NEW YORK UNIVERSITY ANNUAL SURVEY OF AMERICAN LAW

VOLUME 73
ISSUE 1

NEW YORK UNIVERSITY SCHOOL OF LAW
ARTHUR T. VANDERBILT HALL
Washington Square
New York City

PATRON DATA PRIVACY PROTECTION AT PUBLIC LIBRARIES: THE ETHICAL MODEL BIG DATA LACKS

*EMMA TROTTER**

TABLE OF CONTENTS

I. INTRODUCTION

We live in the information age.[1] Defining features include the idyllic: search engines give us instant access to facts on the go, social media helps us stay connected to friends and professional contacts throughout our lives, and online marketplaces quickly deliver

* New York University School of Law, *cum laude*, 2016; Notes Editor, *Annual Survey of American Law*, 2015–2016. I would like to thank Professors Ira Rubinstein, Katherine Strandburg, and Jason Schultz for their advice on this Note. My colleagues in the Notes Writing Program at *Annual Survey*—Harry Black, Ameneh Bordi, Ben Mejia, Georgia Stasinopoulos, and Ashley Sun—also provided extremely valuable support and suggestions.

1. *See generally* MANUEL CASTELLS, THE INFORMATION AGE: ECONOMY, SOCIETY, AND CULTURE (2d ed. 2010) (analyzing contemporary society's use of information).

global products right to our doorsteps. But these benefits have come at a dear, though often unnoticed, price. This Note describes the gradual erosion of user privacy and argues that we should look to public libraries for the solution: big data regulation that more effectively protects that privacy.[2]

In Part I, this Note defines "privacy" and explains how it is vitally important to our intellectual freedom. It also unpacks the term "big data" and explains how the current regulatory framework has allowed for unethical use of such data and the erosion of our privacy.

Part II draws a distinction between library "values" at a theoretical level and library "practices" on the ground day-to-day in order to illustrate the tension between them created by changing technology. Two areas—anonymous Internet browsing for users and libraries' hesitation to utilize data beyond the purpose for which it was originally collected ("secondary use")—illustrate how library values have generally withstood the pressure big data has placed on libraries to erode patron privacy protections in practice.

Part III advances an ethical model that looks to libraries, given this robustness of library values, for big data regulation. In this model, stakeholders maintain the ability to access information while review criteria, applied to each commercial use proposed, also ensure better protection of user privacy. For their part, libraries serve as the safe-keepers of big data and make it available for use by entities other than the provider of the information only if certain procedural safeguards are met (a function which is called a "personal data store"), thereby solving the problem set out in Part I and alleviating the tension raised in Part II. Because of their values, public libraries are ideally positioned to perform this function and help provide the ethical regulation big data lacks.

II.
BIG DATA THREATENS USER PRIVACY, A VITAL PILLAR OF INTELLECTUAL FREEDOM

This Part addresses three questions: what is privacy, how does big data threaten privacy, and why does privacy matter?

2. This Note uses the terms "user," "consumer," "customer," and "patron" interchangeably.

A. What is privacy?

Any discussion of privacy must first acknowledge that the term has no fixed meaning.[3] Definitions attempted by commentators have fallen into three general categories.

One category of attempted definitions seeks to find one essential unifying feature of privacy from which a broader set of rights or obligations flows. Responding to the advent of personal camera use, attorney Samuel Warren and Supreme Court Justice Louis Brandeis defined privacy as the "right to be let alone,"[4] a formulation which seeks to protect emotions and personality by prohibiting the unauthorized use of images.

A second category of attempted definitions seeks to exhaustively list privacy harms. Early on, legal theorist William Prosser enumerated four basic harms: intrusion upon seclusion; public disclosure of private facts; publicity which places the victim in a false light; and, similar to Warren and Brandeis' formulation, appropriation of the victim's name or likeness.[5] While arguably comprehensive at the time, these four torts no longer capture the wide variety of privacy harms that can occur in the information age.[6] Supplementing Prosser, privacy expert Daniel Solove more recently put forth a "taxonomy" of privacy harms, which share no one defining feature, but rather resemble each other, as Solove puts it, based on a set of characteristics—much like an extended family.[7] The characteristics include harm to dignity and broader architectural problems with the ways data systems are structured.[8] One of Solove's harms, "secondary use" of information originally disclosed for an initial, primary purpose, is discussed in Part II.

A third category looks to the interaction of multiple norms. In Professor Helen Nissenbaum's view, privacy regulation should be informed by social values, and such values can be different depend-

3. *See, e.g.*, Daniel J. Solove, *A Taxonomy of Privacy*, 154 U. PA. L. REV. 477, 479–80 (2006).

4. Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 193 (1890).

5. Solove, *supra* note 3, at 482.

6. One reason for this is that tort law protects physical bodily integrity at a more absolute level than it protects against emotional or dignitary harms. *See, e.g.*, *id.* at 487.

7. *Id.* at 486. The sixteen privacy harms Solove identifies in his taxonomy are surveillance, interrogation, aggregation, identification, insecurity, secondary use, exclusion, breach of confidentiality, disclosure, exposure, increased accessibility, blackmail, appropriation, distortion, intrusion, and decisional interference. *Id.* at 490–91.

8. *Id.* at 487.

ing on the context of a particular transaction.[9] Nissenbaum calls this framework "contextual integrity," which calibrates privacy protection based on the norms of specific contexts, demanding "compatibility with presiding norms of information appropriateness and distribution."[10] The framework determines "whether a particular action is determined a violation of privacy" by looking to "several variables, including the nature of the situation or context; the nature of the information in relation to that context; the roles of agents receiving information; their relationships to information subjects; on what terms the information is shared by the subject; and the terms of further dissemination."[11] These variables change depending on the context such that what is a violation within one context would be perfectly normal in another.[12] This vaguer, norms-based formulation of the right to privacy has the effect of defining privacy in a broader and more flexible manner compared to either the theory of Warren and Brandeis or that of Solove, and, as analyzed in Parts II and III, maps well onto the public library context.

In sum, this Note defines privacy as a collection of specific protections against specific violations, as well as an umbrella term encompassing social values we have traditionally held.

B. How does big data—and the way it's currently regulated—threaten privacy?

The information age has coincided with a gradual erosion of user privacy, and a primary reason for this is what we have come to call "big data." Due to ever-increasing analytical capability, big data makes the collection of highly detailed information about our Internet browsing behavior more efficient and allows those datasets to translate much more cheaply and easily into actionable insights, enabling corporations to better target us with advertisements or use scoring algorithms to assess and reduce their risk when taking on

9. Helen Nissenbaum, *Privacy as Contextual Integrity*, 79 WASH. L. REV. 119 (2004). Nissenbaum has also written about the power of search engines to serve as a force for either democratization, much like a library card catalog, or the further entrenchment of mainstream commercial interests. *See, e.g.*, Lucas Introna & Helen Nissenbaum, *Shaping the Web: Why the Politics of Search Engines Matters*, THE INFORMATION SOCIETY 169, 169–70 (2000), http://www.nyu.edu/projects/nissenbaum/papers/Shaping%20the%20Web.pdf.

10. Nissenbaum, *supra* note 9, at 155.

11. *Id.*

12. *Id.*

new customers.[13] The term "big data" encompasses several different aspects of this ongoing process. As Professors Kate Crawford and Jason Schultz write, "First, it refers to technology that maximizes computational power and algorithmic accuracy. Second, it describes types of analyses that draw on a range of tools to clean and compare data. Third, it promotes the belief that large data sets generate results with greater truth, objectivity, and accuracy."[14] In the information age, more data are collected and readily accessed by people, corporations, and governments than ever before. The rest of this Section explores how current privacy laws are ill-equipped to address the harms caused by big data and may even backfire as applied to public libraries.

1. Inadequacy of the FIPs

Privacy regulation has not kept pace with big data. The most influential approach to privacy law in the United States—the Fair Information Practices (FIPs)—percolates through dozens of sectoral laws,[15] is decades old,[16] and inadequately regulates the use of big data. Taken together, the aim of the FIPs is to protect privacy in an increasingly global and impersonal world, where consumers are no longer in contractual privity with each entity that has access to their personal data.[17] The most widely used version was developed in 1980 by the Organisation for Economic Co-operation and Development.[18] The principles are:

---

13. *See generally* Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93 (2014) (describing big data's rise, applications, and drawbacks).

14. *Id.* at 96.

15. In the context of privacy laws, "sectoral" means aimed at one particular sector of the economy. For example, the financial industry has the Fair Credit Reporting Act (FCRA) and the health insurance industry has the Health Insurance Portability and Accountability Act (HIPAA). Two sectoral laws, the Children's Online Privacy Protection Act (COPPA) and the Family Educational Rights and Privacy Act (FERPA), are discussed in the next Section.

16. The FIPs were originally put forward in a 1973 report, U.S. DEP'T OF HEALTH, EDUC. & WELFARE, RECORDS, COMPUTERS, AND THE RIGHTS OF CITIZENS: REPORT OF THE SECRETARY'S ADVISORY COMM. ON AUTOMATED PERSONAL DATA SYSTEMS (1973).

17. *Protection of Privacy and Personal Data*, ORGANISATION FOR ECON. CO-OPERATION & DEV., http://www.oecd.org/internet/ieconomy/protectionofprivacyandpersonaldata.htm (last visited Feb. 9, 2017).

18. *Privacy Principles*, ORGANISATION FOR ECON., CO-OPERATION & DEV., http://oecdprivacy.org/ (last visited Feb. 9, 2017).

- Collection limitation: Requires that personal data be obtained by lawful and fair means, such as through consent of the data subject.
- Data quality: Requires that data collected be appropriate to the purpose for which they are collected and kept up-to-date to ensure accuracy.
- Purpose specification: Requires that the reason for collecting data be clearly expressed at the time of collection.
- Use limitation: Requires that personal data only be disclosed with the consent of the data subject or as required by law.
- Security safeguards: Requires that data be protected against unauthorized use or disclosure.
- Openness: Requires that practices and policies about the use of personal data be reasonably transparent.
- Individual participation: Requires that an individual have the right to access data collected about him and to make corrections if necessary.
- Accountability: Requires that data controllers have the responsibility to successfully implement the principles.[19]

Because of their emphasis on informed consent of the data subject, the FIPs are sometimes summarized as "notice and choice," with roughly the first four FIPs guaranteeing adequate notice and the sixth and seventh FIPs emphasizing choice. Essentially, if a person is told what a data controller plans to do with her data and, based on this knowledge, decides—sometimes simply by using a service—to allow for such usage, the data practice would be unassailable under most of the FIPs, with the possible exception of security, the fifth FIP.

