

Stanford **Technology** Law Review

NOTE

The Mind Gangsters: Why We Should, and How We Can, Limit Surveillance of Digital Reading Habits

THOMAS NOSEWICZ^{*}

CITE AS: 2009 STAN. TECH. L. REV. N1

<http://stlr.stanford.edu/pdf/nosewicz-mind-gangsters.pdf>

¶1 It is not alarmist to say that the Internet is the first truly panoptic system of the mind.¹ Dumbfoundingly dense databanks can—and do—gorge themselves on one’s every move across a webpage. Web tools monitor every specific article a visitor reads, how she was referred to that article, and how long she spent reading it. These tools allow website owners to compile a comprehensive set of statistics about visitors to their websites, including how often they visit, their domains and countries of origin, what pages they view the most, and the operating system and web browser they use to access the website.² This surveillance is omnipresent, all-knowing, and perfectly concealed.

¶2 Some sites go even further and require completion of a registration process that involves relinquishing a zip code, email address, and full name. Compared to the non-wired world, the increase in monitoring capability in these circumstances is exponential.³ The extent of this information gathering is not only annoying, but also increasingly relevant as more and more government data mining programs, some of which rely on information gathered by private companies, are revealed to the public.⁴

¶3 This Note is about how this information should be treated. Part I describes how these systems function to track browsers’ interactions with websites. Part II examines problems with these

^{*} © 2009, Thomas Nosewicz, Stanford-SPILF Public Interest Fellow at the Orleans Public Defenders. Deep thanks to everyone at STLR, Lauren Gelman at Stanford’s Center for Internet and Society, Raha Naddaf, and Brian Wilson for the title.

¹ See, e.g., James Boyle, *Foucault in Cyberspace: Surveillance, Sovereignty and Hardwired Censors*, 66 U. CIN. L. REV. 177 (1997); Shawn C. Helms, *Translating Privacy Values With Technology*, 7 B.U. J. SCI. & TECH. L. 288, 291-94 (2001) (describing the mechanics of the “cyber panopticon”).

² See, e.g., Google Analytics, <http://www.google.com/analytics/> (last visited Sept. 10, 2008); Wiki/Web Analytics – Wikipedia, http://en.wikipedia.org/wiki/Web_analytics (last visited Sept. 10, 2008); Web Log analysis software – Wikipedia, http://en.wikipedia.org/wiki/Web_log_analysis_software (last visited Sept. 10, 2008).

³ See LAWRENCE LESSIG, *CODE AND OTHER LAWS OF CYBERSPACE* 151 (1999).

⁴ See, e.g., U.S. GEN. ACCOUNTING OFFICE, *DATA MINING: FEDERAL EFFORTS COVER A WIDE RANGE OF USES* (2004), available at www.gao.gov/new.items/d04548.pdf. See also Bobby White, *Watching What You See on the Web – New Gear Lets ISPs Track Users and Sell Targeted Ads; More Players, Privacy Fears*, WALL ST. J., Dec. 6, 2007, at B1.

systems, including the vulnerabilities they create for user privacy. Part III investigates techniques available to bypass compulsory registration. Part IV suggests how companies should address this issue, and Part V proposes a federal law, modeled on the Video Privacy Protection Act, to regulate the gathering and sharing of user information.

I. TV ON THE RADIO: USER REGISTRATION TODAY

¶4 The most popular newspaper websites require users to input demographic data to gain access to content.⁵ Companies that provide user registration services, such as Macrovision⁶ (whose clients include *The New York Times*, Knight-Ridder, and *Computerworld*) and Tacoda⁷ (whose clients include *USA Today*, *Los Angeles Times*, and SFGate.com), offer a wide range of incoherently labeled “audience management systems”⁸ that allow “for complete control over exactly who accesses what digital product and in which manner the access is granted.”⁹

¶5 Registration systems differ in the amount and type of information they request. Some websites, such as those that fall under the Advance.net umbrella,¹⁰ do not require an email address, but ask for a zip code, year of birth, and gender.¹¹ Other websites, including the *Los Angeles Times* site, require creation of a user name and disclosure of one’s gender, year of birth and income level.¹² The most intrusive websites, such as *The New York Times* site, request multiple fields of demographic information focusing on income and occupation.¹³

¶6 The bait-and-switch strategy deployed by many sites’ registration process may surprise many users. In these instances, several “teaser” paragraphs of an article are displayed on a main page, while the rest of the article remains “below the fold” and must be clicked for further access. This click brings a browser to a registration prompt, and sunk costs from having begun an article may grease the wheels for the registration process to begin. Many sites also randomly interpose full-screen ads between a reader and subsequent pages of articles.

¶7 Some newspaper sites have embraced a nice compromise between registration and anonymous browsing.¹⁴ For example, SignOnSanDiego.com, the *San Diego Union-Tribune*’s website, used to present a user survey at the top of articles, but did not require completion of the form to read the article. This allowed those who came to the site intending to read a single article to do so without registering. The “nag screen” became annoying only when someone repeatedly viewed the site. This encouraged registration for people who actually used the site’s resources over and over again (and thus presumably put more of a burden on the site), while not blocking off any content behind a registration wall. The problem with such a practice is that the insidiousness of registration lies in the volume and quality of data collected—registering for a site that you visit once does not do much to

⁵ CHRIS JAY HOOFNAGLE, ELECTRONIC PRIVACY INFORMATION CENTER, PRIVACY SELF REGULATION: A DECADE OF DISAPPOINTMENT (2005), <http://www.epic.org/reports/decadedisappoint.html> (EPIC study finding five of the top twenty-five newspaper websites, as determined by circulation, require user registration to view content, and many more require disclosure of other demographic information).

⁶ Macrovision Online Publishing, <http://www.macrovision.com/products/1150.htm> (last visited Sept. 10, 2008).

⁷ Tacoda, <http://www.tacoda.com> (last visited Sept. 10, 2008).

⁸ See Tacoda Audience Management Services Overview, http://www.tacoda.com/ams_overview.htm (last visited May 30, 2006) (“Audience Management Services is an end-to-end marketing application used for analyzing customer interactions, segmenting and monetizing audience members, while providing a directly actionable, closed-loop solution.”).

⁹ eMeta Right Access, http://www.emeta.com/products/prod_rightaccess.html (last visited Aug. 4, 2007).

¹⁰ Advance.net Affiliated Newspapers, http://advance.net/index.ssf?/advance_internet/newspapers.html (last visited Aug. 4, 2007).

¹¹ See, e.g., cleveland.com: Everything Cleveland, Help Us Serve You Better, <http://www.cleveland.com/enter/index.ssf/> (last visited Aug. 4, 2007).

¹² *Los Angeles Times* Registration, <https://www.latimes.com/services/site/registration/show-createprofile.register> (last visited Aug. 4, 2007).

¹³ *New York Times* – Free Registration, <http://www.nytimes.com/gst/regi.html> (last visited Aug. 4, 2007).

