PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: A PROPOSED FRAMEWORK FOR BUSINESSES AND POLICYMAKERS (PRELIMINARY STAFF REPORT)
RESPONSE OF THE ELECTRONIC FRONTIER FOUNDATION

The Electronic Frontier Foundation (EFF) respectfully submits this response to selected questions raised by the Federal Trade Commission (FTC) staff report on a proposed consumer privacy framework, focusing mainly on issues pertinent to the online environment. We view the proposed framework largely as an interpretation of the FTC's existing regulatory authority under sector-specific privacy laws and its general consumer protection mission regarding unfair, deceptive or misleading trade practices. Online privacy, however, raises civil liberties issues above and beyond commercial collection and use of consumer information, because such information in the hands of commercial entities is also readily available to the government for many purposes. Our views on consumer privacy in the commercial realm accordingly reflect our concern for online civil liberties as well.

QUESTIONS FOR COMMENT ON PROPOSED FRAMEWORK

I. Scope

• Are there practical considerations that support excluding certain types of companies or businesses from the framework – for example, businesses that collect, maintain, or use a limited amount of non-sensitive consumer data?

Because we view the framework largely as an articulation of the Commission's basic Section 5 authority over unfair, deceptive or misleading business practices in the area of consumer privacy, we believe it would be inconsistent for any types of entities or businesses to be *per se* excluded from the framework. As explained in more detail below, virtually any kind of consumer data can be used for linking or re-identification. Excluding any category of commercial entity from the framework would still permit that data to flow and be aggregated. As a practical matter, however, it would be reasonable to forbear from regulating businesses that pose little risk to consumers, focusing on those that engage in more egregious privacy abuses.

• Is it feasible for the framework to apply to data that can be "reasonably linked to a specific consumer, computer, or other device"?

We believe that it is generally feasible for the framework to apply to data that can be "reasonably linked to a specific consumer, computer, or other device"; the HIPAA Privacy Rule takes a similar approach. More generally, we agree with the staff report's rejection of the traditional focus on PII (personally identifiable information).[1]

[1] See Arvind Narayanan and Vitaly Shmatikov, *Myths and Fallacies of "Personally Identifiable Information,"* 53 Communications of the ACM 24 (June 2010) and Seth Schoen, *What Information is*

We suggest, however, that the Commission should treat "consumer" as including "home" or "household" as well. In our work on smart grid privacy, for instance, where energy usage is collected per household, we have noted that "household-identifiable information is inherently revealing of household activities and home life, traditionally private domains that are, and should continue to be, protected from observation."[2]

Relatedly, a recent California Supreme Court decision, *Pineda v. Williams-Sonoma Stores, Inc*. (Feb. 10, 2011)[3], criticizes over-reliance on whether information—in that case, ZIP code—identifies specific individuals, noting that much information currently treated as relating to a specific individual actually may not. ("In the case of a cardholder's home address, for example, the information may pertain to a group of individuals living in the same household. Similarly, a home telephone number might well refer to more than one individual.").

For similar reasons, "consumer" should include "account" as well.

• Are there reliable methods for determining whether a particular data set is "linkable" or may become "linkable"?

Using the information-theoretic concept of entropy, one can estimate the extent to which data can uniquely identify a person. As more information is aggregated, it generally becomes easier to resolve to a particular individual.[4]

But while such techniques can help data holders reduce dataset re-identifiability, any estimate of re-identifability is directional at best. There are two basic problems. First, computing resources are extremely large relative to population. Narayanan puts it simply: "a lot of traditional thinking about anonymous data relied on the fact that you can hide in a crowd that's too big to search through. That notion completely breaks down given today's computing power: as long as the bad guy has enough information about his target, he can simply examine every possible entry in the database and select the best match."[5]

From an information-theoretic perspective, only about 33 "bits" of information are needed to uniquely identify a person out of the entire world population. And as

---

"Personally Identifiable" September 11, 2009, *available at* https://www.eff.org/deeplinks/2009/09/what-information-personally-identifiable.

[2] Rulemaking 08-12-009, California Public Utilities Commission, *Joint Comments of the Center for Democracy & Technology and the Electronic Frontier Foundation on Proposed Policies and Findings Pertaining to the Smart Grid* 14 (March 9, 2010), *available at* https://www.eff.org/files/CDTEFFJointComment030910.pdf.

