UNITED STATES DISTRICT COURT
FOR THE SOUTHERN DISTRICT OF NEW YORK

THE AUTHORS GUILD, *et al.,*

Plaintiffs,

v. CIV. NO: 05-CV-8136

GOOGLE, INC.,

Defendant.

BRIEF *AMICUS CURIAE* OF THE CENTER FOR DEMOCRACY & TECHNOLOGY IN SUPPORT OF APPROVAL OF THE SETTLEMENT AND PROTECTION OF READER PRIVACY

Leslie Harris
John B. Morris, Jr.*
David Sohn
Center for Democracy & Technology
1634 I Street, NW, Suite 1100
Washington, D.C. 20006
(202) 637-9800

* Motion for Admission *Pro Hac Vice* Pending

Dated: September 4, 2009

TABLE OF CONTENTS

TABLE OF AUTHORITIES

CASES:

STATE STATUTES & CONSTITUTIONS

FEDERAL AGENCY REPORTS:

OTHER MATERIAL:

UNITED STATES DISTRICT COURT
FOR THE SOUTHERN DISTRICT OF NEW YORK

THE AUTHORS GUILD, *et al.,*

Plaintiffs,

v. CIV. NO: 05-CV-8136

GOOGLE, INC.,

Defendant.

BRIEF *AMICUS CURIAE* OF THE CENTER FOR DEMOCRACY & TECHNOLOGY IN SUPPORT OF APPROVAL OF THE SETTLEMENT AND PROTECTION OF READER PRIVACY

The Center for Democracy & Technology (CDT) respectfully submits this brief *amicus curiae* to support final approval of the Proposed Settlement, and to urge this Court to protect the public's right to reader privacy in conjunction with the approval. CDT believes that the Settlement will have a positive impact on public access to information in our society, and that it should be approved. However, as part of that approval, we believe that the privacy concerns raised by the services enabled by the Proposed Settlement (hereafter referred to as the "New Services") must also be addressed by the Court. We appreciate the opportunity to discuss these concerns and make specific recommendations to address them in this brief.

STATEMENT OF INTEREST

CDT is a 501(c)(3) nonprofit public interest organization dedicated to preserving and promoting openness, innovation, and freedom on the decentralized Internet. CDT is one of the leading organizations in the country advocating for the protection of user privacy in the online

environment. The organization has testified numerous times before the United States Congress on privacy and a range of other issues,[1] and it has filed numerous *amicus* briefs with the U.S. Supreme Court and many other federal and state courts in the country.[2] CDT has been actively engaged in the public debate concerning the privacy implications of the New Services, and it published an extensive report on the issue in late July.[3]

INTRODUCTION AND SUMMARY OF ARGUMENT

The New Services Google intends to offer if the Proposed Settlement[4] is approved will be an extraordinary and valuable resource for all users. Google will at least for the foreseeable future take on a unique role in our society, that of a comprehensive library for research and browsing through books (as well as a major bookstore). However, the New Services also create very significant privacy risks, and could potentially transform a historic haven for reader privacy – the library – into a sweeping new source of data collection and tracking. Libraries have a long history of protecting reader privacy and safeguarding the right to read anonymously. Indeed, patron circulation records are protected by law against undue disclosure in almost all states in the United States. As Google steps into the role of a comprehensive library service provider –

[1] In the past 18 months, for example, CDT staff members testified on privacy issues more than a dozen times before committees of the U.S. Congress. *See* http://www.cdt.org/testimony/.

[2] In the interest of disclosure, CDT does receive a minority of its funding from a broad range of companies involved in the Internet industry, including companies that support the settlement (such as Google) and companies that have announced their intention to oppose the settlement (such as Microsoft and Yahoo!). None these companies (or any other companies) sit on CDT's board of directors or otherwise control the substantive positions that CDT takes on legal or policy issues. CDT is an independent organization that commonly takes legal positions that are adverse to the interests of companies that provide funding to CDT.

[3] *See* CDT, *"Privacy Recommendations for the Google Book Search Settlement*," July 27, 2009, available at http://cdt.org/copyright/20090727_GoogleRecs.pdf.

[4] Settlement Agreement, *Author's Guild v. Google*, Civ. No. 05-CV-8136 (entered Oct. 28, 2008) (hereinafter referred to as "Settlement" or "Proposed Settlement").

through this Court's approval of the Proposed Settlement – it is vital that privacy be considered and protected.

Because the Proposed Settlement is focused specifically on resolving the copyright dispute among the parties, the agreement understandably does not address the full range of reader privacy risks that arise as part of the New Services. Given the sweeping impact of the Proposed Settlement, however, the Court has both the authority and the duty to protect the public's interest in this settlement – including its interest in privacy. We believe that the Court can address the public's interest in privacy – without altering the terms agreed upon by the parties to the settlement – by enforcing the commitments to reader privacy laid out in this brief as part of its ongoing jurisdiction over the implementation of the Settlement.

To its credit, Google has made clear that it recognizes that there are significant privacy issues raised by the New Services, and has taken important steps that address some – but not all – of these concerns. This week Google voluntarily released a Privacy Policy to govern Google's actions with regard to data obtained through the New Services.[5] We strongly applaud these voluntary steps.[6] We nevertheless believe both that Google's privacy commitments do not go far enough, and that any such commitments must in any event be backed up by a mandate and continuing oversight by this Court.

