KEEPING THE INTERNET
OPEN • INNOVATIVE • FREE

www.cdt.org

CENTER FOR DEMOCRACY
& TECHNOLOGY

1634 I Street, NW
Suite 1100
Washington, DC 20006

P +1-202-637-9800
F +1-202-637-0968
E info@cdt.org

WHAT DOES "DO NOT TRACK" MEAN?

A SCOPING PROPOSAL BY THE CENTER FOR DEMOCRACY & TECHNOLOGY VERSION 2.0

April 27, 2011

I. Introduction

"Do Not Track" (DNT) is gaining momentum. In 2007, CDT and a coalition of other public interest groups called on the Federal Trade Commission (FTC) to create a system that would allow consumers to avoid being tracked as they browse the Web.[1] After idling for three years, the idea has reemerged. Last July, FTC Chairman Jon Leibowitz expressed support for DNT in Congressional testimony,[2] and in December, the FTC staff advocated for DNT in its Preliminary Staff Report.[3] We have also seen support for DNT in a House subcommittee hearing,[4] proposed legislation,[5] and submissions to the Internet Engineering Task Force (IETF) and World Wide Web Consortium (W3C).[6]

Perhaps more important, leading browser developers have introduced browser-based DNT mechanisms.[7] Because the browser is the gateway to the Internet for

---

[1] *Consumer Rights and Protections in the Behavioral Advertising Sector* (Oct. 2007), http://www.cdt.org/privacy/20071031consumerprotectionsbehavioral.pdf.

[2] Oral testimony of FTC Chairman Jon Leibowitz. Hearing Before the Subcomm. On Commerce, Trade, and Consumer Prot. Of the H. Comm. On Energy and Commerce, 111th Cong. (July 27, 2010), *available at* http://commerce.senate.gov/public/index.cfm?p=Hearings&ContentRecord_id=0bfb9dfc-bbd7-40d6-8467-3b3344c72235&ContentType_id=14f995b9-dfa5-407a-9d35-56cc7152a7ed&Group_id=b06c39af-e033-4cba-9221-de668ca1978a&MonthDisplay=7&YearDisplay=2010. Leibowitz did not include a discussion of DNT in his written testimony.

[3] Federal Trade Commission (Bureau of Consumer Protection), A Preliminary FTC Staff Report on Protecting Consumer Privacy in an Era of Rapid Change: A Proposed Framework for Businesses and Policymakers, 57-63 (Dec. 1, 2010) available at http://www.ftc.gov/os/2010/12/101201privacyreport.pdf.

[4] Do Not Track Legislation, Is Now the Right Time?: Hearing Before the Subcomm. on Commerce, Trade, and Consumer Prot. Of the H. Comm. on Energy and Commerce, 111th Cong. (Dec. 2, 2010), available at http://energycommerce.house.gov/hearings/hearingdetail.aspx?NewsID=8127.

[5] http://speier.house.gov//index.cfm?sectionid=48&itemid=683

[6] A list of submitted position papers is available at http://www.w3.org/2011/track-privacy/papers/.

[7] Julia Angwin, Web Tool On Firefox To Deter Tracking, THE WALL STREET JOURNAL, January 24, 2011, http://online.wsj.com/article/SB10001424052748704213404576100441609997236.html.

most users,[8] this is a significant development. There is a strong argument that, even without any new legislation or regulation, websites and advertisers could be required to respect a user's request (conveyed through the browser) that they do not want to be tracked.[9]

Accordingly, it is time to define what "tracking" actually means in the context of DNT. Achieving consensus on this question will guide the development and implementation of browser-based DNT tools, serve as the basis for educating users about their options, and guide enforcement bodies, such as the FTC, as they consider the implications of the concept.

Defining "tracking" for the purpose of DNT is not easy. Innovations are difficult to predict and industry practices change quickly. However, clear, flexible guidelines can both empower users and provide clarity for companies. The guidelines presented here are calibrated to promote the efficient development and predictable use of DNT mechanisms.

