

Working Draft: This is our effort to compile the previous discussions and proposed text with regard to market research within the Tracking Protection Group. We apologize for any missed text or viewpoints or for including any issues unrelated to market research.

## **Market Research / Audience Measurement**

This document begins with the options discussed for aggregate reporting in the working group in October 2012. It then provides links and key text on audience measurement and related issues, from oldest to newest.

### **Tracking Compliance and Scope October 2, 2012**

<http://www.w3.org/TR/tracking-compliance/>

#### 6.1.1.8 Aggregate Reporting

##### *6.1.1.8.1 Option 1: Aggregate Reporting*

Regardless of DNT signal, information may be collected, retained and used for aggregate reporting, such as market research and product improvement. Data *MAY* be collected and retained on an individual level, but the use of the data must only be aggregate reporting, and the products of the reporting *MUST* be unlinkable as defined in this document.

##### *6.1.1.8.2 Option 2: Aggregate Reporting*

Regardless of DNT signal, information may be collected, retained and used for aggregate reporting, such as market research and product improvement, if that information is collected and retained for another enumerated permitted use. Data *MAY* be collected and retained on an individual level, but the use of the data must only be aggregate reporting, and the products of the reporting *MUST* be unlinkable as defined in this document. If the operator no longer has another enumerated permitted use for which to use and retain the data, the operator *MAY NOT* use and retain the data for aggregate reporting unless the data has been rendered unlinkable as defined in this document.

##### *6.1.1.8.3 Option 3: No Aggregate Reporting*

There is no permitted use for aggregate reporting outside of the grace period described earlier.

### **Kathy Joe Proposal March 19, 2013:**

#### **New text Issue 25: Aggregated data: collection and use for audience measurement research**

##### **Normative:**

Information may be collected, retained and used by a third party for audience measurement research where the information is used to calibrate or otherwise support data collected from opted-in panels, which in part contains information collected across sites and over time from user agents.

Working Draft: This is our effort to compile the previous discussions and proposed text with regard to market research within the Tracking Protection Group. We apologize for any missed text or viewpoints or for including any issues unrelated to market research.

A third party eligible for an audience measurement research permitted use MUST adhere to the following restrictions. The data collected by the third party:

- Must be pseudonymised, and
- Must not be shared with any other party unless the data are de-identified prior to sharing, and
- Must be deleted or de-identified as early as possible after the purpose of collection is met and in no case shall such retention, prior to de-identification, exceed 53 weeks and
- Must not be used for any other independent purpose.
- In addition, the third party must be subject to an independent certification process under the oversight of a generally-accepted market research industry organization that maintains a web platform providing user information about audience measurement research. This web platform lists the parties eligible to collect information under DNT standards and the audience measurement research permitted use and it provides users with an opportunity to exclude their data contribution.

### **Non-normative: collection and use for audience measurement research**

Audience measurement research creates statistical measures of the reach in relation to the total online population, and frequency of exposure of the content to the online audience, including paid components of web pages.

Audience measurement research for DNT purposes originates with opt-in panel output that is calibrated by counting actual hits on tagged content on websites. The panel output is re-adjusted using data collected from a broader online audience in order to ensure data produced from the panel accurately represents the whole online audience.

This online data is collected on a first party and third party basis. This collection tracks the content rather than involving the collection of a user's browser history.

The collected data is retained for a given period for purposes of sample quality control, and auditing. During this retention period contractual measures must be in place to limit access to, and protect the data, as well as restrict the data from other uses. This retention period is set by auditing bodies, after which the data must be de-identified.

The purposes of audience measurement research must be limited to:

- Facilitating online media valuation, planning and buying via accurate and reliable audience measurement.
- Optimizing content and placement on an individual site.

Audience measurement data must be reported as aggregated information such that no recipient is able to build commercial profiles about particular individuals or devices.

**Kathy Joe Proposal March 6, 2013:**

Working Draft: This is our effort to compile the previous discussions and proposed text with regard to market research within the Tracking Protection Group. We apologize for any missed text or viewpoints or for including any issues unrelated to market research.

