Protecting Reader Privacy in Digital Books

Nicole A. Ozer, Jennifer A. Lynch

American Civil Liberties Union of Northern California; Samuelson Law, Technology & Public Policy Clinic, UC Berkeley Law
39 Drumm Street, San Francisco, California 94111
211 Boalt Hall, Berkeley, California, 94720
nozer@aclunc.org; jlynch@law.berkeley.edu

Abstract
What you choose to read says a lot about who you are, what you value, and what you believe. That’s why you should be able to learn about anything from politics to health without worrying that someone is looking over your shoulder. However, as books move into digital form, new reader privacy issues are emerging. In stark contrast to libraries that retain as little information about readers as possible, digital book services are capturing detailed information about readers: who they are, what books they browse and read, and even how long a given page is viewed, and the notes written in the “margins.” Without strong privacy protections, all of this browsing and reading history can be collected, analyzed, and may end up in the hands of the government or third parties without a reader’s knowledge or consent.

Retaining and strengthening reader privacy in the digital age requires a thorough examination of the potential privacy and free speech implications of digital book services and of the laws and policies that are needed to properly protect readers. Part I of this article discusses the history of strong legal and policy protections for reader privacy. Part II discusses current developments in digital book services. Part III discusses emerging privacy and free speech issues related to digital book services. Part IV proposes some policy and legislative solutions.

History of Strong Protections for Reader Privacy

It has long been recognized that the freedom to read without worrying who is looking over your shoulder plays an essential role in the freedom of thought and speech necessary for a robust democracy. When the right to read has been under assault, pivotal court decisions, state laws, and the ethics of librarians have safeguarded the privacy and free speech of readers, and helped to support the free exchange of ideas and open discourse.

Time and again the government and third parties have tried to collect and use evidence of reading habits in order to identify those with unpopular thoughts and beliefs. During the McCarthy hearings, Americans were questioned on whether they had read Marx and Lenin and even whether their friends had books about Stalin on their bookshelves. (Senate Permanent Subcommittee on Investigations of the Committee on Government Operations 1953). History repeated itself, when following September 11, the F.B.I. targeted libraries, demanding information on patrons. It is estimated that by December 2001, 85 libraries had been approached by the F.B.I. and that over 200 libraries were targeted between 2001 and 2005. (Lichtblau 2005).1

The Courts Have Long Protected Reader Privacy and Free Speech

Decisions by the Supreme Court, federal courts, and state courts have protected reader privacy at crucial junctures. In the 1950s, the Supreme Court held that an individual could not be convicted for refusing to provide the government with a list of individuals who had purchased political books. As Justice Douglas observed, “Once the government can demand of a publisher the names of the purchasers of his publications . . . [f]ear of criticism goes with every person into the bookstall . . . [and] inquiry will be discouraged.” (United States v. Rumley 1953). In the 1960s, the Supreme Court again protected reader privacy, striking down a requirement that individuals must file a written request with the postal service to receive “communist political propaganda” because such a requirement is “almost certain to have a deterrent effect.” The Court especially noted that “[P]ublic officials, like schoolteachers who have no tenure, might think they would invite disaster if they read what the Federal Government says contains the seeds of treason.” (Lamont v. Postmaster General 1965).

Federal and state courts have continued to set important precedent protecting reader privacy. (ACLU of Northern California 2009). When federal investigators attempted to use a grand jury subpoena to obtain Monica Lewinsky’s reading records from Kramerbooks bookstore in Washington, D.C., the court held that the First

1 Actual number not known due to PATRIOT Act “gag order” provisions. Found unconstitutional in Doe v. Gonzales 2007.
Amendment required the government to “demonstrate a compelling interest in the information sought . . . [and] a sufficient connection between the information sought and the grand jury investigation . . . .” (*In re Grand Jury Subpoena to Kramer-books & Afterwords, Inc.* 1998). The Colorado Supreme Court similarly held that book records were clearly protected under the free speech provision of the Colorado state constitution, and in order to obtain such records, the government must meet a “warrant plus” standard, requiring not only a warrant but a prior adversarial hearing, notice to the provider, and showing of a compelling need. (*Tattered Cover v. City of Thornton* 2002). Protection for reading records has not been limited to the brick and mortar world. In a 2007 United States District Court case, the court quashed a government subpoena initially seeking the identities of 24,000 Amazon.com book buyers because

[It] is an unsettling and un-American scenario to envision federal agents nosing through the reading lists of law-abiding citizens...[I]f word were to spread over the Net—and it would—that the FBI and the IRS had demanded and received Amazon’s list of customers and their personal purchases, the chilling effect on expressive e-commerce would frost keyboards across America... well-founded or not, rumors of an Orwellian federal criminal investigation into the reading habits of Amazon’s customers could frighten countless potential customers into canceling planned online book purchases, now and perhaps forever.” (*In re Grand Jury Subpoena to Amazon.com* 2007).

