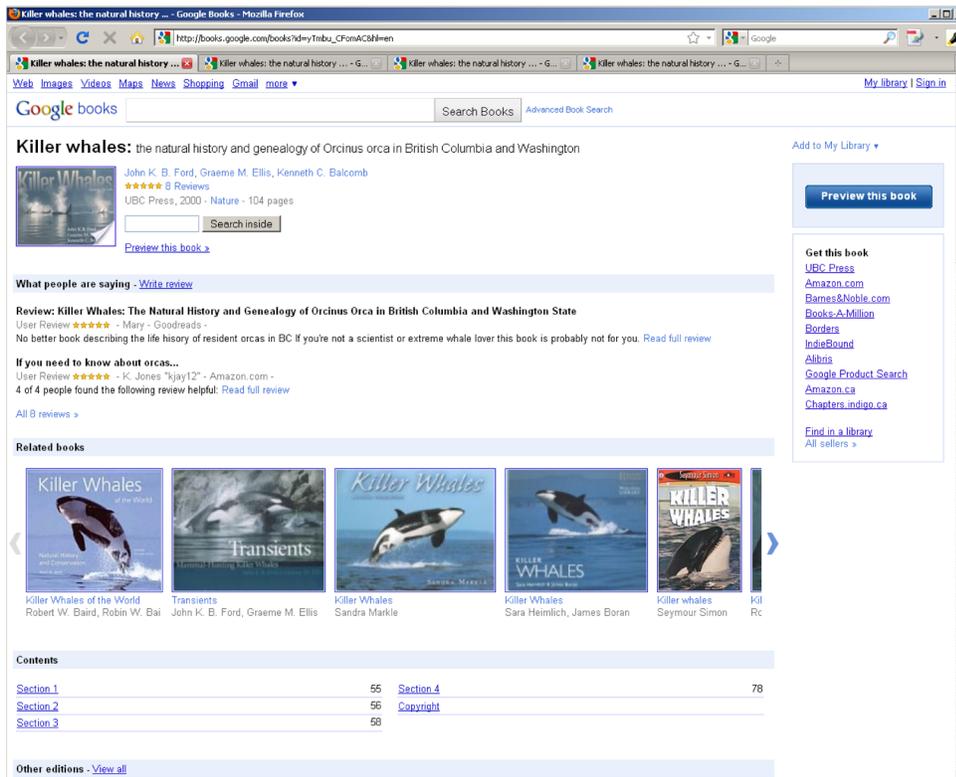
- Google Inc. (n.d.). *About Google Books: History of Google Books*. URL: <http://books.google.com/googlebooks/history.html> [January 15, 2010].
- Google Inc. (2009). *The future of Google Books*. URL: <http://books.google.com/googlebooks/agreement> [January 26, 2010].
- Google Inc. (2010a, February 11). Brief of Google Inc. in support of motion for final approval of amended settlement agreement. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). Doc. 941. URL: [http://thepublicindex.org/docs/amended\\_settlement/google\\_final\\_approval\\_support.pdf](http://thepublicindex.org/docs/amended_settlement/google_final_approval_support.pdf) [March 13, 2010].
- Google Inc. (2010b, January 21). *Google announces fourth quarter and fiscal year 2009 results*. URL: [http://investor.google.com/releases/2009Q4\\_google\\_earnings.html](http://investor.google.com/releases/2009Q4_google_earnings.html) [January 31, 2010].
- Google Inc. (2010c). *Information for publishers and authors about the Library Project*. URL: [http://books.google.com/intl/en-US/googlebooks/publisher\\_library.html](http://books.google.com/intl/en-US/googlebooks/publisher_library.html) [January 26, 2010].
- Grimmelmann, James. (2010, January 28). Letter from the Institute for Information Law and Policy. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). URL: [http://thepublicindex.org/docs/amended\\_settlement/iilp.pdf](http://thepublicindex.org/docs/amended_settlement/iilp.pdf) [February 26, 2010].
- Guthrie, Arlo, Wright, Julia, Hyde, Catherine Ryan, & Linden, Eugene. (2010, January 28). Supplemental objections of Arlo Guthrie, Julia Wright, Catherine Ryan Hyde, and Eugene Linden to the amended class action settlement agreement. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). Doc. 849. URL: [http://thepublicindex.org/docs/amended\\_settlement/Guthrie\\_objection.pdf](http://thepublicindex.org/docs/amended_settlement/Guthrie_objection.pdf) [January 31, 2010].