As currently implemented, the FIPs are inadequate to protect user privacy because, as implemented through sectoral privacy laws in the United States, they reflect a simpler time.[20] Though arguably well-suited to an era in which consumers disclosed data directly to certain businesses and maintained one-on-one relationships with most entities with access to their data, the FIPs as currently implemented have not kept pace with the advent of big data, data bro-

19. *Id.*

20. *See, e.g.*, Florencia Marotta-Wurgler, *Understanding Privacy Policies: Content, Self-Regulation, and Market Forces* 6 (N.Y.U. L. & Econ., Working Paper, Paper 435, 2016), http://lsr.nellco.org/cgi/viewcontent.cgi?article=1439&context=nyu_lewp (describing limitations of current regulatory scheme).

kers, and the interoperability[21] of datasets.[22] The FIPs emphasize notice and consent of the individual user, but in an increasingly interconnected and online world, people generally lack the capacity to make an informed judgment about sharing information with a particular entity. There are many reasons for this paradigm shift, addressed below, including difficulty assessing risk, the enhanced possibility of re-identification, and the growing prevalence of automated data processing.

Consumers often have difficulty assessing the risk of having their privacy violated because of the sheer number of online services a typical consumer uses. For example, a user might search for a product on Google before buying it on Amazon, then open Facebook to share a status update, then click a link to a news article a friend shared. Each of these sites might track the user's activity, and spending hours trying to opt out of data collection site-by-site is not reasonable. A user might also be unaware of the possibility that data anonymously provided to one service can be re-identified though aggregation of data collected across different services, which increases the difficulty of accurately assessing and responding to threats to privacy.[23] Professor Katherine Strandburg explains,

> [B]ecause it is so difficult to assess the marginal expected disutility related to data collection by any single product or service, consumers may well view avoiding data collection by any one particular product or service as a futile gesture in light of continued data collection by other products and services she uses.[24]

Essentially, though a given user may value her privacy across the range of services she uses, she may view updating her preferences on each individual service as more trouble than it's worth from a privacy perspective. The collection of data related to the user from others who do not update their preferences to better protect privacy further complicates this situation.[25] For example,

---

21. Aaron K. Perzanowski, *Rethinking Anticircumvention's Interoperability Policy*, 42 U. CAL. DAVIS. L. REV. 1549, 1553 (2009) (defining interoperability as "a relationship between two or more systems by which they exchange usable information").

22. By some accounts, the increasing irrelevance of FIPs-based regulation has led, or will soon lead, to an overall decrease in privacy suits. *See, e.g.*, Ross Todd, *Wave of Privacy Suits Peters Out*, THE RECORDER (May 29, 2015), http://www.therecorder.com/id=1202727906735/Wave-of-Privacy-Suits-Peters-Out.

23. Katherine J. Strandburg, *Free Fall: The Online Market's Consumer Preference Disconnect*, 2013 U. CHI. LEGAL F. 95, 159 (2013).

24. *Id.*

25. *Id.*

Facebook apps can often access data about a user's friends even if those friends have not consented to use the service.[26] The combined effect of these realities is that even a user who values her privacy may not be incentivized—or may feel powerless—to protect it in a particular transaction.

Another concern is that data a user provides separately to different service providers increasingly can be aggregated and, in some cases, stripped of anonymity protections. The first piece of this equation is data brokers, entities that aggregate and sell user data. For example, Acxiom is a data broker that sells consumer information to its clients so that they can target consumers for advertising campaigns.[27] Though a consumer may think data she provides to a particular service provider will stay there, many service providers sell or share consumer data with partners. Acxiom contracts with many of these service providers to further aggregate data, and then sells these datasets to still more service providers. This is problematic because data which in isolation does not identify the consumer can, in the aggregate, do just that. For example, when AOL in 2006 released supposedly de-identified information[28] consisting of search queries, journalists quickly demonstrated that at least some of the searchers could be identified based on their search terms:

> [S]earch by search, click by click, the identity of AOL user No. 4417749 became easier to discern. There are queries for 'landscapers in Lilburn, Ga,' several people with the last name Arnold and 'homes sold in shadow lake subdivision gwinnett county georgia [sic].' It did not take much investigating to follow that data trail to Thelma Arnold, a 62-year old widow who lives in Lilburn, Ga., frequently researches her friends' medical ailments and loves her three dogs. 'Those are my searches,' she said, after a reporter read part of the list to her.[29]

Similarly, in a recent study researchers were able to identify 80 percent of participants simply by tracking which websites a partici-

---

26. *Social Networking Privacy: How to Be Safe, Secure and Social*, PRIVACY RIGHTS CLEARINGHOUSE, https://www.privacyrights.org/social-networking-privacy (last visited Feb. 19, 2017).

27. Anna Maria Virzi, *Acxiom Names Scott Howe CEO*, CLICKZ (July 28, 2011), http://www.clickz.com/clickz/news/2097642/acxiom-names-scott-howe-ceo.

28. Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86 N.Y.U. L. REV. 1814, 1841 (2011).

29. Michael Barbaro & Tom Zeller, Jr., *A Face Is Exposed for AOL Searcher No. 4417749*, N.Y. TIMES, Aug. 9, 2006, at A1, http://www.nytimes.com/2006/08/09/technology/09aol.html.

pant clicked on through Twitter.[30] The researchers then "crawled through millions of Twitter profiles to see who [participants were] following" and used that information to deduce their identities.[31] The researchers noted that "[a]lthough we happen to use Twitter, it's not like Twitter is uniquely vulnerable . . . . It doesn't take a lot of recorded characteristics to have people become unique."[32] These findings are troublesome for people who think they are engaging in anonymous Internet activity. "Vanity searching," a term for when an individual searches the Internet for her own name in conjunction with various other terms and a pastime in which 47 percent of American adult Internet users reportedly engage, makes the process of re-identification even easier.[33] Even when consumers try to protect their privacy by giving out information sparingly, the prevalence of data brokers and the astonishing ease of re-identifiability mean that consumers can unwittingly expose more personal information than they intended.[34]

Finally, the increasing use of automated data processing and decision-making means that consumers, in addition to inadequately assessing the risks posed by sharing their data, must also contend with a lack of notice about how that data may be used to their detriment. The classic example of automated decision-making is credit score algorithms. When looking at her credit report, a consumer has a general idea that making two late payments in five years may have negatively affected her score. But big data capabilities have drastically increased the number and kinds of variables automated decision-making algorithms, such as credit scoring systems, can consider. Therefore, a consumer may not know that a "credit card company uses behavioral-scoring algorithms to rate consumers' credit risk because they used their cards to pay for marriage counseling, therapy, or tire-repair services," that "[a]utomated systems rank [job] candidates' talents by looking at how others rate their online contributions," that "[t]hreat assessments result in arrests or the inability to fly even though they are based on erroneous information," or that "[p]olitical activists are designated as 'likely' to

30. Vignesh Ramachandran, *You Are Less Anonymous on the Web Than You Think — Much Less*, STAN. ENGINEERING (Oct. 20, 2016), https://engineering.stanford.edu/news/you-are-less-anonymous-web-you-think-much-less.

31. *Id.*

32. *Id.*

33. Duncan Riley, *Do You Use Google for Vanity Searching? You're Not Alone*, TECHCRUNCH (Dec. 16, 2007), http://techcrunch.com/2007/12/16/do-you-use-google-for-vanity-searching-youre-not-alone.

34. *See generally* Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701 (2010).

commit crimes."[35] And she would have no way of finding out due to the proprietary nature of scores, which some claim violates the constitutional right to due process.[36] Yet again, the consumer, who may have consented to the data collection initially, finds herself the victim of a harm that seems to bear little logical relation to the original disclosure. In this manner, big data threatens user privacy in a way that the FIPs have been powerless to prevent.

2. Counter-productivity of COPPA and FERPA in the public library context

Two statutes, the Children's Online Privacy Protection Act (COPPA)[37] and the Family Educational Rights and Privacy Act (FERPA),[38] are part of the sectoral FIPs-based privacy regulation system currently charged with controlling the use of big data.[39] Though libraries are categorically not subject to either law, a pattern of over-compliance[40] threatens library values of privacy protection and access to information[41]—values which, as this Note explores in Part III, are critical to libraries' potential role as the ethical model big data lacks. Therefore, these laws illustrate how big data regulation is currently ineffective and even counterproductive.

As a preliminary matter, it is quite clear from analysis of plain meaning and agency guidance that libraries are not subject to either COPPA or FERPA. COPPA requires operators of commercial websites to obtain parental consent before collecting, using, or disclosing personal information of children under the age of 13.[42] The Federal Trade Commission (FTC), which administers COPPA, states in its guidelines that nonprofit entities are "exempt" from

35. Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 4 (2014).

36. *Id.* at 27–28.

37. 15 U.S.C. § 6502 (2012).

38. 20 U.S.C. § 1232g (2015).

39. COPPA emphasizes notice and choice, while FERPA allows for individual participation. 15 U.S.C. § 6502(b)(1)(A); 20 U.S.C. § 1232g(a)(1)(A).

40. An entity over-complies with a law when it follows a law that does not actually apply to it, either categorically or given a particular set of facts. *See, e.g.*, Rob Silverblatt, *Hiding Behind Ivory Towers: Penalizing Schools that Improperly Invoke Student Privacy to Suppress Open Records Requests*, 101 GEO. L.J. 493, 502 (2013) (discussing schools' over-compliance with FERPA).

41. For more on the connection between privacy and access to information see Part III.

42. 15 U.S.C. § 6502(b)(1)(A)(ii).

COPPA.[43] Since many public libraries are nonprofits or city agencies rather than commercial website operators, they are not subject to COPPA under FTC guidance.[44] Similarly, the text of FERPA and the Department of Education (ED) FERPA guidelines define and discuss covered institutions in a manner that suggests public libraries are exempt. FERPA imposes restrictions on the disclosure of student records that: "(i) contain information directly related to a student; and (ii) are maintained by an educational agency or institution or by a person acting for such agency or institution."[45] According to the statute, covered entities are "any public or private agency or institution which is the recipient of [ED] funds under any applicable program."[46] ED's FERPA guidelines consistently refer to covered institutions as "schools" or "a school" and make no mention of public libraries, so the agency interpretation seems to be that this law applies primarily to schools or school libraries.[47] Any public library not receiving ED funding[48] is, thus, categorically

43. Bureau of Consumer Protection Business Center, *Complying with COPPA: Frequently Asked Questions*, FED. TRADE COMMISSION (July 2013), http://www.business.ftc.gov/documents/0493-Complying-with-COPPA-Frequently-Asked-Questions.