[3] , *Pineda v. Williams-Sonoma Stores, Inc*. (Feb. 10, 2011), *available at* http://www.courtinfo.ca.gov/opinions/documents/S178241.pdf.

[4] Peter Eckersley, *A Primer on Information Theory and Privacy,* January 27, 2010, *available at* https://www.eff.org/deeplinks/2010/01/primer-information-theory-and-privacy.

[5] *About 33 Bits, available at* http://33bits.org/about/.

Narayanan notes, "33 bits is not really a lot. If your hometown has 100,000 people, then knowing your hometown gives me 16 bits of entropy about you, and only 17 bits remain."[6] Indeed, location information not only greatly narrows search, it also decreases the cost of search. It is easier to track down one out of a list of 100 possible persons by brute-force search when they all live in the same city than if they are scattered across California. And if you are trying to associate individuals with an occurrence in a particular place, you focus on those in proximity to that place. Location information is thus extremely useful for re-identification.

Second, any information that distinguishes one person from another can be used for re-identification, and there is an enormous and growing amount of information about individuals available to government and commercial entities in today's information marketplace: consumption preferences, commercial transactions, Web browsing, search histories, and so on.

As the Commission is aware, Narayanan and Shmatikov's initial research used movie-viewing histories to re-identify individuals in the Netflix dataset. EFF senior staff technologist Peter Eckersley showed how browser metadata can be used to distinguish Website visitors.[7] Bradley Malin of Vanderbilt University was able to re-identify a substantial percentage of "de-identified" shared pedigree and genomic records by online newspaper obituaries, which report the name of the deceased and, in many instances, the names of the deceased person's relatives.[8]

The general lesson here is that the linkability problem is a function of the universe of available data, not merely the particular data that one is exchanging

• How should the framework apply to data that, while not currently considered "linkable," may become so in the future?

As explained above, most information about people can be used for "linking" purposes. As Narayanan and Shmatikov put it, "*any attribute can be identifying in combination with others*." Thus, one should generally consider all data to be linkable.

As a practical matter, the framework can provisionally focus on information known to be especially useful in linking. Identified information (including data often found in public records) can help link a unique person to a particular identity. Academic research also provides focus; as noted earlier, location information is extremely critical. For more discussion of the re-identification implications of location, see Narayanan's blog,[9]

---

[6] Ibid.
[7] Peter Eckersley, *Is Every Browser Unique? Results from the Panopticlick Experiment* May 17, 2010, *available at* https://www.eff.org/deeplinks/2010/05/every-browser-unique-results-fom-panopticlick
[8] Testimony of Bradley Malin, Before the U.S. Department of Health and Human Services AHIC, Confidentiality, Privacy, and Security Workgroup, at 4 (June 22, 2007), *available at* http://hiplab.mc.vanderbilt.edu/people/malin/Papers/Malin_CPS_Testimony_6-22-07.pdf.
[9] *Your Morning Commute is Unique: On the Anonymity of Home/Work Location* Pairs May 13, 2009 *available at* http://33bits.org/2009/05/13/your-morning-commute-is-unique-on-the-anonymity-of-

(discussing the implications of Philippe Golle and Kurt Partridge, *On the Anonymity of Home/Work Location Pairs*[10]).

Moreover, because re-identification is a commercial enterprise, the Commission can "follow the market" by monitoring commercial re-identification practices. Certain industries are known to engage in extensive data mining for marketing purposes. See, e.g., *'Scrapers' Dig Deep for Data on Web*[11](using example of website PatientsLikeMe). The pending U.S. Supreme Court case, *Sorrell v. IMS Health*,[12] highlights this issue with respect to data-mining of medical records concerning prescriber-identified prescription information. To the extent that entities in this space claim that they are anonymizing or de-identifying the data they sell or disclose, it would be useful for the Commission to study these ongoing practices and ask how well they mitigate re-identification.

• What technical measures exist to "anonymize" data and are any industry norms emerging in this area?

As noted earlier, it may be possible to make re-identification harder by reducing the entropy or distinguishing power of datasets, but the efficacy of this approach is unclear because of the relevance of other available data, as in the context of IP address logging and cookies.[13]

We do not see any industry norms emerging in this area, although the topic is an active one.[14] Part of the problem may be lack of clarity around what companies mean when they assert that they have anonymized or de-identified data.