- Hyde, Lewis. (2009, October 1). Advantage Google. *The New York Times*. URL: [http://www.nytimes.com/2009/10/04/books/review/Hyde-t.html?\\_r=1](http://www.nytimes.com/2009/10/04/books/review/Hyde-t.html?_r=1) [January 15, 2010].
- International Federation of Library Associations and Institutions. (2009, August 28). *IFLA position on the Google book settlement*. URL: <http://www.ifla.org/en/publications/ifla-position-on-the-google-book-settlement> [July 2, 2010].
- Lazebnik, Ron. (2010, January 28). Exhibit C [5 tools to download any book from Google and save it as PDF]. Declaration of Ron Lazebnik. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). Doc. 865-4. URL: [http://thepublicindex.org/docs/amended\\_settlement/SFWA\\_Ex\\_C.pdf](http://thepublicindex.org/docs/amended_settlement/SFWA_Ex_C.pdf) [January 31, 2010].
- Lessig, Lawrence. (2007, March 8). Larry Lessig on laws that choke creativity. Open for business. Presentation at TED2007, Monterey, CA. URL: [http://www.ted.com/talks/larry\\_lessig\\_says\\_the\\_law\\_is\\_strangling\\_creativity.html](http://www.ted.com/talks/larry_lessig_says_the_law_is_strangling_creativity.html) [January 26, 2010].
- Lessig, Lawrence. (2010, January 26). For the love of culture: Google, copyright, and our future. *The New Republic*. URL: <http://www.tnr.com/article/the-love-culture?page=0,0> [January 26, 2010].
- McCausland, Sally. (2009). Googling the archives: Ideas from the Google Books settlement on solving orphan works issues in digital access projects. *SCRIPTed*, 6(2), 377–393. URL: <http://www.law.ed.ac.uk/ahrc/script-ed/vol6-2/mccausland.asp> [February 24, 2010].
- McGraw-Hill Companies, Inc., et al. v. Google Inc.* (2005, October 19). No. 05 CV 8881 (DC) (S.D.N.Y.). URL: <http://www.authorsguild.org/advocacy/articles/settlement-resources.attachment/mcgraw-hill/McGraw-Hill%20v.%20Google%2010192005.pdf> [January 28, 2010].
- Morrison, Heather. (2009). *Open Content Alliance (OCA) vs. Google Books: OCA as superior network and better fit for an emerging global public sphere*. Unpublished paper, Library Staff Papers, Simon Fraser University Institutional Repository, Vancouver, BC. URL: <http://ir.lib.sfu.ca/bitstream/1892/10651/1/OCAdec09.pdf> [March 16, 2010].
- Murray, Laura J., & Trosow, Samuel E. (2007). *Canadian copyright: A citizen's guide*. Toronto, ON: Between the Lines.
- Nunberg, Geoffrey. (2009a, August 29). Google Books: A metadata train wreck. *Language Log*. URL: <http://languagelog.ldc.upenn.edu/nll/?p=1701> [January 15, 2010].

- Nunberg, Geoffrey. (2009b, August 31). Google's Book Search: A disaster for scholars. *The Chronicle of Higher Education*. URL: <http://chronicle.com/article/Googles-Book-Search-A/48245/> [January 15, 2010].
- O'Reilly, Tim. (2005, September 30). *What is Web 2.0: Design patterns and business models for the next generation of software*. URL: <http://oreilly.com/pub/a/web2/archive/what-is-web-2.0.html?page=1> [January 25, 2010].
- Poulsen, Kevin. (2009a, October 8). Google begins fixing Usenet archive. *Wired*. URL: [http://www.wired.com/epicenter/2009/10/usenet\\_fix/](http://www.wired.com/epicenter/2009/10/usenet_fix/) [January 25, 2010].
- Poulsen, Kevin. (2009b, October 7). Google's abandoned library of 700 million titles. *Wired*. URL: <http://www.wired.com/epicenter/2009/10/usenet/> [January 25, 2010].
- Raj, Kiran. (2009, October 9). *L is for lawsuit*. Panel discussion at D is for Digitize, New York Law School, New York, NY. URL: [http://www.nyls.edu/centers/harlan\\_scholar\\_centers/institute\\_for\\_information\\_law\\_and\\_policy/events/d\\_is\\_for\\_digitize/program](http://www.nyls.edu/centers/harlan_scholar_centers/institute_for_information_law_and_policy/events/d_is_for_digitize/program) [January 26, 2010].