44. The legislative history of COPPA also supports the argument that public libraries are exempt from the law. The American Library Association (ALA) weighed in on an early draft, expressing concern that the current draft "would include many non-profit organizations, including libraries." *Children's Online Privacy Protection Act of 1998: Hearing on S.2326 Before the Subcomm. on Commc'ns of the S. Comm. on Commerce, Science, and Transportation*, 109th Cong. 55 (1998). Subsequent drafts and the final text of the act included the nonprofit exception discussed above.

45. 20 U.S.C. § 1232g.

46. *Id.*

47. *Id. Cf.* Lee S. Strickland et al., *Patriot in the Library: Management Approaches When Demands for Information Are Received from Law Enforcement and Intelligence Agents*, 30 J.C. & U.L. 363, 401 n.195 (2004) (regarding the special case of libraries within schools: "Library records are not specifically mentioned in FERPA . . . yet many universities interpret these records to be covered as educational records."). *Accord* JAMES T. O'REILLY, FEDERAL INFORMATION DISCLOSURE § 19-11 (2d ed. 1990). This argument can be limited to libraries situated within schools, though, so public libraries are still exempt under Strickland's analysis. *See also* DANIEL J. SOLOVE & PAUL M. SCHWARTZ, INFORMATION PRIVACY LAW 520 (5th ed. 2015) (asserting that under FERPA, "which regulates the privacy of student data at schools, parties receiving student data from schools are not directly regulated by [ED]"). This is in contrast with the Department of Health and Human Services rule, which extends liability to business associates of covered entities. 45 C.F.R. § 164.502 (2017).

48. According to the ALA, "the majority of federal library program funds are distributed through the Institute of Museum and Library Services," an agency separate from ED, which means most libraries probably do not receive ED funding. *Appropriations*, AM. LIBR. ASS'N, http://www.ala.org/advocacy/libfunding/fed (last visited Feb. 9, 2017).

not subject to FERPA under the law's plain meaning and as interpreted by ED.[49] Taken together, the plain meaning of the text of COPPA and FERPA and the guidelines provided by the implementing agencies support the conclusion that the laws do not apply to non-university public libraries.

The American Library Association (ALA) has adopted this interpretation and stated in a variety of formats that it does not believe COPPA and FERPA apply to libraries. For example, ALA Deputy Director of the Office for Intellectual Freedom Deborah Caldwell-Stone sums up COPPA as only applying to commercial enterprises: "You're not selling data—that's the last thing you're doing as a library."[50] More formally, in contrast to COPPA's parental consent requirement, the ALA website states that "librarians should not breach a child's confidentiality by giving out information readily available to the parent from the child directly. Libraries should take great care to limit the extenuating circumstances in which they will release such information."[51] Regarding FERPA, the ALA states that "the definition of educational institution . . . excludes a library unless it is incidentally connected to an institution . . . consider[ed] educational. . . . Thus a university library might qualify as an educational institution, but the New York Public Library would not."[52] The ALA's approach is wholly distinct from what some libraries are doing in practice,[53] which has resulted in an erosion of privacy pro-

49. The legislative history of FERPA also lends itself to the proposition that the law does not apply to libraries. Similar to the ED guidelines, the text of hearings and statements reflects that senators treated the broad "agency or institution" language as referring more narrowly to schools. *See, e.g.*, *School Violence and Vandalism: Hearing Before the S. Subcomm. to Investigate Juvenile Delinquency of the S. Comm. on the Judiciary*, 94th Cong. 132 (1975) (characterizing "educational institutions" as "such schools").

50. Sarah Bayliss, *With Tighter COPPA Regulations, Librarians See Hurdles to Kids' Internet Use*, SCH. LIBR. J.: THE DIGITAL SHIFT (July 9, 2013), http://www.thedigitalshift.com/2013/07/k-12/with-tighter-coppa-regulations-librarians-seed-hurdles-to-kids-internet-use.

51. *Questions and Answers on Privacy and Confidentiality*, AM. LIBR. ASS'N., http://www.ala.org/Template.cfm?Section=interpretations&Template=/ContentManagement/ContentDisplay.cfm&ContentID=15347 (last visited Feb. 9, 2017).

52. *Freedom of Information Act Fees*, *in* AM. LIBRARY ASS'N., GOVERNMENT DOCUMENTS ROUNTABLE 137 (1987), http://sul-derivatives.stanford.edu/derivative?CSNID=80000092&mediaType=application/pdf.

53. *See, e.g.*, Benjamin Shmueli & Ayelet Blecher-Prigat, *Privacy for Children*, 42 COLUM. HUM. RTS. L. REV. 759, 784 (2011) (calling the ALA's approach "entirely different . . . from that of COPPA [and] FERPA."). *But see* BJ Ard, *Confidentiality and the Problem of Third Parties: Protecting Reader Privacy in the Age of Intermediaries*, 16 YALE J.L. & TECH. 1, 37 n.177 (2014) (calling the ALA's approach more similar to

tection and could threaten access to certain types of library materials.

Nevertheless some public libraries have chosen to comply with COPPA. For example, the Boston Public Library (BPL) requires that children ages twelve and under apply for a library card in person rather than online, and a staff member confirmed in 2012 that the reason for that restriction was to comply with COPPA.[54] Libraries that voluntarily comply with COPPA by requiring this level of parental involvement are actually failing to protect the privacy of children, thereby subverting the purpose of the law. Professor danah boyd writes, "I've always been under the impression that librarians are also committed to making sure that children have access to information, even information that might upset their parents," such as materials about abuse or homosexuality.[55] Thus over-compliance with COPPA can result in libraries unnecessarily providing children's private reading histories to their parents, therefore resulting in the erosion of privacy protections for children, contrary to COPPA's purpose. Furthermore, this result weakens libraries' historically robust protection of patron privacy, jeopardizing their ability to serve as the ethical model big data lacks.

When it comes to FERPA, schools, not libraries, needlessly comply with the statute, and this can still undermine library values. For example, Stephen F. Austin State University's "Interviewer Release Form" requires oral history interviewers to expressly authorize disclosure of the interviews "to the extent that the [i]nterviews would be considered an education record under federal law."[56] FERPA defines an educational record as "records that are: (1) Directly related to a student; and (2) Maintained by an educational agency or institution or by a party acting for the agency or institution," and excludes "[g]rades on peer-graded papers before they

---

that of COPPA: "Contrast the position articulated by ALA in interpreting the Library Bill of Rights, which affirm[s] the responsibility and the right of all parents and guardians to guide their own children's use of the library and its resources and services." (internal quotation marks omitted)).

54. danah boyd, *Are Librarians Encouraging Libraries to Abide by COPPA?*, SHEKNOWS (Feb. 24, 2012), http://www.blogher.com/are-librarians-encouraging-public-libraries-abide-coppa (assessing that BPL "seems to be going further than similar institutions"). The New York Public Library has similar requirements. *Apply for a Library Card*, N.Y. PUB. LIBR., http://www.nypl.org/help/library-card#apply (last visited Feb. 9, 2017).

55. boyd, *supra* note 54.

56. *Interviewer Release Form*, STEPHEN F. AUSTIN STATE UNIV., http://www.sfasu.edu/heritagecenter/6868.asp (last visited Feb. 9, 2017).

are collected and recorded by a teacher," among other exemptions.[57] Based on these definitions, it is not at all clear that oral history interviews would qualify as an educational record, since they are arguably focused on someone else and not directly related to a student, as well as closer to the assignment status of "peer-graded papers" than the final grade status of a bona fide record. Though it may be more efficient for schools to assume every record is an educational record subject to FERPA, this over-compliance threatens libraries' core value of access to information, a value served by projects like oral history interviews, which libraries also conduct.[58] By weakening access to information, FERPA over-compliance similarly jeopardizes libraries' ability to serve as the ethical model big data lacks.

Libraries are not subject to either COPPA or FERPA. Moreover, unnecessary compliance is actually *less* protective of privacy than noncompliance and curtails legitimate access to information. As such, COPPA and FERPA are examples of big data regulation backfiring.

In sum, big data threatens user privacy by allowing for aggregation and further use of data beyond the point of collection. Current regulation of big data grounded in the FIPs has not been effective at preventing these harms to user privacy and in some cases can even exacerbate that harm.[59] Part III returns to this problem and proposes a solution.

C. Why does privacy matter?

Some people hear the word "privacy" and think, "*I don't need privacy. I have nothing to hide.*" But the point of privacy has nothing to do with having something to hide and everything to do with the freedom to formulate intellectual views free from scrutiny.

Underlying all three definitions of privacy discussed in Part I.A is the premise that privacy is worthy of protection and that its violation constitutes a legally cognizable harm.[60] This idea is well established, perhaps most illustriously in the Fourth Amendment to the U.S. Constitution, which protects against unlawful searches and seizures.[61] However, the realities of technology, widespread government surveillance, and a host of other factors have led some com-

57. 34 C.F.R. § 99 (2017).

58. *See, e.g.*, *Community Oral History Project*, N.Y. PUB. LIBR., http://oralhistory.nypl.org/ (last visited Feb. 9, 2017).

59. *See* Crawford & Schultz, *supra* note 13, at 108.

60. *Id.*

61. U.S. CONST. amend. IV.

mentators to throw up their hands and declare that privacy is "dead."[62] Others insist on privacy's continuing vitality,[63] at least if Americans are to continue to live in a fundamentally free society[64] that frowns upon discrimination based on viewpoint.[65] During the release of the recent Consumer Privacy Bill of Rights, former President Obama stated, "Even though we live in a world in which we share personal information more freely than in the past, we must reject the conclusion that privacy is an outmoded value."[66] A recent study shows that people are approximately equally split between the two camps.[67]

Professor Neil Richards has argued that privacy is essential to the development of ideas in a free society.[68] This intellectual aspect of privacy is "the idea that records of our reading habits, movie watching habits, and private conversations deserve special protection."[69] Richards argues that this type of intellectual exploration is essential to how we formulate our views on political and social issues.[70] The ability to engage in intellectual exploration free from fear of surveillance, or the possibility of having our interests broadcast and being exposed to social judgment—an idea rooted in Warren's and Brandeis's insistence on our right to be let alone—permits us to entertain ideas other people might find offensive.[71] Richards is an absolutist on the subject of intellectual privacy, argu-

62. *See, e.g.*, Matt Hamblen, *McNealy Calls for Smart Cards*, COMPUTER WORLD (Oct. 12, 2001), http:// www.computerworld.com/s/article/64729/McNealy_calls_for_smart_cards_to_help_security.