**State laws and library ethics have provided important protection for reader privacy and free speech**

State laws and library ethics also have provided substantial additional support for anonymous reading. Virtually every state protects public library reading records from disclosure by statute. In many states, violating a public library reading record statute is a misdemeanor criminal violation. In addition to safeguarding library patron records, Rhode Island and Michigan both prohibit book sellers from disclosing information. (R.I. Gen. Laws § 11-18-32 (2009); Mich. Comp. Laws Ann. § 445.1712 (2009)). For example, Michigan requires a warrant or court order before any business selling, renting, or lending books may disclose customer identifying information. With these statutes, states have recognized the importance of having a citizenry that can access and read books without fear of monitoring.

The American Library Association (ALA) has also dedicated itself to protecting reading privacy. The ALA’s Policy Manual guides all librarians that “the freedom to read is essential to our democracy” and “protecting user privacy and confidentiality is necessary for intellectual freedom and fundamental to the ethics and practice of librarianship.” (*American Library Association* 1992). The Library Code of Ethics, first adopted in 1938, now reads:

We protect each library user’s right to privacy and confidentiality with respect to information sought or received and resources consulted, borrowed, acquired or transmitted. (*American Library Association* 1997).

**Digital Book Services Have the Potential to Threaten a Long and Proud History of Reading Privacy**

Now more and more, readers are moving away from reading physical books at bookstores, libraries and in their homes and toward reading on electronic devices and the Internet. Google has spent the last five years digitizing millions of books from the United States’ best libraries and is in the process of finalizing a settlement to digitize and make available the majority of books published in the United States. (*The Authors Guild, Inc. et al. v Google, Inc* 2009). These digital book services have the potential to threaten a long and proud history of reading privacy. Service providers have both the interest and ability to collect vast amounts of information about readers. At the same time, the technological advances in moving books into the digital environment have outpaced existing book privacy laws, leaving few protections currently in place to prevent providers from exposing readers’ information—and by extension, readers’ deepest thoughts—to third parties and to the government.

**Digital Book Service Providers Can Collect Much More Data on Readers Than Libraries or Bookstores Ever Could**

Digital book service providers have the ability to collect and retain very detailed information about readers; tantamount to an offline library or bookstore hiring an agent to follow each individual patron around the stacks,

---


4 On Amazon.com, 35% of sales of books that have a Kindle edition are sold in that format. (Hansell 2009). Four million e-readers were sold worldwide in 2009, with the number expected to grow to twelve million in 2010, eighteen million in 2012. (Ye 2010).

5 More information available at www.aclu.org/googlebooks.
throughout their day, and finally into their homes. Digital book providers can easily track what books an individual considers, how often a given book is read, how long a given page is viewed, and even what notes are written in the “margins.” As reading has moved online, it also has become much easier to link books that are browsed or read with a reader’s other online activities, such as internet searches, emails, stored documents, and social networking. With all of this information, companies can create profiles about individuals, their interests and concerns, and even those of their family and friends.

This tracking is already occurring. For example, Google Books currently tracks: 1) a reader’s initial search query; 2) the specific book browsed and page viewed; 3) the date/time of the search or page view; 4) the reader’s Internet Protocol address, browser and computer operating system; and 5) one or more cookies that uniquely identify the reader’s browser. In a January 2009 New York Times article, a senior member of Google Book Search’s engineering team illustrated the kind of detailed information that the company collects. He admitted that he “was monitoring search queries recently when one...caught his attention.” The engineer could easily tell that the reader spent four hours perusing 350 pages of an obscure 1910 book. (Rich 2009). If a reader has logged in to other Google services such as Gmail at the time he searches for a book, Google can link reading data to the reader’s unique Google account. Google may also combine all this information with information gleaned from its DoubleClick ad service, which tracks users across the Internet. (Google.com 2009b).

Amazon can and does track similar information on readers who use its Kindle. As each Kindle is unique and automatically linked to one particular account holder, the potential for tracking specific reading habits may be even greater than with Google Books. Amazon tracks the time the Kindle is logged onto Amazon’s network and retains information about the books, magazine subscriptions, newspapers and other digital content on the Kindle and the reader’s interaction with that content. This includes an automatic bookmark of the last page read, the content deleted from the device and any annotations, bookmarks, notes, highlights, or similar markups made by the reader. (Amazon.com 2009). The company’s control over its users’ reading habits extends beyond merely tracking them. Amazon’s ability to collect and control content on the Kindle has allowed it to delete whole books without the account holder’s knowledge or consent. (Stone 2009).

Once reader information is collected, it may stay in the companies’ files for an indefinite time. Google has only promised to make a “good faith effort” to provide users with the opportunity to delete personal information. (Google.com 2009b). Amazon customers must contact third party advertisers and websites directly in order to access or opt-out of their information collection practices. (Amazon.com 2008).

