

## What Base Text to Use for the Do Not Track Compliance Specification

The Tracking Protection Working Group was chartered “to improve user privacy and user control by defining mechanisms for expressing user preferences around Web tracking and for blocking or allowing Web tracking elements.” The group reached an important series of decision points this summer, when participants submitted change proposals after the May face-to-face meeting. The co-chairs here offer their determination for the group on one change proposal that presents a fork in the road, after the Digital Advertising Alliance and related participants asked to move to a new draft text. After consideration, the chairs have determined that the group has rejected that change proposal, finding it at odds with our chartered aims and the weight of group consensus.

The question before the group is whether to change its base text for the continued work on the Compliance Specification, to adopt the version proposed by the DAA or to continue addressing issues against the text proposed to the group in June. We conclude, based on the comments submitted, that the June Draft provides a better basis from which to address the criteria for a W3C standard, as understood in the Working Group, than does the DAA Proposal. We thus will continue to use the June Draft as the base text and work through the remaining issues raised. We will not revisit the choices presented in the DAA change proposal and rejected in this decision.

### History

The Working Group was formed in September, 2011. Currently, Matthias Schunter is chair for the Tracking Protection Expression (“TPE”) work, to define the technical mechanisms for expressing a DNT preference. Peter Swire is chair for Tracking Compliance and Scope work (the “Compliance Specification”), to define the meaning of a DNT preference and set out practices for Web sites to comply with this preference. The Working Group currently has 110 participants, representing consumer groups, browsers, first parties, third parties, regulators, subject matter experts, and others. Since its formation, the group has held weekly teleconferences, met face-to-face once a quarter, and exchanged thousands of emails on its mailing list, with all proceedings public at <http://www.w3.org/2011/tracking-protection/>.

Members of the group were asked to submit their comments by July 12 on ISSUE-215, the choice of base text. The first option is to adopt the proposal by the DAA. The DAA is an umbrella organization of online advertising organizations, and its members conduct a large fraction of online advertising. The second option is continuation with the “June Draft,” prepared by co-chair Swire and W3C staff in June, and accepted by the Compliance Specification editors as their draft. Whichever choice is made, the group has submitted over 20 change proposals, many of which will remain open questions for group discussion after today’s decision.

Similarly, whichever choice is made, the precise language of provisions would be subject to polishing and editing.

The question to the group emphasized that the two texts varied importantly on four issues, with the decision today indicating the group's subsequent direction on those issues, plus a topic that differs based on the logical implications of the four issues:

1. Issue 5 – the definition of “tracking.” The DAA text is narrower in what is covered.
2. Issue 16 -- definitions of collecting, retaining, using, and sharing data. The DAA text is narrower in what is covered.
3. Issue 188 – definition of de-identified data. The DAA text would treat data as “de-identified” in situations where the June text would not.
4. Issue 199 – limitations on the use of unique identifiers. The June text would prohibit the use of unique identifiers where alternatives are reasonably available, thus limiting collection of user data in those circumstances.
5. The effects of user choice. Under the June text, the Do Not Track mechanism would opt the user out of its broader definition of tracking. Under the DAA proposal, targeting of advertisements would not be affected by the Do Not Track standard; instead, users would use the separate DAA opt-out mechanism if they wished to limit targeted advertising.

The provisions of the DAA Proposal grow out of months of intense engagement of DAA and its members with the chairs and other members of the Working Group. The goal has been to discover or create an approach that could gain consensus support among the diverse stakeholders in the group. This is the third time this year that the group has considered an initiative from the DAA, following discussions in the February meeting in Boston and the May meeting in California. Since June, a large portion of the group's time has been devoted to clarifying the meaning of the DAA Proposal, providing a detailed record for today's decision. Based on the comments received, the current DAA Proposal is less protective of privacy and user choice than the earlier initiatives.

### **Do Not Track, Do Not Collect, and Data Hygiene**

Based on the comments submitted, the co-chairs conclude that the June Draft will remain the base text. Along with this decision, we are releasing an Explanatory Memorandum that analyzes in detail the 27 comments submitted in response to the call for objections by July 12, and the record of the group's previous work. The Explanatory Memorandum analyzes three main topics, called Do Not Target, Do Not Collect, and Data Hygiene. For each topic, the June Draft is more likely to lead to a standard that reaches the group's objectives than the DAA Proposal under the announced criteria for the group's work.