Until better technical measures are developed, entropy reduction and data minimization as to both collection and retention are probably the most effective approaches.[15] (Microsoft deletes the cookies, the full IP address and any other identifiable user information from its search logs after 18 months).

II. Companies should promote consumer privacy throughout their organizations and at every stage of the development of their products and services

---

homework-location-pairs/.

[10] Philippe Golle and Kurt Partridge, *On the Anonymity of Home/Work Location Pairs*, *available at* http://crypto.stanford.edu/~pgolle/papers/commute.pdf.

[11] Julia Angwin and Steve Stecklow, *'Scrapers' Dig Deep for Data on Web*, October 12, 2010, *available at* http://online.wsj.com/article/SB10001424052748703358504575544381288117888.html

[12] see generally http://www.scotusblog.com/case-files/cases/sorrell-v-ims-health-inc/.

[13] Chris Soghoian, *Debunking Google's log anonymization propaganda*, September 11, 2008, *available at* http://news.cnet.com/8301-13739_3-10038963-46.html.

[14] See, e.g., Grigorios Loukides, Aris Gkoulalas-Divanis, and Bradley Malin, *Anonymization of electronic medical records for validating genome-wide association studies* (2010), *available at* http://www.pnas.org/content/early/2010/04/05/0911686107.full.pdf+html.

[15] See, e.g., Microsoft, *Microsoft Announces Enhanced Privacy Protections for Customers*, July 22, 2008, *available at* http://www.microsoft.com/presspass/press/2007/jul07/07-22EnhancedPrivacyPrinciplesPR.mspx.

• Are there substantive protections, in addition to those set forth in Section V(B)(1) of the report, that companies should provide and how should the costs and benefits of such protections be balanced?

Section V(B)(1) of the report focuses on four substantive protections: data security, reasonable collection limits, sound retention practices, and data accuracy. We do not propose additional substantive protections, but elaborate on those set forth in the report.

First, companies should seek to implement these protections in an integrated fashion and design these protections into their system and product or service architecture. Put another way, we believe that the report's articulation of substantive protections casts "privacy by design" mainly in terms of rules and information practices and tends to ignore what some technologists call "privacy-aware design."

Perhaps the most elementary example is anonymous payment schemes like cash. The classic subway token—a form of cash—makes it hard to compile travel histories of subway users. Similarly, the Bay Area Rapid Transit (BART) system has long used cheap magnetic-stripe fare cards that are not tied to users' identities. Increasingly, however, mass transit systems use persistent fare cards tied to user accounts, creating travel histories as a routine matter.

Similarly, many computing applications can be designed to be more centralized or more distributed, such as in terms of data storage. In privacy terms, it is better for individuals' data to be stored locally on their own devices than on servers somewhere in the cloud. The contingencies of information control are vastly different between a home telephone answering machine and voicemail stored with one's service provider.

In the smart grid context, obfuscation techniques have been suggested. SeeThomas Nicol and Thomas Overbye, *Toward Technological Defenses Against Load Monitoring Techniques* (attaching some form of fast response energy storage between the breaker and the meter may eliminate device signatures without interfering with operations).[16]

Other examples include the careful use of cryptographic techniques into system architectures. Andrew Blumberg and others have done considerable work in the area of transportation and location privacy.[17] Microsoft and others are working on privacy architectures in the smart grid.[18]

Second, companies should bear some responsibility for data that they sell, share or

---

[16] Thomas Nicol and Thomas Overbye, *Toward Technological Defenses Against Load Monitoring Techniques*, Proc. North American Power Symposium (NAPS), 2010, *available at* http://www.iti.illinois.edu/sites/www.iti.illinois.edu/files/docs/tcip/2010_Nicol_Overbye.pdf.

[17] See, e.g., Raluca Ada Popa, Hari Balakrishnan, and Andrew Blumberg, *VPriv: Protecting privacy in location-based vehicular services*, 2009, *available at* http://www.math.utexas.edu/users/blumberg/loc_writing.html.

[18] Alfredo Rial and George Danezis, *Privacy-Preserving Smart Metering*, November 19, 2010, *available at* http://research.microsoft.com/apps/pubs/?id=141726.

otherwise provide to other entities. The report focuses on individual companies' information practices, yet much of today's privacy problems arise from the flow of information between and among companies. The Commission should consider how these protections can be translated into the sharing or exchange context.

