Tracking Transparency

Wendy Seltzer*
Berkman Center for Internet & Society at Harvard University
and Princeton Center for Information Technology Policy
wendy@seltzer.org

March 31, 2011

The Fair Information Practice Principles[1] model of privacy protection, influential in U.S. privacy law, depends upon transparency. Transparency is key to adequate notice, which in turn is necessary to meaningful choice, access, and enforcement. Thus transparency to end-users is a critical component of any do-not-track mechanism. I compare the transparency of server-side header response and client-side request-blocking and suggest that the latter is more directly transparent in its operation.

Tracking itself poses transparency challenges: trackers know more about their prey than is apparent to the typical Internet user. Much tracking happens through back-end correlation, building up server-side profiles outside the view of the user; even what is sent to the browser is often hidden under browser rendering (or non-rendering, in the case of third-party cookies and transparent or zero-pixel images); and users are at information disadvantage to their trackers. Moreover, wariness of the "creepiness" factor, along with simple scale economies, may cause trackers to tune the profiling less finely than their data would make possible. Users are therefore rarely exposed to a direct mirroring of all data collected from them, or the full customization that profiling would make possible. Even as their experience is being customized, users generally have no window into others' experiences for comparison. While advertisers speak publicly of discounts to good customers, the customers worry about price discrimination that charges more to those with demonstrated willingness to pay.

Technological tracking protection measures can give end-users greater control of their privacy choices – but their effectiveness will depend to a great extent on the notice they give users of what is being defended against, and how. While the current Firefox 4 and IE 9 both employ a new HTTP header, DNT, they implement it differently, changing the level of user visibility and control.

In Mozilla's Firefox 4, users navigating to the Advanced, General tab are shown the option "Tell websites I do not want to be tracked." As the online help indicates, this option relies entirely on the recipient for

*Affiliations listed for identification purposes only. Comments reflect personal position, not that of any institution.
[1]FTC Fair Information Practice Principles http://www.ftc.gov/reports/privacy3/fairinfo.shtm.

implementation: "Checking this box will tell websites that you wish to opt-out of tracking by advertisers and other third-parties. Honoring this setting is voluntary – individual websites are not required to respect it."[2] Before this setting can have any impact, websites and their tracking partners must learn to recognize the DNT header and determine a policy response to it.

The DNT policy decisions start with specifying what "tracking" means: Is it following a user across multiple sites or multiple sessions, or does it include watching repeat intra-session visits to the same website?[3] Is it correlating browsing behavior with personally identifying information gained from user input or environment? Could a website assert that it was not "tracking" if it merely collected the information but didn't *use* it? Unless all sites respond using the same definitions, the header will operate differently from site to site.

Thus although the DNT header in Firefox's user-option is relatively simple to exercise, its operation is not very transparent. A user trying to determine the real impact of that toggle would have to visit all of the websites of concern, seek out privacy policies, hope those policies also bound third-party trackers or watch page-loads and requests to ferret out all these "partners" and seek out their policies as well. Even after this sleuthing, the user would still have to rely on the sites to describe and abide by their stated policies, or depend on enforcement authorities such as the U.S. Federal Trade Commission to police against "unfair or deceptive acts or practices." Users will get little or no feedback to let them know whether their do-not-track preferences are being honored on a site-by-site basis.

Microsoft's IE 9 instead offers users a configurable browser-side block-list to control what gets sent to online trackers: "Tracking Protection Lists are like 'Do Not Call' lists for third-party content on a website. By adding a Tracking Protection List or 'TPL,' you can control whether your information is sent to third parties listed on the TPL."[4]

If the user configures this list by choosing from among the third-party TPLs Microsoft links to, he or she can review the contents of the list. These require some investigation: in particular, if the user installs the current TRUSTe list, he should note that "TRUSTe's TRUSTed Tracking Protection List *enables* relevant and targeted ads from companies that demonstrate respectful consumer privacy practices and comply with TRUSTes high standards and direct oversight." (emphasis added) [5] In fact, the current easy.tpl *allows* tracking from 1570 domains, among them freecreditscoreonus.com, cashadvance.com, LipitorLink.com, makingsenseofpainrelief.org, and acxiomdigital.com while blocking only 23 domains. TRUSTE's own site promotes this list "to block companies that offer poor privacy protection, while ensuring that trustworthy companies who

[2] Firefox Help, `http://support.mozilla.com/en-US/kb/Optionswindow-Advancedpanel`

[3] In the age of suspend, how many users keep single sessions open for weeks?

[4] Microsoft, Tracking Protection Lists, `http://ie.microsoft.com/testdrive/Browser/TrackingProtectionLists/faq.html`

[5] `http://ie.microsoft.com/testdrive/Browser/TrackingProtectionLists/`

protect their privacy can continue to provide them with a richer, more personalized browsing experience."[6] Following the link to TRUSTe's "Third Party Data Collection Certification Program Principles," however, one finds that all the TPL buys is protection against the linking of personally identifying information and the option to read another privacy policy and opt out *again* if one wants not to be tracked on any of these 1570 domains.[7] Other lists Microsoft links, from Abine, Easy List, and Privacy Choice, appear to be more straightforward block-lists, but as the spec is designed, if a user chooses multiple lists, the "ALLOW" from any list takes precedence over "BLOCK."

While some of its current implementations may be surprising, however, the operation of the Microsoft-implemented tracking-protection is transparent to the end-user in two ways – its lists are user-side, and some of their effects are directly visible in the browser: a blocked ad does not show up, often leaving a blank space. This mechanism can be bolstered by self-regulatory programs or regulatory policing, but it does not depend on them.

Users can learn to protect privacy – and to determine which aspects of tracking are too invasive – if they get feedback on how their choices change their experiences. If choice is to be meaningful, and if users are expected to understand tracking as part of a bargain for online content, we need to assure that they have a real-time view of what they are exchanging.

I look forward to participating in the W3C workshop to explore how tracking protection technologies can help give end-users visibility into and control of their online experiences.

[6] `http://tracking-protection.truste.com/`

[7] Cashadvance.com, for example, discloses in its linked privacy policy that "All cookies served on the domain, both session and persistent are tied to the Personally Identifiable Information you provide," and that "This privacy statement applies solely to information collected by Company even in cases where we may frame another site with our own." `http://www.cashadvance.com/privacy-policy` It proudly displays a TRUSTe seal and appears with a "+d [ALLOW]" on TRUSTe's TPL.