Evolving Online Privacy –

Advancing User Choice

**Introduction**

*We continue to appreciate all the hard work being put in by the W3C, the co-chairs, and all of the stakeholders participating within the Tracking Protection Working Group. The ultimate objective is a Do Not Track standard that will advance user choice beyond the number of already available options and will be implemented by a significant portion of the ecosystem. A standard that is not adopted does not benefit users and that is everyone’s objective – a practical, easy-to-use tool that will enhance the ability of users to express their online preferences. In order to make that objective possible, the following proposal is put forward as an evolution of the initial proposal built on the important aspects of the DAA Self-Regulatory Principles for Multi-site Data. The proposal covers Parties, Permitted Uses, Explicit and Separate User Choice, and Unlinkability.*

## Part I: Parties

*<Normative>*

Definitions

1. A party is any commercial, nonprofit, or governmental organization, a subsidiary or unit of such an organization, or a person.
2. For unique corporate entities to qualify as a common party with respect to this standard, those entities MUST be commonly owned and commonly controlled (Affiliates) and MUST provide “easy discoverability” of affiliate organizations. An “Affiliate List” MUST be provided within one click from each page or the entity owner clearly identified within one click from each page.
   1. For example, a clear labeled link to the Affiliate List within the privacy policy would meet this requirement or the ownership brand clearly labeled on the privacy policy itself.
3. A First Party is the party that owns the Web site or has control over the Web site the consumer visits. A First Party also includes the owner of a widget, search box or similar service with which a consumer interacts.
   1. NOTE: If a user merely mouses over, closes, or mutes third-party content, that is not sufficient interaction to constitute a First Party widget interaction.
4. Service Providers acting on the behalf of a First Party and with no independent rights to use the First Party’s data outside of the context of that 1st party and Permitted Uses are also considered a First Party.
5. A Third Party is any party other than a First Party, Service Provider, or the user.
6. It is possible to have multiple first parties on a single page but each party must provide clear branding and a link to their respective privacy policy (co-branded experience).

Rules

If a user has not granted an exception (via browser agent or out-of-band consent) AND if an activity is not allowed under Permitted Uses, THEN the following general party rules apply when a user expressly sets their tracking preference to DNT:1:

* First Parties may engage in their normal collection and use, as long as they do not pass profile information on to third parties who could not collect the data themselves under this recommendation. This includes the ability to customize the content, services, and advertising in the context of the 1st party experience.

* 3rd parties MUST NOT use data across multiple, non-affiliated websites.

NOTE: Data collected by a 3rd party MUST be segregated according to the 1st party from which it was collected. A 3rd party MUST NOT aggregate, correlate, or use together data that was collected on different 1st party sites.

* 3rd parties MUST NOT add collected data to a "profile" of a user, whether they own the profile or it is on behalf of the first party..
* 3rd parties MUST NOT leverage previously collected data as a 3rd party to profile a user or to alter a user's experience.
* 3rd parties MUST NOT attempt to personally identify a user (except as necessary to perform an expressly permitted activity).
* A party MUST NOT share (send or receive) collected data or profiles with another party (unless that party is ONLY working on the behalf of that specific party – aka Service Provider relationship).

NOTE: (Outside of DNT Context): Data collected and received from a party MAY be combined with existing 1st party profile data.

* A party MAY choose to remove any previously profiled data.
* All permitted data uses for necessary business activities apply in all cases.
* User granted exceptions (through DNT standard or out-of-band) supersede these rules.

## Part II: Permitted Uses of data for Necessary Business Activities

*<Non-Normative>*

In order to avoid DNT significantly impacted the availability of free content on the Internet a balance is necessary to allow for minimal data use to continue standard site operations – including those that monetize site activities.

*<Normative>*

For each of the Permitted Uses outlined, the following requirements apply:

* Outside of Security, all other Permitted Uses will not allow for altering a specific user’s online experience based on multi-site activity, including:
  + **No profiling**
  + No further alteration to the user experience based on profiled information based on multi-site activity
    - Customization outside of multi-site activity profiles is allowable
* Each party engaging in Permitted Uses and claiming W3C DNT compliance, MUST provide **public transparency of their data retention period**
  + Party MAY enumerate each individually if they vary across Permitted Uses
* **Reasonable technical and organizational safeguards** to prevent further processing.