¹⁴ Steve Outing, *A More Friendly Registration Demand*, POYNTERONLINE, Jan. 7, 2005, <http://www.poynter.org/column.asp?id=31&aid=76616>.

disclose your reading inclinations. Registration that is only triggered by repeated visits presents the same problems as mandatory registration because it enables a data-rich profile to be developed about an individual reader.

¶8 Media companies value registration because it encourages reader loyalty to specific websites and provides demographic data that is attractive to advertisers.¹⁵ Industry figures explain the benefits as “tightly focused content, targeted advertising, and e-mail newsletters that . . . build relationships online between advertisers and users.”¹⁶ Barbara Rice, group director of research at *New York Times Digital*, explains that visitors to the website can be “followed” by ads: “If a person frequently visits the Travel section but that section is sold out [of ads], an advertiser isn’t shut out; they can reach him as he travels throughout the rest of the site.”¹⁷ Collection of this data also allows advertising to be tightly focused: “An advertiser can come to us and say, ‘I want to reach people in these ZIP codes’ . . . For instance, a pharmaceutical company can target women over age 35. eTrade can roll out an online advertising campaign limited to the New York area in conjunction with print and broadcast.”¹⁸ The newspaper frames its use of this registration information as “determin[ing] hidden patterns of uses to [their] website.”¹⁹

¶9 This monitoring may appear innocuous, but a comparison between online registration and analogous experiences—such as subscribing to a newspaper, browsing at a bookstore, watching cable television, and creating a record of borrowings at a library—reveals it is not. Surveillance as intrusive and unflagging as the surreptitious monitoring of Internet reading habits simply does not exist in real space.²⁰

¶10 To begin with, print newspapers maintain information about their subscribers, but often only contact and payment details. Additionally, a print newspaper reader has several methods to maintain anonymity: she can subscribe under a fake name, list her address as a P.O. box, or simply buy the paper at a newsstand. However, with content on a website, a reader does not have to make any commitment, such as the one needed to receive a newspaper for a set amount of time, for the content-provider to know who and where she is and what specific information interests her. The newspaper that was delivered to a reader’s door as a single chunk of information is now atomized and delivered in specific pieces to her web browser, and her usage of each of those pieces is tracked and stored, often without the user’s knowledge.

¶11 Bookstore browsing is also analogous to browsing the web. In a bookstore, an individual can move from magazine to magazine and read a single article from a periodical at her own pace. But unless another human follows her around the bookstore and asks for her name, notes the name of the publication, the article read, and how long she spends reading it, the bookstore browser remains more anonymous than she does on the web.

¶12 Surfing the web is also like surfing cable television and dipping into part of a show between commercial breaks or a news story on a 24-hour cable news channel. Both television shows and websites target ads to regional characteristics. But viewers watch television anonymously (at least before the rise of digital subscription services), and ratings services are voluntary and cover only a small portion of the population—not every single viewer as on the web.

¹⁵ Carl Sullivan, *Newspaper Sites Move to Registration Model*, EDITOR & PUBLISHER, Jan. 23, 2003, http://www.editorandpublisher.com/eandp/news/article_display.jsp?vnu_content_id=1801780.

¹⁶ J.D. Lasica, *Belo: Active and Shifting Audiences*, ONLINE JOURNALISM REV., June 27, 2002, <http://www.ojr.org/ojr/lasica/1025227639.php>.

¹⁷ J.D. Lasica, *The New York Times: Targeting Readers the Old-Fashioned Way*, ONLINE JOURNALISM REV., June 27, 2002, <http://www.ojr.org/ojr/lasica/1025226881.php>.

¹⁸ *Id.*

¹⁹ Keach Hagey, *Having Won a Pulitzer for Exposing Data Mining, Times Now Eager to Do Its Own Data Mining*, THE VILLAGE VOICE, Apr. 24, 2007, <http://www.villagevoice.com/nyclife/0718%2Chagey%2C76522%2C15.html>.

²⁰ See, e.g., Julie E. Cohen, *A Right to Read Anonymously: A Closer Look at “Copyright Management” in Cyberspace*, 28 CONN. L. REV. 981, 1003 (1996) (explaining that “[u]ntil recently, however, the technological means to monitor individuals’ reading habits did not exist”).

- ¶13 The most accurate records kept at a library also do not compare to the records stored by web servers. Though a library has a record of a patron's name, address, and what books she has checked out, the library's records do not contain the fine-grained analysis that web analytics offer, such as the path a reader takes through a website and how long she spent reading an article. Libraries are also unlikely to sell personal information.
- ¶14 Using the Internet is not like any one of these activities. It is like all four: a vast data-stream (newspaper), read at one's leisure (bookstore), pocked with ads (TV), with records of a user's reference stored perpetually (library).
- ¶15 These comparisons highlight the core difference between reading an article online and reading one printed on paper. The paper article exists as a discrete unit, disconnected from the source that provided it, while the website remains linked to its host. To a great extent, the impact of printed newspaper articles is obscure—reader response can be measured in letters to the editor, increased wait times at a restaurant following a positive review, or political fallout, but it is impossible for editors to capture exactly how each individual article is read and even shared.
- ¶16 The solitary process of reading has indeed become communal. A reader's preferences are encoded into a website's greeting scroll, such as the list of "most popular stories" on Yahoo! News.²¹ As an individual reads a newspaper article online, she may also click on related content or advertising, or she may email it to friends, or even comment on its significance. This discursive process allows any monitor to capture an extremely nuanced set of data which can then be combined into marketing profiles and filtered into complex data mining systems.²²

II. THE PROBLEMS WITH USER REGISTRATION

- ¶17 The number of websites requiring registration to access content is increasing without any permanent drop in traffic to these sites.²³ Change is aimed at convenience (such as creating a one-step registration system that allows access to multiple websites) and user acceptance of "complete control,"²⁴ an always ominous phrase, but especially troubling when used to describe mechanisms regulating people's reading. Jay Small, executive director of content and product development for the Scripps Interactive Newspaper Group at E.W. Scripps Co., even shows derision for attempts at scaling back user registration: "Culture-of-the-Internet Utopians will rail against access-control registration until the Web is replaced by that skull-cap, retina-authenticated, holographic interface we're all waiting for. But I know registration works, from experience, so all this philosophizing from outsiders-looking-in will never persuade me otherwise."²⁵
- ¶18 Jay Small and his ilk have at least two defenses for requiring registration. The first is, essentially, that everyone is doing it—that the modern commercial world is already one of pervasive surveillance. A few examples support this point: retail stores commission complicated crowd movement studies to learn how people use their physical environments and this is not seen as unfairly impinging upon privacy.²⁶ Music fans willingly subscribe to services that track their playlists in order to recommend other music they might enjoy.²⁷ Registration can also be useful to people who want to create an

²¹ Yahoo! News, <http://news.yahoo.com/> (last visited Sept. 10, 2008).