The principal focus of this brief is on the need for the Court to protect reader privacy, and to set out the steps the Court should take to do so. But the brief first addresses two threshold

[5] Google Books Privacy Policy, Sept. 3, 2009, available at http://books.google.com/googlebooks/privacy.html.

[6] Google has, for example, committed in its Privacy Policy to (a) only share aggregate, non-identifying information with the Book Rights Registry, (b) allow users to delete information collected in association with their Google Accounts, (c) to allow Institutional Subscribers to authenticate their own constituents without transmitting identifying information to Google, (d) to not require users of the Public Access Service to register with Google, (e) to post detailed privacy policies in connection with the New Services, and (f) to secure user data against unauthorized access.

questions: whether the underlying Proposed Settlement should be approved, and what authority and duty the Court has to consider and address the privacy concerns raised here.

I. THE PROPOSED NEW SERVICES WOULD BE EXTRAORDINARILY VALUABLE, AND THE PROPOSED SETTLEMENT SHOULD BE APPROVED.

As detailed below, we believe that serious privacy concerns must be addressed in conjunction with any approval of the Proposed Settlement. But to be clear, CDT strongly supports the Settlement and thinks it should be approved. The New Services will considerably increase the public's access to millions of books containing much of the world's written knowledge and ideas. Moreover, Google's powerful book search engine and other dynamic tools will transform how the public conducts research, interacts with written text, and shares information and ideas with others. Because it will unquestionably provide a significant public benefit at a size and scale that could not otherwise be replicated in the near term, CDT believes the Proposed Settlement deserves Court approval.

II. THE COURT HAS BOTH THE AUTHORITY AND THE DUTY TO PROTECT THE PUBLIC'S INTEREST IN READER PRIVACY AS PART OF THE COURT'S APPROVAL OF THE PROPOSED SETTLEMENT.

The public's right to reader privacy with respect to the New Services is not a matter that is addressed within the Proposed Settlement. The underlying lawsuit in this case addresses copyright concerns, and thus, the understandable focus of negotiation between the parties was plainly limited to those issues. In simple terms, consideration of privacy protections for the public was irrelevant to the resolution of the specific copyright claims raised here.

The corollary is also true: the privacy concerns and recommendations set out in this brief do not directly impact the terms of the Proposed Settlement, nor do they affect the copyright rights of the parties. Thus, resolution of the privacy concerns would not require any alteration to

the Agreement, nor any further notice and objection period for the class. But the threats to the privacy rights of the unrepresented public flow directly from the system the Proposed Settlement aims to create, and as a collateral consequence of the Proposed Settlement, the privacy implications must be addressed.

The public did not have a seat at the negotiating table leading to this Proposed Settlement. None of the parties represent the public (as might be the case if a government agency were a party, or if the case included a class of *library users*). Nor do any of the parties serve as even rough proxies for the public (as might be the case where, for example, the public's interest in quality medicine would be vindicated by a plaintiff class of victims of tainted drugs). The parties here were – understandably – looking out for their own interests, not the public's interest in reader privacy. That responsibility falls to this Court.

The need for Court action to protect the public in this case is all the more important because it is the comprehensiveness of the New Services that would trigger the real privacy risk here. Without the Settlement, Google simply would not be able to create the extraordinary (and extraordinarily valuable) new resource proposed here – a comprehensive digital library encompassing the vast majority of books in existence in this country – which puts Google in the shoes of a vital American institution, the library. And by transforming the library experience from the publicly-run, decentralized, and largely off-line status quo of today into a private, centralized, and entirely electronic system run by Google, the Proposed Settlement creates the privacy risks raised here. This is not to suggest that the Court should reject the Settlement, but instead to argue that in approving the Settlement, the Court must take action to protect the public's privacy right.

As the Sixth Circuit has concisely stated:

> Judicial approval of a settlement agreement places the power and prestige of the court behind the compromise struck by the parties. . . . In making the reasonableness determination the court is under the mandatory duty to consider the fairness of the decree to those affected, the adequacy of the settlement to the class, and the public interest.

Williams v. Vukovich, 720 F.2d 909, 920-21 (6th Cir. 1983) (citations omitted). In this case, the New Service created by the Proposed Settlement is very much in the public interest, but leaves unresolved critical protections for reader privacy. In order to protect the public interest, those privacy risks must be addressed.

As discussed above, the public is a "third party" whose interests are nonetheless affected by the Proposed Settlement. The Second Circuit has extensively discussed the Court's authority and duty to protect the interest of affected third parties:

> In a class action settlement, the normal focus is on the fairness, reasonableness and adequacy of the settlement to the plaintiff class. Where the rights of third parties are affected, however, their interests too must be considered. *See Williams v. Vukovich*, 720 F.2d 909, 921 (6th Cir.1983) . . .; *see also Donovan v. Robbins*, 752 F.2d 1170, 1176 (7th Cir.1985) ("Even if no third party complains, the judge has to consider whether the decree he is being asked to sign is lawful and reasonable as every judicial act must be.") (citation omitted) . . . . In other words, where the rights of one who is not a party to a settlement are at stake, the fairness of the settlement to the settling parties is not enough to earn the judicial stamp of approval.
>
> Moreover, if third parties complain to a judge that a "decree will be inequitable because it will harm them unjustly, he cannot just brush their complaints aside." *Donovan*, 752 F.2d at 1176 (citations omitted). In fact, section 1.46 of the Manual for Complex Litigation specifically suggests that courts reviewing class action settlements consider "the views of the non-participating parties and counsel." Manual for Complex Litigation (Second) § 1.46, at 53.

