

The Web Privacy Problem is a Transparency Problem

How OpenWPM and the Transparency Census will bring transparency to the web.

Steven Englehardt
@s_englehardt



webtap.princeton.edu



The New York Times – Breaking News, World News & Multimedia

The New York Times

TODAY'S PAPER | TIMES TOPICS

Get Home Delivery-Bay Area | Login | Register | Help | TimesPeople

see what's inside ▾

Advertising

Switch to Global Edition +

JOBS REAL ESTATE AUTOS ALL CLASSIFIEDS

WORLD U.S. POLITICS NEW YORK BUSINESS DEALBOOK TECHNOLOGY SPORTS SCIENCE HEALTH OPINION ARTS Books Movies Music Television Theater STYLE Dining & Wine Fashion & Style Home & Garden Weddings Celebrations TRAVEL

All Blogs Cartoons Classifieds Corrections Crossword Games Education First Look Learning Network Multimedia NYC Guide Obituaries Podcasts Photo Editor Sunday Magazine T Magazine Video

Wednesday, February 1, 2011 Last Update: 10:18 PM ET

Follow Us: [Facebook](#) [Twitter](#) [RSS](#) [Email](#)

Search [electric](#) [orange](#)

Subscribe to Home Delivery-Bay Area | Personalize Your Weather

OPINION ▾

EGYPT

Kristof: The White House should support the protesters in Egypt. [Blog: Mubarak's Speech](#) [Follow on Twitter](#) » [Brookings: In the democratic brave, a quest for dignity.](#) [Cohen: From a culture of entitlement to one of self-empowerment.](#) [Senator John Kerry: A future without Mubarak.](#)

MORE IN OPINION

Editorial: Debt Limit [Bloggingheads: Abortions](#)

MARKETS ▾

JAPAN

Nikkei	15,470.50	23,826.38	+343.44	Closed
Hang Seng	2,798.90	+1.48%	+1.48%	Up

1.30% YIELD SAVING ACCOUNT

GET QUOTES My Portfolio +

Stock, ETFs, Funds Go

Path to Change in Power Still Unclear

By KAREEM FAIRID and ANTHONY SHADID 54 minutes ago

Hundreds of thousands of Egyptians traveled like pilgrims to speak freely and to be heard.

President Hosni Mubarak's vow to step down in the fall was not enough for the hundreds of thousands who poured into Tahrir Square.

[Post a Comment](#) | Read (332)

King of Jordan Dismisses His Cabinet

Antiquities Chief Says Sites Are Largely Secure

New Service Lets Voices From Egypt Be Heard

The Leader: Updates From Day 8

INTERACTIVE FEATURE: A Timeline of Mubarak's Presidency

A chronology of President Hosni Mubarak's 30-year rule in Egypt.

[More Photographs](#)

[Interactive Map: Protests](#)

iMeet built a meeting room just for you.

CHARLES EBBY

30-DAY FREE TRIAL

Advertising

Log In With Facebook

Log in to see what your friends are sharing on nytimes.com.

Privacy Policy | What's This?

WHAT'S POPULAR NOW

The Paradox of Corporate Taxes in America

Why Not Regulate Seriously

Social

1.30% YIELD SAVING ACCOUNT

GET QUOTES My Portfolio +

Stock, ETFs, Funds Go

Path to Change in Power Still Unclear

By KAREEM FAIRID and ANTHONY SHADID 54 minutes ago

Hundreds of thousands of Egyptians traveled like pilgrims to speak freely and to be heard.

President Hosni Mubarak's vow to step down in the fall was not enough for the hundreds of thousands who poured into Tahrir Square.

[Post a Comment](#) | Read (332)

King of Jordan Dismisses His Cabinet

Antiquities Chief Says Sites Are Largely Secure

New Service Lets Voices From Egypt Be Heard

The Leader: Updates From Day 8

INTERACTIVE FEATURE: A Timeline of Mubarak's Presidency

A chronology of President Hosni Mubarak's 30-year rule in Egypt.