- Random House, Inc., v. Rosetta Books LLC*. (2001). 150 F. Supp. 2d 613, (DC) (S.D.N.Y.).
- Random House, Inc., v. Rosetta Books LLC*. (2002). 283 F. 3d 490, (DC) (S.D.N.Y.).
- Samuelson, Pamela. (2009a, August 10). The audacity of the Google Book Search settlement. *Huffington Post*. URL: [http://www.huffingtonpost.com/pamela-samuelson/the-audacity-of-the-google\\_b\\_255490.html](http://www.huffingtonpost.com/pamela-samuelson/the-audacity-of-the-google_b_255490.html) [January 16, 2010].
- Samuelson, Pamela. (2009b, April 14). *Reflections on the Google Book Search settlement* [Slides]. OCLC/Frederick G. Kilgour Lecture in Information and Library Science, University of North Carolina at Chapel Hill, Chapel Hill, NC. URL: <http://www.slideshare.net/naypinya/reflections-on-the-google-book-search-settlement-by-pamela-samuelson> [January 16, 2010].
- Samuelson, Pamela. (2010a). Google Book Search and the future of books in cyberspace. *Minnesota Law Review*, 94(5), 1308–1374. URL: [http://www.minnesotalawreview.org/sites/default/files/Samuelson\\_MLR\\_o.pdf](http://www.minnesotalawreview.org/sites/default/files/Samuelson_MLR_o.pdf) [July 6, 2010].
- Samuelson, Pamela. (2010b, January 6). *Google book settlement as private copyright reform* [Slides]. Presentation at Copyright Culture, Copyright History conference, Tel Aviv University, Tel Aviv, Israel. URL: <http://www.slideshare.net/naypinya/samuelson-gbs-as-copyright-reform> [January 16, 2010].
- Samuelson, Pamela, Aoki, Keith, Armstrong, Timothy K., Auslander, David M., Azzam, Amin, et al. [150 academic authors]. (2010, January 27). Supplemental academic author objections to the Google Book Search settlement. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). URL: [http://thepublicindex.org/docs/amended\\_settlement/Samuelson\\_supplemental\\_objection.pdf](http://thepublicindex.org/docs/amended_settlement/Samuelson_supplemental_objection.pdf) [March 23, 2010].
- Turk, James L. (2010, January 28). Objections of the Canadian Association of University Teachers to the Google Book Search settlement. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). Doc. 900. URL: [http://www.caut.ca/uploads/CAUT\\_Google\\_Settlement.pdf](http://www.caut.ca/uploads/CAUT_Google_Settlement.pdf) [March 27, 2010].
- United States. (2009, October). *Circular 92: Copyright law of the United States and related laws contained in Title 17 of the United States Code*. URL: <http://www.copyright.gov/title17/circ92.pdf> [March 10, 2010].
- United States Department of Justice. (2009, September 18). Statement of interest of the United States of America regarding proposed class action. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). Doc. 720. URL: <http://www.justice.gov/atr/cases/f250100/250180.pdf> [November 23, 2009].
- United States Department of Justice. (2010, February 4). Statement of interest of the United States of America regarding proposed amended settlement agreement. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). Doc. 922. URL: <http://www.justice.gov/atr/cases/f255000/255012.pdf> [March 30, 2010].

