

Analyzing the Impact of Large Scale Online Tracking Measurement

Steven Englehardt
@s_englehardt



The Princeton Web Census

Monthly
1 Million Site Crawl

Collecting:

- Javascript Calls
- All javascript files
- HTTP Requests and Responses
- Storage (cookies, Flash, etc)

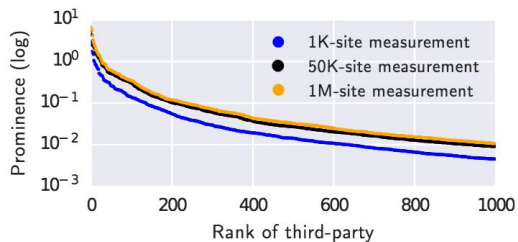
Open Web Privacy Measurement (OpenWPM)

The screenshot shows the GitHub repository page for `citp / OpenWPM`. At the top, there are buttons for `Unwatch` (49), `Unstar` (435), and `Fork` (67). Below these are tabs for `Code`, `Issues` (45), `Pull requests` (0), `Projects` (0), `Wiki`, `Pulse`, `Graphs`, and `Settings`. The repository description is "A web privacy measurement framework <https://webtap.princeton.edu/> — Edit". Below the description, statistics show 480 commits, 4 branches, 12 releases, 13 contributors, and GPL-3.0 license. A progress bar is visible. Below the statistics, there are buttons for `Branch: master`, `New pull request`, `Create new file`, `Upload files`, `Find file`, and `Clone or download`. The commit history table shows the following entries:

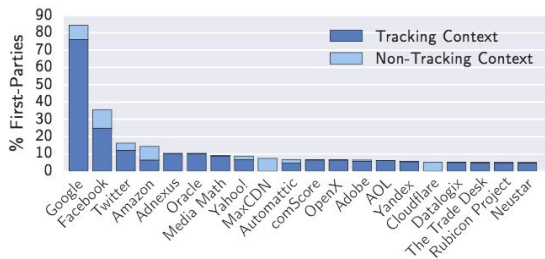
Commit	Message	Time
englehardt	Merge branch 'master' of github.com:citp/OpenWPM	Latest commit 3a14416 7 hours ago
automation	Added comments about new commands	15 days ago
test	disabling audiocontext test for travis CI	15 days ago
.gitignore	Merge branch 'master' of github.com:citp/OpenWPM	10 months ago
.travis.yml	Add travis.yml file to run continuous integration tests.	6 months ago
CHANGELOG	Version bump to 0.6.2. Bugfix in previous version	6 months ago
LICENSE	Removing extra whitespace from all infrastructure files	10 months ago
README.md	Modified readme to only use travis status from master branch	15 days ago

<https://github.com/citp/OpenWPM>

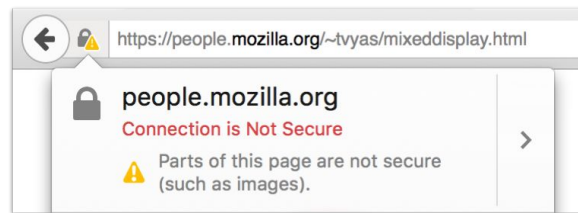
Insights from the Princeton Web Census



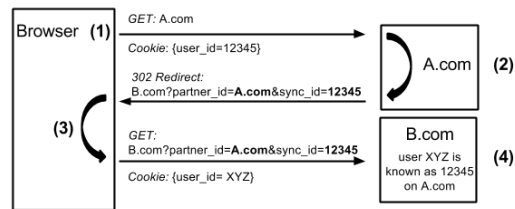
New metric to rank third parties



Consolidation of third-parties



Trackers impede HTTPS adoption

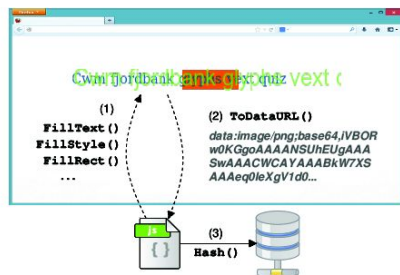


45 out of top 50 3rd parties cookie sync

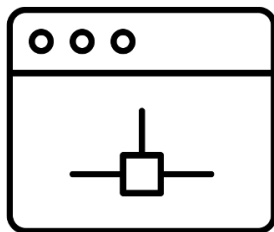
Online Tracking: A 1-million-site Measurement and Analysis (CCS 2016)

How does measurement of new tracking techniques influence trackers and vendors?