63. *See, e.g.*, LAWRENCE LESSIG, CODE 201 (2d ed. 2006) ("There are both changes in law and changes in technology that could produce a much more private (and secure) digital environment.").

64. *See, e.g.*, NEIL RICHARDS, INTELLECTUAL PRIVACY: RETHINKING CIVIL LIBERTIES IN THE DIGITAL AGE, 176–77 (2015).

65. *See, e.g.*, Citron & Pasquale, *supra* note 35, at 27 (arguing that due process should inform basic safeguards in regard to consumer scores and automated decision-making).

66. *See, e.g.*, Kashmir Hill, *Obama Says Privacy Isn't Dead as White House Calls For 'Consumer Privacy Bill of Rights'*, FORBES (Feb. 23, 2012, 9:33 AM), https://www.forbes.com/sites/kashmirhill/2012/02/23/obama-says-privacy-isnt-dead/#2f761cc916a4.

67. *See* Heather Kelly, *Survey: Will We Give Up Privacy Without a Fight?*, CNN (Dec. 18 2014, 10:05 AM), http://www.cnn.com/2014/12/18/tech/innovation/pew-future-of-privacy/.

68. RICHARDS, *supra* note 64 at 176–77.

69. Danielle Citron, *Neil Richards on Why Video Privacy Matters*, CONCURRING OPINIONS (Jan. 4, 2012), http://concurringopinions.com/archives/2012/01/neil-richards-on-why-video-privacy-matters.html.

70. *Id.*

71. *Id.*

ing it is equally vital "whether we're reading communist or anti-globalization books; or visiting web sites about abortion, gun control, cancer, or coming out as gay; or watching videos of pornography, or documentaries by Michael Moore, or even The Hangover 2."[72]

The difficulty of browsing the Internet anonymously and the ease of exploiting consumer data wrought by big data threaten exactly this type of privacy. One example of how intellectual privacy is under threat is the use of automated decision-making discussed in the last Section in which seemingly fragmented data can be re-aggregated and lead to, for example, a consumer's credit score decreasing. Another example is a tactic known as "doxing" or "Kompromat," in which information gleaned from Internet behavior is released online as a form of retaliation;[73] the practice has risen in prevalence as big data tools have gotten more robust and would certainly mortify Warren and Brandeis, who feared mere cameras. Furthermore, all Americans are vulnerable to intrusive government monitoring of Internet and phone use,[74] which may inhibit free inquiry.[75] Given these threats, if a person logically becomes—even ever so slightly—less likely to search online or discuss with friends a certain concept, then, according to Richards, the development of ideas necessary for a free society comes under threat, because people are less likely to formulate or discuss ideas in the first place.

III.
PUBLIC LIBRARY "VALUES" AND "PRACTICES" AS APPLIED TO ANONYMOUS BROWSING AND SECONDARY USE

The remainder of this Note focuses on public libraries as a potential solution to the problems caused by big data. Public libraries

---

72. *Id.*

73. *See* Joseph Cox, *I Was Taught to Dox by a Master*, THE DAILY DOT (Dec. 11, 2015), http://www.dailydot.com/layer8/dox-doxing-protection-how-to/.

74. James Ball et al., *Revealed: How US and UK Spy Agencies Defeat Internet Privacy and Security*, THE GUARDIAN (Sept. 6 2013), https://www.theguardian.com/world/2013/sep/05/nsa-gchq-encryption-codes-security.

75. "[A]fter Edward Snowden revealed the extent of the National Security Agency's spying on citizens in 2013, Google searches for terrorism-related terms such as *al Qaeda* dropped." Matthew Hutson, *Even Bugs Will Be Bugged: Exploring the Next Frontiers in Surveillance*, THE ATLANTIC (Nov. 2016), https://www.theatlantic.com/magazine/archive/2016/11/even-bugs-will-be-bugged/501113/.

are some of America's most trusted institutions[76] and have historically operated as paragons of user privacy.[77] These values make them a natural source of possible solutions to the erosion of privacy caused by big data. But libraries have also not been immune to the pressures of big data. In practice, new technologies available to librarians, such as Amazon e-books and the BiblioCommons card catalog system,[78] exert pressure to loosen historically protective privacy

76. *See, e.g.*, CENTER FOR AN URBAN FUTURE, RE-ENVISIONING NEW YORK'S BRANCH LIBRARIES 35 (2014), https://nycfuture.org/pdf/Re-Envisioning-New-Yorks-Branch-Libraries.pdf (outlining how "libraries are the most trusted institution for immigrants"); 18 U.S.C. § 2709(g) (2015) (excluding libraries from the definition of "wire or electronic communication" covered by the Stored Communications Act). Moreover, as libraries evolve beyond their traditional collection function, a 2015 study showed that over "two-thirds of Americans agree that libraries are important because they improve the quality of life in a community, promote literacy and reading, and provide many people with a chance to succeed." *The State of America's Libraries: 2015*, AM. LIBR., Apr. 2015, at 2, http://www.ala.org/news/sites/ala.org.news/files/content/0415_StateAmLib_0.pdf. Furthermore, in 2012, "there were 92.6 million attendees at the 4 million programs offered by public libraries [representing] a 10-year increase of 54.4% in program attendance." *Id.* at 3. With libraries continuing to play an important role in American communities, it is worth focusing on—and working to address—the privacy issues they face.

77. *See, e.g.*, *Privacy Policy*, S.F. PUB. LIBR., http://sfpl.org/index.php?pg=2000001301 (last visited Feb. 9, 2017) ("Protecting library user privacy and keeping confidential information that identifies individuals or associates individuals with their use of library books, materials, equipment, programs, facilities, and/or staff assistance is an integral principle of the Library."). The privacy of borrowing history has traditionally been protected even by states with lax library privacy laws, in part because the constitutional aspect of the right to privacy, embodied in the First Amendment, has long encompassed the idea of intellectual freedom to explore ideas in private. *See, e.g.*, Stanley v. Georgia, 394 U.S. 557, 565 (1969) ("If the First Amendment means anything, it means that a State has no business telling a man, sitting alone in his own house, what books he may read or what films he may watch. Our whole constitutional heritage rebels at the thought of giving government the power to control men's minds."). This value flies in the face of the modern practice of sharing every last search term with the highest bidding online advertiser in order to provide more targeted search results.

78. *See, e.g.*, April Glaser & Alison Macrina, *Librarians Are Dedicated to User Privacy. The Tech They Have to Use Is Not*, SLATE (Oct. 20, 2014, 12:33 PM), http://www.slate.com/blogs/future_tense/2014/10/20/adobe_s_digital_editions_e_book_software_and_library_patron_privacy.html ("When you check out a library e-book for Amazon Kindle . . . Amazon keeps a list of the library e-book titles checked out, with no option to 'opt-out' of this data collection—in addition to the other personally identifiable information Amazon collects with cookies."); *The Digital Revolution: Tough Challenges and Exciting Possibilities*, AM. LIBR. ASS'N, http://www.ala.org/news/mediapresscenter/americaslibraries/librariestechnology (last visited May 17, 2015) (arguing that "the digital revolution shows no signs of slowing, and the library community is both struggling to keep up and envisioning future library ser-

standards.[79] Libraries also have a long history of combating government incursions into user records,[80] a commitment that these new technologies may not share. The advent of Amazon e-books, BiblioCommons, and other new technologies creates a problem for users.[81]

This Part discusses how library values—primarily ALA guidelines—are, despite the pressure big data has placed on libraries to loosen privacy protections, honored in libraries' day-to-day practices, specifically those regarding anonymous browsing and secondary use.

A. Library values

The ALA[82] places a high value on user privacy. For example, the ALA's patron bill of rights imports many FIPs-like concepts into the library context. First, both the use limitation and collection limitation FIPs are neatly embodied in the following directive: "Limit the degree to which personally identifiable information is collected,

---

vices that incorporate new philosophies, new technologies and new spaces to meet the needs of all users more effectively than ever before").

79. For example, the privacy policy for the popular card catalog software BiblioCommons is much less protective than the privacy policy of a typical public library, and pinging it can involve sending requests to unsecured sites and servers. Eric Hellman, *Analysis of Privacy Leakage on a Library Catalog Webpage*, GO TO HELLMAN (Sept. 15, 2014), http://go-to-hellman.blogspot.com/2014/09/analysis-of-privacy-leakage-on-library.html.

80. *See, e.g.*, Tattered Cover, Inc. v. City of Thornton, 44 P.3d 1044, 1050 (Colo. 2002) (in which "[a]n official from the American Library Association testified about the chilling effect that results from disclosure of library circulation records"); *ALA and National Security Letters 2009*, AM. LIBR. ASS'N, http://www.ala.org/news/mediapresscenter/presscenter/onlinemessagebook/nationalsecurityletters_tp (last visited Feb. 9, 2017) ("ALA has actively opposed the use of National Security Letters since the USA PATRIOT Act was introduced, believing that the protection of library users' privacy and confidentiality is necessary for the protection of intellectual freedom."). Sharing records with an additional party may also be an issue for Fourth Amendment privacy protection because of third-party doctrine. *See, e.g.*, Smith v. Maryland, 442 U.S. 735, 744 (1979).

81. *See, e.g.*, Bonnie Tijerina, *Developing a Consensus Framework for Patron Privacy*, MEDIUM (Apr. 7, 2015), https://medium.com/@bonlth/developing-a-consensus-framework-for-patron-privacy-7668094ad4f8 ("When there's so much to gain from translating raw patron data into meaningful and useful information to learn about our communities or improve services and products, how do [you] know where to draw the line on use or non-use of our patron's information?").

82. The ALA is "the oldest and largest library association in the world." *About ALA*, AM. LIBR. ASS'N, http://www.ala.org/aboutala/ (last visited Nov. 3, 2016). Its mission is "to provide leadership for the development, promotion and improvement of library and information services and the profession of librarianship in order to enhance learning and ensure access to information for all." *Id.*

monitored, disclosed, and distributed."[83] Second, different policies exemplify the FIPs as summarized by the ideas of notice and choice. The ALA recommends placing "the user in control of as many choices as possible," a directive that necessarily implies a user would also have notice about those choices.[84] Finally, the openness and individual participation FIPs are reflected in user access to their borrowing histories through the online systems employed by many libraries, a transparent practice which enables users to know what data are being collected about them.[85] Measured by the FIPs, libraries protect patron privacy at a high level.