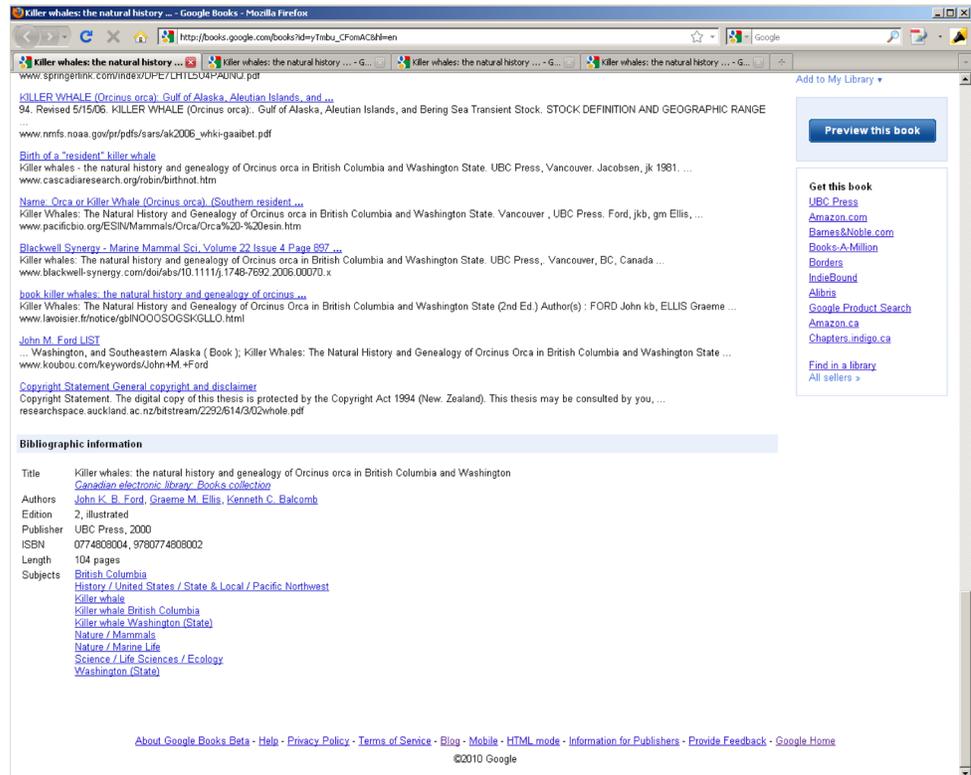
**Scholarly and Research  
Communication**

VOLUME 2 / ISSUE 1 / 2011

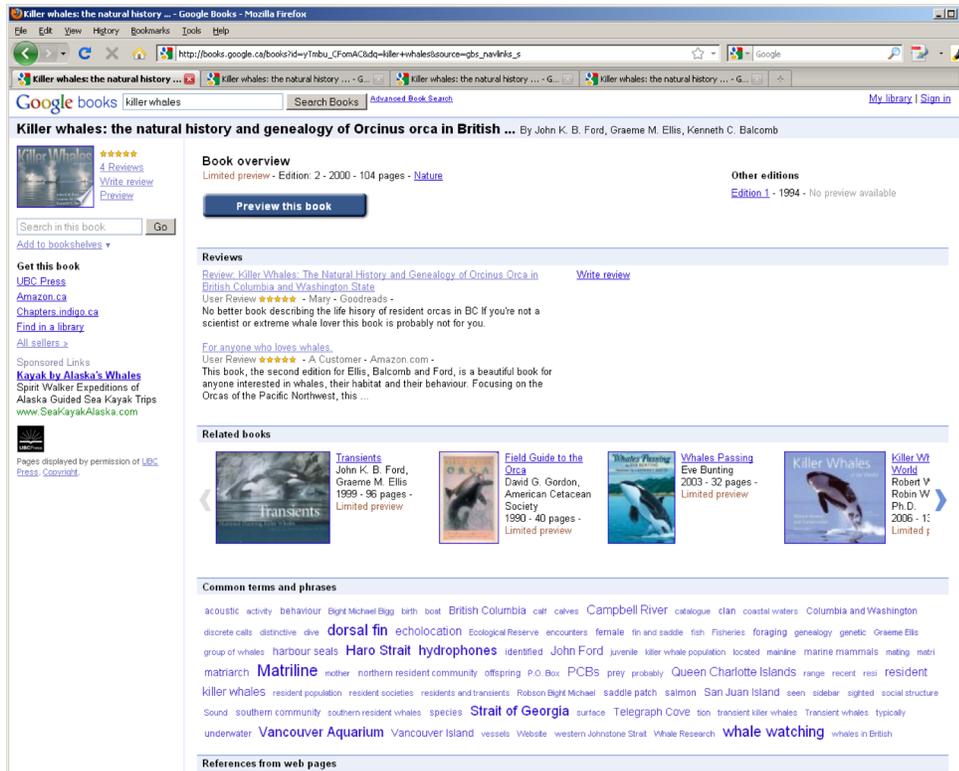
- Westcott, Grace. (2009, February 20). Googleopoly. *The Globe and Mail*. URL: <http://www.theglobeandmail.com/news/arts/books/googleopoly/article972380/singlepage/#articlecontent> [December 3, 2009].
- Willinsky, John. (2006). *The access principle: The case for open access to research and scholarship*. Cambridge, MA: The MIT Press.
- Yahoo! Inc. (2009, September 8). Objection of Yahoo! Inc. to final approval of the proposed class action settlement. *Authors Guild, Inc., Association of American Publishers, Inc., et al. v. Google Inc.* Case No. 05 CV 8136 (DC) (S.D.N.Y.). Doc. 288. URL: <http://thepublicindex.org/docs/letters/yahoo.pdf> [January 21, 2010].