In re Masters Mates & Pilots Pension Plan and IRAP Litigation, 957 F.2d 1020, 1025-26 (2d Cir. 1992) (footnote and some citations omitted). As the Second Circuit makes clear, the Court must look out for the public's interests in this case.

The following sections of this brief discuss the public's interest in reader privacy, and propose concrete steps the Court should take – as part of an approval of the Proposed Settlement – to protect reader privacy.

III. THE PUBLIC HAS A STRONG INTEREST IN READER PRIVACY.

As a legal, policy, and practical matter, readers have long enjoyed a high level of anonymity and privacy with respect to their reading habits, as one element of their broader – and constitutionally-based – interest in privacy over their personal information. *See Nat. Ass'n of Letter Carriers v. U.S. Postal Service*, 604 F. Supp. 2d 665, 673 (S.D.N.Y. 2009) ("The Second Circuit has held that there is a constitutional right to privacy in personal information . . . ."), *citing Doe v. City of New York*, 15 F.3d 264, 267 (2d Cir. 1994). The First Amendment protects the right to receive information anonymously,[7] and true to that value, libraries have a longstanding commitment to intellectual freedom and patron privacy. The American Library Association *Code of Ethics* states: "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."[8]

Furthermore, state laws protecting library patron records from public disclosure have for many years reinforced individuals' right to privacy with respect to what they read. Forty-eight states have enacted statutes that either expressly protect these records or exempt them from public disclosure rules.[9] In several states, violating a library privacy statute is a misdemeanor

---

[7] *See, e.g.*, Julie Cohen, "*A Right to Read Anonymously: A Closer Look at 'Copyright Management' In Cyberspace*," 28 Conn. L. Rev. 981 (1996).

[8] Code of Ethics of the American Library Association, Jan. 22, 2008, available at http://www.ala.org/ala/aboutala/offices/oif/statementspols/codeofethics/codeethics.cfm.

[9] *See* American Library Association, "State Privacy Laws Regarding Library Records," available at http://www.ala.org/ala/aboutala/offices/oif/ifgroups/stateifcchairs/stateifcinaction/stateprivacy.cfm.

offense.[10] With respect to book purchases, courts have generally been reluctant to compel bookstore owners to reveal information about their customers' purchases.[11]

Beyond constitutional and statutory protections, as a practical matter the average public library affords a very high level of privacy protection. A reader can usually enter a library anonymously; browse at will; pull any book off of a shelf without notice; and read the book or gather sought-after information – all without leaving any record whatsoever. If the reader checks out a book, that transaction is typically protected by library policy, state law, or both.

In contrast with this broad legal and practical protection of reader privacy, the New Services represent a sea change with respect to the treatment of access to material that has historically been held by libraries. Google is in many ways taking on the role of the public library as a gateway to information, only on a much larger and more comprehensive scale. Providing such breadth of electronic access to so many published books will give Google an unparalleled view of people's reading and information-seeking habits. By hosting the book search service and closely managing user access, Google will have the capability to collect data about individual users' searches, preview pages visited, books purchased, and even time spent reading particular pages. Whereas in the offline world such data collection is either impossible or widely distributed among disconnected libraries and bookstores, Google will hold a massive centralized repository of books and of information about how people access and read books online.

Furthermore, Google is likely to be the only comprehensive source for digitized out-of-print books. Although many parts of the Proposed Settlement are non-exclusive, Google alone

[10] *See, e.g.*, Ariz. Rev. Stat. § 41-1354 (2008); Ark. Code. Ann. § 13-2-702; D.C. ST § 39-108; Fla. Stat. Ann. § 257.261; Mont. Stat. Ann. § 22-1-1111.

[11] *See, e.g., Tattered Cover, Inc. v. City of Thornton*, 44 P.3d 1044 (Colo. 2002).

will have the broad right to scan and display the works of any class members that do not register with the Registry (unless a future class action yields a similar settlement). The ability to offer a comprehensive collection, together with the magnitude of the task of scanning and hosting millions of books, is likely to make Google the dominant provider of online universal library functions for at least the immediate future, and likely for years to come.

Under the terms of the Proposed Settlement, Google would be able to collect – and indeed may be required to collect – a broad range of data about readers. Detailed user information will be collected and used to differentiate among the services offered, to calculate payments to rightsholders, and to prevent unauthorized access to the scanned books.[12] For the purposes of limiting the portion of a book available to particular users in Previews, Google will track page-by-page access using personal information such as "IP address, cookies, and similar signals that may be available."[13] Google will manage and track individual consumer purchases using "account login or other equivalent method."[14] For controlling access to Institutional Subscriptions, Google will authenticate users using "IP address authentication, user login, and/or leveraging authentication systems already in place at an individual institution."[15] Google may offer a "Book Annotation" feature to allow readers to provide their own commentary and other content on individual pages of a digital book. For annotations made on purchased books, the settlement requires Google to have each reader "identify (e.g., by name, login or user id) each individual with whom such Book Annotation will be shared" (for up to 25 individuals).[16]

[12] *See, e.g.,* Settlement Attachment D ("Settlement Security Standard") § 3.9; Settlement § 6.6(v).

[13] Settlement Security Standard § 3.9.1.

[14] *Id.* § 3.9.2.

[15] *Id.* § 3.9.3.

[16] Settlement § 3.10(c)(ii)(5)(d).