This issue is especially appropriate to analyze in computing and online contexts. For instance, familiar mobile applications are to some extent subject to the governance of the mobile device's operating system. Vendors of mobile operating systems therefore have some role in how the applications collect information. Similarly, applications commonly found on social media platforms have some kind of relationship with the platform, and the information flows among user, application and platform deserve attention in a way that goes beyond the user's relationships with the platform and each application, viewed individually.

• Should the concept of "specific business purpose" or "need" be defined further and, if so, how?

As we read the report, the concept of "specific business purpose" or "need" serves at least two functions. First, it is part of disclosure or transparency. It follows that the notion of specificity should be tied to consumer expectations of meaningful disclosure. Our experience with privacy policies in the smart grid context indicates that policies were often underspecified—lacking, for example, definitions for critical terms, such as the types of energy usage data protected.[19]

Second, the concept of "specific business purpose" or "need" is part of the calculus over the reasonableness of business practices such as information collection or retention. In the smart grid context, policies often list purposes for which data will be used that are so broadly stated (e.g., "to provide you with a better experience") as to allow virtually limitless uses of data.[20]

While we are wary of government definitions of terms that may be used in many different contexts, the concept of specificity cannot serve either of these intended functions without some standardization in terminology. If the meanings of terms like "affiliate" or "anonymization" vary too widely, consumers can easily be confused.

Finally, information disclosure must be sufficient to permit evaluation of whether a given practice, e.g. collection or retention, is reasonable given that purpose specification. As an

[19] See Rulemaking 08-12-009, California Public Utilities Commission, *Proposed Smart Grid Privacy Policies and Procedures—Opening Response of the Center for Democracy & Technology and the Electronic Frontier Foundation to Assigned Commissioner's Ruling of September 27, 2010* 8 (Oct. 15, 2010); *id*., n.31 (noting that one utility's privacy policy interchangeably uses the terms "customer information," "personal information," "personally identifiable information," and "personal customer information" without defining those terms), *available at* https://www.eff.org/files/PoliciesandProcedures_15Oct2010_OpeningComment.pdf.

[20] *Id*., n. 32.

illustrative example, CDT and EFF proposed the following language on purpose specification to the California Public Utilities Commission in the smart grid privacy context:

"3. PURPOSE SPECIFICATION The notice required under section 2 shall provide—
(a) an explicit description of—
(1) each category of covered information collected, used, stored or disclosed by the covered entity, and, for each category of covered information, the specific purposes for which it will be collected, stored, used, or disclosed, and
(2) each category of covered information that is disclosed to third parties, and, for each category, (i) the purposes for which it is disclosed, (ii) the identities of the third parties to which it is disclosed, and (iii) the value of the disclosure to the customer…" *Id.*, Appendix A, at 2.

• Is there a way to prescribe a reasonable retention period?

• Should the retention period depend upon the type or the sensitivity of the data at issue? For example, does the value of information used for behavioral advertising decrease so quickly that retention periods for such data can be quite short?

We believe that reasonable retention periods can be prescribed and agree with the report's default principle of minimal retention: that businesses should retain consumer data for only as long as they have a specific and legitimate business need to do so. Empirical data will be needed here, but it should be possible for industry to show how the value of specific types of information to advertising or to fraud prevention changes over time.

Such analysis, however, relates primarily to the business case for retention and not its sensitivity for consumers. Clearly, some kinds of data have clear potential to harm consumers (such as a Social Security number being abused for the purpose of identity theft), and numerous statutes give special protection to medical and financial information.

Our general position with respect to online data collection, however, is that the general importance of protecting this information is so high that further distinctions about sensitivity may be unnecessary. Online data collection typically is both information about one's communications and about one's reading and viewing habits. Such information should be treated as the digital equivalent of one's reading choices in a public library, and generally treated as sensitive.

III. Companies should simplify consumer choice

Commonly accepted practices

• Is the list of proposed "commonly accepted practices" set forth in Section V(C)(1)

of the report too broad or too narrow?