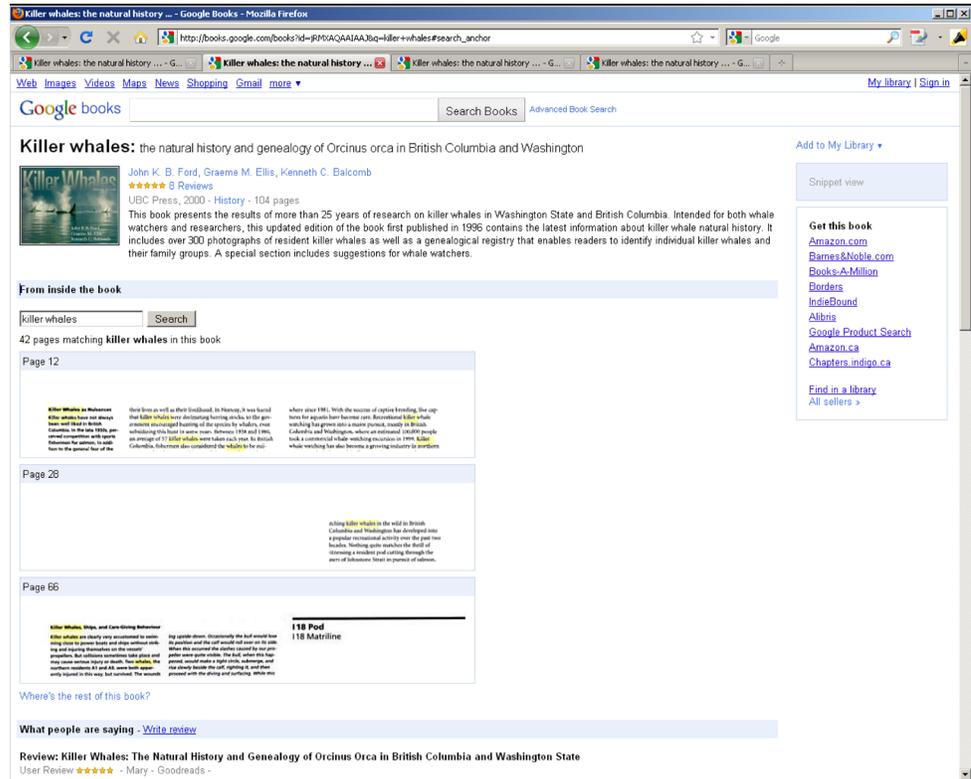
Google Book Search record for the 2000 edition of *Killer Whales: The Natural History and Genealogy of Orcinus orca in British Columbia and Washington State*, by John K. B. Ford, Graeme M. Ellis, and Kenneth C. Balcomb (unless otherwise noted, all screen shots accessed August 20, 2010). Google Books has offered both limited preview and no preview at different times; this screen shot indicates preview is available. UBC Press submitted the book to the Partner Program but has changed its preview preferences for this title. On the right, there is a “Get this book” link to UBC Press as well as other sellers and libraries.