²² Cohen, *supra* note 20, at 986 (detailing the techniques of profiling and data mining).

²³ Carl Sullivan, *Newspaper Sites Move to Registration Model*, EDITOR & PUBLISHER, Jan. 23, 2003, http://www.editorandpublisher.com/eandp/news/article_display.jsp?vnu_content_id=1801780 (reporting that "newspaper sites that built up registration walls last year have found that readers didn't leave in droves. In fact, several of these papers now have more online visitors than they had before requiring registration"). *But see* Is User Registration Passe?, THE LOCAL ONLINER, Dec. 8, 2005, available at <http://web.archive.org/web/20060207095758/http://localonliner.com/?p=45>.

²⁴ eMeta Right Access, http://www.emeta.com/products/prod_rightaccess.html (last visited Aug. 4, 2007).

²⁵ Posting of Jay Small, http://www.poynter.org/article_feedback/article_feedback_list.asp?id=67655 (Jun. 28, 2004, 1:24:39 PM).

²⁶ See Malcolm Gladwell, *The Science of Shopping*, THE NEW YORKER, Nov. 4, 1996, available at <http://www.gladwell.com/pdf/shopping.pdf>.

²⁷ See Jeremy Atkinson, *Free Music Recommendation Services*, EXTREME TECH, May 25, 2005, <http://www.extremetech.com/article2/0,1697,1967383,00.asp>.

identity for an online forum or blog.²⁸ The registration process to access content on a newspaper site does not require much more than these services and is usually a minor, one-time inconvenience. Registration screens often do not even appear until a user tries to access older articles on a site, or until she loads content that does not belong to a wire service.

¶19 The second defense is that, though registration requires relinquishing private information, users know exactly what they are disclosing, and choose to do it anyway. Companies requiring registration view it as an equitable exchange—they are, after all, providing content, bandwidth, and web design. Barbara Rice of *New York Times Digital* explains: “Yes, it’s an extra step the user has to take. But it’s a *quid pro quo*. You’re receiving a premium product for free.”²⁹

¶20 But it is not actually free. The price of access to these websites is one’s reading privacy. A content provider’s sale of personal data to third-parties belies this “it’s free” rhetoric,³⁰ as does the wide range of content available on the Internet that does not require registration. It is typically only print media migrating to the Internet that requires registration, a defensive move that may reflect a misunderstanding and fear of the Internet’s ability to provide open and immediate access to information.³¹ Blogs, wikis, and news sites associated with broadcast media usually do not require registration to read articles, but these content providers still make money based on advertisements. This further undermines the idea that registration is a fair exchange for access to newspapers’ websites.

¶21 If website owners are merely concerned about increasing the effectiveness of their advertisements, registration does not add as much value to the process of provisioning advertising as they think. Rice offers an example of how registration data is used: “The editorial side likes to know who’s reading what articles and packages. Are stories being read by a New York City audience or international audience?”³² There is a far less intrusive way to gather this data. IP addresses, which are logged as a matter of course by most websites, provide geographical data. Anyone wanting to get a sense of the depth of information contained in an IP address can load Geobytes’ IP Address Map³³ which provides an uncanny profile including a map of a browser’s current location, zip code, nationality, area code, longitude, and latitude. If newspapers are truly concerned with seeing whether locals access their website, this information is available without the hassle—and costs to the company—of registration. IP addresses do not present the same privacy problems as user registration because linking an IP address to an individual user’s name, while possible, requires more effort from a content provider, as well as coordination with an Internet service provider.

¶22 The content providers’ justifications for registration are unsatisfactory. As noted above, much more information is collected through online registration than from other entities gathering information about our consumption habits. But even assuming that users know how much of their privacy is traded for access—and there is no reason to assume that people read the fine print in privacy policies—this information travels much further than a reader would suppose. This highlights the real problem here: widespread trafficking in private information.

A. Problems of Sharing

¶23 Privacy policies of major media websites do little to protect user’s privacy. Instead, the policies often act as waivers and grant the websites license to distribute user information. For example, the

²⁸ *But see* Shiichan Anonymous BBS, <http://wakaba.c3.cx/shii/shiichan> (last visited Aug. 4, 2007) (explaining the benefits of anonymity for online forums).

²⁹ Lasica, *supra* note 17.

³⁰ And users who want to estimate what their data is worth can use Swipe’s Data Calculator. *See* Swipe – Data Calculator, http://turbulence.org/Works/swipe/swipe_data_cal.html (last visited Sept. 20, 2008).

³¹ *See also* John C. Dvorak, *Registration? For What?*, PC MAG., Oct. 5, 2004, available at <http://www.pcmag.com/article2/0,1759,1646213,00.asp> (moving content to the Internet “seems to be just something that newspaper people feel they have to do because everyone else is doing it”).

³² Lasica, *supra* note 17.

³³ Geobytes IP Address Locator, <http://www.geobytes.com/IpLocator.htm?GetLocation> (last visited Aug. 4, 2007).

Thomas Nosewicz: The Mind Gangsters: Why We Should, and How We Can, Limit Surveillance of Digital Reading Habits

privacy policy for advance.net allows the company to “share e-mail addresses and sell or share all other information with our affiliates and with carefully selected companies who we think can offer you services and products of interest to you.” A user is allowed to opt out of this service by emailing the company.³⁴

¶24 Such a policy offers no real protection to an individual user. The choice of who receives the user information (even if it is merely “shared” and not sold) is left to advance.net. Opting out is allowed, but this presumably will happen for most users *after* the information has already been given to another party, unless she is one of those rare types who reads online privacy policies before signing up for the service. No mechanism for purging one’s information is specified in the privacy policy. Other privacy policies do not even allow for this minimal level of opting out.³⁵

¶25 It should be noted that not all registration or data collection necessarily leads to spam. *The New York Times* has an easily understood, if lengthy, privacy policy that indicates users’ e-mail addresses will only be shared with third parties if their permission is obtained, implying that users will only be sent spam from third parties if they have affirmatively asked to receive it.³⁶ *The Times* also admits that it uses registration data to target advertising, but that it will not disclose user data to third parties except in aggregated form, or unless the affected user has specifically authorized it.³⁷ Its website also logs IP addresses, but only for “systems administration and troubleshooting purposes.”³⁸ The IP addresses logged are also used in an “aggregate fashion” to track access to the website.³⁹

¶26 As this information percolates into the personal profiles maintained by consumer profiling companies, privacy concerns will become more acute.⁴⁰ Google has already changed the dynamic of job interviews and other personal relationships⁴¹—imagine if Google searches began to include someone’s reading habits.

¶27 However, if this data is sold to, or shared with, third-party spammers, the effect is usually merely annoying. And, as spam filters improve or spam becomes commonplace, most people will not even notice the extra mail. The most alarming aspect of the data collected by websites via user registration is its use by law enforcement.⁴² Privacy policies generally contain exceptions allowing companies to disclose information to comply with the law.⁴³ Though an in-depth exploration of the criminal procedure implications of the compilation of these records about the reading habits of U.S. citizens is beyond the scope of this Note, a brief skimming of the issue is appropriate.

