Do Not Track — Compromise Proposal

Unofficial Draft 06 June 2012

Editors:
Peter Eckersley, Electronic Frontier Foundation
Tom Lowenthal, Mozilla
Jonathan Mayer, Stanford University

Abstract

Abstract, version, and status information are not relevant in this partial draft.

Status of This Document

This document is merely a public working draft of a potential specification. It has no official standing of any kind and does not represent the support or consensus of any standards organisation.

Table of Contents

1. User Agents

1.1 Explicit Consent Requirement

Note: This section was recently added and has not been extensively discussed with stakeholders. Please consider it a preliminary position.

An ordinary user agent MUST NOT send a Tracking Preference signal without a user's explicit consent.

Example: The user agent's privacy preferences pane includes controls for configuring the Tracking Preference signal.

Example: On first run, the user agent prompts the user to configure the Tracking Preference signal.

2. Parties, First Parties and Third Parties

2.1 Parties

2.1.1 Definitions

A ***functional entity*** is any commercial, nonprofit, or governmental organization, a subsidiary or unit of such an organization, or a person.

Functional entities are ***affiliated*** when they are related by both common majority ownership and common control.

A ***party*** is a set of functional entities that are affiliated.

2.1.2 Transparency

2.1.2.1 Requirement

A functional entity must make its affiliated functional entities easily discoverable by a user.

2.1.2.2 Non-Normative Discussion

Affiliation may be made easily discoverable by a user in many ways, including but not limited to: prominent and common branding on pages, one click away within a privacy policy, or a machine-readable format in a well-known location.

2.2 Network Interaction

2.2.1 Definition

A ***network interaction*** is an HTTP request and response, or any other set of logically related network traffic.

2.2.2 Non-Normative Discussion

Determination of a party's status is limited to a single transaction because a party's status may be affected by time, context, or any other factor that influences user expectations.

2.3 First Parties and Third Parties

2.3.1 Definitions

A ***first party*** is any party, in a specific network interaction, that can infer with high probability that the user knowingly and intentionally communicated with it. Otherwise, a party is a third party.

A ***third party*** is any party, in a specific network interaction, that cannot infer with high probability that the user knowingly and intentionally communicated with it.

2.3.2 Non-Normative Discussion

2.3.2.1 Overview

We draw a distinction between those parties an ordinary user would or would not expect to share information with, "first parties" and "third parties" respectively. The delineation exists for three reasons.

First, when a user expects to share information with a party, she can often exercise control over the information flow. Take, for example, Example Social, a popular social network. The user may decide she does not like Example Social's privacy or security practices, so she does not visit examplesocial.com. But if Example Social provides a social sharing widget embedded in another website, the user may be unaware she is giving information to Example Social and unable to exercise control over the information flow.

Second, we recognize that market pressures are an important factor in encouraging good privacy and security practices. If users do not expect that they will share information with an organization, it is unlikely to experience market pressure from users to protect the security and privacy of their information. In practice, moreover, third parties may not experience sufficient market pressure from first parties since increasingly third parties do not have a direct business relationship with the first party websites they appear on. We therefore require a greater degree of user control over information sharing with such organizations.

Last, third parties are often in a position to collect a sizeable proportion of a user's browsing history – information that can be uniquely sensitive and easily associated with a user's identity. We wish to provide user control over such information flows.

We recognize that, unlike with a bright-line rule, there can be close calls in applying our standard for what constitutes a first party or a third party. But we believe that in practice, such close calls will be rare. The overwhelming majority of content on the web can be classified as first party or third party, with few cases of ambiguity in practice.

We require a confidence at a "high probability" before a party can consider itself a first party. Where there is reasonable ambiguity about whether a user has intentionally interacted with a party, it must consider itself a third party. Our rationale is that, in the rare close cases, a website is in the best position to understand its users' expectations. We therefore impose the burden of understanding user expectations on the website. We also wish, in close cases, to err on the side of conforming to user expectations and protecting user privacy. If the standard is insufficiently protective, ordinary users have limited recourse; if the standard imposes excessive limits, websites retain the safety valve of explicitly asking for user permission.