Do Not Target. Many public discussions about the Do Not Track standard have stated the main disagreement as “Do Not Target vs. Do Not Collect.” The DAA Proposal avoids this disagreement. It does not address Do Not Target, as that term has been generally understood, and it also does not include Do Not Collect. These twin omissions go far toward explaining the clear opposition to the DAA Proposal from essentially all the formal comments, except for those associated with the DAA itself.

The idea of Do Not Target is that the user should have some choice about seeing ads targeted to that user based on the user’s browsing history. For instance, a user who reads about a Hawaii golf vacation may then see ads on other web sites for Hawaii golf vacations (often called “retargeting”) and may see ads targeted for the sorts of people who typically go on such golf vacations (often called “profiling”). Currently, the DAA operates a self-regulatory program called Ad Choices, offering the users the ability to opt out of the sorts of targeting covered by that program. In February, 2012, the DAA announced in connection with a White House event that it would, within nine months, “add browser-based header signals to the set of tools by which consumers can express their preferences under the DAA Principles.” The “browser-based header signals” are the DNT signals being developed by the Working Group. Based on the group’s discussion of the current DAA Proposal, however, a user who turns on Do Not Track would have no effect on this sort of retargeting or profiling. No browser-based header signal would affect which ads a user received. The DAA Proposal thus does not meet the widely-understood meaning of Do Not Target.

Do Not Collect. Many participants in the process, including the Federal Trade Commission, have said that DNT should go beyond simply a limit on targeting. On this view, a reduction in “tracking” should mean a reduction in the information collected about an individual. As a corollary, collection that takes place, such as to serve a web page, would be minimized and retained only so long as necessary for a permitted use.

Consumer groups and others in the Working Group have emphasized a closely related idea, that advertising and other web activities should take place, so far as reasonably possible, without use of unique identifiers. For instance, targeted advertising on the Internet today often occurs through use of a unique cookie, on each computer or other device. These participants have emphasized the feasibility of other approaches to serving ads based on user interest, and the importance of shifting the online advertising eco-system toward next generation practices that combine advertising revenues with user privacy and choice.

The June Draft provides that third parties “must not rely on unique identifiers for users or devices if alternative solutions are reasonably available.” The DAA Proposal deletes reference in the standard to alternative, privacy-protecting solutions if they are reasonably available.

The DAA Proposal thus does not address the often-expressed goal of Do Not Collect. Indeed, one representative of the DAA informed the group this month that limits on collection would be a “non-starter.”

Data hygiene. In the group’s discussions, the term “data hygiene” refers to a range of controls that a company may apply to de-identify data and reduce the risk that data is revealed without authorization. Information about users is collected for a variety of legitimate reasons; once that data is collected, however, a broad range of stakeholders agrees that the holders of data should exercise good hygiene, preventing data spills and generally treating users’ data securely and appropriately.

The topic of de-identification has been a major focus of the group’s work in 2013. An important reason for this focus has been a consensus in the group that data at some point is scrubbed enough so that data becomes out of scope of the DNT specification. Put another way, data at some point is de-linked or de-identified enough that use of it does not count as “tracking.” In drawing the line between in-scope and out-of-scope, a major privacy concern is that data might be re-identified. Therefore, more scrubbing creates less risk for individuals. A major industry concern has been to retain the utility of data, and more scrubbing often means lower usefulness for industry.

The DAA Proposal calls for a combination of technical, administrative, and operational controls to move tracking data as originally collected (“raw” or “red” data) to what it calls a de-identified state (“yellow” data). It then defines additional technical measures that would move the data to a de-linked state (“green” data). The exact meaning of the text, and its interpretation, remain subject to considerable uncertainty, based on the official comments and discussion on the public mailing list.

With respect to data hygiene, commenters highlighted at least two concerns about the DAA Proposal. First, the DAA Proposal, apparently consistent with the term’s use in the DAA’s self-regulatory code, uses the term “de-identified” at a stage where a company retains the information needed to re-identify an individual. This threshold for “de-identification” is less strict than usage of the term in other privacy regimes, such as the U.S. medical privacy rule or in usage in the European Union.