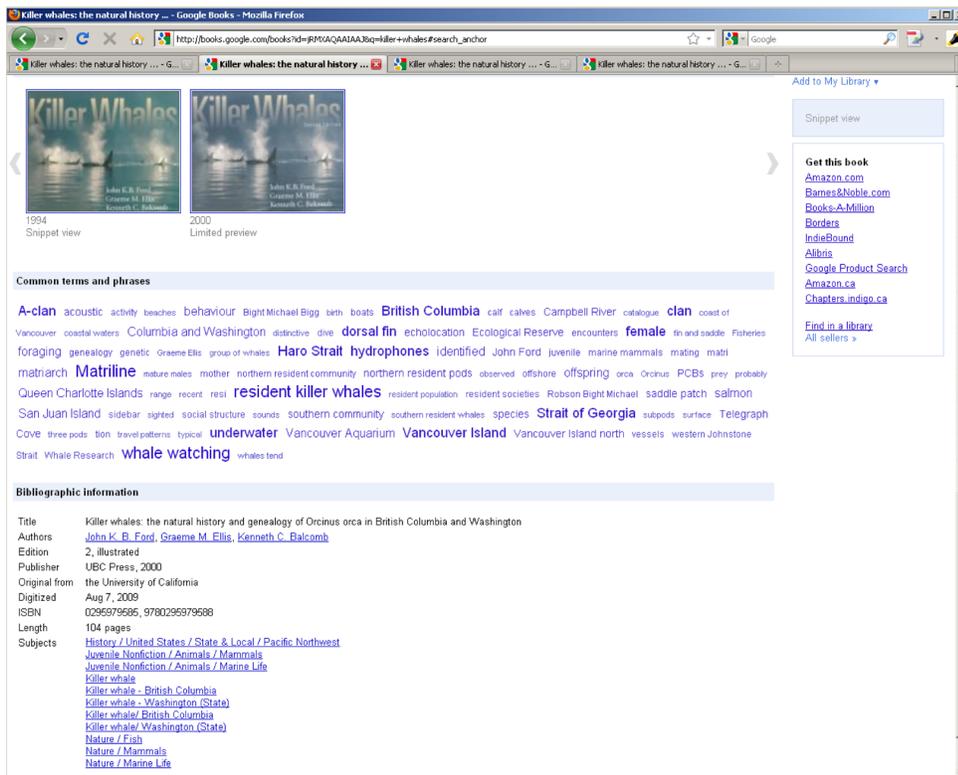
This is the same online book page as the previous image (2000 edition of *Killer Whales*), but this image shows the bibliographic data at the bottom of the page. The publisher is listed as UBC Press, and the ISBN is a UBC Press ISBN. This book was digitized without authorization, but the digitization status is only available from the settlement claims website and does not appear in Google Books. The preview button and links to sellers float as the user scrolls up and down the page.



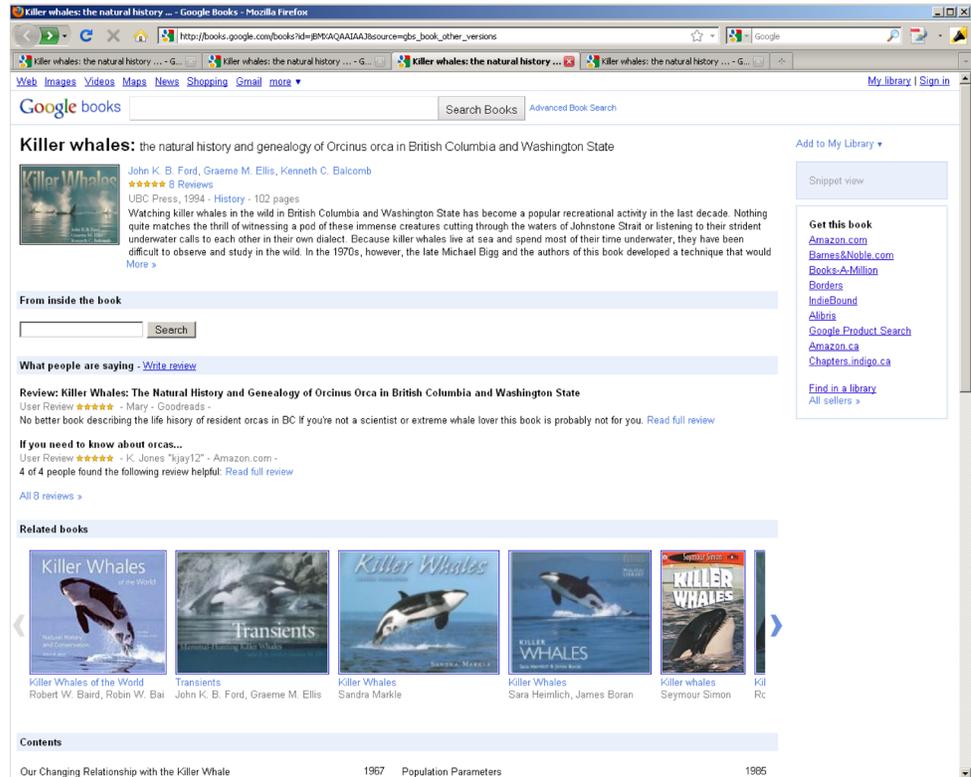
This image shows an earlier version of a book overview page in Google Books (screen shot accessed April 26, 2010). This is the same book and edition as the previous images, but note that the “Get this book” links are on the left and are accompanied by a revenue-generating text ad and the UBC Press logo. Google Books still includes the text ad and publisher logo on the book preview pages, but it has now removed these elements from the book overview page. This means that rights-holders who do not allow any preview use will not be able to earn any revenues from text ads either, and that search users who land on the “About this book” page will not generate any ad revenues unless they click through to the preview pages. This structural change may be Google’s way of encouraging rights-holders to allow previews, which allows Google to offer more book content to its search users.