¶28 In general, introducing someone’s reading material is allowed by the rules of evidence in a criminal prosecution.⁴⁴ There is also not much law restraining the government’s ability to obtain information about people’s online reading habits in the first place. The Fourth Amendment currently

³⁴ advance.net Privacy Policy, <http://www.advance.net/privacypolicy/> (last visited Aug. 4, 2007).

³⁵ Hoofnagle, *supra* note 5, at 9 (explaining that “both the LA Times and Chicago Tribune websites do not allow users to opt out of information sharing, advertising and communications from the newspapers and their ‘affiliates’ [although you can opt out of sharing of your information with their advertisers and other third parties]”).

³⁶ *The New York Times*, Member Center – Site Help – The New York Times Privacy Policy, <http://www.nytimes.com/ref/membercenter/help/privacy.html> (last visited Sept. 28, 2008).

³⁷ *Id.*

³⁸ *Id.*

³⁹ *Id.*

⁴⁰ See Robert O’Harrow Jr., *LexisNexis to Buy Seisint for \$775 Million*, WASH. POST, Jul. 15, 2004, at E01 (“Information giant LexisNexis Group said yesterday it will pay \$775 million in cash for Seisint Inc., a privately held data service that created a controversial tool called the Matrix, which gave state and federal authorities new power to analyze records about Americans after the Sept. 11, 2001, terror attacks.”); LexisNexis acquires Seisint, Inc., <http://www.seisint.com/> (last visited Sept. 28, 2008) (noting that “Securint, the most powerful background screening product on the market today, can be accessed through the pull-down menu at the top of the [LexisNexis] page”).

⁴¹ See Neil Swidley, *A Nation of Voyeurs*, BOSTON GLOBE MAG., Feb. 2, 2003, at 10.

⁴² Hagey, *supra* note 19 (describing similar concern expressed by the Cato Institute).

⁴³ *New York Times* Privacy Policy, *supra* note 37 (“We may occasionally release personal information as required by law, for example, to comply with a court order or subpoena.”).

⁴⁴ See, e.g., *United States v. Curtin*, 489 F.3d 935 (9th Cir. 2007).

Thomas Nosewicz: The Mind Gangsters: Why We Should, and How We Can, Limit Surveillance of Digital Reading Habits

plays no role in protecting this information: a media company's careful storage and frequent access to visitor logs would render them business records and thus easily subpoenaed by the government on a showing of relevance. Unless a visitor's web transactions were being recorded as they happened—analogue to a wiretap of voice communications—the government's access to this information is allowed by *United States v. Miller*, which held that information voluntarily conveyed to another party and subsequently turned over by that party to the government is free of Fourth Amendment protection.⁴⁵

¶29 Extraconstitutional statutory protections for electronic communications do not provide any protection either.⁴⁶ Indeed, under the PATRIOT Act, federal law enforcement has increased subpoena-like powers to request “business records” from companies in the context of intelligence gathering overseas.⁴⁷

¶30 Unsurprisingly, government agents have access to this information even if they do not subpoena it.⁴⁸ The U.S. government has contracted with the largest compilers of personal data about American citizens, such as ChoicePoint, to access their voluminous databases⁴⁹ in pursuit of law enforcement targets.⁵⁰ It is conceivable that the government might seek similar arrangements with newspapers, or their data collecting partners, to determine who has been reading “suspicious” or politically unpopular material.

¶31 Systems like Seisint's Accurint use sophisticated computer systems that have labeled “almost every American adult with a unique identifier [Accurint,] drawing on billions of records, can deliver dossiers online in an instant, including addresses, jobs, assets, voter registration and associates.”⁵¹ This information “enable[s] investigators to rapidly pull together lists of suspects, based on characteristics such as age, race and an individual's associates.”⁵²

¶32 Law enforcement's access to this data is troubling, and not just for fears of “Big Brother.” Law professor Kathleen Sullivan has written about the dangers of government access to large amounts of personal information about its citizens, even if it does not know what to do with it. Such access allows “unauthorized snooping, leaking of information, blackmailing by employees, bureaucratic error, and hacking and identity theft by enterprising high school students or criminals.”⁵³

¶33 Even if one thinks those risks are acceptable, use of the information in high-value intelligence gathering is not clearly valuable. Security expert Bruce Schneier notes that government resources

⁴⁵ *United States v. Miller*, 425 U.S. 435 (1976).

⁴⁶ 18 U.S.C. § 2703(c)(2)(C) (2006) (“records of session times and durations . . . [shall be disclosed to a governmental entity] . . . when the governmental entity uses an administrative subpoena authorized by a Federal or State statute or a Federal or State grand jury or trial subpoena or any [other lawful] means available”).

⁴⁷ 50 U.S.C. § 1861 (2006). *But see Doe v. Gonzales*, 500 F.Supp.2d 379 (S.D.N.Y. 2007) (declaring unconstitutional 18 U.S.C. § 2709, another PATRIOT Act provision with a non-disclosure element similar to 50 U.S.C. § 1861).

⁴⁸ Glenn R. Simpson, *FBI's Reliance on the Private Sector Has Raised Some Privacy Concerns*, WALL ST. J., Apr. 13, 2001, available at <http://www.atgpress.com/privacy/pri004.htm>.

⁴⁹ EPIC ChoicePoint Page, <http://www.epic.org/privacy/choicepoint/> (last visited Aug. 4, 2007) (noting that ChoicePoint operates a number of websites devoted to law enforcement access to personal information); Robert O'Harrow Jr., *In Age of Security, Firm Mines Wealth of Personal Data*, WASH. POST, Jan. 20, 2005, at A01..

⁵⁰ See Richard L. Fricker, *The INSLAW Octopus*, WIRED, Mar./Apr. 1993, available at <http://www.wired.com/wired/archive/1.01/inlaw.html> (historical background on the government's attempts at linking myriad databases).

⁵¹ See O'Harrow, *supra* note 40.

⁵² *Id.*

⁵³ Kathleen Sullivan, *Under a Watchful Eye: Incursions on Personal Privacy*, in *THE WAR ON OUR FREEDOMS: CIVIL LIBERTIES IN AN AGE OF TERRORISM* 128, 132 (Richard C. Leone & Greg Anrig, Jr. eds., 2003); see, e.g., *Welfare Records Leaked to Insurers*, DIGITAL RTS. IR., Jul. 16, 2007, <http://www.digitalrights.ie/2007/07/16/welfare-records-leaked-to-insurers/>; Roy Mark, *Federal Agent Indicted for Cyber-Stalking*, E-WEEK.COM, Sept. 21, 2007, <http://www.eweek.com/c/a/Database/Federal-Agent-Indicted-for-CyberStalking/>; *Civil Servant Mole Leaked Intelligence to Criminal*, INDEPENDENT.IE, Oct. 15, 2007, <http://www.independent.ie/national-news/civil-servant-mole-leaked-intelligence-to-criminal-1166835.html>.