[More Photographs](#)

[Interactive Map: Protests](#)

iMeet built a meeting room just for you.

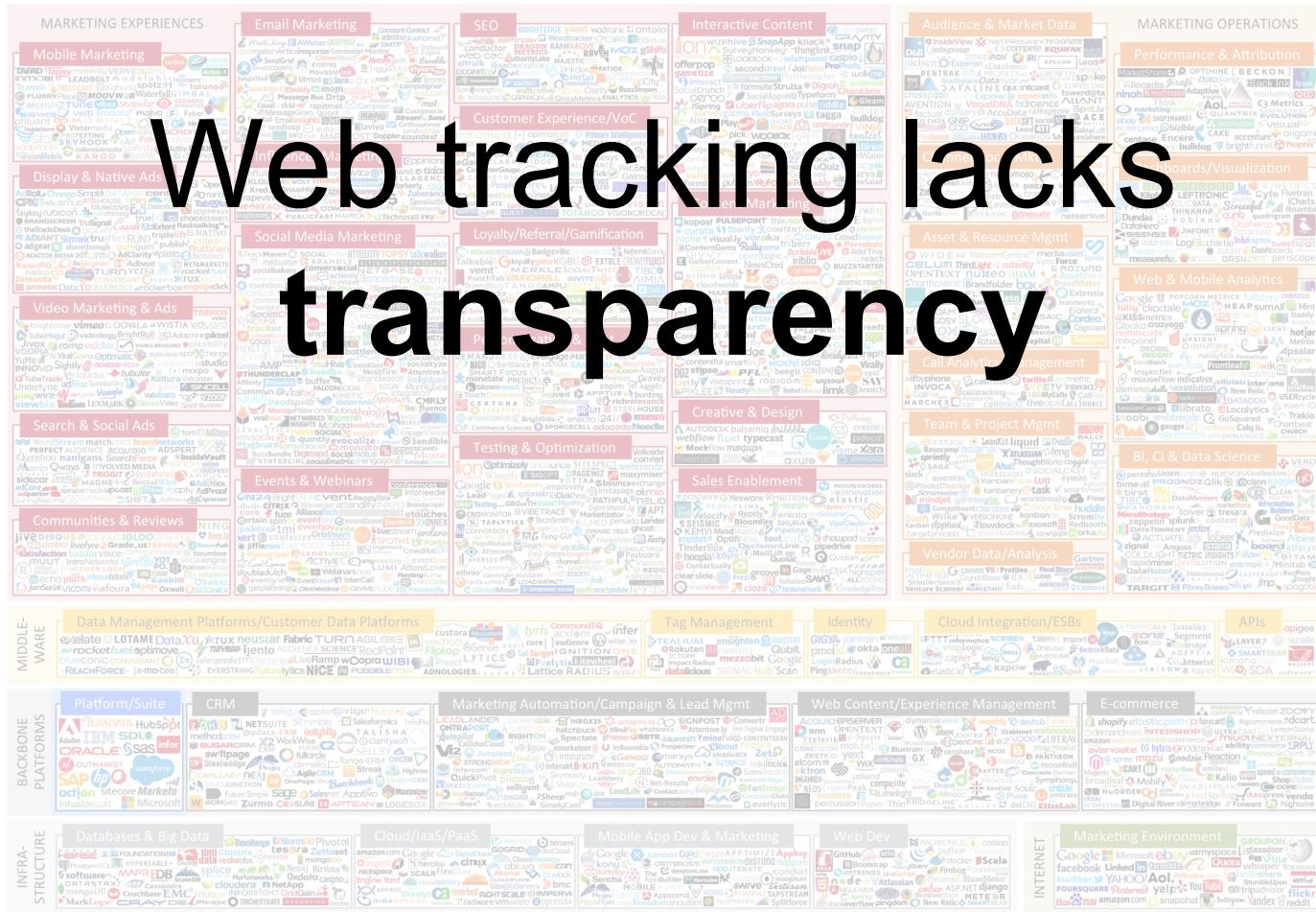
CHARLES EBBY

30-DAY FREE TRIAL

Advertising

Source: Mayer & Mitchell; Third-Party Web Tracking: Policy and Technology





Web tracking lacks transparency

...but we are changing that

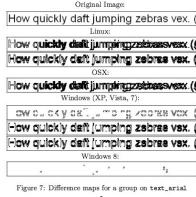


Web tracking lacks transparency

...but we are changing that
(and I'll show you how we already have)