In short, the Proposed Settlement gives Google the potential, and in some instances the need, to collect substantial quantities of sensitive reader information. Google will also need to share some usage data with the Registry. Specifically, Google will share sales and subscription usage data for calculating and distributing payment to rightsholders,[17] market research data concerning various Preview options,[18] and data pertaining to audits and security breaches.[19] Although some collection and sharing of *aggregate* data is of course necessary to effectuate the settlement, the Proposed Settlement does not contain a broad restriction on the sharing of *individually identifiable* user data. The agreement does state that the Registry cannot force Google to disclose "confidential or personally identifiable information except as compelled by law or valid legal process" in the case of a security breach, but it does not address *voluntary* disclosure by Google.[20] More generally, the Proposed settlement does not address Google's collection, use, retention, and sharing of user data outside the specific contexts of security breaches and audits.

The risk to reader privacy is compounded by the fact that Google currently offers dozens of *other* services and software products on the Web and on end-user devices. In the absence of binding limits on what Google can do with the data it collects about readers through the New Services, Google would remain free to combine that data with other data that Google collects, adding a rich and personal dimension to the profiles that Google already maintains about individuals' searching and Web surfing habits. Reading habits add an intimate element to

[17] Settlement § 6.6(v); Settlement Attachment C ("Plan of Allocation").

[18] Settlement § 4.3(e)(i).

[19] Settlement §§ 8.2, 8.3.

[20] Settlement §§ 7.3(b), 8.6(a); *see also* §§ 4.6(e), 15.1 (establishing confidentiality requirements in the case of audit).

profiles that may already be attractive for a variety of uses, from marketing (by Google and its affiliates) to litigation (by those serving subpoenas on Google).

As valuable as the New Services will be, without action by this Court to protect reader privacy, they have the potential to transform the typical library experience from one of anonymity and privacy to one of data collection and tracking.

IV. AS PART OF THE SETTLEMENT APPROVAL, THE COURT SHOULD MANDATE THAT GOOGLE COMPLY WITH "FAIR INFORMATION PRACTICES" WITH REGARD TO ITS NEW SERVICES.

Part V below discusses specific aspects of privacy that the New Services put at risk, and offers recommendations on appropriate responses to those risks. But a key preliminary question is what, exactly, the Court can do to protect reader privacy in the context of an approval of the Settlement. Although in the vast majority of cases courts will accept or reject proposed class action settlements in their entirety, those are not the only choices. In considering the impact of a proposed settlement on a third party, "[a] federal judge has the full powers of an equity judge." *Donovan v. Robbins*, 752 F.2d 1170, 1176 (7th Cir.1985). To resolve this litigation – and to facilitate the very significant public benefits that the Proposed Settlement would allow – this Court can craft additional requirements that avoid harm to third parties (in this case, the public).

Ideally, these requirements – both the substantive requirements and the Court's continuing supervision – would be something to which the parties would consent. As noted above, Google has already voluntarily taken some important actions to address privacy concerns raised here, and we hope that Google will consider the specific requirements recommended below. But the Court has the power, in the public interest, to protect reader privacy even if such protection is *not* already in the Proposed Settlement. In *New England Carpenters Health v. First Databank*, 602 F. Supp. 2d 277 (D. Mass. 2009), the district court approved a class action

settlement, but modified a term of the settlement to address the harmful impact of the proposed settlement on third parties. *See id.* at 285. In the instant case, we do not seek any alteration to any term of the Proposed Settlement itself, but instead we urge the Court to impose some collateral conditions on Google as part of an approval order. Such additional conditions would allow the public to receive the benefits of the New Services without suffering the accompanying risks to privacy.

Concretely, we urge the Court to impose a set of mandatory privacy requirements on the New Services, and for the Court to include oversight of these requirements in ongoing Court supervision that the Proposed Settlement already envisions. The specific privacy requirements we urge the Court to consider are based on "Fair Information Practices" ("FIPs"), a set of principles that have been widely accepted over the past several decades as the preeminent framework for protecting digital privacy.[21] In the United States, FIPs have been the basis for legislative and administrative approaches to protecting consumer privacy, and are thus the appropriate framework within which to evaluate the New Services. FIPs, for example, form the basis of the Privacy Act of 1974, which regulates the federal government's collection, use, and disclosure of personal information.[22] In 1980, the members of the Organization for Economic Cooperation and Development (OECD), including the United States, approved a set of FIPs for

---

[21] *See* Robert Gellman, *Fair Information Practices: A Basic History*, Dec. 31, 2008, available at http://bobgellman.com/rg-docs/rg-FIPshistory.pdf; *see also* CDT, *Guide to Online Privacy*, available at http://www.cdt.org/privacy/guide/basic/fips.php.

[22] The earliest version of the FIPs comprised five principles. *See* Department of Health, Education, and Welfare (DHEW), *Records, Computers and the Rights of Citizens: Report of the Secretary's Advisory Committee on Automated Personal Data Systems*, July 1973, available at http://aspe.os.dhhs.gov/datacncl/1973privacy/tocprefacemembers.htm. An expanded set of eight principles formed the basis Privacy Act of 1974. *See* Privacy Protection Study Commission, *Personal Privacy in an Information Society*, June 1977, available at http://aspe.hhs.gov/datacncl/1977privacy/toc.htm ("These five principles and the findings of the DHEW Committee, published in July 1973, are generally credited with supplying the intellectual framework for the Privacy Act of 1974, though in drafting the statute the Congress, influenced by its own inquiries, refined the five principles to eight.").

both commercial and government practices.[23] The Federal Trade Commission has on numerous occasions expressed its support for FIPs as the basis for governing commercial online privacy.[24]