*<non-normative>*

* + Examples of acceptable safeguards: collection limitations, data siloing, authorization restrictions, k-anonymity, unlinkability, retention time, anonymization, pseudonymization, and/or data encryption.

NOTE: While definitely a Permitted Use, **compliance with local laws** and public purposes, such as copyright protection and delivery of emergency services, is not listed separately.

*<Normative>*

1. **Security -** Data MAY be collected and used for the express and limited purpose of security and fraud detection and defense. This includes data reasonably necessary for enabling authentication/verification, detecting hostile transactions and attacks, providing fraud prevention, or bolstering site and system security.

*<Non-Normative>*

Restricting security and fraud detection and defense efforts could harm users. We do not want to mistakenly turn Do Not Track into a signal for user vulnerability.

* Resources: For additional details on the uses of Security data please see the DAA Self-Regulatory Principles for Multi-site Data: Authentication, Verification, Fraud Prevention and Security & Compliance, & Public Purpose and Consumer Safety

*<Normative>*

2. **Financial Purposes -** Data MAY be collected and used for the limited purpose of financial fulfillment such as billing and audit compliance. This purpose is strictly necessary for the continued operation of most websites and requires uniqueness to prove user interactions (ad impression and ad click) were indeed achieved as billed for.

*<Non-Normative>*

NOTE: Typically all relevant advertising order criteria is necessary for retention of ad interactions.

* Examples of data uses include, but are not limited to:
  + Ad Impression verification (CPM)
  + Ad Click verification (CPC)
  + Site Conversion associated with Ad Impression or Ad Click (CPA)
  + Quality Measures such as ad position (location on page, above/below fold) and site the ad was served on (high quality vs. low quality content association)
* Reference: For additional examples and associated details please see the IAB Financial Audit Guidelines.

*<Normative>*

3. **Frequency Capping -** Data MAY be collected and used for the limited purpose of frequency capping – the practice of keeping “count” of the number of times a user or device has seen a specific ad and then halting further display of that ad once the designated threshold has been reached.

*<Non-Normative>*

NOTE: Restricting the number of times a user agent displays ads prevents a user from having to see repetitive ads, prevents publishers from displaying repetitive ads, and prevents advertisers from harming the reputation of their clients.

* Examples of important data uses include, but are not limited to:
  + Reach and frequency metrics
  + Ad performance
  + Logging the number and type of advertisements served on a particular Web site(s).
  + Reporting
* Reference: For additional examples and associated details please see the DAA Self-Regulatory Principles for Multi-site Data:

*<Normative>*

**4. Product Debugging –** Data MAY be collected and used for the limited purpose of identifying and repairing site errors to intended functionality (“Debugging”).

*<Non-Normative>*

Detailed information is often necessary to replicate a specific user’s experience to understand why their particular set of variables is resulting in a failure of expected functionality or presentation.

These variables could include items such as:

* cookie IDs
* URL of the page
* device (UA) details
* content specifics
* activity/event specifics to narrow in on the cause of the discrepancy

*<Normative>*

**5. Aggregate Reporting -** Data MAY be collected and used for the express and limited purpose of aggregate reporting. Aggregate reporting end-points should meet the objectives of “unlinkability” (see below) and therefore are outside of the scope of the DNT standard. There is a time interval necessary to retain event level records to aggregate across the necessary time spans accurately (daily, weekly, monthly, quarterly, etc.).

*<Non-Normative>*

While detailed event level data is not present at the outcome of aggregated reporting it is a necessary ingredient to arrive there.

Examples of uses of aggregate reporting:

* Product Improvement (via Site Analytics)
  + Navigation (Referring sites, Exit Points, Internal Navigation used or not used)
  + Visitor Counts (New, Returning, High Activity, Low Activity)
* Market Research
* the characteristics of a market or group of consumers; or
* the performance of a product, service or feature, in order to improve existing products or services or to develop new products or services.