Information may be collected to create statistical measures of the reach in relation to the total population, and frequency of exposure of the content to the online audience, including paid components of web pages. One such method is through using a panel of users who have affirmatively agreed to have their media consumption and web surfing behavior measured across sites.

The panel output is calibrated by counting actual hits on tagged content and re-adjusting the results in order to ensure data produced from the panel accurately represents the whole audience. The counts must be pseudonomised. Counts are retained for sample, quality control, and auditing purposes during which time contractual measures must be in place to limit access to, and protect the data from other uses. A 53 week retention period is necessary so that month over month reports for a one year period may be re-run for quality checking purposes, after which the data must be de-identified. The counted data is largely collected on a first party basis, but to ensure complete representation, some will be third party placement. This collection tracks the content rather than involving the collection of a user's browser history.

The purposes must be limited to:

- facilitating online media valuation, planning and buying via accurate and reliable audience measurement.
- optimizing content and placement on an individual site.
- Audience measurement data must be reported as aggregated information such that no recipient is able to build commercial profiles about particular individuals or devices.

**Issues:**

- ISSUE-25: Possible exception for research purposes
  - <http://www.w3.org/2011/tracking-protection/track/issues/25>
- ISSUE-34: Possible exemption for aggregate analytics (Closed)
  - <http://www.w3.org/2011/tracking-protection/track/issues/34>
- ISSUE-74: Are surveys out of scope? (Closed)
  - <http://www.w3.org/2011/tracking-protection/track/issues/74>

MRC handles this in the US whilst the JICWEBs reporting standards of ABC handles this in the UK and AGMA is the German audit body. Here is

a longer list <http://www.i-jic.org/index.php?PHPSESSID=55143f172846ed39c7958cbeb837a85a>

and here is

ABC <http://www.abc.org.uk/PageFiles/50/Web%20Traffic%20Audit%20Rules%20and%20Guidance%20Notes%20version2%20March%202013%20master.pdf>

Working Draft: This is our effort to compile the previous discussions and proposed text with regard to market research within the Tracking Protection Group. We apologize for any missed text or viewpoints or for including any issues unrelated to market research.

### **Ed Felton Questions about Kathy Joe Proposal:**

Kathy, thanks for providing this language.

Echoing Rob, I would find it useful to have a clearer understanding of how your proposal would change what is allowed under the standard. The existing or likely safe harbors for consent, first-party practices, and unlinkable data would apply regardless of whether there is a special permitted use for research. To what extent does your proposal change the normative requirements of the standard relating to collection, retention and use of data?

The text as provided seems to mix normative requirements with non-normative descriptions of the current practices of some companies. But I am not always clear on where the boundary lies. For example, I take the 53-week retention as a description of current practices, rather than a requirement that data must be kept for at least 53 weeks. On the other hand, the discussion of de-identification later in the same sentence is probably meant as normative. It would be nice to separate the text into normative and non-normative sections.

Finally, if we are going to have a research exemption, I don't see why it should be limited to commercial research. Non-commercial research should be allowed as well, such as independent research on privacy practices and the effects of DNT, as long as the same normative requirements are met.

### **Kathy Joe Response to Ed Felton Questions:**

The proposal for audience measurement research falls outside the 'consent, first-party practices, and unlinkable' categories because it involves collecting and retaining pseudonomised data for a limited period for limited purposes (ie quality checks) with contractual measures in place to restrict access to, and protect the data from other uses.

The 53-week retention period is a normative maximum. 'Audience measurement data must be reported as aggregated information' is also normative. As to your final point, ESOMAR is a worldwide association gathering market, social and opinion researchers including academics and non-commercial researchers all of whom sign up to the same self regulatory code. At the European level, we would not make a distinction between commercial and non-commercial research since public body, academic and university researchers play a role in commercial research and vice versa.

### **Kathy Joe Proposal Oct. 4, 2012:**

<http://lists.w3.org/Archives/Public/public-tracking/2012Oct/0089.html>

This text is intended to complete the section on aggregated data which contained a reference to market research, until the last version issued just before the f2f meeting in Amsterdam.