Fair Information Practices today center around eight principles (as recently restated by the U.S. Department of Homeland Security)[25]:

> **Transparency**: Organizations should be transparent and provide notice to the individual regarding the collection, use, dissemination, and maintenance of personal information;
>
> **Individual Participation**: Organizations should involve the individual in the process of using personal information and, to the extent practicable, seek individual consent for the collection, use, dissemination, and maintenance of personal information. Organizations should also provide mechanisms for appropriate access, correction, and redress regarding use of personal information;
>
> **Purpose Specification**: Organizations should specifically articulate the authority that permits the collection of personal information and specifically articulate the purpose or purposes for which the personal information is intended to be used;
>
> **Data Minimization**: Organizations should only collect personal information that is directly relevant and necessary to accomplish the specified purpose(s) and only retain personal information for as long as is necessary to fulfill the specified purpose(s);
>
> **Use Limitation**: Organizations should use personal information solely for the purpose(s) specified in the notice. Sharing personal information outside the organization should be for a purpose compatible with the purpose for which the personal information was collected;
>
> **Data Quality and Integrity**: Organizations should, to the extent practicable, ensure that personal information is accurate, relevant, timely, and complete;
>
> **Security**: Organizations should protect personal information through appropriate security safeguards against risks such as loss, unauthorized access or use, destruction, modification, or unintended or inappropriate disclosure; and
>
> **Accountability and Auditing**: Organizations should be accountable for complying with these principles, providing training to all employees and contractors who use personal information, and auditing the actual use of personal information to demonstrate compliance with these principles and all applicable privacy protection requirements.

---

[23] OECD, *Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*, Sept. 23, 1980, available at http://www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.html.

[24] Federal Trade Commission, *Privacy Online: A Report to Congress*, June 1998, available at http://www.ftc.gov/reports/privacy3/toc.shtm; Federal Trade Commission, *Privacy Online: Fair Information Practices in the Electronic Marketplace*, May 2000, available at http://www.ftc.gov/reports/privacy2000/privacy2000.pdf.

[25] Department of Homeland Security Guidance Memorandum, "Fair Information Practice Principles: Framework for Privacy Policy," Dec. 29, 2008, available at http://www.dhs.gov/xlibrary/assets/privacy/privacy_policyguide_2008-01.pdf.

As a privacy framework that has been recognized in both the legislative and administrative arenas, FIPs establish the appropriate basis for privacy protections in the context of the New Services. As detailed below, we urge the Court to adopt a FIPs-based set of requirements on the provision of the New Services pursuant to the Proposed Settlement.

V. SPECIFIC RECOMMENDED PRIVACY REQUIREMENTS.

We detail below specific requirements – based on the above principles of Fair Information Practices – that we urge the Court to impose on the delivery of New Services as a component of the approval of the Settlement Agreement. Many – but not all – of these requirements are practices that Google has announced that it has or will voluntarily adopt. Indeed, this week, Google published online a preliminary Privacy Policy to govern the New Services if they are approved.[26] We strongly applaud Google both for issuing the policy, and for many of the commitments made within the policy.

Google's voluntary Privacy Policy, however, does not fully resolve the privacy concerns raised by the New Services – and would not resolve them *even if* Google voluntarily agreed to comply with all of the recommendations listed below (which Google has not yet done). This is simply because Google's commitments made in its Privacy Policy are *voluntary*, and could be changed or withdrawn at any time. Thus, we believe that it is critical for the Court to make the privacy requirements mandatory on Google or any successor provider of the New Services.

We urge the Court to consider the following specific privacy concerns, and to adopt the suggested privacy requirements listed below[27]:

---

[26] Google Books Privacy Policy, Sept. 3, 2009, available at http://books.google.com/googlebooks/privacy.html.

[27] For a more full discussion of the issues and recommendations, *see* "Privacy Recommendations for the Google Book Search Settlement," July 27, 2009, http://cdt.org/copyright/20090727_GoogleRecs.pdf.

Privacy Concern #1 – Lack of Notice (Transparency/Purpose Specification FIP Principles): The Proposed Settlement contains no provision requiring Google to notify readers about the data it collects in connection with New Services. Although Google does voluntarily provide some notice, we believe that it should be required to clearly and prominently disclose the following:

(a) What information Google collects in connection with the New Services, including information that can be used to identify individual readers (IP addresses, cookie information, and account information, for example);
(b) What information Google collects about individuals' use of the New Services (search terms, book selections, page selections, or length of stay on a particular page, for example);
(c) The purpose for which this information is collected;
(d) How long each type of data is retained;
(e) What technical mechanisms (such as cookies) Google uses to track readers on the site;
(f) How readers can exercise choice about having their data collected and used in connection with the New Services; and
(g) How reader data is safeguarded against theft or misappropriation.

In light of the special sensitivity of readership information and library browsing, a link to this notice should be displayed more prominently than the usual privacy notice associated with other Google services.

- **Google Books Privacy Policy:** Google's newly released privacy policy largely conforms to this recommendation. Google commits to posting, for any of the New Services, a privacy policy that, in combination with Google's main Privacy Policy, will include most of the basic elements of the consumer notice we recommend immediately above, although Google does not commit to provide significant notice of retention periods or the specific purposes for which user data is collected.
- **Recommended Action:** The Court should required Google to provide appropriate notice to users about its information practices. We urge the Court to include the following provision in an order approving the Proposed Settlement:

Proposed language: For all products and services enabled by the agreement, Google shall publicly post in a prominent manner a privacy policy (or policies) that shall disclose what information Google receives and stores when users use the product or service, including any unique identifiers; how Google uses such information; how such information is protected against unauthorized access; what choices users have about disclosing such information; what technical mechanisms (such as cookies) are used to track users; and the time period(s) for which such information is retained.