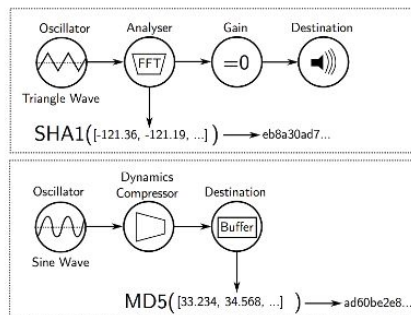
Canvas



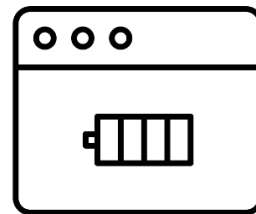
WebRTC



Audio

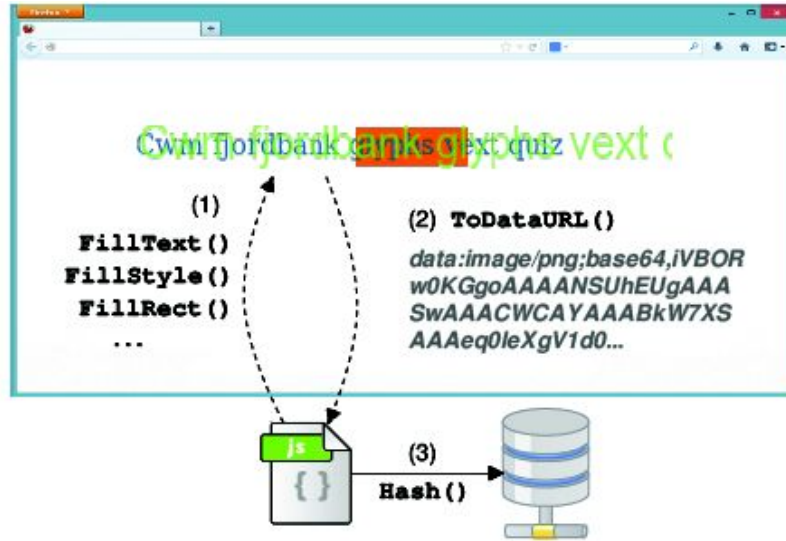


Battery



<https://webtransparency.cs.princeton.edu/webcensus/>

Canvas Fingerprinting



The Web Never Forgets: Persistent Tracking Mechanisms in the Wild (Acar, et al.)