In this spirit, CDT offers this proposal as a preliminary effort to scope what "tracking" should (and should not) communicate in the context of browser-based DNT mechanisms.

We have drawn on definitions and ideas found in a diverse set of sources, including the FTC's online behavioral advertising self-regulatory guidelines,[10] the Interactive Advertising Bureau's online behavioral advertising self-regulatory guidelines,[11] Rep. Bobby Rush's 2010 consumer privacy bill (the BEST PRACTICES Act),[12] CDT's online advertising threshold analysis,[13] and documents that CDT has produced through its work in technical standards bodies.[14] Since CDT released the first version of this document in February, we have met with a wide range of stakeholders, including industry, advocates, and regulators, in refining this definition.
Finally, while this draft discusses DNT solely in the context of data generated by web-based activities, the implementation of DNT should ultimately not be limited to web-based activities. Global "opt out" mechanisms such as DNT could be extended to other industries as well, such as mobile operating systems. We urge the makers of these sorts of products to explore means to empower users to persistently communicate DNT preferences.

---

[8] Center for Democracy & Technology, *Browser Privacy Features: A Work in Progress*, Version 3.0 (Dec. 2010) available at http://www.cdt.org/files/pdfs/20101209_browser_rpt.pdf.

[9] Section 5 of the FTC Act (and comparable state laws) prohibit deceptive or unfair business practices. If a consumer reasonably expects that a website that responds to DNT-tagged web requests will not track the user, the violation of the user's "terms of use" could well be interpreted as a deceptive or unfair practice.

[10] Federal Trade Commission (Bureau of Consumer Protection), *Self-Regulatory Principles For Online Behavioral Advertising: Behavioral Advertising Tracking*, Targeting & Technology (Feb. 2009), available at http://www.ftc.gov/opa/2009/02/behavad.shtm.

[11] Interactive Advertising Bureau, *Self-Regulatory Principles for Online Behavioral Advertising* (July 2009), available at http://www.iab.net/about_the_iab/recent_press_releases/press_release_archive/press_release/pr-070209.

[12] BEST PRACTICES Act, H.R. 5777, 111th Cong. (2009).

[13] Center for Democracy & Technology, *Threshold Analysis for Online Advertising Practices* (Jan. 2009), *available at* http://www.cdt.org/privacy/20090128threshold.pdf

[14] Alissa Cooper, John B. Morris, and Erica Newland. *Privacy Rulesets: A User-Empowering Approach to Privacy on the Web*. In W3C Workshop on Privacy for Advanced Web APIs, London, UK, July 2010, *available at*ww.w3.org/2010/api-privacy-ws/papers/privacy-ws-12.html.

II. What Should "Do Not Track" Mean?

The user experience online involves the unintentional disclosure and commercial compilation of many different kinds of user data among different entities, comprising a wide range of practices that could be called "tracking." At the most basic level, online communication requires the exchange of IP addresses between two parties. Completion of e-commerce transactions normally involves the sending of credit card numbers and user contact information. Social networking sites often revolve around user-provided profiles. Much web content is supported by advertising and much of this advertising is linked to either the content of the page visited or to a profile about the particular user or computer. Complex business models have arisen around the online data flows.

CDT believes that DNT mechanisms should, at their core, empower users to prevent the collection and correlation of data about Internet activities that occur on different sites. Users expect control over who is tracking them and how tracking data may be shared. To that end, CDT offers the following provisional definition of "tracking":

> **Tracking is the collection and correlation of data about the web-based activities of a particular user, computer, or device across non-commonly branded websites, for any purpose other than specifically excepted third-party ad reporting practices, narrowly scoped fraud prevention, or compliance with law enforcement requests.**

This definition of "tracking" frames the ideas and descriptions in this paper.