Nissenbaum's emphasis on context[86] helps explain why libraries protect privacy at such a high level; libraries' core mission of access to information and their lack of a profit motive create and constantly reinforce norms that are highly protective of patron data privacy. At the heart of the mission of a public library is access to information. Libraries exist as repositories of information, and staff members stand by to assist patrons in their research questions. Layering on Richards's argument that freedom from scrutiny of research is essential to the development of new ideas, it becomes clear that a commitment to information access carries with it an implied commitment to anonymous information access. In order to fully explore ideas and make up their minds on complicated issues, people need to feel free from the type of scrutiny created by search engines that log every query and allow re-identification of even purportedly anonymous data. Richards writes,

> For generations, librarians have understood this. Libraries were the Internet before computers—they presented the world of reading to us, and let us as patrons read (and watch) freely for ourselves. But librarians understood that intellectual privacy matters. A good library lets us read freely, but keeps our

83. *Resolution on the Retention of Library Usage Records*, AM. LIBR. ASS'N, http://www.ala.org/Template.cfm?Section=ifresolutions&Template=/ContentManagement/ContentDisplay.cfm&ContentID=135888 (last visited Feb. 9, 2017).

84. *Privacy: An Interpretation of the Library Bill of Rights*, AM. LIBR. ASS'N, http://www.ala.org/PrinterTemplate.cfm?Section=interpretations&Template=/ContentManagement/ContentDisplay.cfm&ContentID=132904 (last visited Feb. 9, 2017).

85. *See, e.g.*, *Quick Start Guide*, N.Y. PUB. LIBR., https://www.nypl.org/sites/default/files/BCQSGGettingStartedPRIVACY.pdf (last visited Feb. 9, 2017) ("Your borrowing history is visible only to you."). In contrast, patron browsing histories are typically not stored at all. *See, e.g.*, *Your Privacy on Public Computers*, NORTH OLYMPIC LIB. SYS., http://www.nols.org/about-nols/public-computer-privacy.html (last visited Mar. 8, 2017).

86. *See* Nissenbaum, *supra* note 9.

> records confidential in order to safeguard our intellectual privacy.[87]

Through libraries, patrons are free to read whatever they want and are therefore free to think and express whatever they want. In contrast, in the heavily scrutinized and minutely tracked world of online search, this latter kind of freedom is curtailed.

Additionally, public libraries have a core public-service mission that is unadulterated by a profit motive, which could lead libraries to use information for purposes other than those for which it was collected, something commercial data brokers like Acxiom do routinely.[88] With no need to seek out ever-more-profitable customers, libraries are free to adopt the approach of doing right by everyone and honoring the commitments made in their privacy policies.[89] Richards explains, "[S]haring has to be done on [patrons'] terms, not on those that are most profitable for business."[90] Honoring this belief has made libraries trusted above and beyond other institutions in the United States.[91]

Because the twin goals of providing the public service of access to information while ethically limiting the use of patron information define the library context as used by Nissenbaum,[92] certain actions are off-limits for libraries. For example, contextually, it would be inappropriate for libraries to use patron data for reasons other than those for which it was collected, and it would therefore violate the integrity of the library context for libraries to sell data to third parties, like commercial data brokers do. Patrons provide their personal data to apply for a library card and borrow books, and both the nonprofit framework and libraries' commitment to intellectual freedom dictate that this data—from name and email address to browsing history—not be shared elsewhere. Similarly, it would be

87. Danielle Citron, *Neil Richards on Why Video Privacy Matters*, CONCURRING OPINIONS (Jan. 4, 2012), http://concurringopinions.com/archives/2012/01/neil-richards-on-why-video-privacy-matters.html. *See also* Melissa Morrone, *How Your Local Library Can Help You Resist the Surveillance State*, WAGING NONVIOLENCE (July 8, 2014), http://wagingnonviolence.org/feature/local-library-can-help-resist-surveillance-state/ (characterizing libraries as "community anchor institutions").

88. Jason Morris & Ed Lavandera, *Why Big Companies Buy, Sell Your Data*, CNN (Aug. 23, 2012), http://www.cnn.com/2012/08/23/tech/web/big-data-acxiom/.

89. See *infra* Table 1.

90. Citron, *supra* note 87.

91. *See, e.g.*, CENTER FOR AN URBAN FUTURE, *supra* note 76 (outlining how "[l]ibraries are the most trusted institution for immigrants"); 18 U.S.C. § 2709(g) (2015) (excluding libraries from the definition of "wire or electronic communication" covered by the Stored Communications Act).

92. Helen Nissenbaum, *Privacy As Contextual Integrity*, 79 WASH. L. REV. 119 (2004).

contextually inappropriate for librarians to deny access to information based on disagreeing with the content of certain materials such as political manifestos or religious texts.

Despite pressures wrought by big data to erode people's privacy as discussed in Parts I.B and III, library values retain a strong commitment to the protection of patron privacy as illustrated by two library practices that embody their values: allowing anonymous browsing and minimizing secondary use.

B. Practice #1: Anonymous browsing

Many libraries allow users to conduct research and browse the Internet anonymously, which is a dying commodity in the information age. This practice is very closely aligned with library values.

Libraries enable anonymous browsing in a several ways. First, many public library computers do not require logins,[93] and those that do often do not keep track of browsing history.[94] If a patron uses the computer solely to do research and refrains from logging into any services tied to his identity,[95] he can effectively browse anonymously.[96] Though search engines can still log queries associated with the computer's IP address, the risk of re-identification, as described in Part I.B, is comparatively low because multiple users access the computer. Libraries do not allow anonymous browsing by mere happenstance; rather, ALA policies point to the important reasons behind allowing such intellectual exploration: "In a library (physical or virtual), the right to privacy is the right to open inquiry without having the subject of one's interest examined or scrutinized

93. *See, e.g.*, *Library Services: Computer and WiFi Access*, DALL. PUB. LIBR., https://dallaslibrary2.org/services/wifi.php (last visited Feb. 9, 2017) (specifying that "[c]atalog and database computers at the library do not require any card to use").

94. *See, e.g.*, *Acceptable Use Policy*, THE PUB. LIBR. OF BROOKLINE, https://www.brooklinelibrary.org/wp-content/uploads/2016/06/AcceptableUsePolicy.pdf (last visited Feb. 9, 2017) (revealing that "although the library keeps no records of your Internet activity, it does record login and logout times linked to the barcode on your library card"). The fact that Patron X logged in at 2:15 and logged out at 3:15 reveals very little information and is almost as good as not requiring a login.

95. *See, e.g.*, Samantha Felix, *This Is How Facebook Is Tracking Your Internet Activity*, BUSINESS INSIDER (Sept. 9, 2012), http://www.businessinsider.com/this-is-how-facebook-is-tracking-your-internet-activity-2012-9?op=1/#-started-off-as-just-a-normal-day-1 (describing how Facebook tracks browsing activity on other websites while a user is logged in).

96. Because of big data capabilities, this feat is nearly impossible on a personal computer. *See, e.g.*, Barbaro & Zeller, *supra* note 29.

by others."[97] Similarly, "[p]rotecting user privacy and confidentiality is necessary for intellectual freedom and fundamental to the ethics and practice of librarianship."[98] As discussed in Part I.C, Richards has argued that this kind of privacy in accessing information is essential to the development of ideas in a free society.[99] By enabling anonymous research and consequently the exploration and formation of intellectual ideas free from scrutiny or judgment, libraries provide what Richards considers an essential service in our democracy.

Other commentators view untraceable browsing with more fear. For them, with anonymity comes a lack of accountability that could allow Internet users to carry out crimes, such as those against children, without getting caught.[100] But the Children's Internet Protection Act (CIPA) requirement that public library computers filter out Internet content that could be harmful to children, such as pornographic or violent material, somewhat tempers this concern.[101] A full analysis of CIPA is beyond the scope of this Note, but it suffices to say that, although patrons may not access the full range of the Internet's dark corners on public library computers, libraries provide a vital service to our democracy by allowing patrons to browse anonymously. By allowing anonymous browsing, libraries are honoring in practice the privacy values espoused by the ALA.

---

97. *An Interpretation of the Library Bill of Rights*, AM. LIBR. ASS'N, http://www.ala.org/PrinterTemplate.cfm?Section=interpretations&Template=/ContentManagement/ContentDisplay.cfm&ContentID=132904 (last visited Mar. 8, 2017).

98. *Resolution on the Retention of Library Usage Records*, AM. LIBR. ASS'N, http://www.ala.org/Template.cfm?Section=ifresolutions&Template=/ContentManagement/ContentDisplay.cfm&ContentID=135888 (last visited Feb. 9, 2017).

99. *See supra* Part I.C; RICHARDS, *supra* note 64 at 176–77.

100. *See, e.g.*, Solon Barocas & Helen Nissenbaum, *Big Data's End Run Around Anonymity and Consent*, *in* PRIVACY, BIG DATA, AND THE PUBLIC GOOD: FRAMEWORKS FOR ENGAGEMENT 50 (Victoria Stodden et al. eds., 2014); McIntyre v. Ohio Elections Comm'n, 514 U.S. 334, 377 (1995) (Scalia, J., dissenting) (arguing that anonymity in the election context should be secondary to the "universal and long-established American legislative" tradition of disclosure).

101. Children's Internet Protection Act (CIPA), Pub. L. No. 106-554, § 1701, 114 Stat. 2763A-335 (2000). As a condition of federal funding, CIPA requires libraries and schools to use Internet filtering, among other measures, to protect children from viewing harmful content online. CIPA was deemed not to violate the First Amendment in *United States v. American Library Ass'n, Inc.*, 539 U.S. 194 (2003).

C. Practice #2: Secondary use

In many contexts, data are transmitted from one entity to another for a particular, limited purpose. For example, a person might consent to provide her doctor with information about her recent symptoms for the purpose of allowing the doctor to make an informed diagnosis. She would be less likely to consent to publication of the same information in her local newspaper. Big data, however, makes exactly this type of "secondary use" attractive, because automated scoring or targeted advertisements become easy and cheap to accomplish using larger and larger datasets.[102] But secondary use is a violation of contextual integrity as defined by Nissenbaum[103] because information can be shared across contexts without regard for why and to whom it was originally shared.