Companies Have Strong Business Incentive to Collect Information

There are strong incentives for companies to collect detailed information about readers and to retain it for as long as possible. It is no secret that information about readers means more lucrative targeted advertising. An additional byproduct may also be to attract some customers who appreciate customized recommendations for books based on past reading history or interests.

According to industry experts, companies can expect up to ten times the revenue for advertisements based on behavioral data. (Federal Trade Commission 2009). Google is no stranger to the concept that the more it knows about individuals, the higher the advertising profits, both for Google and for its advertising partners. Google’s CEO Eric Schmidt has explained that “the ads are worth more if they’re more targeted, more personal, more precise.” (Dowd 2009).

Schmidt has made clear the company’s interest in amassing and analyzing as much information as possible about users of all of its services. And according to Schmidt, Google’s current extensive data collection is just the tip of the iceberg:

We are very early in the total information we have within Google. The algorithms will get better and we will get better at personalization. The goal is to enable Google users to be able to ask the question such as ‘What shall I do tomorrow?’ and ‘What job shall I take?’ [...] We cannot even answer the most basic questions because we don’t know enough about you. That is the most important aspect of Google’s expansion. (Daniel and Palmer 2007).

Google has not spent millions of dollars digitizing the world’s books out of the goodness of its heart. Detailed information about the identity and interests of millions of readers collected through the Google Book Search product could provide an important additional source of information for its algorithms and further aid in targeted advertising.

Safeguards for Reader Privacy Are Important for Protecting Readers and Companies’ Bottom Line

While there are economic incentives to collect information about readers, there are also important reasons, both for the public good and the bottom line, to ensure that there are robust reader privacy protections for digital books. The tracking and retention of data on digital book services will chill people from accessing particular sites and purchasing particular books if they have to fear how that information

---

* In 2007, the New York Times estimated that Google had already spent more than five million dollars. (Hafner 2007).
could end up being used or abused. These chilling effects are not hypothetical. In In re Grand Jury Subpoena to Kramerbooks & Afterwords, Inc., the court found that as a result of a grand jury subpoena for a patron’s book purchases, “[m]any customers have informed Kramerbooks personnel that they will no longer shop at the bookstore because they believed Kramerbooks to have turned documents over . . . that reveal a patron’s choice of books.”

Studies have also documented the chilling effect government surveillance has on reading and use of the Internet. One survey found that 8.4% of Muslim-Americans changed their internet usage because they believed their habits were being tracked by the government. (Sidhu 2007). Indeed, without anonymity, it is clear that many individuals will not access certain types of Internet speech. (ACLU v. Gonzales 2007) (finding that “[m]any people wish to browse and access material privately and anonymously, especially if it is sexually explicit,” that “[a]s a result of this desire to remain anonymous, many users who are not willing to access information non-anonymously will be deterred from accessing the desired information,” and that “web site owners such as the plaintiffs will be deprived of the ability to provide this information to those users”).

Consumers have also expressed dissatisfaction with targeted advertising business models that customize the advertisements shown based on the websites a user has visited and the content she has viewed, even when assured that the tracking is “anonymous.” In a recent nationwide survey, 68% of adults said that they “definitely would not allow it,” while 19% would probably not allow it. (Turow et al. 2009). Further, 72% of Americans “want the right to opt out when companies track their online behavior.” (Turow et al. 2009, citing 2008 Consumers Union national telephone survey). Further, 69% feel there should be a law that gives people the right to know everything that a website knows about them and 92% believe there should be a law that requires “websites and advertising companies to delete all stored information about an individual, if requested to do so.” (Turow et al. 2009). If there are not adequate protections in place to limit collection and retention of detailed book records and to safeguard that information from disclosure to third parties and the government, consumers are not going to feel confident using digital book services. Reducing the number of books that people feel safe accessing is not good for public discourse or for a company’s bottom line.

Many of the state book privacy laws were written for the library context and did not anticipate online services that can collect vast amounts of information about reading habits. Although federal and state court precedents protect privacy and free speech rights inherent in book records, some digital book providers argue book records should be treated like basic business records, and that they should be able to disclose this sensitive information without any judicial oversight. While the government would have to comply with the Fourth Amendment and obtain a warrant based on a showing of probable cause to come into an individual’s home and search a bookshelf or seize a book, neither Amazon or Google promises to demand a warrant or even a court order if asked to turn over customers’ digital book records to the government. Google’s Privacy Policy reserves the right to disclose user information whenever it has a “good faith belief” that disclosure is reasonably necessary to “satisfy any applicable law, regulation, legal process or enforceable governmental request.” (Google.com 2009b). Amazon, reserves the right to disclose subscriber information whenever “release is appropriate to comply with the law,” necessary to enforce its Conditions of Use or other agreements or to protect the rights, property, or safety of Amazon.com, its users, or others. (Amazon.com 2008).