We generally agree with the list of proposed "commonly accepted practices" in terms of consent requirements. We do not think, however, that practices "commonly accepted" by consumers should be equated with practices based on either operational requirements (such as fraud prevention) or public policy reasons. In the latter two categories, where the consumer almost by definition lacks any well-formed expectations, greater transparency will help educate consumers about how their information is actually used. Google's Government Requests tool is a good example of how businesses can promote transparency about such matters.[21]

• Are there practices that should be considered "commonly accepted" in some business contexts but not in others?

• What types of first-party marketing should be considered "commonly accepted practices"?

More research is needed to determine how actual consumer acceptance of practices varies across business context, but we believe that consumers primarily expect that their data will be collected and used for the purposes of completing their transactions.

IV. Practices that require meaningful choice

Special choice for online behavioral advertising: Do Not Track (DNT)

Many consumers understand that the websites they visit collect information about them, often for advertising purposes. But most consumers do not understand that when they visit those websites, other entities also collect information about them. The modern website typically brings together content from many different web servers and your browser assembles those pieces of content to display what looks like a single page from a particular branded entity. In this situation, however, your browser is actually requesting data from both that branded entity and many other "third party" servers—and all of those servers can get data from your browser at the same time. [22]

Research has shown that the average popular website incorporates 64 independent mechanisms for tracking visitors over time and across other websites.[23] According to a 2009 national survey by researchers at the University of Pennsylvania and University of California, more than 80% of Americans believe websites should not track their behavior for advertising.[24] And more than 90% believe advertisers should be required by law to

[21] *Greater transparency around government requests*, the Official Google Blog, April 20, 2010, *available at* http://googleblog.blogspot.com/2010/04/greater-transparency-around-government.html.
[22] Peter Eckersley, *How Online Tracking Companies Know Most of What You Do Online (and What Social Networks Are Doing to Help Them)*, September 21, 2009, *available at* https://www.eff.org/deeplinks/2009/09/online-trackers-and-social-networks
[23] Julia Angwin, *The Web's New Gold Mine: Your Secrets*, WALL ST. J., July 30, 2010.
[24] Joseph Turow et al., *Americans Reject Tailored Advertising*, at 15, *available at*

stop tracking on request.[25]

To be clear, our concern here is not advertising but privacy against online tracking: protecting consumers against the largely invisible, poorly understood, and continually escalating surveillance of their online activities. As we noted in our introduction, online surveillance raises significant civil liberties concerns given the potential for government access to information about consumers held by businesses. And our earlier discussion of the vicissitudes of anonymization and re-identification made clear that this information can be used to identify online users and discover their reading, viewing, associational and consumption choices. DNT thus implicates a "wider debate about the monitoring of user activity online, and even more widely, the aggregation of personal information for a variety of purposes."[26]

We do not articulate a precise definition of "tracking" here, but believe that public opposition to tracking is based on consumer expectations regarding surveillance of online activities that results in the compilation of records that can be used to connect records of a person's online activities across space, websites, or time. We hope that businesses will respect DNT voluntarily, such as by adopting more limited logging and retention practices for users who enable DNT.

Mozilla is publicly discussing many of the issues surrounding the meaning of "tracking" and the characterization of users' relationships to websites.[27] As a provisional matter, we note that consistency with the Commission's proposed framework implies that some kinds of tracking—while still tracking—may not be categorically subject to DNT. For instance, under the staff report tracking that is limited to a single "first-party"[28] website would probably fall outside DNT.[29] DNT may not cover commonly accepted practices for which consent is not required, such as tracking necessary to complete a user's intended online transaction. DNT may not cover tracking that is necessary to prevent fraud or respond to security incidents, provided such data is minimized, only kept for as long as necessary, and not used for other purposes.

The above approach to defining "tracking" has some virtues with respect to detection. Tests of compliance could begin with the largest domains, checking to see whether they continue to use technologies such as tracking cookies, super-cookies or fingerprinting when users enable DNT. If so, those firms could be asked to explain how their activities

---

http://ssrn.com/abstract=1478214.

[25] *Id.* at 23.

[26] Arvind Narayanan, *Do Not Track isn't just about Behavioral Advertising,* December 10, 2010*, available at* http://cyberlaw.stanford.edu/node/6573

[27] Michael Hanson (Mozilla Labs), *Thoughts on Do Not Track*, January 23, 2011, *available at* http://www.open-mike.org/entry/thoughts-on-do-not-track.