Thomas Nosewicz: The Mind Gangsters: Why We Should, and How We Can, Limit Surveillance of Digital Reading Habits

spent on massive “data mining” techniques are misspent because terrorist attacks are so rare that meaningful profiles to sift data through cannot be developed.⁵⁴

B. Problems Beyond Sharing

¶34 User registration also creates problems beyond those presented by data-sharing. First, registration threatens a constitutionally rooted right of anonymous reading.⁵⁵ Law professor Julie Cohen has traced this right⁵⁶ and finds its clearest articulation in *Stanley v. Georgia*, which held that the government could not prohibit the mere possession of obscene materials in the home. Justice Marshall wrote that one has a “right to be free from state inquiry into the contents of [one’s] library”⁵⁷ and that the United States’ “whole constitutional heritage rebels at the thought of giving government the power to control men’s minds.”⁵⁸

¶35 As a matter of law, this holding has little effect on the registration issue since most websites requiring registration are private entities, and thus not bound by the government’s constitutional obligations. But pervasive, private registration still chills anonymous reading. Just because a registration scheme is run by a private actor does not necessarily make people more comfortable with it. In fact, private sector registration schemes may raise more concerns than government-run registration because of profit-driven motivations and the lack of regulation and oversight.

¶36 Registration also limits the scope of reading material available online. Search engines cannot easily archive information that is behind registration walls.⁵⁹ As more people turn to the Internet for their primary source of news and research, and more articles are locked behind registration walls, the number of people registering to access this content will increase. This creates a losing tradeoff as privacy is sacrificed to access information that was formerly available anonymously. Website registration also funnels people into using the same sites to avoid the inconvenience of multiple registrations, which stifles the attractive polyvalence of the Internet.⁶⁰

¶37 The large user base at sites such as *The New York Times*⁶¹ suggests that registration is seen by users as a mere speed-bump on the path to accessing content. This perception is reinforced by persistent cookies and the ability of browsers to store passwords, which makes logging in at each website a one-time occurrence. The convenience of persistent login can promote website loyalty, but registration requirements hinder the one-off linking practices⁶² popular within the blogosphere.⁶³ For

⁵⁴ Bruce Schneier, *Why Data Mining Won’t Stop Terror*, WIRED, Mar. 9 2006, available at http://wired.com/news/columns/0,70357-0.html?tw=wn_index_3.

⁵⁵ See *Kleindienst v. Mandel*, 408 U.S. 753, 762-63 (1972) (collecting cases identifying a First Amendment right to receive information).

⁵⁶ Julie E. Cohen, *A Right to Read Anonymously: A Closer Look at “Copyright Management” in Cyberspace*, 28 CONN. L. REV. 981, 1003-19 (1996).

⁵⁷ *Stanley v. Georgia*, 394 U.S. 557, 565 (1969).

⁵⁸ *Id.* at 565.

⁵⁹ Google’s guidelines for webmasters offer this advice: “Allow search bots to crawl your sites without session IDs or arguments that track their path through the site. These techniques are useful for tracking individual user behavior, but the access pattern of bots is entirely different. Using these techniques may result in incomplete indexing of your site.” Google, Webmaster Guidelines – Webmaster Help Center, <http://www.google.com/support/webmasters/bin/answer.py?hl=en&answer=35769> (last visited Oct. 3, 2008).

⁶⁰ See YOCHAI BENKLER, *THE WEALTH OF NETWORKS 3* (Yale University Press 2006) (theorizing that what “characterizes the networked information economy is . . . decentralized individual action—specifically, new and important cooperative and coordinate action carried out through radically distributed, nonmarket mechanisms”).

⁶¹ Lasica, *supra* note 17 (reporting that *The New York Times* has required registration since it began offering content, and had 10 million registered users as of 2002); see also Sullivan, *supra* note 15 (noting that *The New York Times* has recently “beefed up its registration questionnaire to ask users more about their occupations and about their print subscriptions”).

⁶² Adam L. Penenberg, “Searching for The New York Times,” WIRED, July 14, 2004, <http://www.wired.com/culture/lifestyle/news/2004/07/64110/> (“[O]nly half of the *Times* Web users enter through the homepage. The rest come via links provided by mass e-mails, blogs and other publications.”).

⁶³ J.D. Lasica, *Privacy, Personal Data and Taking Users for Granted*, ONLINE JOURNALISM REV., June 27, 2002, <http://www.ojr.org/ojr/lasica/1025226464.php> (noting the consequences of registration for blogs).

example, Slashdot, a prominent links-based discussion site, will not link to websites that require registration, with the major exception of *The New York Times*.⁶⁴

¶38 We should worry about the long-term effects of habitual registration. As more and more mainstream websites require registration, registering becomes more and more mainstream.⁶⁵ This slow Mithridatization defangs registrations of all sorts and makes them seem a normal part of modern life. It also causes security vulnerabilities: as people are forced to create more usernames and passwords, “password fatigue” may lead to the reuse of passwords. A single password can become a master key for that user’s accounts across the board, so if the password is compromised at one website, it may lead to vulnerability at others, such as banks or retailers.⁶⁶

¶39 The combination of the problems described above—the transfer of intimate information by content providers to private and governmental databases, combined with registration’s chilling effect on anonymous reading and disfiguration of the Internet’s intellectual landscape—shows that registration regimes implicate important issues regarding how our society treats access to knowledge.⁶⁷

III. USER-DRIVEN PRIVACY PRESERVATION

¶40 So what is a persnickety reader to do? The mere existence of privacy policies offers no guarantee of actual privacy,⁶⁸ so the burden of protecting privacy typically falls upon the user. There are several methods by which users may avoid disclosing private information even in the face of unfavorable privacy policies.

A. Alternate Sources

¶41 The first technique is to circumvent the registration process entirely by accessing the same information via an alternate source. For example, blogs and other website post entire articles or large chunks of them.⁶⁹

B. Bring Tha Noise

¶42 The easiest and most obvious way to preserve privacy while still accessing content requiring registration is to input false information into user registration systems.⁷⁰ For instance, a browser can input a birth date of January 1, 1900, and a zip code consisting of 66666 and still access a website because the data is not verified. Anonymous email services also enable the convenient creation of temporary email addresses for use during the registration process.⁷¹ These techniques allow users to

⁶⁴ J.D. Lasica, *Getting To Know You*, ONLINE JOURNALISM REV., June 27, 2002, <http://www.ojr.org/ojr/lasica/1025227718.php> (quoting a Slashdot editor: “If we link to a site where most of our readers have to fill out an intrusive registration process to read the story we linked to, what’s going to occur? Most of them aren’t going to do it; they’ll come back to Slashdot and instead of writing some sort of useful comment they’ll write a complaint. Someone will cut-and-paste the story text into a comment. So everyone loses. The newspaper doesn’t get the readership. Slashdot readers write complaints instead of commenting on the story. And finally, everyone ends up reading the story in a comment posted on our site instead of the original site.”).