2.3.2.2 Common Examples and Use Cases

1. A user accesses an Example News article. The page includes an advertisement slot, which loads content from many companies other than Example News. Those companies are third parties.
2. A user accesses an Example News article. The page includes an analytics script that is hosted by Example Analytics, an analytics service. Example Analytics is a third party.
3. A user accesses an Example News article. It includes a social sharing widget from Example Social, a popular social network. Example Social is a third party.
4. A user visits Example Diary, which is hosted by the free blogging service Example Blog Hosting but located at examplediary.com. Example Blog Hosting is a third party.
5. A user launches Example Application, an app on a mobile device. The app includes a library from Example Advertising Network that displays ads. Example Advertising Network is a third party.

2.3.2.3 Multiple First Parties

There will almost always be only one party that the average user would expect to communicate with: the provider of the website the user has visited. But, in rare cases, users may expect that a website is provided by more than one party. For example, suppose Example Sports, a well known sports league, collaborates with Example Streaming, a well known streaming video website, to provide content at www.examplesportsonexamplestreaming.com. The website is prominently advertised and branded as being provided by both Example Sports and Example Streaming. An ordinary user who visits the website may recognize that it is operated by both Example Sports and Example Streaming.

2.3.2.4 User Interaction with Third-Party Content

A party may start out as a third party but become a first party later on, after a user interacts with it. If content from a third party is embedded on a first party page, the third party may become an additional first party if it can infer with high probability that the average user knowingly and intentionally communicated with it. If a user merely moused over, closed, or muted third-party content, the party would not be able to draw such an inference.

2.3.2.4.1 Examples and Use Cases

Example: Example Weather offers an unbranded weather widget that is embedded into websites, including Example News. The widget contains small links to Example Weather's website and privacy policy. A user visits Example News and scrolls through the weekly forecast in the Example Weather widget.

Discussion: Example Weather is a third party. The user has interacted with Example Weather's widget, but an ordinary user would not expect that scrolling through the widget involves communicating with Example News.

Example: Example Social, a popular social network, hosts a social sharing button that other websites can embed. The button is colored and styled in the same fashion as Example Social's website, contains descriptive text that is specific to Example Social, includes Example Social's logo, and very frequently appears on Example Social's website. Example News embeds the Example Social button, and a user clicks it.

Discussion: Example Social is a first party once the user clicks its embedded social sharing button. The average user would understand that by clicking the button she is communicating with Example Social.

3. Information Practices

3.1 Reception, Retention, Use, and Sharing

A party ***receives*** data if the data comes within its control.

A party ***retains*** data if the data remains within the party's control.

A party ***uses*** data if the party processes the data for any purpose, including for storage.

A party ***shares*** data if the party enables another party to receive the data.

3.2 First Party

A first party MUST NOT share information with a third party that the third party is prohibited from receiving itself.

Best Practice 1: Additional Voluntary Measures

Best Practice 1: Additional Voluntary Measures

A first party may voluntarily take steps to protect user privacy when responding to a Do Not Track request.

3.3 Third Party

3.3.1 General Rule

A third party MUST NOT receive, retain, use, or share any information related to communication with a user or user agent. There are exceptions to this general rule as defined in the following sections. In case of ambiguity, an exception MUST be construed narrowly. Each exception operates independently; exceptions cannot be combined except where explicitly noted otherwise.

3.3.2 Exceptions

3.3.2.1 Protocol Information

3.3.2.1.1 Definition

Protocol information includes:

- any information that a user agent necessarily shares with a web server when it communicates with the web server (e.g. IP address and User-Agent), and
- the URL of the top-level page, communicated via a Referer header or other means, unless the URL contains information that is not unlinkable (e.g. a username or user ID).

Protocol information does not include:

- any information that a web server could cause to not be sent but still communicate with the user agent (e.g. a cookie or a Request-URI parameter generated by the user agent), except the URL of the top-level page, and
- any data added by a network intermediary that the operator of a web server has actual knowledge of (e.g. a unique device identifier HTTP header).

3.3.2.1.2 In General

A third party MAY receive and use protocol information for any purpose, subject to a two-week retention period.

3.3.2.1.3 Non-Normative Discussion: Contextual Personalization

Under the general rule on protocol information a third party MAY temporarily use a top-level page URL for the purpose of contextually personalizing content.

3.3.2.1.4 Additional Limit on Geolocation

Under the general rule a third party MAY temporarily use an IP address for geolocation. The geolocation MUST be coarse.