[28] Hanson, *ibid*., usefully discusses some issues associated with the concepts of "first party" and "third party" in the tracking context.

[29] Staff report, at 54-56. Because there is no sharp technical distinction between first- and third-party websites, and websites can be first parties in one context but third parties in another, first- and third-party status should generally correspond to consumer expectations.

fit within one of the exceptions.

• How should a universal choice mechanism be designed for consumers to control online behavioral advertising?

The design of these mechanisms will evolve over time, but the basic idea is to express the user's intent. Because the browser is the main "window" through which consumers are tracked on the web, using the browser to say, "don't track me" will be heard by every web-tracking entity you encounter—even if you don't know about them.

At this point, we believe the best design is a simple device configuration setting that allows a consumer to set a preference in her web browser that she does not want to be tracked. In slightly more technical terms: When a web browser requests content or sends data using HTTP, the protocol that underlies the web, it can optionally include extra information, called a "header." A Do Not Track setting could simply add a header indicating the user wishes to not be tracked, e.g. "DNT," which the browser would automatically send to HTTP servers.

This could be augmented by server-to-client responses acknowledging compliance with the DNT request, or asking for opt-in permission to track the user anyway. Client software could offer various kinds of user interface responses to requests to opt-back-in.

• How can such a mechanism be offered to consumers and publicized?

Technical discussion of DNT issues is ongoing, and reference implementations have been available for some time.[30] Now that Mozilla has incorporated header-based DNT into its latest release of the Firefox browser,[31] we expect more rapid DNT development; researchers are working on an Internet Engineering Task Force Internet draft for DNT.[32]

Microsoft's Tracking Protection is a distinct privacy mechanism.[33] The general idea is that the user subscribes her browser to a regularly updated list of servers to which her browser will block all incoming third-party connections. If that list contains the addresses for ad networks, the browser will block ad network connections unless you actually click on a link (like an ad belonging to that network) or enter the address.

---

[30] See Do Not Track: Universal Web Tracking Opt-Out, http://donottrack.us.
[31] See http://firstpersoncookie.wordpress.com/2011/01/23/more-choice-and-control-over-online-tracking/.
[32] Jonathan Mayer, *Minor Updates to the Do Not Track Header*, January 27, 2011 *available at* http://cyberlaw.stanford.edu/node/6597.
[33] See *IE9 and Privacy: introducing Tracking Protection* December 7, 2010 *available at* http://blogs.msdn.com/b/ie/archive/2010/12/07/ie9-and-privacy-introducing-tracking-protection-v8.aspx; *Update: Effectively Protecting Consumers from Online Tracking* January 25, 2011, *available at* http://blogs.msdn.com/b/ie/archive/2011/01/25/update-effectively-protecting-consumers-from-online-tracking.aspx.

We think this remains a useful tool, but it's not a substitute for DNT; it's more of a complement. In list-based approaches, the onus is fully on the browser—equipped with this list—to protect the user from being uniquely identified. Meanwhile, online advertisers could still try any method they wish to track user behavior, so long as it happens from authorized domains.

• How can such a mechanism be designed to be clear, easy-to-find, usable, and understandable to consumers?

As noted above, Mozilla is already offering DNT. The mechanism can be included in the privacy settings of devices and browsers, and added as a feature to the existing "private browsing" modes of these devices.[34]

• How can such a mechanism be designed so that it is clear to consumers what they are choosing and what the limitations of the choice are?

It is important to be clear with consumers that there are exceptions to DNT, such as when completing a transaction or in order to detect fraud. However, the exceptions to DNT should not overwhelm the messaging regarding this feature, in the same way the various exceptions surrounding the Do Not Call Registry are available for interested consumers but are not the overriding message when consumers opt-out of receiving telemarketing calls.

• What are the potential costs and benefits of offering a standardized uniform choice mechanism to control online behavioral advertising?

The benefits of standardization are significant: consumers get a single point of control, and service providers get a single channel through which to receive consumer requests. Without standardization, ordinary consumers are simply unable to make effective privacy choices, and service providers are forced to deal with an unpredictable technical environment. We emphasize again that the main goal of DNT is to control online behavioral tracking, and not advertising itself.

• How many consumers, on an absolute and percentage basis, have utilized the opt-out tools currently provided?