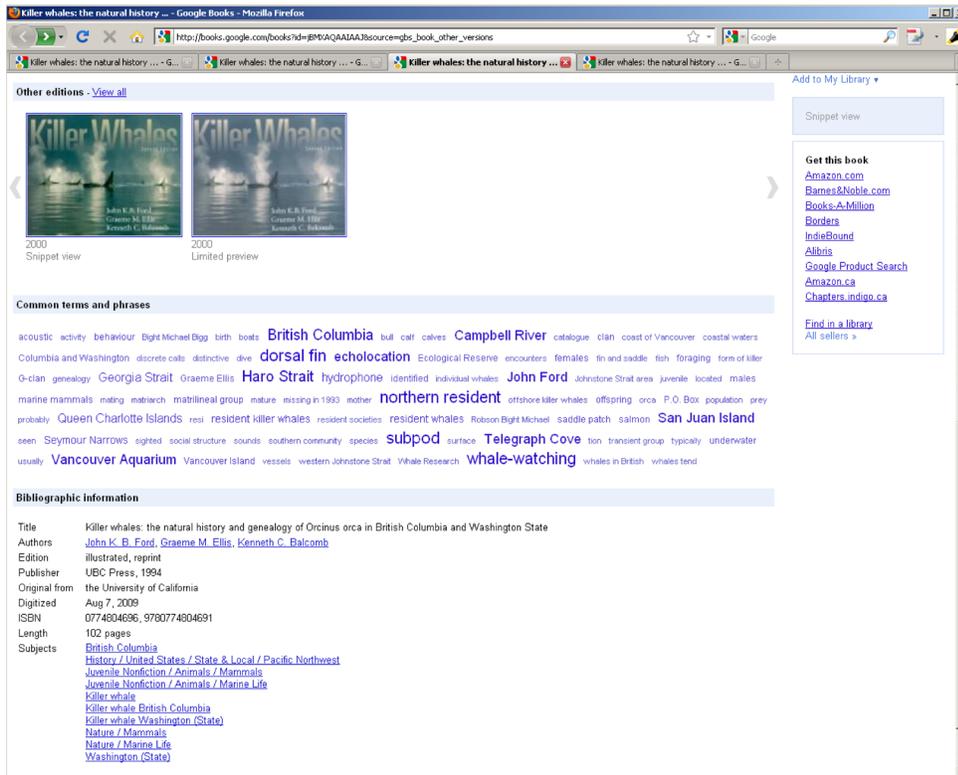
*Killer Whales* originated as a UBC Press title and was subsequently co-published with the University of Washington Press. The above Google Book Search record is for the University of Washington Press ISBN of the same 2000 edition. Only snippet view is available (no preview). This edition is not covered by the Partner Program, so there is no link to purchase the book from either publisher.



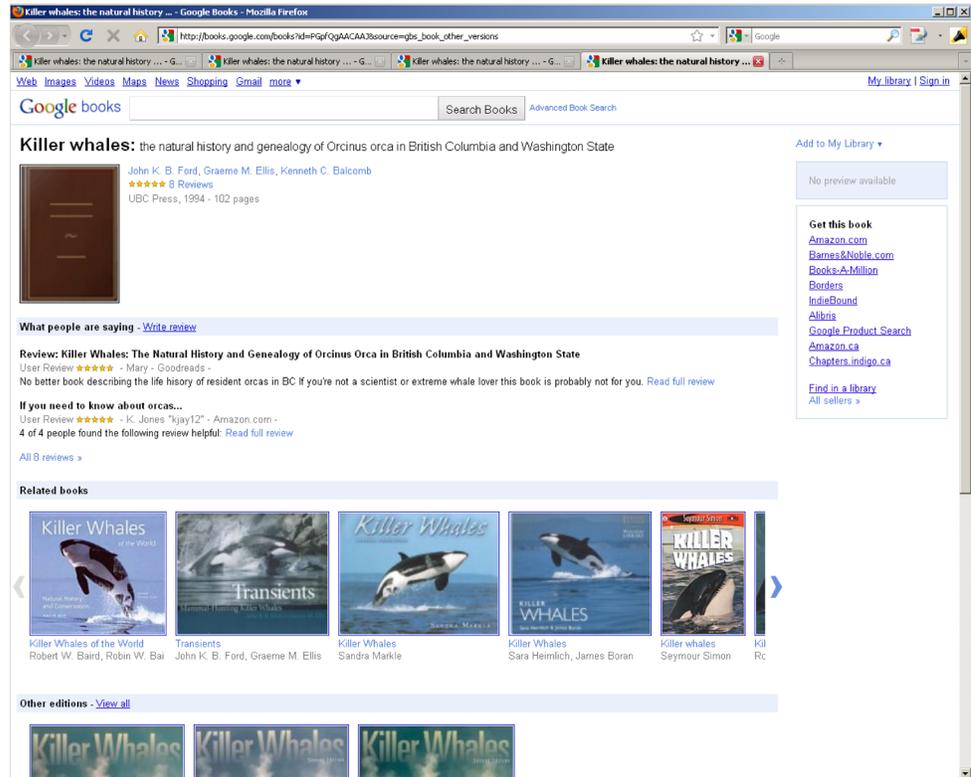
This is the same online book page as the previous image. The publisher is still listed as UBC Press, but the ISBN does not belong to UBC Press. This ISBN was digitized without authorization. Note the additional data fields indicating it was digitized from the University of California on August 7, 2009.