Working Draft: This is our effort to compile the previous discussions and proposed text with regard to market research within the Tracking Protection Group. We apologize for any missed text or viewpoints or for including any issues unrelated to market research.

We propose that it be included under issue 25

Aggregated data

6.1.1.1 Short Term Collection and Use for market research

Note

Information may be collected and used for market research and research analytics, so long as the information is only retained for the time necessary to complete the research study. This is providing that the raw information is not transmitted to a third party, the information is not used to build a commercial profile about individual users or alter any individual's user experience, and there is no return path to an individual.

A key method for ensuring privacy while collecting and processing large amounts of data is removing any link to a device identifier. Raw data for market research may contain for example an IP address or a marker for a cookie, which may be temporarily retained for sample and quality control as well as auditing purposes. No individual can be identified in the subsequent aggregated statistical report.

**Meeting Sept. 26, 2012:**

<http://www.w3.org/2012/09/26-dnt-minutes#item03>

<aleecia> Changes from the editors' draft: Remove "Aggregate Reporting" section. Ensure that unlinkable data is prominently declared out of scope of these requirements earlier in the document. Ensure that the "Short Term" permitted use makes it clear that retaining identifiable data for the short term is allowed for creating aggregate reports.

**Agenda for 5 September 2012 call**

<http://lists.w3.org/Archives/Public/public-tracking/2012Sep/0000.html>

(e) Issue-25, Possible exemption for research purposes

In Seattle, this dropped from our list of permitted uses. Research can, of course, always be done with consent.

PROPOSAL: we close this issue as outdated.

**Kathy Joe Proposal Feb. 13, 2012:**

<http://lists.w3.org/Archives/Public/public-tracking/2012Feb/0354.html>

As stated in our text for market research (Issue 34: see below), market research identifiable data will be held as long as the campaign runs to provide consistent data after which all identifiers will be removed after a reasonable period.

The period needs to be flexible as it depends on how long the campaign runs and necessary quality controls to check the integrity of the data as agreed with the client.

Working Draft: This is our effort to compile the previous discussions and proposed text with regard to market research within the Tracking Protection Group. We apologize for any missed text or viewpoints or for including any issues unrelated to market research.

Note, under no circumstances is identifiable data provided to the client or used for profiling purposes.

#### Issue 34: Exemption for aggregated data

- Aggregated data is permissible for purposes such as research, industry trends, and analytics. Parties wishing to use aggregated data must take reasonable steps to ensure that data does not reveal information about individual users, user agents, or devices and it must not be possible to identify an individual with aggregated cross site data.

Description: The research client wants statistical measurements of how many users have been exposed to their campaigns in broad categories across different sites. The client will for instance place the research company's tags on their ads on one or more sites that count viewers based on cookies. Any identifiers are removed as soon as the data has been sorted into broad categories eg country.

Suggestion/Example: ExampleResearch collects data for ExampleProducts Inc. which is running an ad campaign online on various sites. It gathers cross-site data on how often a user views a relevant ad but none of their other web behaviour. The purpose is to fulfil a request by a first party (the advertiser), and the results are shared only with the first party.

The output is restricted to aggregated and unidentifiable data, will not impact a user's experience, use is only for the statistical research purpose and cannot be linked to a specific user, computer or device and cannot be used for profiling. Identifiable data will be held as long as the campaign runs to provide consistent data and then all identifiers will be removed after a reasonable period.

### **DAA Market Research Definition:**

#### D. Market Research

Market Research means the analysis of: market segmentation or trends; consumer preferences and behaviors; research about consumers, products, or services; or the effectiveness of marketing or advertising. A key characteristic of market research is that the data is not re-identified to market directly back to, or otherwise re-contact a specific computer or device. Thus, the term “market research” does not include sales, promotional, or marketing activities directed at a specific computer or device.

*Any contact back to a computer or device that is based on an aggregate use of data that may have been collected from such computer or device is not disqualified from being “market research” because data collected from such computer or device was included in the aggregate use.*