Our understanding is that approximately 10 million users have installed AdBlock Plus. Research suggests that the rate of "private browsing mode" is between 6 to 8 percent across the most common browsers.[35]

---

[34] For discussion of "private browsing," see Gaurav Aggarwal et al, *An Analysis of Private Browsing Modes in Modern Browsers, available at* http://crypto.stanford.edu/~dabo/pubs/abstracts/privatebrowsing.html.

[35] *Id*., at 9.

• What is the likely impact if large numbers of consumers elect to opt out? How would it affect online publishers and advertisers, and how would it affect consumers?

As noted above, advertising opt-out tools like AdBlock Plus are already widely available; Microsoft's Tracking Protection may increase use of such tools. DNT, by contrast, is aimed at tracking and not advertising *per se*, and under the staff report's approach, it would seem that most forms of online advertising— contextual advertising, demographic advertising, search advertising, placement advertising, and social network advertising— would not be affected by DNT.[36] We therefore believe that the impact of DNT on online publishers and advertisers will not be as great as industry seems to fear.

We also question the relevance of a recent study of the effects of European Union (EU) privacy regulation.[37] According to that study, online behavioral tracking techniques cause a roughly 2.3 percent increase in advertising effectiveness. The magnitude of this effect, however, is based on stated purchased intentions, rather than actual purchases. Because stated purchase intentions do not always correlate to actual purchases,[38] the true magnitude (or the very existence) of the purported effect cannot be established from this analysis.

Moreover, the study further found that EU regulation had no statistically significant negative impact on advertising effectiveness for the vast majority of the advertisements considered: that is, larger advertisements, dynamic and/or media-rich advertisements, contextual advertisements that are targeted to consumers based on the content of the site, e.g. car advertisements on car websites. In short, the study should be read as showing that regulation had no impact on advertising effectiveness, except for a very specific subset of advertisements.

Even if the impact is more significant, however, we believe that the industry would be able to innovate around these problems, by inventing new metrics techniques and targeting methods that are privacy protective and consistent with the rules of the opt-out.[39]

---

[36] Jonathan Mayer, *Do Not Track Is No Threat to Ad-Supported Businesses*, January 20, 2011, *available at* http://cyberlaw.stanford.edu/node/6592.

[37] *See* Avi Goldfarb and Catherine Tucker, *Privacy Regulation and Online Advertising*, *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1600259.

[38] *E.g.*, Thomas Juster, "Consumer Buying Intentions And Purchase Probability: An Experiment In Survey Design," *Journal of the American Statistical Association*, 61(315), 658-696 (1966); Charles Manski, "The Use Of Intentions Data To Predict Behavior: A Best Case Analysis," *Journal of the American Statistical Association*, 85(412), 934-940 (1990).

[39] For instance, it is possible to count unique visitors to a website using first-party tracking cookies, rather than third-party tracking cookies, and many websites already have this infrastructure in place (see Wikipedia: Web_analytics#Problems with cookies, as of 2/3/2011) . An even better approach is to use a time-limited, non-unique first-party cookie that simply says "this user has been counted for the month."

• In addition to providing the option to opt out of receiving ads completely, should a universal choice mechanism for online behavioral advertising include an option that allows consumers more granular control over the types of advertising they want to receive and the type of data they are willing to have collected about them?

We believe that such mechanisms can and should be developed over time, but it may be premature to require compliance with any such mechanisms at this point. If fine-grained choice mechanisms have been proposed, deployed, and shown to work well for consumers and service providers, the FTC could consider standardizing them over time.

• Should the concept of a universal choice mechanism be extended beyond online behavioral advertising and include, for example, behavioral advertising for mobile applications?

Yes. Rather than being a setting in a browser, such choices might be a setting for the entire phone operating system, especially on phone platforms for which "apps" are widely installed and used.

• If the private sector does not implement an effective uniform choice mechanism voluntarily, should the FTC recommend legislation requiring such a mechanism?

Given Mozilla's adoption of a header-based DNT mechanism, and ongoing technical elaboration of the DNT header standard, we believe that browser vendors are likely to integrate DNT tools into browsers, and that such a requirement will be unnecessary. Consumers will be able to express their intent about being tracked quite clearly.

If we are right, the critical question is whether tracking entities will respect consumers' clearly expressed intent. If they do not, we believe the FTC should recommend legislation, subject to two provisos: (1) the legislation should be narrowly tailored, such as to standardizing Do Not Track mechanisms or crafting safe harbor rules; and (2) avoids mandating particular technical methods of compliance.