Transparency encourages best practices

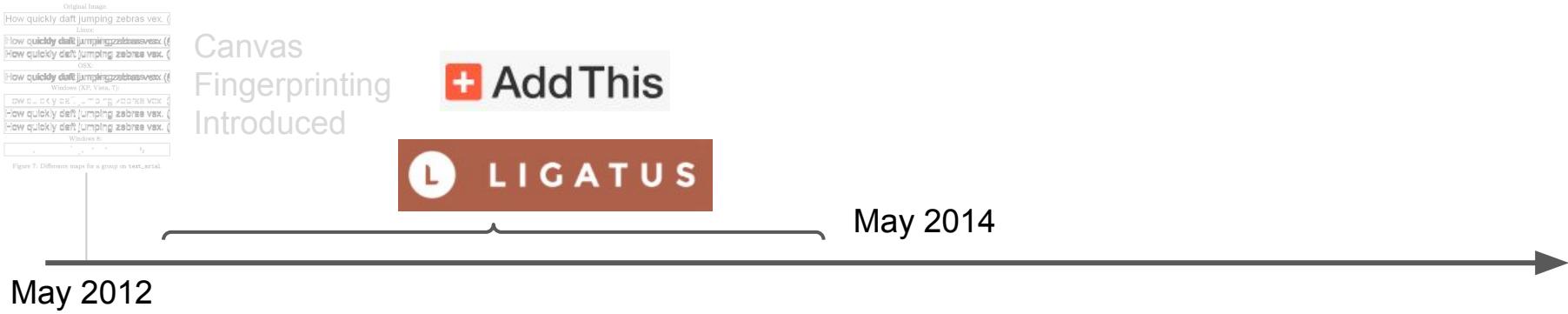


Canvas
Fingerprinting
Introduced

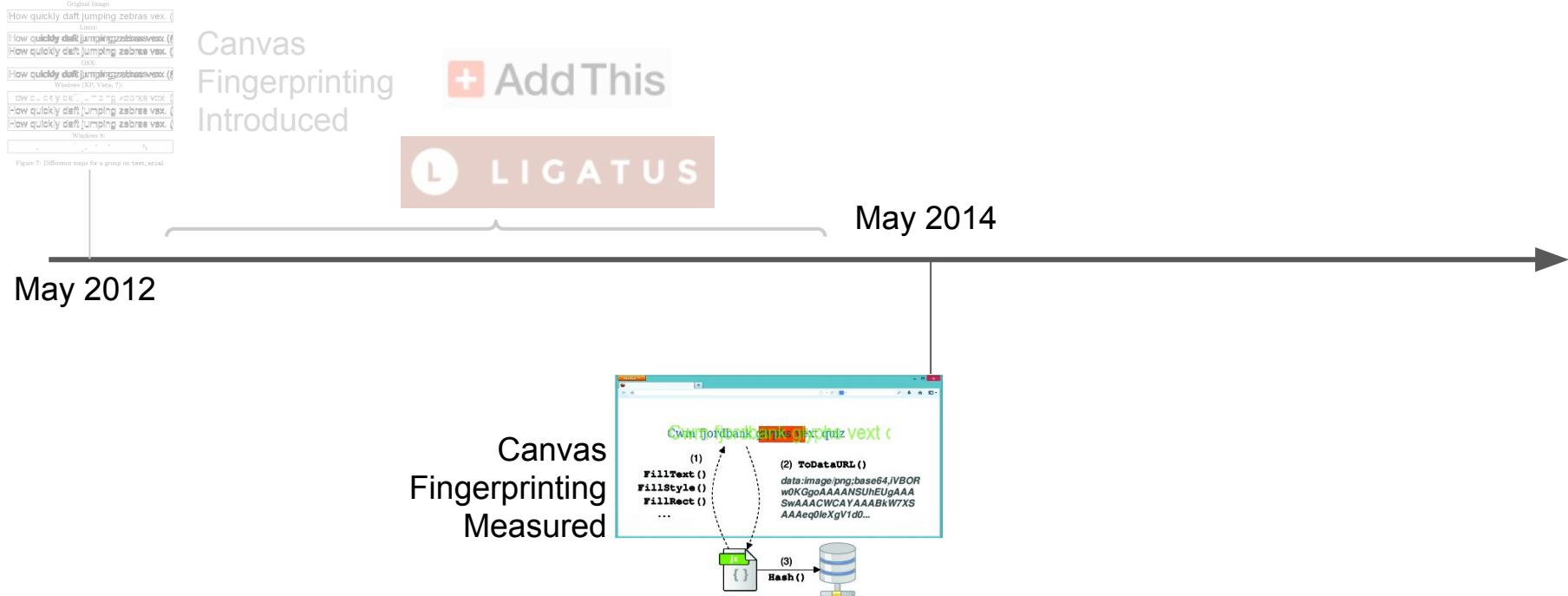
May 2012



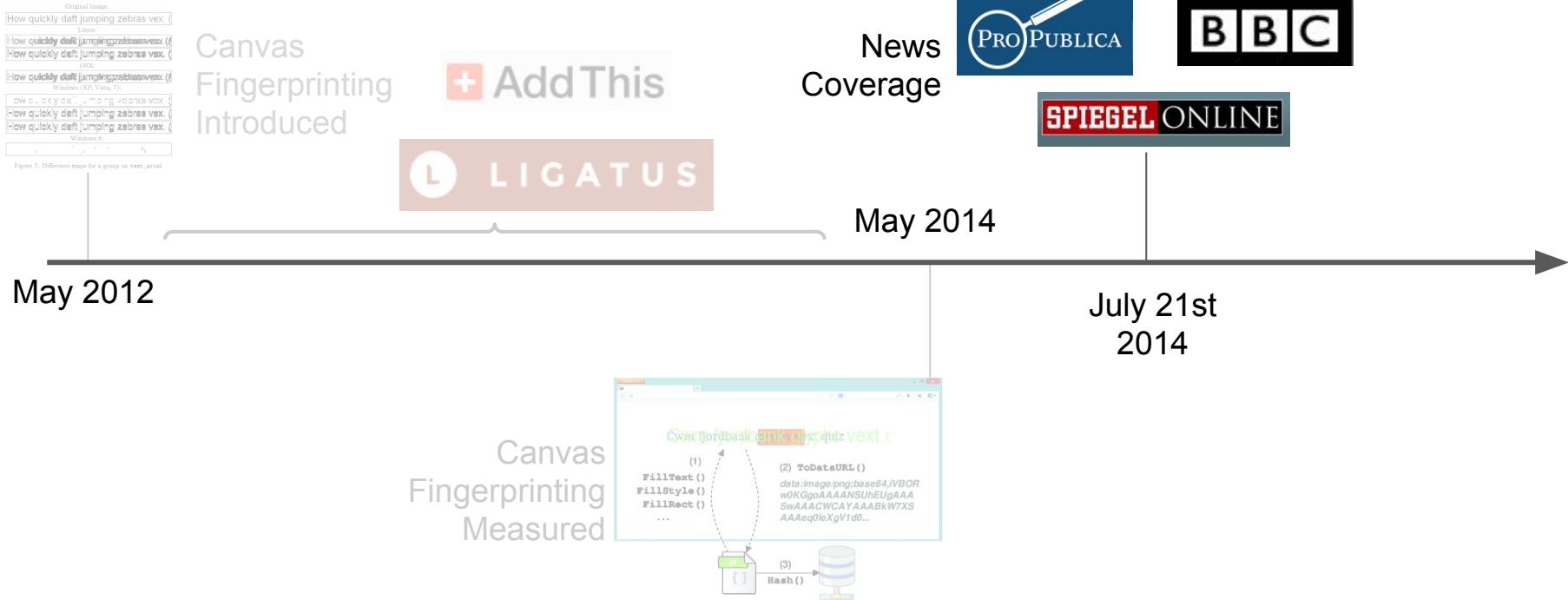