References: For additional examples and associated details see the DAA Self-Regulatory Principles for Multi-site Data: Market Research & Product Development

## Part III: Explicit and Separate User Choice

*<Non-Normative>*

It is important that Do Not Track remain a user choice that is expressly and separately activated by a user. In those cases where a User Agent is known to have not allowed an explicit choice by a user, Servers have the option to let the user know they will not honor the DNT signal from that User Agent. Any attempt to mislead a user to activate Do Not Track would also allow for a Server to have the option to communicate they will not honor that specific User Agent’s DNT implementation.

*<Normative>*

1. A User Agent must obtain explicit, informed consent to turn on the DNT header\*
2. The User Agent must also make available via a link in explanatory text where DNT is enabled to provide more detailed information about DNT functionality
3. Any User Agent claiming compliance must have a functional implementation of the browser exceptions in this specification
4. Servers MAY respond to users that their UA is “non-compliant” if they believe this to be the case
   1. User Agents MUST relay Server responses to users to ensure transparency
   2. Servers SHOULD be prepared to defend why they have reached this conclusion
   3. Servers that respond to 100% of DNT requests regardless of User Agent details ARE NOT compliant with this recommendation
   4. Servers MAY offer users additional information through a resource link
5. Efforts to misled users to activate Do Not Track MAY also be seen as “non-compliant”

\*NOTE – The TPWG already agreed on this point

## Part IV: Unlinkability

*<Normative>*

Un-linkable Data is outside of the scope of the Tracking Preference standard as information is no longer reasonably linked to a particular user, user agent, or device.

Definition:  A dataset is un-linkable when commercially reasonable steps have been taken to de-identify data such that there is confidence that it contains information which could not be linked to a specific user, user agent, or device in a production environment, and which the entity will commit to make no effort to re-identify, and prohibit downstream recipients of un-linkable data from re-identifying it

*<Non-Normative>*

There are many valid and technically appropriate methods to de-identify or render a data set "un-linkable".  In all cases, there should be confidence the information is unable to be reverse engineering back to a "linkable" state.  Many tests could be applied to help determine the confidence level of the un-linking process.  For example, a k-anonymous test could be leveraged to determine if the mean population resulting from a de-linking exercise meets an appropriate threshold (a high-bar k-anonymous threshold would be 1024).

As there are many possible tests, it is recommended that companies publically stating W3C Tracking Preference compliance provide transparency to their delinking process (to the extent that it will not provide confidential details into security practices) so external experts and auditors can assess if they feel these steps are reasonable given the risk of a particular dataset.

* Information That Is Un-linkable When Collected:  A third party may collect non-protocol information if it is, independent of protocol information, un-linkable data. The data may be retained and used subject to the same limitations as protocol information.

Example: Example Advertising sets a language preference cookie that takes on few values and is shared by many users.

* Information That Is Un-linkable After Aggregation:  During the period in which a third party may use protocol information for any purpose, it may aggregate protocol information and un-linkable data into an un-linkable dataset. Such a dataset may be retained indefinitely and used for any purpose.

Example: Example Advertising maintains a dataset of how many times per week Italy-based users load an ad on Example News.

* Information That Is Un-linkable After Anonymization:  At some point after collection, a unique ID from a product cookie has a one-way salted hash applied to the identifier to break any connection between the resulting dataset and production identifiers.  To further remove dictionary attacks on this method, its recommended that "keys" are rotated on a regular basis.

## Closing

We believe this proposal advances the existing opt-out cookie structure (albeit that is a fair and good one) and evolves online privacy in several ways:

* Users gain a consistent, local tool to communicate their opt-out preference
  + Avoids property specific opt-out pages
* The user's choice is persistent for each device/UA
  + Avoids accidental deletion
* Outside of Security purposes, the user will no longer experience alterations to their online experiences derived from multi-site activity
  + No profiling <aka "tracking"> and extends beyond online behavioral advertising
* Only minimal data is retained for necessary business operations and retention periods are transparent to users
  + The Internet remains a viable ecosystem for free content
* All real “harms” are removed
  + Outside of government intrusion risk where there are no documented cases of this occurring with 3rd party anonymous log file data

We hope you agree this clearly advances the already established opt-out program offered to users with many key new additions and evolved perspectives to what was already a long and hard negotiated outcome for users of the current Internet. The online privacy debate did not begin with the W3C Tracking Protection Working Group and most definitely won't end with it but through collective agreement with this proposal we will have marked an important date in the positive progression of the online privacy debate.