We recognize the inevitable difficulty in defining "commonly branded websites" (first parties, under our formulation) in this context. It was once generally presumed that any domain name other than the one from which the user explicitly requested a webpage was a third party. However, sometimes first-party sites now employ separate domains for reasonable design, security, or commercial purposes, and conversely, some third parties provide services from first-party domains. Accordingly, we suggest that two parties (a first and a third) be considered distinct if they do not share "common branding"—a concept that is an approximation for a consumer's reasonable expectations.[15]

While there are certainly considerable privacy concerns associated with the collection and usage of consumer data by first parties, the idea of "Do Not Track" was originally conceived as a means to prevent the correlation of information across multiple, unrelated websites by third-party advertising companies.[16] As such, we do not believe that first parties, with whom an ordinary consumer is more likely to understand that they are sharing information, should be required to adhere to a global DNT instruction. However, as we have long advocated,[17] we

---

[15] For a more in-depth discussion of the distinction between first and third parties, *see* Alissa Cooper and Hannes Tschofenig, Overview of Universal Opt-Out Mechanisms for Web Tracking, http://tools.ietf.org/html/draft-cooper-web-tracking-opt-outs-00#section-2.

[16] See Alissa Cooper, Do Not Track. No, Seriously., CDT Blog, November 8, 2007, http://cdt.org/blogs/alissa-cooper/do-not-track-no-seriously.

[17] See Testimony of Leslie Harris, President and Chief Executive Officer of the Center for Democracy and Technology, before the House Committee on Energy and Commerce, Subcommittee on Commerce, Trade, and

continue to believe that first parties should offer consumers the right to opt out of secondary uses of their data on an individualized basis.

CDT does not believe that DNT is intended to enable users to block advertising or prevent all data collection. CDT also believes that the collection and use of "actively shared" data–data that users knowingly and voluntarily provide in web forums, on social networking profiles, or on blogs or microblogs–is out of scope for DNT. Similarly, DNT does not resolve important questions about how to ensure appropriate protection for sensitive information (such as health or financial information) that can be discerned from our web-based activities.[18] While the consolidation and unexpected uses of such data can raise serious privacy concerns, CDT recommends that the collection and use of this data be addressed in other ways, such as through comprehensive consumer privacy legislation.

CDT recommends that DNT be understood to address the collection and use of transactional, or "passively shared" data. Users do not typically expect that records are being collected of the various sites and pages they visit across the web, particularly because such collection is often performed by companies that are not consumer-facing. While a user might reasonably expect that individual websites can track her across that website, many users do not expect or want companies or their industry partners to be able to track what they browse and read across multiple, unrelated sites.[19] It is this concern that should be the focus of DNT.

CDT also believes that a user's decision to enable a DNT mechanism can be overridden upon a grant of affirmative and specific permission from the user in response to a clear and conspicuous[20] prompt from a service. For example, a news site could present a dialog box and request that the user grant permission to a certain ad network to track her on that site as a condition of service. Or a photo-sharing site could offer users a choice during registration to host limited amounts of data for free or to host more if the users allow certain third-party tracking. As long as the request for permission is clear and prominent and a user is given the opportunity to make an informed choice about the value proposition, companies should be able to obtain a user's affirmative permission to override the generic DNT instruction.

In short, when the user has chosen to use a DNT mechanism, all future tracking (as defined above) becomes opt-in.

A summary of our preliminary recommendations for what should and should not be considered "tracking" activities for the purposes of DNT can be found in the chart below. These recommendations elaborate on our core definition, above.

Consumer Protection on "The BEST PRACTICES Act of 2010 and Other Federal Privacy Legislation" (July 22, 2010) available at www.cdt.org/files/pdfs/CDT_privacy_bill_testimony.pdf.

[18] *Id.* CDT has long advocated that the collection and use of sensitive information by third parties require opt-in consent.