Transparency encourages best practices



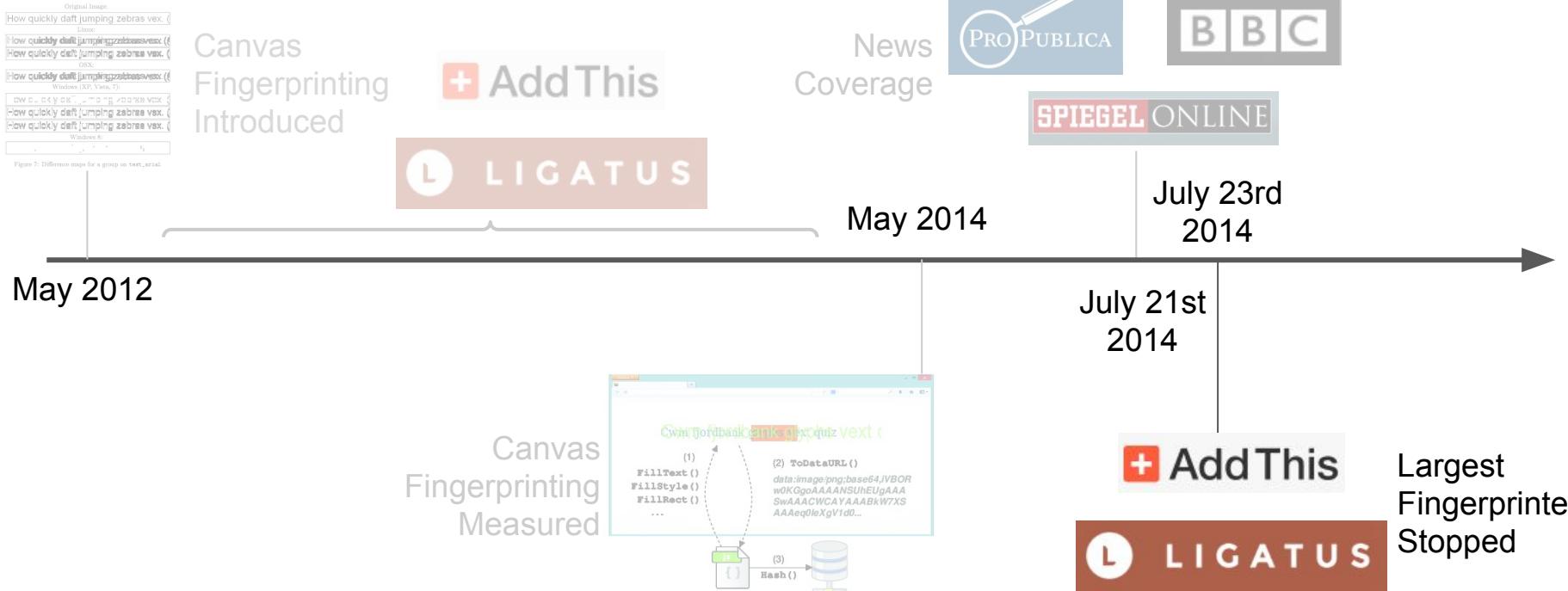
Transparency encourages best practices



Transparency encourages best practices



Transparency encourages best practices



Canvas Fingerprinting was a known
technique for **2 years**.



In just **2 months** following our
measurement work the largest users of
canvas fingerprinting stopped.



Why?