3.3.2.1.5 Security and Fraud Prevention

A third party MAY receive and use protocol information for the detection and prevention of security breaches and fraudulent activity, subject to a six-month retention period and the restrictions imposed in the subsequent sections on security and fraud prevention.

3.3.2.2 Unlinkable Data

3.3.2.2.1 Definitions

A dataset is ***unlinkable*** when there is a high probability that it contains only information which could not be linked to a particular user, user agent, or device by a skilled analyst.

N-unlinkability is the special case of K-anonymity where all values are considered part of the pseudo-identifier.

3.3.2.2.2 Validation

Third parties that receive, retain, or use unlinkable data MUST either:

1. publicly publish information that is sufficiently detailed for a skilled analyst to evaluate the implementation, or
2. ensure that any datasets are at least 1024-unlinkable.

3.3.2.2.3 Information That Is Unlinkable When Received

A third party MAY receive non-protocol information if it is, independent of protocol information, unlinkable data. The data MAY be retained and used subject to the same limitations as protocol information. Such data MUST be disassociated from protocol information when it is first used or within two weeks, whichever is sooner.

Example: Example Advertising sets a language preference cookie that takes on few values and is shared by many users. Log entries containing this preference cookie as well as protocol information are collected on each of Example Advertising's webservers. When Example Advertising processes its logs, it computes unlinkable datasets using the protocol logs and language cookies. After that process, Example Advertising no longer stores log files that associate protocol log entries with the language cookies.

3.3.2.2.4 Information That Is Unlinkable After Aggregation

During the period in which a third party may use protocol information for any purpose, it may aggregate protocol information and unlinkable data into an unlinkable 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 each of its ads on Example News.

3.3.2.3 Outsourcing

A first party MAY outsource website functionality to a third party, in which case the third party may act as the first party under this standard with the following additional restrictions.

3.3.2.3.1 Technical Precautions

3.3.2.3.1.1 Operative Text

Throughout all data reception, retention, and use, outsourced service providers MUST use all feasible technical precautions to both mitigate the linkability of and prevent the linking of data from different first parties.

Structural separation ("siloing") of data per first party, including both

1. separate data structures and
2. avoidance of shared unique identifiers

are necessary, but not necessarily sufficient, technical precautions.

3.3.2.3.1.2 Non-Normative Discussion

3.3.2.3.1.2.1 Siloing in the Browser

Outsourcing services should use browser access control features so that stored data specific to one first party is never accessed or received when the user visits another first party.

3.3.2.3.1.2.1.1 Same-Origin Policy

The same-origin policy silos stored data by domain name. An outsourcing service can use a different domain name for each first party.

Example: Example Analytics provides an outsourced analytics service to Example News and Example Sports, two unrelated websites. Example Analytics stores its cookies for Example News at examplenews.exampleanalytics.com, and it stores its cookies for Example Sports at examplesports.exampleanalytics.com.

3.3.2.3.1.2.1.2 Cookie Path Attribute

The HTTP cookie path can be used to silo data to a first party.

Example: Example Analytics stores its cookies for Example News with "Path=/examplenews", and it stores its cookies for Example Sports with "Path=/examplesports".

3.3.2.3.1.2.1.3 Storage Key

For key/value storage APIs, such as Web Storage and Indexed Database, an outsourcing service can use a different key or key prefix for each first party.

Example: Example Analytics stores data for Example News at window.localStorage["examplenews"] and data for Example Sports at window.localStorage["examplesports"].

3.3.2.3.1.2.2 Siloing in the Backend

3.3.2.3.1.2.2.1 Encryption Keys

An outsourcing service should encrypt each first party's data with a different set of keys.

3.3.2.3.1.2.2.2 Access Controls

An outsourcing service should deploy access controls so that only authorized personnel are able to access siloed data, and only for authorized purposes.

3.3.2.3.1.2.2.3 Access Monitoring

An outsourcing service should deploy access monitoring mechanisms to detect improper use of siloed data.

3.3.2.3.1.2.3 Retention in the Backend

An outsourcing service should retain information only so long as necessary to provide necessary functionality to a first party. If a service creates periodic reports, for example, it should delete the data used for a report once it is generated. An outsourcing service should be particularly sensitive to retaining protocol logs, since they may allow correlating user activity across multiple first parties.