Canvas Fingerprinting

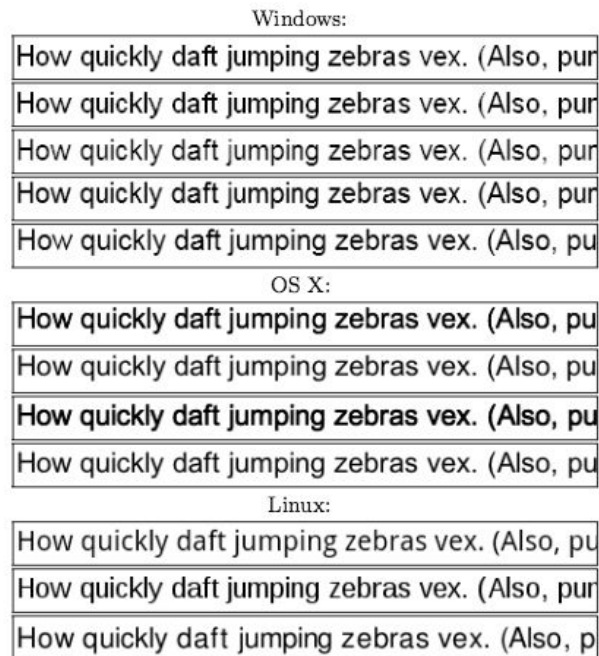


Figure 6: 13 ways to render 20px Arial

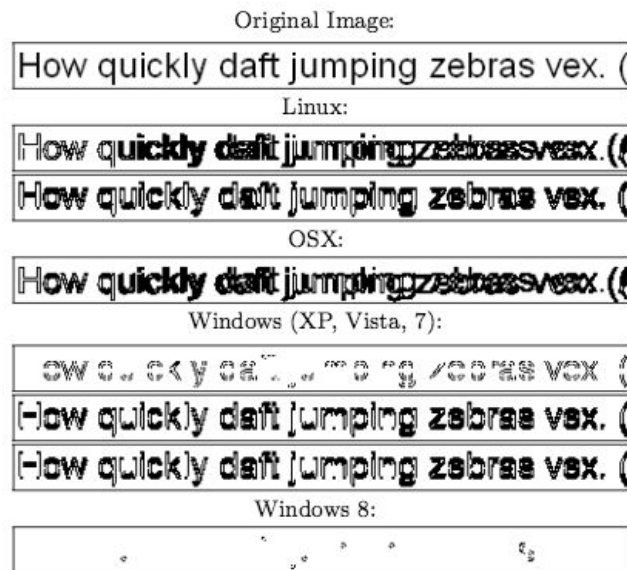


Figure 7: Difference maps for a group on text_arial

Canvas fingerprinting returns in the absence of measurement

May 2014: 5% of sites

Aug 2014: ~0.1% of sites

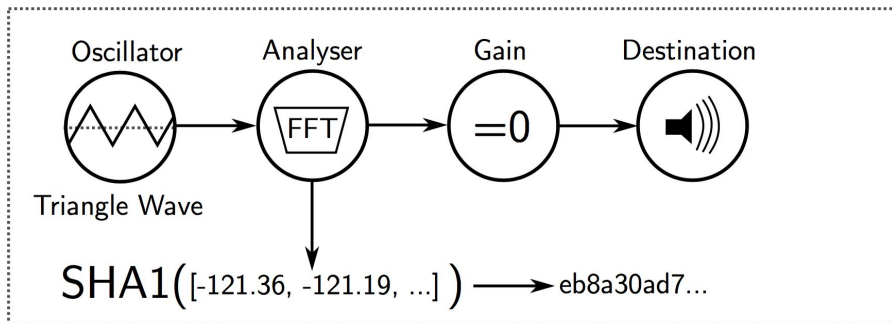
Jan 2016: 2.6% of sites

Percentage of the Alexa top 100k sites

Using AudioContext for fingerprinting

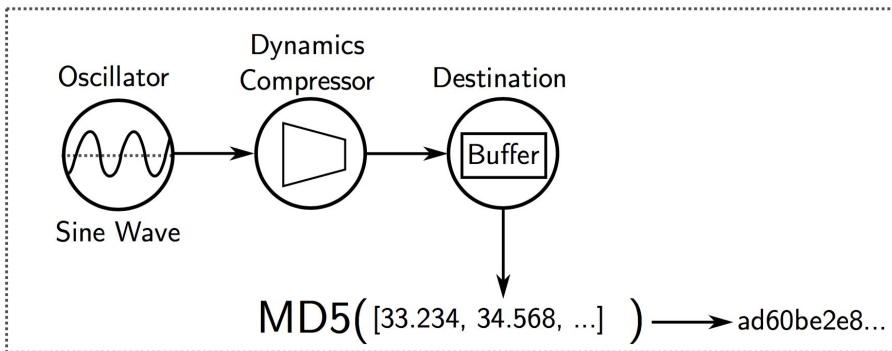
Used by:

cdn-net.com script



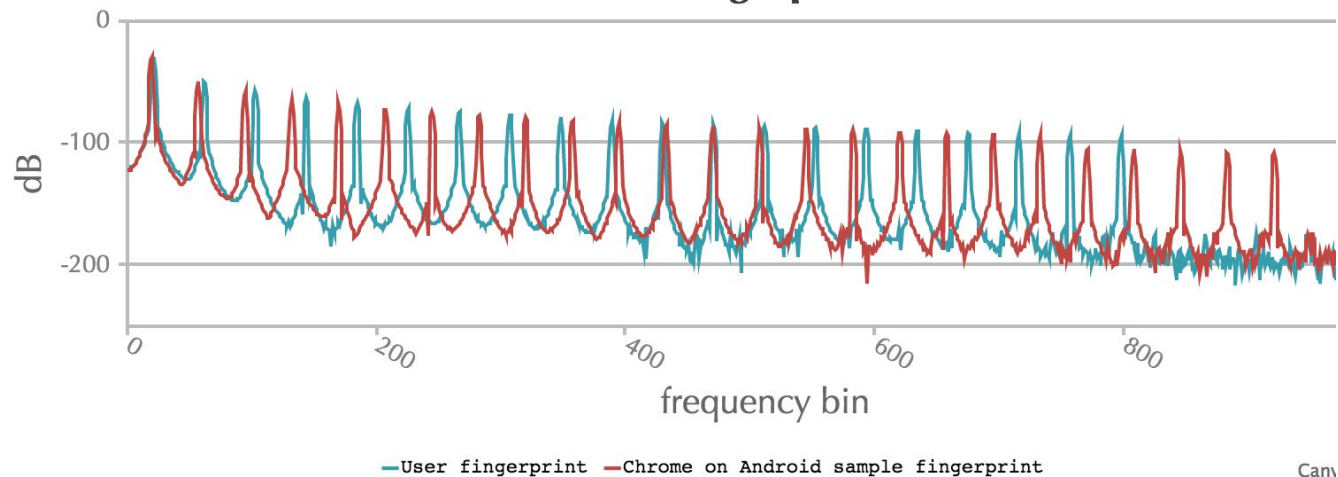
Used by:

pxi.pub and
ad-score.com scripts



Using AudioContext for fingerprinting

Audio Fingerprint

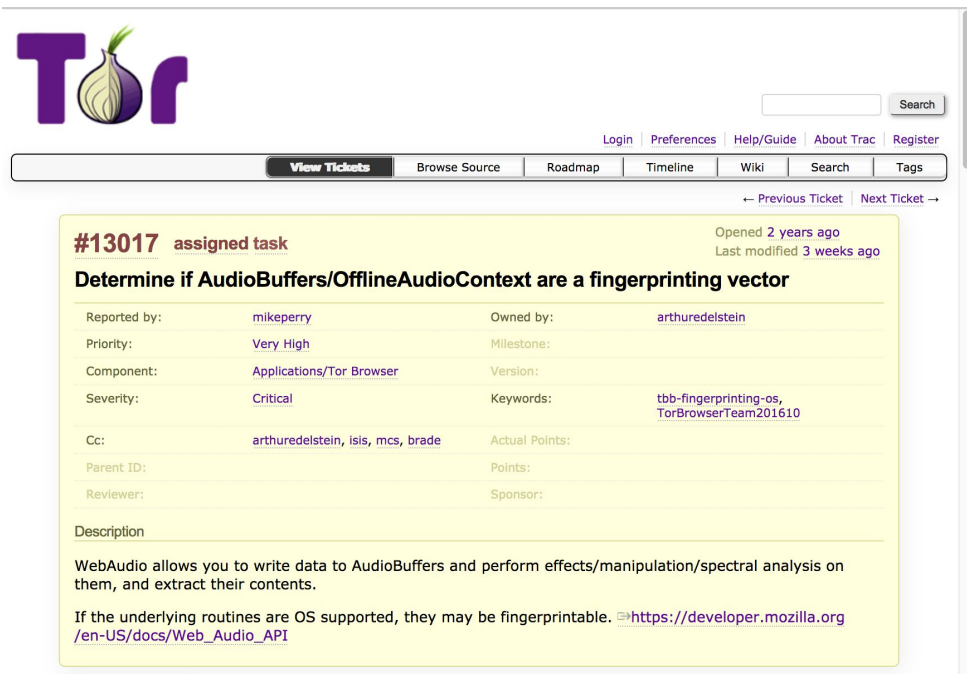


Live test page: <https://audiofingerprint.openwpm.com/>

AudioContext fingerprinting the Tor Browser

271 samples from the Tor Browsers

- 7 distinct fingerprints (2 fingerprints account for 80% of samples)
- Overlap with fingerprints from Firefox shows these largely reveal OS of device



The screenshot shows a Trac ticket interface. At the top left is the Tor logo. To the right is a search bar and a 'Search' button. Below these are navigation links: 'Login', 'Preferences', 'Help/Guide', 'About Trac', and 'Register'. A horizontal bar contains 'View Tickets' and other navigation options. Below this bar are links for 'Previous Ticket' and 'Next Ticket'. The main content area displays ticket #13017, labeled 'assigned task'. It shows the ticket was opened 2 years ago and last modified 3 weeks ago. The title is 'Determine if AudioBuffers/OfflineAudioContext are a fingerprinting vector'. The ticket details include: Reported by: mikeperry, Owned by: arthuredelstein, Priority: Very High, Milestone: (empty), Component: Applications/Tor Browser, Version: (empty), Severity: Critical, Keywords: tbb-fingerprinting-os, TorBrowserTeam201610, Cc: arthuredelstein, isis, mcs, brade, Actual Points: (empty), Parent ID: (empty), Points: (empty), Reviewer: (empty), Sponsor: (empty). The description states: 'WebAudio allows you to write data to AudioBuffers and perform effects/manipulation/spectral analysis on them, and extract their contents. If the underlying routines are OS supported, they may be fingerprintable. =>https://developer.mozilla.org/en-US/docs/Web_Audio_API'.