Privacy Concern #2 – No Collection Limitation (Data Minimization FIP Principle):

As noted above, the Proposed Settlement requires Google to collect reader data for a variety of purposes. The Proposed Settlement, however, does not significantly limit what *other* information Google might collect. Nor does it say whether Google is permitted to collect details about how individual readers interact with books, but press reports demonstrate that – in its existing book search service – Google can already track which pages of a book users view and how long they view each page.[28]

Google's potential technical capability to intimately track reader behavior should not trump individuals' long-standing ability to read books anonymously. Thus, CDT believes that Google should be permitted to collect *only* the data necessary to provide the services described in the settlement, and that Google should limit collection of detailed data connected to readers' use of books (for example, pages read or time spent reading) to situations in which such usage data is necessary to account for Preview uses (or to provide services chosen by the user where the user has expressly given consent for the collection of the data).

- **Google Books Privacy Policy:** Google's policy does not make any commitments on this issue.

[28] *See* Motoko Rich, "Google Hopes To Open a Trove of Little-Seen Books," New York Times, Jan. 4, 2009, available at http://www.nytimes.com/2009/01/05/technology/internet/05google.html (revealing that a Google employee knew how many book pages a particular user had viewed, and for how long).

- **Recommended Action:** The Court should require Google to collect only information necessary to effectuate the Proposed Settlement or as expressly permitted by the user. We urge the Court to include the following provision in an order approving the Proposed Settlement:

  > **Proposed language:** Google shall limit its collection of data about users' use of the products and services authorized by the agreement to that which is necessitated by the terms of the agreement or is essential to the provision of the products and services. Google shall not collect additional data concerning any user's use of the products and services authorized by the agreement without obtaining prior consent from the affected user.

Privacy Concern #3 – Institutional Users (Data Minimization FIP Principle): Under the terms of the Proposed Settlement, institutions can obtain an "Institutional Subscription" that allows many users to use the New Services through the institution (thus, a college might allow its students access to the New Services through this approach).[29] There is no need under the Proposed Settlement for *Google* to be involved in the authentication of individual users within institutions. Institutional Subscribers alone should be responsible for authenticating their own end users without sharing authentication credentials or other personal information with Google.

- **Google Books Privacy Policy:** Google's policy commits to fully addressing this concern.[30]
- **Recommended Action:** The Court should require Google to permit Institutional Subscribers to authenticate their own users. We urge the Court to include the following provision in an order approving the Proposed Settlement:

  > **Proposed language:** Google shall permit Institutional Subscription institutions to authenticate their users locally, without transmitting authentication credentials or other identifying information to Google.

[29] Settlement § 4.1

[30] *See* Privacy Policy ("Schools or other institutions that sign up for subscriptions will be able to authenticate users . . . without [Google] knowing who that user is.").

Privacy Concern #4 – Book Annotations (Use Limitation FIP Principle): Under the Proposed Settlement, Google is permitted to offer a Book Annotation feature with which users can create and share annotations.[31] The Settlement does not, however, restrict Google's use of the contents of annotations, which may contain especially sensitive user information (because they involve users generating their own personal content, and identifying other users with whom to share it).

- **Google Books Privacy Policy:** Google's policy does not make any commitments on this issue.
- **Recommended Action:** Google should not be permitted use book annotation information without clear user consent or as necessary to comply with the Proposed Settlement. We urge the Court to include the following provision in an order approving the Proposed Settlement:

  > **Proposed language:** Google shall limit its use of book annotation information (a) to provide the annotation service, (b) to comply with the terms of the agreement, or (c) for uses specifically authorized by the user.

Privacy Concern #5 – User Access (Individual Participation FIP Principle): A key element of Fair Information Practices is that individuals be able to learn (and if needed, correct or delete) what information is held about them. Google should provide users the ability to access and correct (or delete) the information that has been collected in connection with their user accounts, including purchase histories, annotations and records of annotation sharing, and, to the extent they are tied to user accounts, search histories.

[31] Settlement § 3.10(c)(ii)(5)

- **Google Books Privacy Policy:** Google's policy commits to allowing users access and control over some types of information ("My Library," Web History, and book purchases), but not over other data collected in connection with the New Services.
- **Recommended Action:** Google should be required to give users access and control over their information. We urge the Court to include the following provision in an order approving the Proposed Settlement:

  > **Proposed language:** Google shall provide users of the products and services authorized by the Agreement with the ability to review all information (including, for example, search and annotation records) collected in connection with a Google account or similar user-identifier, and to delete such information unless it is essential to comply with the terms of the agreement. With regard to purchases, Google shall provide users the ability to delete individual purchases and/or their entire purchase histories at any time should they decide that they no longer need online access to those purchased books, to the extent such deletion is possible given the accounting obligations necessary to provide the service.

Privacy Concern #6 – Data Integration (Use Limitation FIP Principle): The terms of the Proposed Settlement place nearly no restrictions on Google's potential use of data collected in connection with the New Services. Given the potential sensitivity of reading habits, Google should refrain from using such data for purposes other than to provide and secure the New Services. By default, information collected through the New Services should not be used in connection with any other Google services or combined with data from other Google services, such as Web search or advertising.