3.3.2.3.2 Internal Practices

3.3.2.3.2.1 Operative Text

Throughout all data reception, retention, and use, outsourced service providers MUST use sufficient internal practices to prevent the linking of data from different first parties.

3.3.2.3.2.2 Non-Normative Discussion

3.3.2.3.2.2.1 Policy

An outsourcing service should establish a clear internal policy that gives guidance on how to receive, retain, and use outsourced data in compliance with this standard.

3.3.2.3.2.2.2 Training

Personnel that interact with outsourced data should be familiarized with internal policy on compliance with this standard.

3.3.2.3.2.2.3 Supervision and Reporting

An outsourcing service should establish a supervision and reporting structure for detecting improper access.

3.3.2.3.2.2.4 Auditing

External auditors should periodically examine an outsourcing service to assess whether it is in compliance with this standard and has adopted best practices. Auditor reports should be made available to the public.

3.3.2.3.3 Use Direction

An outsourced service

1. MUST use data retained on behalf of a first party ONLY on behalf of that first party, and
2. MUST NOT use data retained on behalf of a first party for their own business purposes, or for any other reasons.

3.3.2.3.4 First-Party Requirements

3.3.2.3.4.1 Representation

A first party's representation that it is in compliance with this standard includes a representation that its outsourcing service providers comply with this standard.

3.3.2.3.4.2 Contract

A first party MUST enter into a contract with an outsourcing service provider that requires that outsourcing service provider to comply with these requirements.

3.3.2.4 User Permission

A website my engage in practices otherwise prohibited by this standard if a user grants permission. Permission may be attained through the browser API defined in the companion Tracking Preference Expression document. A website may also rely on "out-of-band" consent attained through a different technology. An "out-of-band" choice mechanism has the same effect under this standard as the browser exception API, provided that it satisfies the following bright-line requirements:

1. **Actual presentation:** The choice mechanism MUST be actually presented to the user. It MUST NOT be on a linked page, such as a terms of service or privacy policy.
2. **Clear terms:** The choice mechanism MUST use clear, non-confusing terminology.
3. **Independent choice:** The choice mechanism MUST be presented independent of other choices. It MUST NOT be bundled with other user preferences.
4. **No default permission:** The choice mechanism MUST NOT have the user permission preference selected by default.

An "out-of-band" choice mechanism must additionally satisfy the following high-level standard:

An ordinary user would know that the choice overrides his or her privacy protections under this standard.

3.3.2.5 Security

3.3.2.5.1 Operative Text

A third party MAY receive, retain, and use data about a particular user or user agent for the purpose of ensuring its security, provided that there are reasonable grounds to believe the user or user agent was attempting to breach the party's security at the time the data was received.

Note: This draft does not address the extent to which technical and business precautions are required for security data.

3.3.2.5.2 Non-Normative Discussion

This exception grants third parties (e.g. advertising networks) some latitude to mitigate security risks. Websites that users store sensitive personal information on (e.g. financial services and webmail) are all first-party; they are able to receive, retain, and use information about all users for security purposes.

3.3.2.6 Fraud Prevention

3.3.2.6.1 Operative Text

A third party MAY receive, retain, and use data about a particular user or user agent for the purpose of preventing fraud, provided that there are reasonable grounds to believe the user or user agent was attempting to commit fraud at the time the data was received.

Note: This draft does not address the extent to which technical and business precautions are required for fraud prevention data.

3.3.2.6.2 Non-Normative Discussion

When a user meaningfully interacts with third-party content (e.g. clicking an ad), the third party can receive, retain, and use data for fraud prevention. Third parties can also use protocol information for fraud prevention. This exception provides an additional capability to, in certain circumstances, track impressions for fraud prevention.

3.3.2.7 Unknowing Information Practices

Note: This section was recently added and has not been extensively discussed with stakeholders. Please consider it a preliminary position.

A party MAY receive, retain, and use data as otherwise prohibited by this standard, so long as is unaware of such information practices and has made reasonable efforts to understand its information practices. If a party learns that it possesses information in violation of this standard, it MUST delete that information at the earliest practical opportunity.

A. References

A.1 Normative references

No normative references.

A.2 Informative references

No informative references.