Second, the scope of the DAA Proposal’s data hygiene provisions is narrower than under the June Draft, and they may not be effective in practice. For instance, the DAA would allow “aggregate scoring” outside of the data hygiene provisions. Although there is some uncertainty about precisely how the DAA would define aggregate scoring, the basic idea is that a company could keep and update a user profile linked to a unique identifier, so long as the specific URLs seen by the user were discarded. For instance, a DAA representative gave the example that “cookie ID 1234 could have an interest score of 4 in off-road vehicles and an interest score of 14 for flower purchase intender.”. This cookie and associated information would be outside of the scope of the DAA’s data hygiene provisions. A number of

comments raised credible technical objections about the quality of protection against re-identification provided by the DAA proposal, both for aggregate scoring and more generally.

The chairs recognize the considerable work that members of the DAA have devoted to crafting data hygiene proposals. Indeed, the DAA Proposal includes new details for how DAA members could address the market research and product development exceptions to the current DAA Principles. The data hygiene proposals state that market research and product development would occur in the intermediate “yellow” state. In contrast, the current DAA Principles permit those activities to use raw or “red” data. In this way, the DAA could consider changing its own self-regulatory program to fulfill its previous statements that it would tighten its market research and product development exceptions.

This improvement in treatment of those exceptions, however, does not constitute a general regime of data hygiene about the data that consumers would reasonably expect to be included in the scope of Do Not Track. The comments support the conclusion that the June Draft provides a better platform for building an overall approach to data hygiene.

### **Criteria for a Do Not Track Standard**

Based on this record, the co-chairs conclude that the June Draft satisfies the stated criteria for the Compliance Specification better than the DAA Proposal. Swire proposed these criteria in the Boston meeting, and they have been widely cited since by group members. As explained in more detail in the Explanatory Memorandum, three criteria were that the Tracking Protection Compliance Specification should be:

1. Consistent with the group’s charter. The charter says that a standard should define “mechanisms for expressing user preferences around Web tracking and for blocking or allowing Web tracking elements.” The DAA Proposal does not use the DNT signal to address either Do Not Target or Do Not Collect, and so does not fulfill the charter.
2. Significant change from the status quo. The DAA Proposal data hygiene provisions address how to conduct market research and product development, but multiple comments state there is no significant change from the status quo. The overall comments indicate that the June Draft more clearly meets this criterion.
3. Easy to explain why DNT:1 reduces tracking for participating sites. Based on discussions in the Group, and comments submitted, it is difficult to explain to users how the DAA Proposal reduces tracking for users who select DNT. Retargeting and profiling would continue unchanged. Collection would be unchanged, and the principal changes would be to how data is handled internally by companies after it is collected.

Advertising industry comments stress how adopting the DAA Proposal as base text would meet the other criterion set forth in Boston: adoption. In Boston, however, Swire specifically noted that adoption of a standard with no or little change from the status quo would not be sufficient. Other commenters have emphasized that there would be widespread confusion if consumers select a Do Not Track option, only to have targeting and collection continue unchanged. Data hygiene provisions are surely worth pursuing. By themselves, however, they do not address the “tracking” in Do Not Track.

In conclusion, the decision today is based on the comments submitted by July 12, as well as the emails and other public records established for the Working Group. ISSUE 215, the choice of base text, is hereby closed, and the June Draft will be the base text for the group’s continued work. As previously noted, this decision also substantially affects ISSUE 5 (tracking), 16 (definition of collection, etc.), 188 (unique identifiers), and 191 (de-identification). Having considered the points above, we will not accept changes that are merely re-statements of these elements from the DAA proposal.

The Working Group will turn to examination of the other change proposals to the June Draft, as announced previously to the group and as listed on the group’s web page. We plan to work on these immediately in the weekly call on July 17, and will seek to close as many as possible this month. Before the end of July, the group will discuss whether and how to proceed in light of the current Last Call deadline scheduled for the end of July.

Matthias Schunter  
Peter Swire

Co-Chairs, Tracking Protection Working Group  
World Wide Web Consortium