The remainder of this Section demonstrates that libraries, with some exceptions, generally act in keeping with their values by employing an opt-in approach to secondary use. Libraries collect data, such as name, address, e-mail address, and phone number, when users apply for library cards, and they typically will not use this data for any other purpose, such as fundraising, unless the user opts in. Though the FIPs require the consent of the data subject, they do not specify whether that consent must be opt-in rather than opt-out.[104] But the opt-in approach generally taken by libraries provides a superior level of protection to data subjects, because they have knowledge of and control over purposes for which their data may be used. Therefore, libraries' approach to secondary use in practice is generally consistent with their values.

The difference between the opt-in and opt-out approaches turns on whether the default is consent or lack thereof. While an opt in means that inaction by the customer leads to no data processing conducted, "[a]n 'opt out' means that a consumer's information will be processed unless she takes action to contact the data processing entity and indicate her contrary wishes."[105] For example, if a library card applicant at San Jose Public Library (SJPL), when confronted with the below choice, takes no action (i.e., leaves the

102. *See supra* Part I.C; William Hersch, *Secondary Use of Clinical Data from Electronic Health Records*, https://dmice.ohsu.edu/hersh/secondary-use-trec.pdf.

103. *See generally* Nissenbaum, *supra* note 9.

104. *But see, e.g.*, J. C. Bruno & Elsa Crozatier, *Compliance with the European Union Directive in the Transfer of Employee Personal Data*, 83 MICH. B.J. 48, 49 (2004) (discussing the EU Directive's requirement of opt-in consent for "sensitive" information).

105. DANIEL J. SOLOVE & PAUL M. SCHWARTZ, INFORMATION PRIVACY LAW 791 (2015).

box unchecked) she will not receive fundraising requests from SJPL's "supporting organizations."

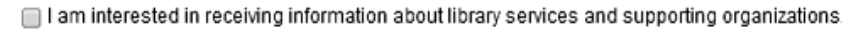

Figure 1: San Jose Public Library,[106] Opt-In

In contrast, if a library card applicant at Seattle Public Library (SPL), when confronted with the below choice, takes no action (i.e., leaves the YES button selected) she will probably receive fundraising requests from SPL's Foundation at least a few times a year.

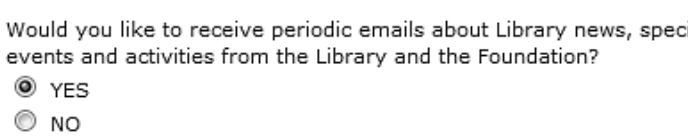

Figure 2: Seattle Public Library, Opt-Out[107]

Whether the default is set to opt-in or opt-out is particularly important because studies show that people are very likely to simply accept the default presented to them rather than taking any action.[108] Thus the opt-out approach is less protective of user privacy but attractive to businesses who may enter the big data marketplace and re-identify or sell user data without further consent.

Supporters of the opt-out approach generally emphasize that opt-in would be expensive for businesses to comply with, but this argument is flawed. For example, Professors Michael Staten and Fred Cate hypothesize "higher prices, reduced benefits, diminished service, and higher acceptance standards for new credit products" if financial institutions were required to use opt-in consent.[109] But this argument takes as a given that the benefit of having access to yet another credit card offer at a potentially cheaper rate trumps the harm to consumers of turning over their data to a seemingly endless array of service providers looking to market to them. This is partly because the argument about higher expenses inherently favors the side best able to quantify costs, and privacy is hard to pin

106. *Online Library Card Application: Adult*, San José Libr., https://catalog.sjlibrary.org/selfreg/public (last visited Feb. 9, 2017).

107. *Library Card Application*, Seattle Pub. Libr., https://www.spl.org/using-the-library/get-started/library-card-application (last visited Feb. 9, 2017).

108. Eric J. Johnson et al., *Defaults, Framing and Privacy: Why Opting In-Opting Out*, 13 Marketing Letters 5, 13 (2002), https://www0.gsb.columbia.edu/mygsb/faculty/research/pubfiles/1173/defaults_framing_and_privacy.pdf.

109. Michael E. Staten & Fred H. Cate, *The Impact of Opt-In Privacy Rules on Retail Markets: A Case Study of MNBA*, 52 Duke L.J. 745, 776 (2003).

down to a specific monetary value.[110] The argument is also overly simplistic, because it is expensive for businesses to perform or seek any service, but some such expenses must be deemed acceptable costs of doing business.[111] Staten and Cate have not demonstrated that an opt-in requirement is economically irrational based on a weighing of the costs and benefits that encompasses the value of user privacy.

Supporters of the opt-in approach, on the other hand, argue that the unrestricted data sharing facilitated by the opt-out approach constitutes harm to consumers and, as such, data processors should be required to bear additional costs to mitigate that harm. For example, Professor Jeff Sovern points out that businesses have little incentive to use opt-in systems when, under current regulation, the opt-out approach allows them to "engage[ ] in an activity that imposes costs on others but [does not] require[ ] [them] to take those costs into account when deciding whether to pursue the activity."[112] He continues, "An opt-in system . . . can shift costs and thereby 'internalize' this externality."[113] Similarly, Professors Edward Janger and Paul Schwartz point out the information asymmetry problem inherent in the opt-out approach. Specifically, opt-out default laws "fail[ ] to create any penalty on the party with superior knowledge" (i.e., businesses) and "leave[ ] the burden of bargaining on the less informed party, the individual consumer."[114] For this reason, proponents of opt-in argue that the burdens and costs of placing consumer data in jeopardy should fall squarely on service providers, who are the primary beneficiaries of consumer data and the cheapest cost-avoiders of the potential harm to consumers.

Some critics of notice and choice believe that the advent of predictive big data analytics, which use data that consumers provide to guess other personal information about them, makes informed opt-in impossible. Crawford and Schultz argue that these "predictive privacy harms," such as guessing who is gay or who carries a certain disease, in circumstances where users would not disclose that information directly, cannot be addressed by opt-in because consumers cannot have full knowledge of what they are opting in

110. *See, e.g.*, James P. Nehf, *Incomparability and the Passive Virtues of Ad Hoc Privacy Policy*, 76 U. COLO. L. REV. 1, 29 (2005).

111. *Cf.* Diana Farrell, *The Real New Economy*, HARV. BUS. REV. (Oct. 2003), https://hbr.org/2003/10/the-real-new-economy.

112. Jeff Sovern, *Opting In, Opting Out, or No Options at All: The Fight for Control of Personal Information*, 74 WASH L. REV. 1033, 1106 (1999).

113. *Id.*

114. Edward J. Janger & Paul M. Schwartz, *The Gramm-Leach-Bliley Act, Information Privacy, and the Limits of Default Rules*, 86 MINN. L. REV. 1219, 1241 (2002).

to.[115] Even if a website discloses that data collected will be used to predict other attributes about a user, a user cannot know what exactly will be predicted about her and thus cannot accurately assess the risk. But this argument fails to account for how granular opt-in consent could be, based on, for instance, authorized types of data, authorized recipients of the data, authorized uses of the data, or any combination thereof. If consent is made granular enough, thereby allowing users to opt in to none, some, or all of these options, and information is stored centrally so that a person only has to update privacy preferences in one location, the opt-in approach can overcome this type of critique. Part III will continue this analysis of the unexplored potential of transparency, centrality, and individual access in the big data context.

The ALA comes down strongly in favor of the opt-in approach, noting "[a]ny use of [personally identifiable information] beyond circulation or administration should be authorized only on an opt-in basis."[116] Following the ALA's lead, the majority of public libraries do not use patron directory information for fundraising purposes. Of the seventeen U.S. public libraries examined,[117] only four use a fully opt-out approach; one uses a hybrid approach, and the rest operate as opt-in:

Name	Opt-in or opt-out?	Location of policy
Atlanta-Fulton PL	Opt-in.	"Ask a Librarian" chat session, July 9, 2014.[118]
Boston PL	Opt-out.	Disclaimer.[119]
Chattanooga PL	Opt-in.	Email exchange, July 23, 2014.
Cleveland PL	Opt-in.	Checkbox (default is unchecked).[120]
Chicago PL	Opt-in.	"Ask a Librarian" chat session, July 9, 2014.

115. Crawford & Schultz, *supra* note 13, at 95.

116. Am. Libr. Ass'n, *supra* note 51.

117. These seventeen libraries were chosen as industry leaders in consultation with public library staff.

118. In a typical exchange, I would ask if the library ever used the e-mail address I provided to register for a library card to contact me for fundraising purposes. More often than not, I got responses such as "Of course not," or "No, never," reflecting among librarians a high degree of sensitivity to privacy issues and an understanding of their responsibility to protect user privacy.

119. *Patron Privacy Policy*, Bos. Pub. Libr., http://www.bpl.org/general/policies/privacy.htm (last visited Feb. 9, 2017) ("We may use your email address for notices and library promotions.").

120. *Library Card Application*, Clev. Pub. Libr., https://onlinereg.cpl.org/ (last visited Feb. 9, 2017).

Name	Opt-in or opt-out?	Location of policy
County of Los Angeles Library System	Opt-in.	Disclaimer.[121]
Free Library of Philadelphia	Opt-in.	Disclaimer.[122]
Hennepin County Library	Opt-in.	Disclaimer.[123]
Houston PL	Opt-in.	"Ask a Librarian" text message exchange, July 23, 2014.[124]
King County PL	Opt-in.	Checkbox (default is unchecked).[125]
Los Angeles PL	Opt-in.	"Ask a Librarian" text message exchange, July 9, 2014.
Multnomah County Library	Hybrid.	Disclaimer.[126]
New York PL	Opt-out.	Checkbox (default is checked).[127]
San Francisco PL	Opt-in.	Checkbox (default is unchecked) and disclaimer.[128]

121. *Privacy Policy*, County L.A. Pub. Libr., http://www.colapublib.org/privacy.php (last visited Feb. 9, 2017) ("We do not permit this information to be used for marketing purposes.").

122. *Site Privacy*, Free Libr. Phila., https://libwww.freelibrary.org/policies/privacy/%20htm (last visited Feb. 9, 2017) ("Our primary use of the personal information you volunteer is to contact you regarding service issues, or if you opt for it, news information about Library services and partnerships.").