V. Companies should increase the transparency of their data practices

Improved privacy notices

• What is the feasibility of standardizing the format and terminology for describing data practices across industries, particularly given ongoing changes in technology?

With respect to advertising, as opposed to web metrics, numerous proposals for non-tracking targeted advertising are under development. See for example, V. Toubiana, A. Narayanan, D. Boneh, H. Nissenbaum, S. Barocas. *Adnostic: Privacy-Preserving Targeted Advertising,* NDSS 2010, *available at* http://crypto.stanford.edu/adnostic/.

Much more can be done to standardize descriptions of data practices. For example, companies could provide consumers with a simple way to be notified if privacy policies change or, better yet, before changes take effect. As a proof-of-concept project, EFF created the TOSBack page[40] which tracks changes in terms of service for numerous consumer-facing websites such as Google, Netflix, Facebook and PayPal. Consumers can stay up-to-date on changes to the terms of service of these companies by subscribing to an RSS feed or visiting the TOSBack website. Though EFF keeps minimal information about visits to the site, a recent day saw about 5,500 hits to the TOSBack site, not including images. We consider the popularity of this site to be a direct reflection of strong consumer interest in staying aware of changes made to privacy policies.

Reasonable access to consumer data

• Where companies do provide access, how should access apply to information maintained about teens? Should parents be able to access such data?

We believe that teenagers require the option of privacy against their parents.

Material changes

• What types of changes do companies make to their policies and practices and what types of changes do they regard as material?

Facebook.com is a case study in how material changes can be implemented to a website's privacy policies in ways that conflict with users' privacy expectations. Facebook's decision to implement sweeping changes to their service in 2010 affected millions of users. These users shared data with Facebook under the aegis of a particular privacy policy, but Facebook then changed that privacy policy in ways that made private data elements public without adequate notice and consumer education. EFF created a timeline[41] of the changes in Facebook's privacy policy from 2005 to 2010, which showcased Facebook's move from a policy under which a user's data elements were only available to other Facebook users belonging to groups that the user specifically chose, to a policy under which those same data elements were available to people without Facebook accounts.

This policy transition was a plainly material change that resulted in considerable public outcry. Facebook's transition from a more private profile to a less private profile was a material change that did not adequately respect the privacy concerns of the millions of affected individuals.

• What is the appropriate level of transparency and consent for prospective changes to data-handling practices?

---

[40] *Available at* http://www.tosback.org/timeline.php
[41] Kurt Opsahl, *Facebook's Eroding Privacy Policy: A Timeline,* April 28, 2010. *available at* https://www.eff.org/deeplinks/2010/04/facebook-timeline/

Websites that plan to change their data-handling practices should provide consumers with adequate prior notice. Optimal changes to policies are best integrated slowly. For example, Twitter.com adopted a consumer-focused approach when it launched "New Twitter." Users were given the option of switching to New Twitter or keeping their traditional Twitter profile, and also offered an opportunity to try New Twitter while being able to go back to the earlier version of the profile for several months. This provided users with a chance to explore the changes on their own timeline, clearly aware of the fact that a change was taking place. Twitter also invested in consumer outreach and education about the change throughout the process, including videos and articles.[42] While Twitter's changes were minimal in regards to data handling, the method of introducing consumers to changing policies should be considered a model for any substantial policy changes.

Consumer education

• What role should government and industry associations have in educating businesses?

One role for government is to lead by example. With respect to DNT, the Government could, without any legislation, "realistically embrace the header as an improved mechanism for tracking opt outs on government sites," thus "[a]voiding the chaos of 100+ different federal agency opt out cookies" and "providing early support for the Do Not Track header at a time when the technology proposal could very much use a boost."[43]

[42] Evan William, *A Better Twitter*, September 14, 2010, *available at* http://blog.twitter.com/2010/09/better-twitter.html and *Twitter: discover what's new in your world*. Posted September 14, 2010, *available at* http://www.youtube.com/watch?v=rIpD7hfffQo.

[43] Christopher Soghoian, *What the US government can do to encourage Do Not Track*, *available at* http://paranoia.dubfire.net/2011/01/what-us-government-can-do-to-encourage.html.