[19] Aleecia M. McDonald, User Perceptions of Online Advertising, Yale ISP Conference, March 25-26, 2011, *available at* http://www.law.yale.edu/documents/pdf/ISP/Aleecia_McDonald.pdf; *See* Joseph Turrow, Jennifer King, Chris Jay Hoofnagle, Amy Bleakley & Michael Hennessey, *Contrary to What Marketers Say, Americans Reject Tailored Advertising and Three Activities that Enable It* (Sept. 2009), *available at* http://graphics8.nytimes.com/packages/pdf/business/20090929-Tailored_Advertising.pdf.

[20] FTC Privacy 2010.

Tracking	Not Tracking
Third-party online behavioral advertising	Third-party ad and content delivery
Third-party behavioral data collection for first party uses	Third-party analytics
Third-party behavioral data for other uses	Third-party contextual advertising
Behavioral data collected by first parties and transferred to third parties in identifiable form	First-party data collection and first-party use
Demographic information appended to the user's device	Federated identity transaction data
	Specifically excepted third-party ad reporting
	Data collection required by law and for legitimate fraud prevention purposes

A. What is "Tracking"?

CDT proposes that the following activities are illustrative of "tracking":

- Third-party online behavioral advertising
- Third-party behavioral data collection for first party uses
- Third-party behavioral data collection for other uses
- Behavioral data collected by first parties and transferred to third parties in identifiable[21] form
- Demographic information appended to the user's device

We describe these activities below.

1. Third-party online behavioral advertising

Today, many websites (commonly described as "first-party sites" or publishers) contract out both advertising and content to third party advertising and content syndication networks. These networks have the ability to place a unique identifier on a user's computer, which the network can subsequently recognize as the user moves from site to site. Using this identifier, the network can amass a profile of a range of sites visited by the user. For the purposes of this document, third-party online behavioral advertising means the collection of data about a particular user, computer, or device, regarding activities across non-commonly branded websites for the purpose of using such data to predict user preferences or interests and to deliver advertising to that individual or her computer or device based on the preferences or interests inferred from such web-based activities.

Under this definition, the following would be considered third-party online behavioral advertising:

[21] Data is in identifiable form if it can be reasonably linked to a specific consumer, computer, or other device. This is the definition provided by the FTC. *See e.g., supra* note 3.

Example 1: An advertising network contracts with a number of websites to place web bugs (also known as tracking pixels) on these websites in order to place HTML cookies on the devices of visitors to these websites. These cookies allow the advertising network to collect information about some of the websites the user visits in order to compile a profile to associate with that user's device. The advertising network uses this information to target advertisements to the user. Alternatively, the advertising network sells this information, along with an IP address, unique device ID, or other potentially identifying information to another advertising network that uses this information to target advertisements to the user.

Example 2: A social networking site employs iframes[22] on a wide range of non-commonly branded websites to embed content customized for the user. In order to display the customized content on each website the user visits, the social network must receive the URL of each of these websites. The social network collects these URLs and uses them to target advertisements to the user. Alternatively, the social network collects these URLs and sells them to – or shares them with – an advertising network, along with user data that can be reasonably linked to a specific user, computer, or other device. The advertising network uses this information to target advertisements to the user.

Example 3: A user visits the website of a furniture retailer and examines the specifications for a particular sofa. Alternatively, the user places this sofa in her online shopping cart but does not proceed to purchase the sofa. As the user navigates around the website, the furniture retailer permits a number of third-party advertising firms to place tracking cookies in the browser of the user. The fact that the user nearly purchased the sofa is used to target advertisements for the sofa on websites that are not commonly branded with the furniture retailer or on advertisements that show up when the user opens the web browser on her mobile device (This is one form of a practice commonly known as "re-targeting.").

2. Third-party behavioral data collection for first-party uses

This practice refers to the collection of data about a particular user, computer, or device regarding activities across non-commonly branded websites, by a particular company for the purpose of using such data to advertise to or customize the products or services that the said company (or a commonly branded site) provides to the user.

Under this definition, the following would be considered third-party behavioral data collection for first-party uses:

Example 1: An online portal website contracts with other websites to place web bugs on these websites and to place HTML cookies or other unique identifiers on the devices of visitors to these websites. These cookies allow the company to collect information about some of the websites the user visits. The company uses this information to customize the content on the online portal it provides for the user, the search results it presents when the user uses its search engine, or the advertisements it shows alongside those search results.