123. *Privacy and Security*, Hennepin County Libr., http://www.hennepin.us/your-government/open-government/accessibility-privacy-security (last visited Feb. 9, 2017) ("The county won't collect personal information about you unless you choose to provide it. The county doesn't give, share, sell, or transfer any personal information for commercial purposes.").

124. "Houston Library Foundation will not send you any unsolicited information, including email regarding any commercial offers or advertisements at any time." Text message from Houston Library Found. to author (July 9, 2014) (on file with author).

125. *Get A Library Card*, King County Libr. Sys., https://w3.kcls.org/get-a-library-card (last visited Feb. 9, 2017).

126. *Privacy and Confidentiality of Library Records*, Multnomah County Libr., https://multcolib.org/policies-manuals/statement-privacy-and-confidentiality-library-records (last visited Feb. 9, 2017). ("The Library will not collect or retain your . . . information without your consent.).

127. *Apply for a Library Card*, N.Y. Pub. Libr., https://catalog.nypl.org/selfreg/patonsite (last visited Feb. 9, 2017).

128. *Privacy Policy*, S.F. Pub. Libr., http://sfpl.org/index.php?pg=2000001301 (last visited Feb. 9, 2017) ("Library users may choose to opt in and enable My Check-out History. . . . Any information the library user chooses to provide will be used only to provide or improve library services, such as information gathered through voluntary library user surveys.").

Name	Opt-in or opt-out?	Location of policy
San Jose PL	Opt-in.	Checkbox (default is unchecked).[129]
Seattle PL	Opt-out.	Checkbox (default is checked) and disclaimer.[130]
Topeka and Shawnee County PL	Opt-out.	Phone call, Feb. 27, 2017.

By employing an opt-in approach to secondary use, most libraries are honoring in practice the privacy values espoused by the ALA.

Given the robustness of library commitments to user privacy, as underscored by library practices of anonymous browsing and secondary use, libraries are well-positioned to serve as a new ethical model for big data, primarily due to their context: lack of a profit motive and inherent expertise in handling and ensuring access to information. The next Part advances one idea for how libraries could achieve this goal.

IV.
PUBLIC LIBRARIES CAN PROVIDE THE ETHICAL MODEL BIG DATA LACKS BY SERVING AS PERSONAL DATA STORES IN TOMORROW'S BIG DATA MARKETPLACE

Because libraries' privacy-protective values, non-profit context and expertise in handling information have inspired the kind of consumer trust that could generate better data and, ultimately, better, more ethically-driven research, libraries are well-positioned to help regulate big data. Richards reminds us that in order to fully explore ideas and make up their minds on complicated issues, people need to feel free from scrutiny, as they traditionally have at libraries.[131] While, as Richards acknowledges, "librarians aren't often thought of as particularly imaginative or innovative . . . this stereotype is wrong. Librarians are our first and oldest information pro-

129. *Online Library Card Application: Adult*, SAN JOSÉ LIBR., https://catalog.sjlibrary.org/selfreg/public (last visited Feb. 9, 2017). *See also* Figure 1, *supra*.

130. *Library Card Application*, SEATTLE PUB. LIBR., https://www.spl.org/using-the-library/get-started/library-card-application (last visited Feb. 9, 2017). *See also* Figure 2, *supra*.

131. *See* Danielle Citron, *Neil Richards on Why Video Privacy Matters*, CONCURRING OPINIONS (Jan. 4, 2012), http://concurringopinions.com/archives/2012/01/neil-richards-on-why-video-privacy-matters.html. *See also* Melissa Morrone, *How Your Local Library Can Help You Resist the Surveillance State*, WAGING NONVIOLENCE (July 8, 2014), http://wagingnonviolence.org/feature/local-library-can-help-resist-surveillance-state/ (characterizing libraries as "community anchor institutions").

fessionals, with special expertise in the issues intellectual records raise."[132] This Part explains how libraries could serve as "personal data stores" once this technology is fully developed.

One of the most salient problems with the FIPs, addressed in Part I.B, is that they place the onus on the user to become informed about her choices and take steps to change default settings to better protect her privacy. For example, the "individual participation" FIP[133] allows users to find out what information exists about them and how that information is used. But this approach works best for people who already care about privacy and are therefore willing to take the time to educate themselves about how their data are being used and then implement the preferences they develop during this process. Some users lack this inherent interest, and thus the incentive to take any steps away from the defaults given. Additionally, as discussed in Part I.B.1, even users who want to protect their privacy may decide that it is simply not efficient to do so under the current regulatory framework.

In the library context, this problem is compounded by structural incentives to over-comply with privacy laws like COPPA and FERPA, as discussed in Part I.B.2. Regarding COPPA, intended to protect children's privacy and safety online, libraries could of course simply stop over-complying, but this is unlikely to happen without broader structural change. Economist John E. Calfee and Professor Richard Craswell explain that a rational actor is incentivized to over-comply when the cost of over-compliance is minimal and the risk of incurring liability is still present despite compliance given the ambiguous norm.[134] Acting rationally, librarians might determine that the additional costs of requiring parental consent for children to obtain library cards is relatively small compared to the risk of either the legal norm shifting (i.e., the FTC changing its guidelines and deciding to enforce the law against nonprofits that operate websites for children) or societal norms coming to strongly

132. RICHARDS, *supra* note 64, at 176–77. Furthermore as libraries evolve beyond their traditional collection function, a 2015 study showed that over "two-thirds of Americans agree that libraries are important because they improve the quality of life in a community, promote literacy and reading, and provide many people with a chance to succeed." *The State of America's Libraries: 2015*, AM. LIBR., Apr. 2015, at 2, http://www.ala.org/news/sites/ala.org.news/files/content/0415_StateAmLib_0.pdf. Furthermore in 2012, "there were 92.6 million attendees at the 4 million programs offered by public libraries [representing] a 10-year increase of 54.4% in program attendance." *Id.* at 3.

133. ORGANISATION FOR ECON. CO-OPERATION & DEV., *supra* note 17.

134. John E. Calfee & Richard Craswell, *Some Effects of Uncertainty on Compliance with Legal Standards*, 70 VA. L. REV. 965, 966–67 (1984).

favor parental consent and access over children's privacy. Regarding FERPA, intended to protect students from unwanted disclosure of their academic records, schools may have simply found over-compliance a convenient rationale for keeping hidden various other kinds of records, from ethical lapses in athletic management to violent crimes among students.[135] Therefore, neither libraries nor schools are properly incentivized by the existing FIPs structure to stop over-complying, which is detrimental to the library values of both privacy and access to information.

Because of the inadequacy of the FIPS, there is a void between how much we as a society may value privacy and how much we as individual users can do to protect it. If libraries are willing to fully embrace the potential of big data—and the responsibility that comes with it—they have a chance to position themselves at the center of the action by becoming personal data stores (PDS), a term for systems that could provide consumer data access to both consumers and, after consent, entities seeking to conduct research using that data.[136] According to Professor Ira Rubinstein, PDS are a still-developing concept that in execution could "provide both a secure data store for a wide variety of personal information (including official records like birth and marriage certificates, licen[s]es, and passports, transaction records, online profiles, and social media content, and user names and passwords) as well as a new class of user-driven services."[137] Library-managed PDS could offer user-friendly and privacy-protective access to information in keeping with libraries' historically high privacy standards, allowing the archaic FIPs approach—and with it the incentive to over-comply—to be phased out.

A necessary first step to implementing library-managed PDS would be convincing libraries to rise to this challenge. There is

135. *See* Silverblatt, *supra* note 40 at 502–04.

136. *See* Richards, *supra* note 64, at 169. ("[O]ften law is not enough. The law has limits, and law alone cannot solve all of the problems of privacy and free speech. If we care about these values as a society, we must protect them beyond the legal system, as part of our culture and social norms."). Already, intermediaries like Facebook and Twitter play an important role in mediating our ability to speak freely, and such entities "need to recognize that they have a special responsibility in the twenty-first century to safeguard expression." *Id.* at 174. "In determining the content of their ethical rules, information professionals in the digital age should also look to older kinds of professionals[,] most importantly to librarians." *Id.* at 176.

137. Ira Rubinstein, *Big Data: The End of Privacy or a New Beginning?*, 3 INT'L DATA PRIVACY L. 74, 82 (2013), http://idpl.oxfordjournals.org/content/early/2013/01/24/idpl.ips036.full.

some evidence that libraries are already willing to move away from traditional book-loaning functions. For example, many libraries have already moved in the direction of offering educational services, a step outside their traditional function just as becoming PDS would be. The New York Public Library (NYPL) offers after-school programs for kids as well as English conversation groups, literacy classes, and technology trainings geared toward adult learners.[138] The Los Angeles Public Library (LAPL) recently expanded on a similar portfolio of services with a 2014 launch of "a library-based program [aimed at adults] that will confer accredited high-school diplomas on city residents" to combat the city's high dropout rate of sixty to seventy percent in some neighborhoods.[139] Because of these recent forays into new services, libraries may be swayed to take further action in the tech sector. Richards argues that PDS "serve a compelling need—they help individuals organize and manage their daily lives and give them tools for realizing the inherent value of their own data. Thus, they are both convenient and a source of insight (via a new class of apps for monitoring and analysing one's own behaviour)."[140] This fits well with libraries' recent increased activity in the education sector and may consequently be viewed as a natural next step. Public libraries are the original data aggregators, and continuing to be so in the future as PDS thus fits with their mission and values.[141]

The fact that consumers trust libraries to protect their interests,[142] in part because they lack a profit motive, also would make libraries the ideal stewards of big data. This trust could lead to consumers providing better data and enriching the big data universe.

---

138. *Classes & Workshops*, N.Y PUB. LIBR., http://www.nypl.org/events/classes/calendar (last visited Feb. 9, 2017).

139. Sommer Mathis, *Los Angeles Public Library/ High* School, *in Five Creative Solutions*, THE ATLANTIC (June 25, 2014), http://www.theatlantic.com/magazine/archive/2014/07/creative-solutions/372285/.

140. RICHARDS, *supra* note 64, at 176–77.

141. Though libraries may appear to some to be hopeless relics of the past, librarians are in fact "thriving in a technology fueled world. . . . Libraries today house more than books, and librarians are more than good stewards of materials. Both have morphed and evolved to meet the changing needs of their patrons, by embracing technological advancements." Frankie Rendon, *How Innovation and Technology Are Shaping Libraries of Today*, HUFFINGTON POST (May 1, 2014), http://www.huffingtonpost.com/frankie-rendon/how-innovation-and-techno_b_5244601.html. Many open-minded librarians may embrace the PDS project as their next role in enhancing meaningful information access. After all, "[s]earch engines do provide a plethora of information, quickly and easily, but there is no guarantee of the quality of the information." *Id.*

142. *See, e.g.*, CENTER FOR AN URBAN FUTURE, *supra* note 76.