And if a digital book service does disclose records, the reader may never know. Google says that it will provide notice to individuals when it receives third party demands for information. (Google.com 2009a). But Google has refused to disclose even general numbers related to government requests for information. (Kravets 2009). Amazon says that it will provide notice for disclosure to third parties, but this promise exempts all of the disclosure situations described above (to comply with the law, conditions of use, and protecting rights of Amazon/users). (Amazon.com 2008).

Next Steps to Protect Digital Reader Privacy

Digital book services are growing and the safeguards necessary to ensure the freedom to read are largely being left out of the story. The time is now for companies, policymakers and public interest groups to work together to take important first steps to: (1) develop robust privacy policies that take into account the sensitivity of reading records; (2) aggressively defend reader privacy; and (3) update and develop new laws to ensure that reader privacy is safeguarded in the digital age.

Companies Should Develop Robust Digital Book Privacy Policies and Aggressively Protect the Rights of Readers

Any digital book service should include at least the following four areas of basic reader protections:

Current Privacy Laws May Not Adequately Protect Digital Book Readers

Digital book services promise to revolutionize access to books, but due in part to outdated privacy laws, when individuals read these digital books, the government and third parties may be able to read over their shoulders.

Companies Should Develop Robust Digital Book Privacy Policies and Aggressively Protect the Rights of Readers

Any digital book service should include at least the following four areas of basic reader protections:
1) Reader Transparency: Readers should know what information is being collected and maintained about them and when and why reader information has been disclosed. Digital book service companies should develop robust privacy policies that take into account the sensitivity of book records information and publish annually the number and type of demands for reader information that are received.

2) Protection Against Disclosure: Readers should be able to use a digital book service without worrying that the government or a third party is reading over their shoulder. The digital book service should promise that it will protect reader records by responding only to properly-issued warrants from law enforcement and court orders from third parties. It also should promise that whenever it is legally possible, it will tell readers if anyone demands access to information about them before that information is disclosed so the reader has the opportunity to fight disclosure. The digital book services also should promise not to reveal any information about the use of its services to credit card processors or any other third parties.

3) Limited Tracking: Just as readers can anonymously browse books in a library or bookstore, they should have the ability to anonymously browse, search, and read books on a digital book service. Logging information for digital book services should not be kept for longer than necessary to complete a transaction, and never longer than 30 days. In addition, digital book services should not link any information about a reader's use of the book service with any information about that reader's use of other online services without specific, informed consent.

4) User Control: Readers should have complete control of their purchases and purchasing data. Readers must be able to review and delete their records and have extensive permissions controls for their "bookshelves" or any other reading displays. Readers also must be able to "give" books to anyone, including to themselves, without tracking.

Update and Develop New Laws to Safeguard Reader Privacy in the Digital Age

In order to ensure that reader privacy is safeguarded in the digital age, updating existing privacy laws and developing new laws to incorporate these basic protections will be necessary. States around the country have long recognized the importance of protecting the privacy and confidentiality of reading records. Now that digital book service providers are in a position to collect vast amounts of information about the reading habits of individuals, it would be appropriate for states to expand the scope of existing laws to address this new area and safeguard reading privacy.

Federal book privacy law also should be explored. Congress has already recognized the privacy interests of users of expressive material by enacting privacy protections for video and cable viewing records. The Video Privacy Protection Act prohibits disclosure of video viewing records without a warrant or court order, requires notice prior to any disclosure of personally identifiable information to a law enforcement agency, and requires the destruction of personally identifiable information one year after it becomes unnecessary. (Video Privacy Protection Act 1988). The Cable Communications Policy Act similarly prohibits disclosure of cable records absent a court order. (Cable Communications Policy Act 1984). Book records should be similarly protected. A federal law would help to ensure uniformity and clear standards for companies, individuals, and third parties making information requests.

Conclusion

The United States has a long history of protecting reading privacy. As the popularity of digital book services grows, we must ensure that these protections extend to digital reading records. Forcing individuals to choose between digital books and keeping their reading interests private is not good for business or for the public good. With strong economic incentives for digital book providers to collect detailed information about reading habits, and many book privacy laws outdated or incomplete, the time is now for businesses to pledge to build robust protections into digital book services and policymakers to update privacy law to safeguard the freedom to read.

References


The Authors Guild, Inc. et al. v Google, Inc., Case No. 05 CV 8136. (S.D.N.Y. 2009).


Dowd, Maureen 2009, Dinosaur at the Gate, New York Times, Apr. 14, 2009,


In re Grand Jury Subpoena to Amazon.com, 246 F.R.D. 570, 573 (W.D. Wis. 2007).


Tattered Cover v. City of Thornton, 44 P.3d 1044, 1059 (Colo. 2002).