[22] Using an iframe tag, a website can display an html document that is hosted on a third-party website. In essence, this first-party website grants the third-party website an "embassy" on its page. *See* http://en.wikipedia.org/wiki/HTML_element#Frames.

Example 2: A social platform uses a cookie tied to a user's profile on the platform to render social "widgets" that display personalized content on websites outside the platform's domain. Rendering these widgets requires that the platform receive both the user's unique identifier and the address of the webpage user is visiting. The platform uses the information about the third party domains visited by the user to serve advertisements on its own website.

3. Third-party behavioral data collection for other uses

This practice refers to the collection of data generated by or derived from a particular computer or device regarding activities across non-commonly branded websites, by a particular entity. Collection for other uses may include offline marketing or market research based on aggregated tracking of a population of users. While aggregate market research may raise fewer privacy risks than individualized targeting and profiling, many users may object to the tracking of their web activities for research purposes, and persons who use a browser-based DNT mechanism would reasonably expect to be opting out of such tracking.

Under this definition, the following would be considered third-party behavioral data collection for other uses:

Example 1: A company contracts with popular websites to place web bugs on these websites and to use "browser fingerprinting"[23] to uniquely identify of visitors to these websites. This allows the company to collect information about some of the websites the user visits and associate that information with an identifier linked to the user's device. The company then sells this behavioral data.

Example 2: A company contracts with a range of websites to place web bugs on websites and uses HTML5 web storage[24] store unique identifiers onto users' devices. This allows the company to collect information about some of the websites the users visit and associate that information with identifiers linked to those users' devices. The company eventually aggregates this data into a market research report detailing how a large population of web users surf the web.

Example 3: A company provides authentication services for publishers that run commenting systems. When the company authenticates a user, it retains a copy of the URL of the site that the user was logging in to. The company combines information from across sites to create a record of the sites at which it has authenticated the user. The company uses this data for market research purposes.

---

[23] Peter Eckersley, *How Unique is Your Web Browser?*, Electronic Frontier Foundation, https://panopticlick.eff.org/browser-uniqueness.pdf.

[24] Ian Hickson, ed., *Web Storage*, W3C Editor's Draft, http://dev.w3.org/html5/webstorage/.

4. *Behavioral data collected by first parties and transferred to third parties in identifiable form*

This category covers the collection of data generated by or derived from a particular user, computer, or device regarding activities on one website or across commonly-branded websites, by a particular company. This company then transfers that data to a non-commonly branded company, for uses other than the rendering of the service for which the user provided the data, in a form such that the data can be reasonably linked to a specific user, computer, or other device.

Under this definition, the following would be considered behavioral data collected by first parties and transferred to third parties in identifiable form:

Example 1: A large e-retailer collects transactional data from users as they peruse its website and commonly-branded websites. The company runs an algorithm to determine which IP addresses it receives are "static" — that is, they persistently identify return users to that site. The company then associates passive web browsing activity of its site with the static IP addresses of those users. The company then sells this data to another company without aggregating the information. The second company then uses this data to deliver targeted advertisements or content to devices using those IP addresses on other websites.

Example 2: A user visits the website of a furniture retailer and sets up an account with the retailer, providing her email address: example@example.com. She then examines the specifications for a particular sofa or perhaps places this sofa in her online shopping cart but does not proceed to purchase the sofa. The furniture retailer sells to an ad network or data aggregator the fact that the person with email address example@example.com is interested in this particular sofa. (This is another form of "re-targeting.").

5. *Demographic information appended to the user's device*

This category covers instances when a vendor might place on the user's device a cookie with personal or demographic information that is designed to be read by one or more ad networks (or other entities) as the user browses the web. This information essentially tags the user as matching certain demographic characteristics as she browses the web. We recognize that this category of activities does not fall squarely under the definition of "tracking" provided above. However, based on considerable input in the development of this revised definition, because a user would certainly not expect that as they traverse the web they are broadcasting demographic or personal information, CDT recommends that the placement of these cookies be considered tracking.