Tor

Search

Login | Preferences | Help/Guide | About Trac | Register

View Tickets | Browse Source | Roadmap | Timeline | Wiki | Search | Tags

← Previous Ticket | Next Ticket →

#13017 assigned task Opened 2 years ago
Last modified 3 weeks ago

Determine if AudioBuffers/OfflineAudioContext are a fingerprinting vector

Reported by:	mikeperry	Owned by:	arthuredelstein
Priority:	Very High	Milestone:	
Component:	Applications/Tor Browser	Version:	
Severity:	Critical	Keywords:	tbb-fingerprinting-os, TorBrowserTeam201610
Cc:	arthuredelstein, isis, mcs, brade	Actual Points:	
Parent ID:		Points:	
Reviewer:		Sponsor:	

Description

WebAudio allows you to write data to AudioBuffers and perform effects/manipulation/spectral analysis on them, and extract their contents.

If the underlying routines are OS supported, they may be fingerprintable. =>https://developer.mozilla.org/en-US/docs/Web_Audio_API

**Bug 1288359****Pref to disable Web Audio API****RESOLVED FIXED** in Firefox 51▼ **Status** (bug RESOLVED as FIXED for Firefox 51)

-

Actual results:

AudioContext, DynamicsCompressor and OscillatorNode are being used in the wild to "fingerprint" web users, as reported by researchers: <https://techcrunch.com/2016/05/19/audio-fingerprinting-being-used-to-track-web-users-study-finds/>

(online tests here: <https://browserprint.info>)

Expected results:

An option to disable Web Audio is needed to protect users privacy. It appears this preference was introduced as "media.audio_data.enabled" per this request https://bugzilla.mozilla.org/show_bug.cgi?id=665598 but it's not available anymore in the latest versions of Firefox.

Firefox adds a preference to
disable Web Audio API

Fingerprinting Protection Mode

Lloyd Dewolf edited this page on Jan 29 · 11 revisions

Fingerprinting methods blocked in Fingerprinting Protection Mode

- [Canvas fingerprinting](#)
- [WebGL fingerprinting](#)
- [AudioContext fingerprinting](#)
- [WebRTC IP leakage](#)
- [Battery Status fingerprinting](#) (disabled in general, not just when FP mode is turned on)

<> Code

! Issues 23

🔗 Pull requests 2

📊 Projects 0

📖 Wiki

⚡ Pulse

📈 Graphs

Added Audio key #156

**Closed**

cberescu wants to merge 3 commits into Valve:master from cberescu:master



Conversation 14



Commits 3



Files changed 3



cberescu commented on May 23, 2016

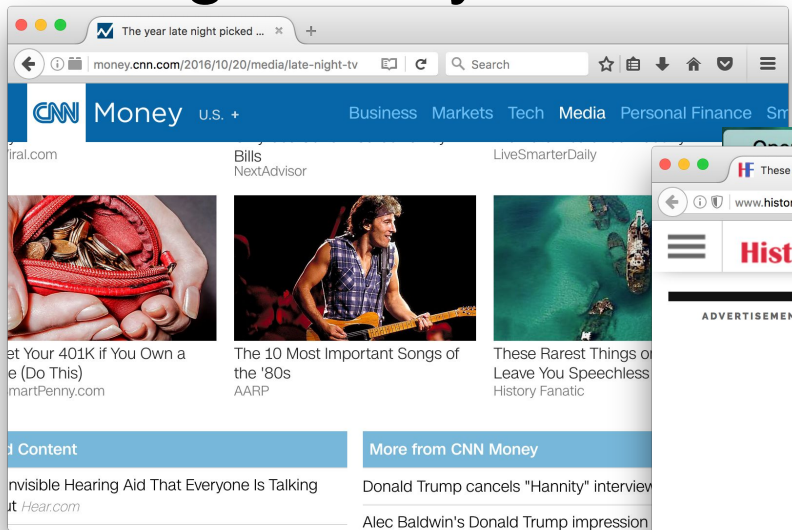


Using the OfflineAudioContext API created a key for the fingerprint. The mobile support for the OfflineAudioContext is not great so it jumps over if mobile is detected.

Tested also the AudioContext API but does not work all the time.