- **Google Books Privacy Policy:** Google's policy does not make any commitments on this issue. Indeed it suggests that data collected in connection with the New Services *will* be integrated with existing services.[32]

[32] *See* Privacy Policy ("When you use Google Books, we receive log information similar to what we receive in Web Search. This includes: the query term or page request (which may include specific pages within a book you are browsing), Internet Protocol address, browser type, browser language, the date and time of your request and one or

- **Recommended Action:** Google should be required to limit use of data collected in connection with the New Services to those purposes necessary to effectuate the Proposed Settlement, or those purposes for which affected users have given their express consent. We urge the Court to include the following provision in an order approving the Proposed Settlement:

  > **Proposed Language:** Google shall not use data collected in connection with any products or services authorized by the agreement for any purposes other than those necessary to provide and secure such products and services without obtaining prior consent from the affected user.

Privacy Concern #7 – Disclosure Standard (Use Limitation FIP Principle): The Proposed Settlement is silent with respect to general law enforcement and civil litigant access to the data that Google collects in connection with the New Services. Given the unique comprehensiveness of the New Services, the strong tradition and policy of library privacy, and the special sensitivity associated with reading, it is vital that Google ensure that warrants and subpoenas for data comply with appropriate standards. At a minimum, in response to government requests, Google should take reasonable steps to require that the government obtain a court order or warrant issued upon probable cause to compel disclosure of information that could be used to identify a user or to associate a user with access to particular books.[33] With respect to access by civil litigants, Google should not, unless otherwise required by law, disclose any information about users to a third party in a civil or administrative action absent a judicial determination and order that the party seeking the information has a compelling interest in the

---

more cookies that may uniquely identify your browser"; "Google Books operates a lot like Web Search and other basic Google web services.").

[33] The probable cause standard was drawn from the decision of the Colorado Supreme Court in *Tattered Cover, Inc. v. City of Thornton*, 44 P.3d 1044 (2002), a case involving a subpoena for bookstore records. In that case, the court required not only a warrant, but also additional protections. *See also* Colorado Constitution, Article 2, §10 (protecting book records from disclosure; California Constitution, Article 1, § 2(a).

information, that the information cannot be obtained by less intrusive means, that the case has *prima facie* validity, that the user has been given the opportunity to object (where allowable by law), and that the First Amendment right to anonymously speak and receive information has been appropriately considered.[34]

- **Google Books Privacy Policy:** Google's policy commits to complying with any applicable law that gives books special protection, but does not commit to seeking a particular protective standard in other cases.
- **Recommended Action:** Google should be required to seek a strongly protective standard in all cases unless binding precedent requires a different standard. We urge the Court to include the following provision in an order approving the Proposed Settlement:

  > **Proposed language:** Unless precluded by law or binding precedent, in response to government requests, Google shall take reasonable steps to require that the government obtain a court order or warrant issued upon probable cause to compel disclosure of information that could be used to identify a user of the products and services authorized by the agreement or to associate a user with access to particular books. With respect to access by civil litigants, Google shall not, unless otherwise required by law, disclose any information about users to a third party in a civil or administrative action absent a judicial and order determination that the party seeking the information has a compelling interest in the information, that the information cannot be obtained by less intrusive means, that the case has *prima facie* validity, that the user has been given the opportunity to object (where allowable by law), and that the First Amendment right to anonymously speak and receive information has been appropriately considered.

Privacy Concern #8 – Notice to Users About Data Requests (Transparency/Use Limitation FIP Principles): It is vital that readers be notified before they are identified to the government or another third party based on a legal request. The Proposed Settlement is silent on whether Google will notify users of data requests about their identity or other data.

[34] These tests are drawn from *Tattered Cover, In re Grand Jury Subpoena to Kramerbooks & Afterwords Inc.*, 26 Med. L. Rptr. 1599 (D.D.C. 1998), and *Dendrite Int'l v. Doe*, 775 A.2d 756 (N.J. App. Div. 2001).

- **Google Books Privacy Policy:** Google's policy commits to notify users of data requests unless prohibited by law.[35]
- **Recommended Action:** Google should be required to notify users of data requests unless prohibited by law. We urge the Court to include the following provision in an order approving the Proposed Settlement:

> **Proposed language:** Unless prohibited by law, Google shall seek to notify affected users of the products and services authorized by the agreement of third-party demands for information that could be used to identify them or to associate them with access to particular books. Such notification shall occur in a timely manner sufficient to allow the affected users to object to the disclosure.

Privacy Concern #9 – Registry and Rightsholder Requests for Information (Use Limitation FIP Principle): Under the terms of the Proposed Settlement, Google must share with the Registry significant *aggregate* information about readers' use of the New Services, but the Settlement does not prevent Google from sharing individualized information about users. There is no reason under the Settlement for the Registry to receive any such information, and requests for individualized information from the Registry or from Rightsholders should be treated by Google the same as requests from any other third party.

- **Google Books Privacy Policy:** Google's policy commits to only sharing aggregate information with the Registry, and to treating individualized requests from the Registry as third-party requests.[36]
- **Recommended Action:** Google should be required to share only aggregate information with the Registry as made necessary by the Proposed Settlement.

---

[35] *See* Privacy Policy ("We are committed to notifying the affected user if we receive such a request that may lead to disclosure of their information; if we are permitted to do so by law and if we have an effective way to contact the user, we will seek to do so in time for the user to challenge the request.").