⁶⁵ See ROBERT ELLIS SMITH, BEN FRANKLIN’S WEB SITE: PRIVACY AND CURIOSITY FROM PLYMOUTH ROCK TO THE INTERNET 49-73, 284-309 (Privacy Journal 2000).

⁶⁶ Adrian Holovaty, OJR Article on User Registration, <http://www.holovaty.com/blog/archive/2002/06/28/1047> (June 28, 2002) (describing “the potential for severe password security breaches” in these situations).

⁶⁷ A decrease in user privacy also has economic effects, including the cost of spam, identity theft, telemarketing and higher prices paid by consumers who forego registration or “loyalty” schemes. See ROBERT GELLMAN, PRIVACY, CONSUMERS, AND COSTS: HOW THE LACK OF PRIVACY COSTS CONSUMERS AND WHY BUSINESS STUDIES OF PRIVACY COSTS ARE BIASED AND INCOMPLETE (March 2002), <http://www.epic.org/reports/dmfprivacy.html>.

⁶⁸ Hoofnagle, *supra* note 5.

⁶⁹ Lasica, *supra* note 66.

⁷⁰ See also Hoofnagle, *supra* note 5 (describing “privacy self-defense”).

⁷¹ Bugmenot.com, Frequently Asked Questions, <http://www.bugmenot.com/faq.php#12> (last visited Aug. 4, 2007).

complete even the more intrusive registration forms, which require authentication emails, without actually relinquishing any truthful demographic data.

¶43 Inputting noise into the system like this is a commonly used technique, with estimates of false user data ranging from a negligible amount,⁷² to ten percent,⁷³ to over a quarter,⁷⁴ of all information submitted through registration forms.⁷⁵

C. Bugmenot.com

¶44 Noise can provide pretty good anonymity. But the transaction costs are high because a user must sit at her terminal and generate fake data each time she wants to read a news link. There is a simpler solution—bugmenot.com.

¶45 Bugmenot is a service that collects user logins and passwords for websites that are generally available to the public after registration.⁷⁶ For example, someone who wants to access the *Los Angeles Times* website, but does not want to undergo the registration process, can access Bugmenot, search for a user-submitted account for the *Los Angeles Times*, and log in to the paper's website with one of the user-submitted accounts.⁷⁷ This account name and password will often be stored by a user's browser, which enables repeat viewing of the *Los Angeles Times* without having to access Bugmenot again.⁷⁸

¶46 Though Bugmenot's name indicates that it was created to make a user's experience more convenient,⁷⁹ it also ends up preserving a user's privacy because it cloaks an individual user's trail through a website by combining it with the trails of everyone else who used the same login.⁸⁰

⁷² Small, *supra* note 26 (indicating that “[r]eports of widespread registration fraud are . . . apparently quite exaggerated . . . [nowhere] near the order of magnitude research companies seem happy to report”).

⁷³ *Web Newspaper Registration Stirs Debate*, CNN.COM, June 14, 2004, available at http://chnm.gmu.edu/digitalhistory/links/cached/chapter6/6_28a_registration.htm (reporting that “[a]bout 10 percent to 15 percent of the 300,000 registrations [to the *Philadelphia Inquirer* website] to date have had e-mail addresses”).

⁷⁴ Posting by Vin Crosbie to Poynter Online, http://www.poynter.org/article_feedback/article_feedback_list.asp?id=60149&DGPCrSrt=&DGPCrPg=4 (Jan. 30, 2004, 6:27:59 PM) (noting that “Odyssey President Nick Donatello told a UC/Berkley [sic] Graduate School of Journalism that his firm’s objective research shows that about 27 percent of people intentionally falsify registration data”).

⁷⁵ Noise has been used as a security feature in other contexts. In World War II, the Allies used a device known as SIGSALY which added random noise to each end of a phone conversation. This noise could then be removed by someone with a matching machine at the receiving end. See Wikipedia, SIGSALY, <http://en.wikipedia.org/wiki/SIGSALY> (last visited Aug. 4, 2007). Garbage is added to encrypted emails to make them a uniform length and thus harder for interceptors to reckon which emails should be targeted for decryption. Logging in to computer systems using pre-programmed logins like “admin,” “anonymous” or “cyberpunk” created noise in system logs because no unique identifier could be associated to particular actions. Interview with Lauren Gelman, Executive Director, Stanford Law School Center for Internet and Society, in Stanford, Cal. (2007).

⁷⁶ Bugmenot's website claims to have account information for more than 224,641 websites. Bugmenot.com, Frequently Asked Questions, <http://www.bugmenot.com/faq.php#10> (last visited Oct. 11, 2008).

⁷⁷ There is also a *New York Times*-specific login generator at <http://www.majcher.com/nytview.html>.

⁷⁸ Bugmenot is also available via a browser plug-in for Firefox, which automatically inputs multiple user names and passwords into a website until a working account is found, making accessing the website a seamless experience. Eric Hamiter, BugMeNot, <http://erichamiter.com/firefox/bugmenot/> (last visited Jan. 26, 2008).

⁷⁹ Though privacy does rank first on the reasons for not registering an account, as offered on Bugmenot's FAQ:

“Why not just register?

- It's a breach of privacy.
- Sites don't have a great track record with the whole spam thing.
- It's contrary to the fundamental spirit of the net. Just ask Google.
- It's pointless due to the significant percentage of users who enter fake demographic details anyway.
- It's a waste of time.
- It's annoying as hell.
- Imagine if every site required registration to access content.”

Bugmenot.com, Frequently Asked Questions, <http://www.bugmenot.com/faq.php#03>.

⁸⁰ But using a Bugmenot account (or inputting false information) will not hide a user's IP address. To conceal IP address, users must employ an anonymizing service. These programs cloak a user's true IP address with another one. These services are not

¶47 Bugmenot is not infallible. Anyone can cycle through the user names, including webmasters, who could defeat the circumvention by disabling the accounts listed on Bugmenot.⁸¹ But this countermeasure would be costly and time-consuming for webmasters, and thus unlikely to occur (at least until the process is automated). The account-disabling technique could also probably be stymied by Bugmenot's user base, which could submit new logins for popular websites more quickly than a webmaster could take them down.

¶48 Wide adoption of Bugmenot also has a potential downside.⁸² Popularization of a registration avoidance scheme like Bugmenot could eventually incite the ire of the website owners, whose valuable user data is threatened. As more people take advantage of these anonymity-preserving solutions, the content providers may escalate their registration regimes. EPIC warns that "providing 'bad' or incorrect information might result in an increased tendency on the part of newspapers to require more invasive information from users."⁸³

IV. INDUSTRY-INITIATED SOLUTIONS

¶49 The proactive solutions offered above should not stand alone. Privacy should not be the privilege of the technologically savvy, but should be built into systems that all people use to access content. The best methods for preserving user privacy come from the content side of the transaction and apply to all users, in contrast to the incomplete protection provided by the user-driven techniques described above.