Example: A user fills out a web form news site, providing her email address and city of residence. A third-party network also receives that information, and queries a data broker's database to determine the user's identity and associated demographic information. A third-party cookie is then placed on her device, identifying her as a female age 35-40 who lives in Washington DC. This cookie is read by web bugs on sites across the web, allowing other

publishers and networks to customize advertisements or content for a female user who is 35-40 years old and lives in Washington DC.

B. What is *Not* "Tracking"?

DNT should not be conceived as a blanket prohibition against collection or use of user data; nor should it be construed as a prohibition against all third-party advertising. As discussed above, CDT believes that "actively shared" data—such as data users provide on social networking profiles, web forums,[25] and through registering for various accounts—is largely out of scope, even though the use of this data raises a separate set of privacy concerns. Instead, CDT recommends that DNT be implemented to focus on the collection of the transactional, or "passively shared," data that is created as users navigate the web.

A number of difficult distinctions remain. For example, web analytics and product improvement services may place unique cookies on users' devices and therefore seem to employ passive tracking. However these products can be designed so as to "silo" data and avoid tracking "across sites" (and instead just measure the number of times a unique user visits commonly-branded site(s)). In these cases, the practice falls outside of our core definition and raises fewer privacy concerns. Accordingly, DNT may not cover certain practices even though they involve employing a unique identifier.

In these cases especially, it will be difficult for users to determine whether or not companies are respecting their DNT preferences.[26] Thus, we suggest that both the entity collecting data and the entity on whose website data is collected (if this entity has contracted with a data collector) should be required to provide in their privacy policy a clear, affirmative, accountable statement describing the purpose of their data collection and specifically disclaiming any "tracking" behavior.[27]

CDT proposes that the following activities are illustrative of activities that are "not tracking" for the purposes of DNT:

- Third-party ad and content delivery
- Third-party analytics
- Third-party contextual advertising
- First-party data collection and first-party use
- Federated identity transaction data

---

[25] While we believe this is outside the scope of DNT, actively shared data can still be collected in ways that pose significant privacy risks. *See* Julia Angwin and Steve Stecklow, *'Scrapers' Dig Deep for Data on Web*, THE WALL STREET JOURNAL, October 12, 2010, http://online.wsj.com/article/SB10001424052748703358504575544381288117888.html

[26] For example, with the "do not track" HTTP header, consumers have to trust that the parties they are interacting with are respecting the header's direction. With the "block list" approach, consumers have to trust that sites not on the block list (or on a "white list" or "allow list") are not actually tracking them.

[27] We further recognize that an audit requirement, perhaps in a baseline consumer privacy law or related rulemakings, may ultimately be necessary in some situations – such as for those companies that drop unique identifiers – to ensure that tracking preferences are being honored.

- Specially excepted third-party ad reporting
- Data collection required by law and for legitimate fraud prevention purposes

1. Third-party ad and content delivery

Most modern websites import content from other domains when rendering a page for a visitor. This content could be a widget displaying the weather or stock prices, or it could be advertising optimized and delivered by a third party. Even if a user has enabled a DNT mechanism, third parties should be allowed to deliver content and advertisements on first party sites so long as the third parties do not engage in the behavior described above as "tracking."

Under this definition, the following would be considered third-party ad and content delivery:

Example: The front page of a sports blog contains an iframe displaying scores from a partner (but third-party) site, and a banner ad delivered by a third-party ad network. The banner ad shows advertisements that are targeted to the content of the page and not any particular visitor. Both third parties need the IP address and other basic information about the device requesting the content so that the ad can be delivered to the user. However, neither third party correlates the information received about the device with any other information about non-commonly branded websites that the user visits.