Code and info about Audio Fingerprint found here: <https://audiofingerprint.openwpm.com/>

Using Battery Status to Track

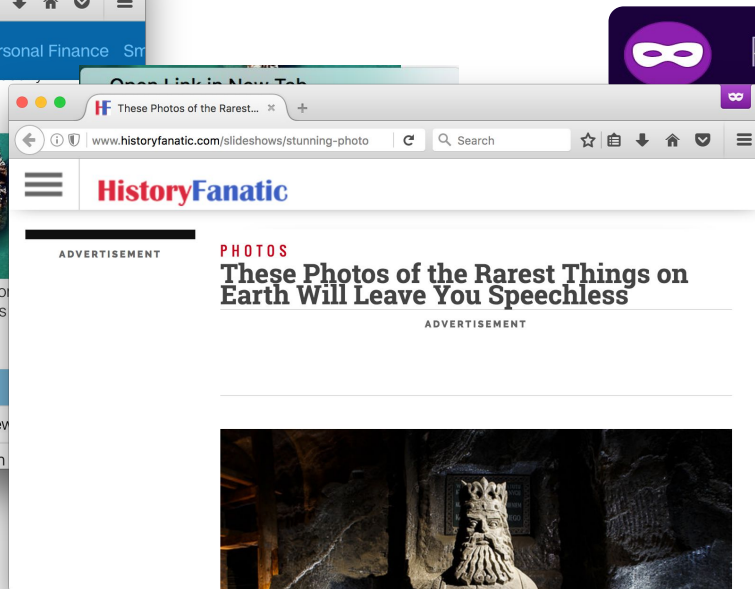


Battery Status:

level: 0.11

dischargeTime: 12867

The Leaking Battery, Olejnik et. al. (2015)

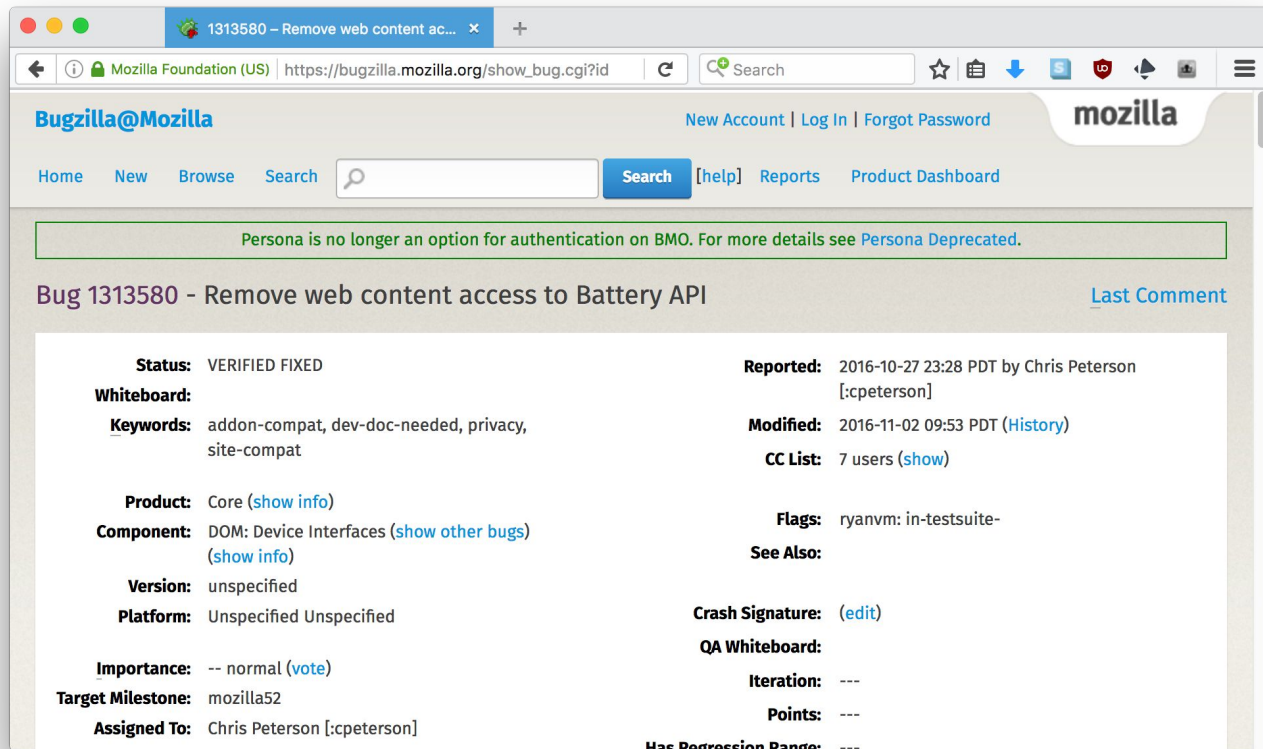


Battery Status:

level: 0.11

dischargeTime: 12867

Browsers remove BatteryStatus API citing privacy



The screenshot shows a web browser window displaying the Mozilla Bugzilla page for bug 1313580. The browser's address bar shows the URL `https://bugzilla.mozilla.org/show_bug.cgi?id=1313580`. The page header includes the Mozilla logo and navigation links like 'New Account', 'Log In', and 'Forgot Password'. Below the header is a search bar and a navigation menu with links for 'Home', 'New', 'Browse', 'Search', '[help]', 'Reports', and 'Product Dashboard'. A green banner message states: 'Persona is no longer an option for authentication on BMO. For more details see [Persona Deprecated](#).' The main heading of the bug is 'Bug 1313580 - Remove web content access to Battery API', with a link to 'Last Comment' on the right. The bug details are organized into two columns. The left column contains: 'Status: VERIFIED FIXED', 'Whiteboard:', 'Keywords: addon-compat, dev-doc-needed, privacy, site-compat', 'Product: Core (show info)', 'Component: DOM: Device Interfaces (show other bugs) (show info)', 'Version: unspecified', 'Platform: Unspecified Unspecified', 'Importance: -- normal (vote)', 'Target Milestone: mozilla52', and 'Assigned To: Chris Peterson [cpeterson]'. The right column contains: 'Reported: 2016-10-27 23:28 PDT by Chris Peterson [cpeterson]', 'Modified: 2016-11-02 09:53 PDT (History)', 'CC List: 7 users (show)', 'Flags: ryanvm: in-testsuite-', 'See Also:', 'Crash Signature: (edit)', 'QA Whiteboard:', 'Iteration: ---', 'Points: ---', and 'Has Regression Range: ---'.

Bugzilla@Mozilla [New Account](#) | [Log In](#) | [Forgot Password](#) **mozilla**

[Home](#) [New](#) [Browse](#) [Search](#) [Search](#) [\[help\]](#) [Reports](#) [Product Dashboard](#)

Persona is no longer an option for authentication on BMO. For more details see [Persona Deprecated](#).

Bug 1313580 - Remove web content access to Battery API [Last Comment](#)

Status: VERIFIED FIXED	Reported: 2016-10-27 23:28 PDT by Chris Peterson [cpeterson]
Whiteboard:	Modified: 2016-11-02 09:53 PDT (History)
Keywords: addon-compat, dev-doc-needed, privacy, site-compat	CC List: 7 users (show)
Product: Core (show info)	Flags: ryanvm: in-testsuite-
Component: DOM: Device Interfaces (show other bugs) (show info)	See Also:
Version: unspecified	Crash Signature: (edit)
Platform: Unspecified Unspecified	QA Whiteboard:
Importance: -- normal (vote)	Iteration: ---
Target Milestone: mozilla52	Points: ---
Assigned To: Chris Peterson [cpeterson]	Has Regression Range: ---

Browsers remove BatteryStatus API citing privacy

The image shows two overlapping browser windows. The background window displays the Mozilla Bugzilla page for Bug 1313580, titled "Remove web content ac...". The foreground window displays the Mozilla Bugzilla page for Bug 164213, titled "Remove Battery Status API from the tree".

Bug 164213 - Remove Battery Status API from the tree

Status: RESOLVED FIXED

Product: WebKit

Component: WebKit Misc.

Version: WebKit Nightly Build

Platform: Unspecified Unspecified

Importance: P2 Normal

Assigned To: Alex Christensen

URL:

Keywords:

Depends on:

Blocks:

Reported: 2016-10-30 20:26 PDT by Brady Eidson

Modified: 2016-11-02 14:32 PDT ([History](#))

CC List: 8 users ([show](#))

See Also: [129040](#)

Show dependency [tree](#) / [graph](#)

Browsers remove BatteryStatus API citing privacy

The screenshot shows a web browser with two tabs. The active tab is titled 'Bug 164213 - Remove Battery Stat...' and displays a bug report from the Brave browser. The browser's address bar shows the URL 'https://bugs.webkit.org/show_bug.cgi?id=164213'. The bug report page has a header with the Brave logo and navigation links. Below the header, there are statistics for the bug: 2,040 issues, 21 pull requests, 7 projects, and a Wiki link. The main content of the bug report is titled 'Disable support for Battery Status API' and describes a fix for issue #1885, stating that battery status is disabled for all users due to privacy concerns, citing Firefox and Chromium. The bug report also lists the auditor as @bbondy and mentions that the test plan is covered by automated tests. The bottom of the screenshot shows a commit history for the bug, with a commit by diredeltas on Dec 2, 2016.