A. Faceless Data

¶50 Websites could move to an advertising system that does not require registration but still targets advertisements. The targeting could be based on the content of an article, like Google's advertisements,⁸⁴ and the reader's location as determined by IP address. When compared to registration, this location- and content-based targeting would better preserve user privacy because the relevant disclosure would be limited to the content of a single article and not the accumulated interests of a particular reader. Local newspapers might object to this suggestion because they do not have relationships with as wide a range of advertisers as Google does. For example, if a New Orleans newspaper runs an article about a record-breaking snowfall in Michigan, it is unlikely any of its local advertisers would have anything to offer related to the article.

¶51 If a new advertising system is not workable, user-tracking systems should be developed that do not store any personally identifiable information. This is the regime that is currently used by TiVo, a service that allows users to digitally record television.⁸⁵ Information about reading habits should be stored so that it is not possible to link it to a particular person, or so that individually-marked data destructs if it is extracted from a general matrix. Even if names and email addresses are not retained, IP addresses should also be scrubbed from a particular reading profile. Data should be heavily

user-friendly or widely used, but could be more widely popularized if built into browsers or available as browser plug-ins or extensions.

⁸¹ *The New York Times* is aware of Bugmenot. See Dan Mitchell, *A Broadband Beat-Down*, N.Y. TIMES, June 25, 2005, at C5 ("If newspaper marketers think they are receiving reliable user information via those annoying site registrations, they should run their Web addresses through bugmenot.com").

⁸² The creator of a Bugmenot login may also open herself up to liability, if she could ever be tracked down. *The New York Times*' "Registration and Security" guidelines note that "[y]ou are responsible for all usage or activity on your NYTimes.com account, including use of the account by any third party authorized by you to use your Member ID and password." *The New York Times* Member Agreement, <http://www.nytimes.com/ref/membercenter/help/agree.html#g>.

⁸³ Hoofnagle, *supra* note 5.

⁸⁴ Google AdSense Tour, http://www.google.com/services/adsense_tour/ (last visited Oct. 14, 2008) (explaining that "AdSense delivers relevant text and image ads that are precisely targeted to your site and your site content").

⁸⁵ TiVo Privacy Policy, <http://www.tivo.com/abouttivo/policies/tivoprivacypolicy.html> (last visited Oct. 14, 2008) ("TiVo does not collect or access any Personally Identifiable Viewing Information (as defined below) from your TiVo DVR without your prior consent. Absent your consent, TiVo does not keep track of what shows you-as an individual or household- have watched, recorded, or rated . . .").

encrypted or kept in a unique proprietary database. This would protect accidental or illicit venting of the user information, and also help forestall any unforeseen “function creep”⁸⁶ of the registration technologies.

B. Data Tracking

¶52 Databases can also be designed to increase accountability about how the data is used. Each datum describing a person’s visit should have its own identifying code that can be tracked.⁸⁷ This system can piggy-back onto the structure already in place to organize the information gathered by the registration process. The tracking number would act like a homing beacon and allow users to follow where their data was sent. Though this technology would not limit how the information was distributed, the new level of transparency might make websites hesitant to share user’s personal information with scurrilous third-parties if they know people will be able to see exactly who they are doing business with and how the data is used. Anyone who has surfed the Internet could ask where information about her interests have been scattered and how it is being used.

¶53 At first glance, this seems to be an unworkable idea in the days of international scoff-law spam-lords, but a similar tracking technique has been used by domain owners who create site-specific email addresses that include the name of a website requiring registration. For example, if I receive all email to any account at stopmindgangersters.com, I can register my email as lat@stopmindgangersters.com with the *Los Angeles Times*. Then, if I ever receive email addressed to “lat@stopmindgangersters.com,” I know it is because the *Los Angeles Times* used or transferred my address to a third party. By monitoring what email addresses get spam, the spam has effectively become watermarked with the site that gave up the information and a savvy Internet user can see what website passed on the email address. In the same way, letting a registered user see where information about them travels creates a new layer of accountability that may encourage reform of data sharing practices.

C. Blanket Licenses

¶54 Finally, content providers could secure financial arrangements with Internet service providers (ISPs). If old media companies are worried about monetizing their content for the Internet, an agreement with telecommunications companies and other service providers would ease this fear. If each ISP paid a set amount to the major newspapers so that their individual customers would not need to register, the money from this arrangement could offset the increased value that the user profiles add for advertisers. Educational institutions, which also provide access to the Internet for many in the U.S., could enter their own arrangements as they do for many of the medical and other databases available to students and faculty.

The old registration regimes (and corollary work-arounds) could remain in place for those not accessing the sites through a partner ISP, so that those not able to take advantage of such an arrangement could still access the content. Perhaps as the amount of users that were not required to register because of their ISP increased, the cost of maintaining a user registration system would become prohibitive and these systems would be scrapped for all users, thus allowing people not accessing the site via one of the partner ISPs to read the site for “free.”

¶55 This plan’s most serious flaw is that it could not completely compensate the sites for the amount of revenue drawn in from advertising. So this plan would not seek to provide advertising-free access to content (as “premium memberships” to sites like Salon.com do), but merely to limit the information procured by registration—advertisements would still exist, just not be targeted based

⁸⁶ M. Granger Morgan and Elaine Newton, *Protecting Public Anonymity*, 22 ISSUES IN SCI. & TECH. 83, 86 (Fall 2004), available at http://www.issues.org/21.1/granger_morgan.html (explaining how a powerful system can be used in unexpectedly “beneficial but perhaps also pernicious ways”).

⁸⁷ This is architecturally similar to Lessig’s proposed P3P protocol. LESSIG, *supra* note 3, at 160. However, unlike P3P, the tracking information proposed here does not prospectively define the privacy relationship with a website.

on individual user characteristics. This sort of arrangement could be marketed by the ISPs as “instant access to the top 25 newspapers!” But given the relatively easy access Internet users already have to these websites once they register, the actual additional marketing sway this would have for an ISP is probably negligible. This arrangement would also merely shift the tracking data from the media sites to the ISP, concentrating it in one place instead of being split piecemeal over many different databases.

V. LEGISLATIVE SOLUTION: THE MIND GANGSTER ACT

¶56 Both of the above Parts propose solutions to the problem of user registration that could be implemented by non-government actors. But the federal government should also address this issue because regulation is unlikely to come from within the content industry. As mentioned, websites requiring registration regard problems with the schemes as a matter of “market acceptance,” not privacy.⁸⁸

¶57 A new legal regime need not forbid user registration altogether. It does not even have to forbid selling user data and traffic patterns. Instead, a new federal law should limit what is collected and corral the sharing of data. The law could have three parts: a notice requirement, best practices for maintaining information, and proscriptions against sharing the data. Call it the MIND GANGSTER Act (Means and Intent to Not Distribute Great Amounts of Necessarily Gathered Statistics That are Enabled by Registration).