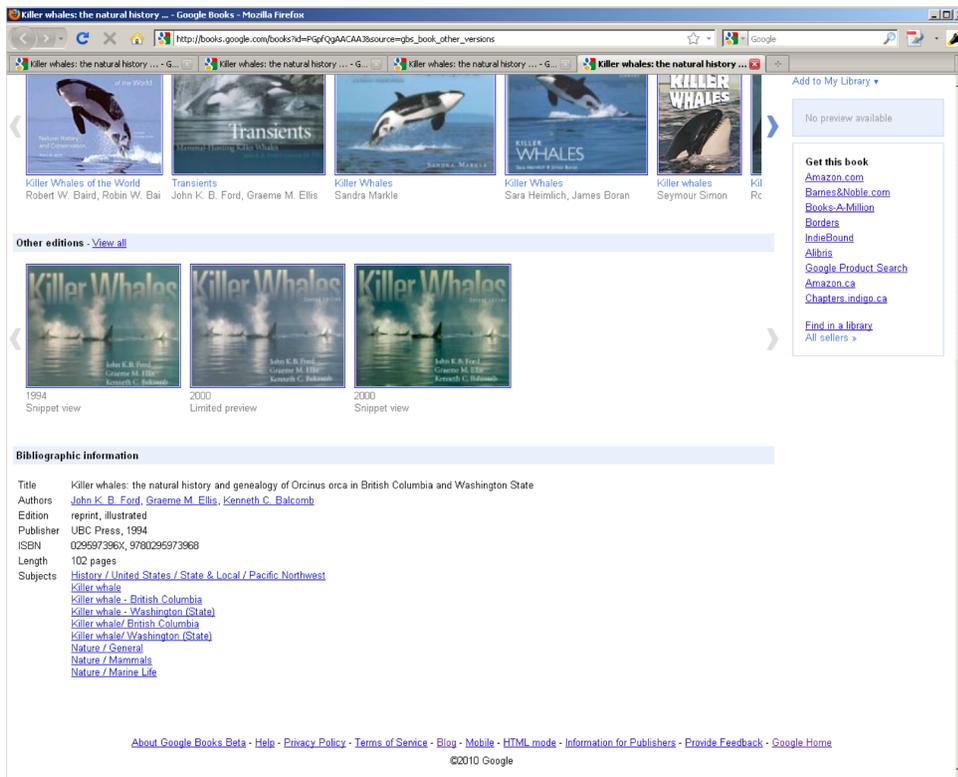
Google Book Search record for the 1994 UBC Press edition of *Killer Whales*. Note the book overview differs from those that appeared on the 2000 edition's page. Snippet view is available by entering a query in the "From inside the book" search box.



This is the same online book page as the previous image. The publisher is listed as UBC Press, and the ISBN belongs to UBC Press. The settlement claims database indicates this ISBN was not digitized without authorization. This edition was digitized from the University of California on August 7, 2009; this is after the settlement cut-off of May 5, 2009.



This is the 1994 edition published by the University of Washington Press. It does not allow preview or snippet view and has very little information displayed; even the cover image is missing.



This is the same online book page as the previous image. Although the publisher is again listed as UBC Press, this ISBN belongs to the University of Washington Press. This ISBN was not digitized without authorization.