1313580 - Remove web content ac... x +

Mozilla Foundation (US) | https://bugzilla.mozilla.org/show_bug.cgi?id=1313580 | Search

Bugzilla@Mozilla

Home New Browse

Bug 1313580 - Remove web content ac...

Status: VERIFIED

Whiteboard:

Keywords: addo site-

Product: Core

Component: DOM

Version: unsp

Platform: Unsp

Importance: -- no

Target Milestone: mozi

Assigned To: Chris

bugzilla.mozilla.org

Home | New | Browse | Search

« First Last » | « Prev Next »

Bug 164213 - Remove Battery Status API

Status: RESOLVED

Product: WebKit

Component: WebKit

Version: WebKit

Platform: Unspecified

Importance: P2 Normal

Assigned To: Alex Ch

URL:

Keywords:

Depends on:

Blocks:

Show details

brave / browser-laptop

Watch

<> Code

Issues 2,040

Pull requests 21

Projects 7

Wiki

Pulse

Disable support for Battery Status API

Fix #1885. Battery status is disabled for all users (not just those with Fingerprinting Protection) because Firefox and v c.f. <https://bugs.chromium.org/p/chromium/issues/detail?id=661792>.

Auditors: @bbondy

Test Plan: covered by automated test

master v0.14.2dev-RC2 ... v0.12.15dev-RC1

diredeltas committed on Dec 2, 2016

1 parent b0d4e0e

Battery Status Not Included: Assessing Privacy in Web Standards

Lukasz Olejnik
University College London
lukasz.w3c@gmail.com
<https://lukaszolejnik.com>

Steven Englehardt
Princeton University
ste@cs.princeton.edu
<http://senglehardt.com>

Arvind Narayanan
Princeton University
arvindn@cs.princeton.edu
<http://randomwalker.info>

Abstract—The standardization process is core to the development of the open web. Until 2013, the process rarely included privacy review and had no formal privacy requirements. But today the importance of privacy engineering has become apparent to standards bodies such as the W3C as well as to browser vendors. Standards groups now have guidelines for privacy assessments, and are including privacy reviews in many new specifications. However, the standards community does not yet have much practical experience in assessing privacy.

In this paper we systematically analyze the W3C Battery Status API to help inform future privacy assessments. We begin by reviewing its evolution — the initial specification, which only cursorily addressed privacy, the discovery of surprising privacy vulnerabilities as well as actual misuse in the wild, followed by the removal of the API from major browser engines, an unprecedented move. Next, we analyze web measurement data from late 2016 and confirm that the majority of scripts used the API for fingerprinting. Finally, we draw lessons from this affair and make recommendations for improving privacy engineering of web standards.

I. INTRODUCTION

The Battery Status API offers an interesting and unusual case study of privacy assessment in the web standardization

formally defined recommendations for privacy assessments during the development of internet protocols [1]. The W3C has also invested considerable resources into the creation of specialized methodologies for such privacy assessments [2], [3]. This includes taking public stances on privacy expectations [4] and defining new groups and processes to evaluate privacy [5], [6], [7]. Even the frameworks used during specification, development, and publishing have been updated to encourage all specification authors to include privacy and security review sections [8], [9].

These recent advances in privacy review are timely, as new and proposed web features will provide websites with much deeper access to the user’s device and environment, especially on smartphones and Internet-of-Things (IoT) devices. Examples include Bluetooth connectivity [10], low-level device sensors such as ambient light, acceleration, and vibration [11], [12], and even the user’s interpupillary distance, in the context of Virtual Reality [13].

But why should standards consider privacy at all, rather than leave it to browser vendors? Perhaps the market will then allow each user to choose a browser that provides his

Summary:

1. Transparency helps good and bad actors

- a. Vendors prioritize fixes
- b. Major trackers react to public pressure
- c. Less known trackers start using the technique

2. Non-technical users may be left behind

- a. Privacy protection often ends up in products for technical users
 - i. Tor Browser
 - ii. Brave Browser
- b. Solutions for non-technical users may take years, leaving them at a potential disadvantage

A Path Forward

1. **Take continual measurements**

- a. Princeton Web Census data collected monthly
- b. Data interface under development (ask us for an invite)

2. **Provide up-to-date results**

- a. Make updated fingerprinting script lists available

3. **Research measurement-informed privacy solutions**

- a. Sandboxing trackers and fingerprinters (my talk tomorrow)
- b. Automated fingerprint detection