Example 2: A social platform uses a cookie, tied to a user's profile on the platform, to render social "widgets" that display personalized content on websites outside the platform's domain. Rendering these widgets requires that the social network receive both the user's unique identifier and the address of the webpage the user is visiting. The social network renders the widget and then immediately disassociates the URL of the page the user is visiting from other information about the user, such as the user's name, demographic data and the URLs of other sites visited by the user.

2. Contextual advertising

Contextual advertising refers to the delivery of advertising based on the content of a webpage, a search query, or a user's contemporaneous behavior on a website without regard to activities of the user or her computer or device on non-commonly branded websites.

Under this definition, the following would be considered contextual advertising:

Example 1: A user is reading an article on the web about new smart phones and an advertisement for a smart phone is displayed alongside the article. The decision to display the advertisement was not influenced by the user's activity on other, non-commonly-branded websites.

Example 2: A user initiates a search engine query for "movie theaters in Washington, DC." Alongside the search results, advertisements are shown for an Academy Award-nominated movie and restaurants that are located near a popular movie theater in Washington, DC.

Example 3: A user in Washington, DC uses a search engine to search for "movie theaters." The search engine is able to guess, from the users IP address, that the user is located in Washington, DC and displays content and advertising that is tailored to the DC area.

3. First-party data collection and first-party use

Generally speaking, first-party data collection and use should not be considered "tracking" for the purposes of DNT. First-party data collection refers to the collection of data generated by or derived from a particular computer or device regarding activities across commonly-branded website or websites. Sometimes, this data may be used, by the first party for the purposes of delivering first-party behavioral advertisements or otherwise customizing content on a website that is under the first party's common branding.

However, when the first party sells user data that can be reasonably linked to a specific user, computer, or other device, this activity does not fall under the category of "first-party data collection and first-party use." Instead it falls under the category of "behavioral data collected by first parties and transferred to third parties in identifiable form," which we classify as a tracking activity. Similarly, when the first party combines data acquired through a non-tracking activity with data obtained through a tracking activity, any use of that combined data becomes a "tracking activity."

Under this definition, the following would be considered first-party data collection and first-party use:

Example 1: An e-retailer recognizes return visitors using cookies or account logins. The e-retailer uses only past purchases the visitor has made on its website, past webpages she visited within its website, log data,[28] and gender information the visitor provided in her user profile to customize the content displayed to the visitor on the sites, homepage, or the advertisements the visitor sees as she traverses the website.

Example 2: A company offers a search engine and a social networking site; the services are commonly branded. The search engine customizes search results and the advertisements adjacent to the search results based on the user's social networking profile, location, past uses of the search engine, and – of course – the search terms themselves. The search engine does not use data from third parties to customize the results or ads, nor does it use data collected by tracking the user on non-commonly-branded websites.

[28] Log data includes data such as: IP address and port number; browser type, version, and operating system; screen size; technologies, fonts, and audio formats supported by the browser; URL of the page that directed the visitor to the site (the referrer); whether the visitor has bookmarked the website on the web browser; the webpages within the site that the visitor visited, the webpage the visitor visited first on the site (the entry page), and the webpage the visitor visited last on the site (the exit page); bandwidth used; the amount of time the visitor spent during a visit to the site; the time and date of the site visit.

4. Federated identity transaction data

Websites are increasingly outsourcing user registration and authentication processes to third-party identity providers. These identity providers have a unique vantage point from which to passively log users' registration and authentication activities over time and across an array of contexts.[29] Use of this data by an identity or authentication provider should be limited to statistical reporting to a relying party (here, the website) in connection with the activity on that website. As with third-party analytics, as long as the data collected by the third-party are not aggregated or merged across non-commonly branded sites and domains, that activity should not be considered tracking.

Under this definition, the following would be considered a federated identity transaction:

Example: A popular web portal offers to its registered users the ability to use their portal account to authenticate with a wide range of other, third-party sites. The portal's authentication service generates data regarding the sites its users have accessed. However, this data is not correlated or used to create a profile about the user. The data remains siloed for the purposes of statistical reporting to individual third-party sites.