[36] *See id.* ("The . . . Registry . . . will receive aggregate, non-personally identifiable information about usage of Google Books. Like any other third party, the Registry will not have access to individual user information unless it goes through proper legal processes or in other narrow circumstances set out in the Privacy Policy.").

Additionally, Google should be required to treat Registry and Rightsholder requests for individualized information as third party requests. We urge the Court to include the following provision in an order approving the Proposed Settlement:

> **Proposed language:** Google shall share with the Registry only aggregate usage information as shall be necessary to comply with the terms agreed upon in the agreement. Such information shall not directly identify any individual user of the products and services authorized by the agreement. Google shall treat Registry and Rightsholder requests for user-identifying information in the same manner as it treats such requests coming from any third party.

Privacy Concern #10 – Data Request Summary (Transparency/Accountability and Auditing FIP Principles): To gauge the impact on privacy flowing from New Services, it is very important that Google provide on a regular basis data about the number, types, and dispositions of government or third-party subpoenas, warrants, or other data requests.

- **Google Books Privacy Policy:** Google's policy does not make any commitments on this issue.
- **Recommended Action:** Google should be required to disclose to this Court and to the public the number, types, and dispositions of data requests. We urge the Court to include the following provision in an order approving the Proposed Settlement:

> **Proposed language:** Unless prohibited by law, Google shall publicly post aggregate information regarding requests for disclosure of user information related to the products and services authorized by the agreement to government or other third-party requestors. Such disclosure shall include the number of requests by government and civil litigants for user information or user-identifying data Google has received, the types of information sought, the types of legal action underlying the requests, Google's response to each request, and the types of information, if any, that were in fact disclosed.

Privacy Concern #11 – Data Retention Limitation (Data Minimization FIP Principle): The Proposed Settlement requires Google to retain information about users in identifiable form for a number of purposes. The Settlement, however, does not impose any

restrictions on how long Google may retain the information collected in connection with the New Services. It appears there is seldom if ever a need under the Proposed Settlement for Google to retain individual data about users for longer than 90 days, and 90 days has become the industry-leading standard for retention of identifiable data in the Web search industry.[37]

- **Google Books Privacy Policy:** Google's policy does not disclose how long Google intends to retain information about users.
- **Recommended Action:** Google should be required to retain user data no longer than 90 days. We urge the Court to include the following provision in an order approving the Proposed Settlement:

  > **Proposed language:** With respect to all products and services authorized by the Agreement, Google shall retain user-identifying data only as long as is reasonably necessary for the purpose for which such data is collected, but in no event shall such data be retained in identifiable form or in association with any user identifier for longer than 90 days without the express consent of the user.

Privacy Concern #12 – Security and Compliance (Security FIP Principle): The Proposed Settlement articulates detailed security and compliance requirements to ensure the security and confidentiality of information relating to the Rightsholders,[38] but the Settlement does not place any requirement on Google to secure *users'* information. Google should extend equivalent security protections to both types of information.

- **Google Books Privacy Policy:** Google's policy commits to applying strong security protections to users' information.

[37] *See* Jessica Guynn, "Yahoo to Purge User Data after 90 Days," Los Angeles Times, Dec. 18, 2008, available at http://articles.latimes.com/2008/dec/18/business/fi-yahoo18; Anne Toth, "Your Data Goes Incognito," Yahoo! Corporate blog, Dec. 17, 2008, available at http://ycorpblog.com/2008/12/17/your-data-goes-incognito/.

[38] Settlement "Security Standard".

- **Recommended Action:** Google should be required to secure user data collected in conjunction with the New Services. We urge the Court to include the following provision in an order approving the Proposed Settlement:

  > **Proposed language:** Google shall ensure the security of data concerning users' use of products and services authorized by the agreement, using measures that are substantially similar to the Security Standard for digitized files described in Attachment D to the Settlement Agreement.

CONCLUSION

The New Services enabled by the Proposed Settlement will be extraordinarily valuable, and will make available to the public a vast amount of knowledge and information that is largely inaccessible today. The Settlement should be approved. But the New Services create serious privacy concerns, and the Court must take affirmative action – as part of the settlement approval – to protect reader privacy.

Respectfully Submitted,

/s/ John B. Morris, Jr.

Leslie Harris
John B. Morris, Jr.*
David Sohn
Center for Democracy & Technology
1634 I Street, NW, Suite 1100
Washington, D.C. 20006
(202) 637-9800

* Motion for Admission *Pro Hac Vice* Pending

Dated: September 4, 2009

CERTIFICATE OF SERVICE

I, John B. Morris, Jr., hereby certify that the BRIEF *AMICUS CURIAE* OF THE CENTER FOR DEMOCRACY & TECHNOLOGY IN SUPPORT OF APPROVAL OF THE SETTLEMENT AND PROTECTION OF READER PRIVACY was transmitted to the chambers of Judge Denny Chin by overnight delivery service, and is being served by e-mail pursuant to the Notice and Preliminary Approval Order of the Court on:

Michael J. Boni
Counsel for the Author Sub-Class
Bookclaims@bonizack.com

Jeffrey Cunard
Counsel for the Publisher Sub-Class
Bookclaims@debevoise.com

Daralyn Durie
Counsel for Google
Bookclaims@kvn.com

as well as all non-party counsel who counsel has been able to identify as having entered appearances or been granted *pro have vice* status in this matter.

/s/

_____

John B. Morris, Jr., Esq.
Center for Democracy & Technology
1634 I Street, NW, Suite 1100
Washington, D.C. 20006
(202) 637-9800

Dated: September 4, 2009