Rubinstein writes, "Data mining and ad targeting are based on guesswork, whereas in [PDS], potential customers would knowingly and intentionally reveal data that are likely more to be detailed, accurate, complete, and up-to-date than any inferred data."[143] Putting trusted libraries in the PDS role would facilitate this type of information transfer and benefit consumers as well as industry with a freer flow of more accurate data. It is worth acknowledging, however, that libraries' lack of a profit motive or potential inability to support PDS technology could also be impediments to the successful implementation of the PDS project. Libraries and the ALA might need to explore private partnerships or look to other privacy-protective technologies like the search engine DuckDuckGo, which protects user privacy and avoids personalized results, in order to run the PDS project on at least a cost-recovery basis.[144] But libraries' trustworthiness and lack of a profit motive would likely be crucial to the project getting off the ground initially.

Furthermore, with personnel on staff who have studied information system design, libraries are well set up to selectively grant access based on different kinds of user needs.[145] The library PDS model could grant three kinds of access: an individual's access to her own personal data, a limited public dashboard only displaying data subject to stringent re-identifiability safeguards, and a robust business or government user platform with access granted subject to a review process similar to Institutional Review Boards (IRBs) at research universities, which vet and approve research projects based on ethical standards. The IRB could consider the following criteria, derived from Cate's suggested updates to the FIPs, in deciding whether or not to allow the use of information by businesses for specific projects: "the degree and likelihood of benefits resulting from such uses," "the degree and likelihood of harm posed by such uses," and "the measures in place to guard against such harm."[146] This process could allow the approval of commercial projects that maximize benefit while minimizing harm and disallow projects that do not satisfy Cate's criteria. Meanwhile the individual, having already opted in by providing data, could be granted a gran-

143. Rubinstein, *supra* note 137, at 86.

144. DuckDuckGo is supported through advertisements. Jackie Chou, *DuckDuckGo Startup Profile*, CHOU PROJECTS (Jan. 29, 2015), http://chouprojects.com/duckduckgo-search-engine/.

145. RICHARDS, *supra* note 64, at 176–77. Whether or not all users of the PDS should be registered library patrons is another question.

146. Fred H. Cate et al., *Data Protection Principles for the 21st Century: Revising the 1980 OECD Guidelines* 17 (2014), http://www.oii.ox.ac.uk/publications/Data_Protection_Principles_for_the_21st_Century.pdf.

ular opt-out right, while the IRB-like process could serve as an additional check against uses of consumer data that are against the public interest.[147] The public dashboard, in contrast, would provide a baseline level of access to any user looking to conduct research and pursue innovation. These distinctions among different types of users would enable many benefits of big data to proceed in a more controlled, privacy-protective manner compared to the current free-for-all state[148] of regulation, and libraries would be the natural choice to design and implement this new type of information system.[149]

Ideally, libraries would serve as PDS after the FTC, which regulates privacy,[150] and industry stakeholders, who are happy with currently less effective regulation, get on the same page about the need to protect consumer privacy, rendering the current ineffective, FIPs-based approach to privacy regulation obsolete. There is some evidence that this, too, could be accomplished with relative ease. The FTC has indicated support for the kind of interests, such as security and individual access, served by PDS. For instance, a recent speech by the FTC Chair suggested that "[a]ny system based on trading of property rights [like that currently underpinning the big data marketplace] further requires service providers providing a safe trading infrastructure and services to individuals."[151] As stated above, libraries are well-positioned to provide that infrastructure. This policy statement may mean that the FTC would be interested

147. This may require vetting third party service providers, and libraries are equipped to do that too. *See, e.g.*, *Privacy Guidelines for Electronic Resources Vendors*, INT'L COALITION LIBR. CONSORTIA, http://icolc.net/statement/privacy-guidelines-electronic-resources-vendors (last visited Feb. 9, 2017) (asserting that "the ICOLC issues these guidelines with respect to the privacy interests of our member libraries' users in the interest of informing the companies with which we do business about what is acceptable in the products and services that we license").

148. *See supra* Part I.B.

149. An IRB-like process for big data would flip typical U.S. privacy regulation on its head. "Most information privacy law focuses on collection or disclosure and not use. Once data ha[ve] been legitimately obtained, few laws dictate what may be done with the information." Paul Ohm, *Changing the Rules: General Principles for Data Use and Analysis*, *in* PRIVACY, BIG DATA, AND THE PUBLIC GOOD: FRAMEWORKS FOR ENGAGEMENT 96 (Victoria Stodden et al. eds., 2014).

150. *A Brief Overview of the Federal Trade Commission's Investigative and Law Enforcement Authority*, FED. TRADE COMMISSION, https://www.ftc.gov/about-ftc/what-we-do/enforcement-authority (last visited Feb. 9, 2017).

151. Lokke Moerel, Professor of Glob. ICT Law, Tilburg Univ., Lectured Delivered During Acceptance of Professorship, *in* Big Data Protection: How to Make the Draft EU Regulation on Data Protection Future Proof 25 (Feb. 14, 2014), http://www.debrauw.com/wp-content/uploads/NEWS%20-%20PUBLICATIONS/Moerel_oratie.pdf.

in getting involved in or even leading the IRB-like process. Benefits could inure to industry, too. PDS "quite possibly would lower compliance costs for firms that rely on data in PDS as opposed to collecting and storing data in their own data stores."[152] More research is necessary on this point, but public libraries serving as PDS could provide a foundation for big data regulation that the FTC and industry find satisfactory.

If successfully implemented, libraries serving as PDS would help safeguard the privacy of non-patrons and patrons alike. Public libraries—along with pretty much every industry and sector of the global economy—now face additional challenges due to the fast-paced evolution of technology. Pressed to stay relevant in the digital age when many books, movies, and other sources of information and entertainment can be easily located online, libraries are facing pressures created by big data to depart from their historically strong commitment to user privacy, which could threaten the credibility of libraries to serve as an ethical model for the regulation of big data. A primary reason for this is that librarians are using more technology. Even if librarians and library policies remain as committed as ever to patron privacy, new contracts with vendors such as Amazon and BiblioCommons (an online version of a card catalog with various interactive features) may be overriding these efforts.[153] For example, when a user checks out a library e-book using an Amazon Kindle reader, she completes that action using an Amazon account.[154] Amazon then stores her reading history—often alongside a wealth of other consumption habits, financial information, and personal information—and offers her no way to opt out of that storage.[155] Similarly, as recently as fall 2016, the default setting of BiblioCommons, which is in use in over 200 public libraries in the United States, Canada, Australia, and New Zealand,[156] was to share user content, including ratings, lists, and comments, across all li-

152. Rubinstein, *supra* note 137, at 86.

153. *See, e.g.*, Glaser & Macrina, *supra* note 78 ("Adobe's Digital Editions e-book software collects and transmits information about readers in plain text. That insecure transmission allows the government, corporations, or potential hackers to intercept information about patron reading habits, including book title, author, publisher, subject, description, and every page read.").

154. *Id.*

155. *Id.*

156. *BiblioCommons Features Local Library Staff Recommendations and Reviews*, LIBR. TECH. GUIDES, https://librarytechnology.org/news/pr.pl?id=19506 s (last visited Feb. 9, 2017).

braries that utilize the software,[157] thereby violating, in Richards' view, an important aspect of their intellectual privacy.[158] Though BiblioCommons promises to forgo sharing "your information or activity with ad networks or other entities that are not directly involved in the services you choose to use"[159] and now offers users better control over which parts of their borrowing history are visible,[160] BiblioCommons can revert, at will, back to a less privacy-protective default. The faster processing and slick new sharing features of these products can lure librarians,[161] who may be unknowingly subjecting their patrons' data to uses to which they did not consent. Requiring patrons to use Amazon, BiblioCommons or similar products to access library materials means that patron data—including highly sensitive aspects of intellectual freedom, such as reading history—are being tracked much more than they were historically. But as PDS, libraries could address the threats technologies like Amazon and BiblioCommons pose for patron privacy—as well as the threats posed by similar technologies in other industries—by subjecting consumer data to the IRB-like process and granular user opt-out right discussed above. Under that process, no user would find her borrowing history—or any other information—suddenly made public without her informed consent.

V.
CONCLUSION

Big data poses a threat to privacy the likes of which Warren and Brandeis could hardly have imagined when they first spoke out

157. *Our Platform*, BIBLIOCOMMONS, https://web.archive.org/web/20150609212240/http://bibliocommons.com/how-we-work/our-platform (last visited Feb. 9, 2017).

158. *See supra* Part I.C.

159. *See, e.g.*, *BiblioCommons US Privacy Statement*, BROOKLYN PUB. LIBR., https://brooklyn.bibliocommons.com/info/privacy (last visited Feb. 9, 2017). Biblio Commons appears to have a privacy policy specific to each library with which it contracts.

160. *BiblioCore Features Sheet*, BIBLIOCOMMONS https://static1.squarespace.com/static/586d7efa2994cab071cbb4ae/t/5878f25b6b8f5bb797912ab3/1484321373369/BiblioCore+Feature+Sheet+2017.pdf (last visited Feb. 9, 2017) ("Privacy controls allow patrons to choose to share, or keep their content private.").

161. *See, e.g.*, Bonnie Tijerina, *Developing a Consensus Framework for Patron Privacy*, MEDIUM (Apr. 7, 2015), https://medium.com/@bonlth/developing-a-consensus-framework-for-patron-privacy-7668094ad4f8 ("When there's so much to gain from translating raw patron data into meaningful and useful information to learn about our communities or improve services and products, how do [you] know where to draw the line on use or non-use of our patron's information?").

against the growing use of cameras in 1890. In the information age, our privacy laws and commitment to the importance of privacy in a free society have not kept pace with technology's increased capacity to inflict harm. Despite facing their own challenges given the FIPs structure and new products with problematic sharing features, public libraries remain well-suited to reverse that trend. Libraries' enablement of anonymous browsing and refusal to condone secondary use reflect a robust understanding of the importance of privacy. Librarians' wealth of knowledge regarding system design and their commitment to providing access to information would make them ideal stewards of big data. By becoming PDS, public libraries could provide the ethical model big data lacks and reorient us to a better privacy future.