¶58 The first part of the law could impose a disclosure requirement on websites. This requirement would disallow smoke screens of “improving the user experience” when the website is actually selling user data to other parties. This section of the law would mandate clear, unambiguous notice—like the Surgeon General’s warning on cigarette packets—that personal information will be sold for a profit or handed over to law enforcement agencies.⁸⁹ A law requiring this type of warning (and restrictions on placement, font and size) would finally place a real decision into a visitor’s hands because the Internet user would have full knowledge of just what is at stake in giving up her name, email and occupation.⁹⁰

¶59 The second part of the MIND GANGSTER Act could prescribe the methods by which this information is collected. Since having any one individual’s reading habits on file is not particularly useful for advertisers, but could be particularly damaging to that one person, best practices for data collecting should be inscribed via law. Email and IP addresses and other identifying information should not be retained or in any way linked to data about reading habits.

¶60 Finally, there should be strict restrictions on data sharing.⁹¹ This law should forbid disclosure of reading habits and personally identifying information, including emails, to third-parties without the express consent of the person involved.⁹² And because of the important First Amendment interests

⁸⁸ Cohen, *supra* note 20, at 989; *see also* Scott Foster, *Online Profiling Is On The Rise: How Long Until the United States and the European Union Lose Patience With Self-Regulation?*, 41 SANTA CLARA L. REV. 255, 257–58 (2000). *But see* Network Advertising Initiative, <http://www.networkadvertising.org/> (last visited Oct. 16, 2008).

⁸⁹ Such a law would extend the type of FTC enforcement seen in *GeoCities*, FTC Docket No. C–3849 (Feb. 12, 1999) (consent order) (proscribing misrepresentation about data collection on websites). For an example of such clear wording, *see* AllThingsD – About Us – Tracking Cookie Information, <http://allthingsd.com/trackingcookies/> (last visited Dec. 30, 2008).

⁹⁰ Some commentators object to protecting privacy via a “property regime.” *See* LESSIG, *supra* note 3, at 161 (outlining both sides of this argument).

⁹¹ This law must be finely crafted or it will hurt civil liberties in other areas. *See* Bill Stuntz, *Privacy and Transparency, Continued*, THE NEW REPUBLIC ONLINE, Apr. 25, 2006, <http://www.tnr.com/doc.mhtml?i=w060424&s=stuntz042506> (last visited Apr. 25, 2006) (on file with author) (noting that “[s]egregationists attacked civil rights legislation on the same [privacy] ground: enforcing it would require small business owners to disclose too much private information That tells me privacy, at least in some of its forms, is dangerous: an impediment to the kinds of government action that are needed to protect minorities and the poor.”).

⁹² Some state laws already offer similar protections. For example, Michigan has a stronger law protecting customers’ identity regarding sale, rental or borrowing of books, written material, sound or video recording. MICH. COMP. LAWS ANN. § 445.1712 (West 2008). An Internet website that allows readers to view written material might fall within the purview of this law.

at stake in this information, this law should also amend the rules of criminal procedure for subpoenas by elevating the standard beyond mere relevance, as currently required.⁹³

¶61 This type of regulation is not unprecedented. In one cranny of the U.S. Code lurks a strange golem—the Video Privacy Protection Act (VPAA).⁹⁴ The VPAA was passed after the media revealed some of Judge Robert Bork’s salacious video rentals during his failed Supreme Court confirmation hearing. The law protects the disclosure of video rental records without the renter’s consent, except in a few defined circumstances.⁹⁵ Disclosure is allowed pursuant to a court order in a civil proceeding,⁹⁶ and to law enforcement, provided a grand jury subpoena.⁹⁷ Some information may be disclosed to other business partners in two circumstances: the first is the so-called “genre exception,” which allows marketers access to the name and address of a renter as well as the “subject matter” of material she rented, so long as the consumer has been given a chance to opt out of this disclosure.⁹⁸ The second allows any disclosure if it is “incident to the ordinary course of business of the video tape service provider.”⁹⁹

¶62 The MIND GANGSTER Act would offer many of the same protections, but should not have any exceptions for disclosures to marketers, such as the “genre exception” in the VPAA. This new law should also more precisely define the “ordinary course of business” to not include conversations or meetings with advertisers—personally identifiable information should never be shared in these settings. Creating these holes in the protection would allow for manipulation of the rules by content providers who could reconfigure their practices to comply with the letter of the law but still collect and disseminate sensitive information. As mentioned, statutory liquidated damages should be high because liability limited to an individual’s harm will not effectively incentivize content providers to protect user information.

¶63 One pitch-perfect provision of the VPAA that should be brought into the online context is the destruction of customer information after one year,¹⁰⁰ as this would minimize any damage from database security being compromised. It may cause some frustration to itinerant users of websites requiring registration who can no longer log-in, but there are probably few people who access one site a few times over the course of years and can still remember their user names and passwords.

VI. CONCLUSION

¶64 Now is the time to protect the bounty of user registration gathered in the larders of the world’s content companies. If the huge amount of data about the Internet’s reading habits leaves the hands of the companies that have collected it, it will be impossible to sanitize. It is far easier to leave a genie homeless than tell it to get back in its bottle.

¶65 Content providers should design privacy-protecting systems now before such protections are rendered moot. The increasing interest in data mining by law enforcement and intelligence agencies should motivate Congress to protect their constituents’ privacy now. In the meantime, users should practice techniques that protect their reading choices from capture by the sleepless sentinels of the world wide web.

⁹³ See *In re Grand Jury Subpoena to Amazon.com Dated August 7, 2006*, 246 F.R.D. 570, 572 (W.D. Wis. June 26, 2007) (finding a “legitimate First Amendment concern” in letting the government “peek into the reading habits of specific individuals without their prior knowledge or permission.”). See also Daniel J. Solove, *The First Amendment as Criminal Procedure*, 82 N.Y.U. L. REV. 112 (2007). But see 50 U.S.C. § 1861(a)(1) (2006) (National security investigations shall “not [be] conducted solely upon the basis of activities protected by the first amendment to the Constitution.”).

⁹⁴ 18 U.S.C. § 2710 (2006).

⁹⁵ 18 U.S.C. § 2710(2) (2006).

⁹⁶ 18 U.S.C. § 2710(b)(2)(F) (2006).

⁹⁷ 18 U.S.C. § 2710(b)(2)(C) (2006).

⁹⁸ 18 U.S.C. § 2710(b)(2)(D) (2006).

⁹⁹ 18 U.S.C. § 2710(b)(2)(E) (2006).

¹⁰⁰ 18 U.S.C. § 2710(e) (2006).