5. Specifically excepted third-party ad reporting

Third-party ad reporting refers to the logging of ad views by a third party for the purposes of identifying when a user interacts with a particular advertising campaign, limiting the number of times a particular ad is shown to a particular user, and preventing click fraud. All of these activities should be narrowly tailored to that which is reasonably necessary to accomplish the specific ad reporting purpose.

In order to optimize third-party ad and content delivery, a third-party company may place a unique identifier on a user's device in order to record data about the user's engagement with the third-party. As long as this unique identifier is only used to collect information about the user's views or interactions with the advertisements (and not the first-party content), and about the user's interactions with the entity sponsoring the advertisement (and no other entities), this would not be considered a tracking activity.[30]

Click-fraud monitoring, which involves data collection in order to prevent misrepresentation of the number of interactions with an ad, should be conducted in a carefully, tailored manner that minimizes the amount of data collected. CDT would welcome a discussion about the innovations necessary to prevent click-fraud in a minimally invasive fashion.

---

[29] The availability of this data to the identity provider is not inevitable; it depends upon the implementation of the underlying authentication protocols.

[30] Ad impressions can be tracked using methods other than cookies with unique identifiers. Some could argue that DNT should never allow third-party cookies with unique identifiers to be placed on users' computers. CDT is interested in exploring ways to prevent fraud and to track ad delivery and impressions without using third-party cookies with unique identifiers.

The following would be considered specifically excepted third-party ad reporting:

Example 1: A news website contracts with a third-party ad network to deliver non-behaviorally targeted ads to site visitors. The ad network places a unique HTML cookie on visitors' computers in order to count unique views and in order to ensure that visitors do not see the same ad over and over. As long as the data collected about the user is exclusively about the advertisements themselves and is not tied to the first-party site that the user visited (does not include, for example, the URL of the site), this activity would not be considered "tracking."

Example 2: An advertising network retains for 90 days the logs of the IP addresses of all users who have clicked on a particular ad for click fraud prevention purposes. The advertising network does not combine the IP addresses with information about individual site visitors (or the websites they have visited) in order to create profiles about these visitors. The advertising network also does not sell or otherwise transfer these IP addresses.

6. Data collection required by law and for legitimate fraud prevention purposes

The collection and retention of data for fraud prevention purposes (e.g., services that help prevent financial or identity fraud) or for purposes required by law should not be considered tracking, provided that collection and retention are narrowly tailored to that which is reasonably necessary to accomplish these specific purposes. Click-fraud related activities fall under the category of "specifically excepted third-party ad reporting."

Under this definition, the following would be considered data collection for legitimate fraud prevention purposes:

Example: A bank uses browser fingerprinting methods to determine whether users are logging in from the same computer from which they usually log in. Those who are logging in from a different computer are asked to use two-factor authentication.

III. Conclusion

With browser developers unveiling DNT mechanisms, it is now essential to seek consensus as to what an affirmative consumer statement—"do not track me"—actually means. This consensus will aid consumers in understanding what to expect from DNT and will assist companies that are working to implement it. This proposal is a contribution to this discussion.

Although this proposal focused specifically on web interactions, DNT should be considered in other contexts. For example, there is a strong argument that consumers would benefit if the mobile ecosystem adopted similar mechanisms.

It is also important to emphasize that DNT does not "fix" privacy. DNT is a bundle of privacy-enhancing technologies and policies that can help return some control to users with respect to certain types of tracking behaviors. A baseline, comprehensive privacy bill that provides

substantive protections for users is still necessary to fully protect users, promote trust in the online environment, and position the U.S. as a global leader as other countries rapidly work to update their own privacy regimes.

CDT will continue consulting with relevant stakeholders as we work toward making DNT a reality. We welcome any comments, questions, or concerns.

For further information, contact:
Justin Brookman
Director, Consumer Privacy Project
202-637-9800
justin@cdt.org