Our measurement work removed **information asymmetry** between trackers and the rest of the web.

Our measurement work removed **information asymmetry** between trackers and the rest of the web.

The screenshot shows the ProPublica website. At the top is a dark blue header with the ProPublica logo (magnifying glass icon) and the text "PRO PUBLICA | Journalism in the Public Interest". Below the header is a light gray navigation bar with links for "Home", "Our Investigations", "Data", "MuckReads", "Get Involved", and "About Us". To the right of the navigation bar is a search bar labeled "Search ProPublica".

Dragnets

Tracking Censorship and Surveillance

Meet the Online Tracking Device That is Virtually Impossible to Block

A new kind of tracking tool, canvas fingerprinting, is being used to follow visitors to thousands of top websites, from WhiteHouse.gov to YouPorn.

by [Julia Angwin](#)

ProPublica, July 21, 2014, 8 a.m.

This is part of an ongoing

[Dragnets](#)

Our measurement work removed **information asymmetry** between trackers and the rest of the web.

The screenshot shows a news article from BBC News. At the top, there's a navigation bar with the BBC logo, a sign-in link, and categories like News, Sport, Weather, Shop, Earth, Travel, and more. Below that is a red banner with the word "NEWS". On the left, there's a sidebar for "ProPublica" with links for "Home" and "Our Investigations". A main headline reads "Meet the Dragnets Tracking Censorship and". The main content starts with "Virtually Browser 'fingerprints' help track users". Below the headline, it says "A new kind of tracking WhiteHouse.gov to Y" and "by Julia Angwin ProPublica, July 21, 2014, t". There's also a small image of the White House dome at the bottom.

Our measurement work removed **information asymmetry** between trackers and the rest of the web.

Xavier Ashe @XavierAshe · 24 Jul 2014
Remove AddThis from your website right now - > Clear Your Cookies? You Can't Escape Canvas Fingerprinting bit.ly/1pKLih

fingerprinting scripts
avias fingerprinting scripts on their homepages. Please consult canvas fingerprinting and other advanced tracking mechanisms

Clear Your Cookies? You Can't Escape Canvas Fing...

The web really never forgets, canvas fingerprinting code allows websites you track you regardless of your privacy settings or if you clear cookies/cache.
darknet.org.uk

1

Our measurement work removed **information asymmetry** between trackers and the rest of the web.

Xavier Ashe @XavierAshe · 24 Jul 2014

James King @JamesKingMe · 24 Jul 2014
Canvas fingerprinting is a technological way of stalking

allows websites you track you regardless of your privacy settings or if you clear cookies/cache.
darknet.org.uk

Our measurement work removed **information asymmetry** between trackers and the rest of the web.

Xavier Ashe @XavierAshe · 24 Jul 2014

James King @JamesKingMe · 24 Jul 2014

tom @tomquas · 24 Jul 2014
oh look, t-online.de uses **canvas fingerprinting** and violates my privacy. the company behind de-mail... source: securehomes.esat.kuleuven.be/~gacar/persist... #wtf

Our measurement work removed **information asymmetry** between trackers and the rest of the web.



Our measurement work removed **information asymmetry** between trackers and the rest of the web.

Xavier Ashe @XavierAshe · 24 Jul 2014

James King @JamesKingMe · 24 Jul 2014

tom @tomquas · 24 Jul 2014

Jstheater @jstheater · 22 Jul 2014

Alex Sherer @cymcyms · 22 Jul 2014

I feel gross, because I used @addthis to share this article. But, everyone should know about #canvasfingerprinting news.genius.com/Gizmodo-what-y...

Tor Works: 2

Tor client n path to over. Green: splend, red: clear:

Gizmodo – What You Need to Know About the Sne...

What do the White House and YouPorn have in common? Their websites both use canvas fingerprinting, a newer form of online tracking designed to make it har...

genius.com

Our measurement work removed **information asymmetry** between trackers and the rest of the web.

The screenshot shows a Mozilla Support Forum page. At the top, there's a sidebar with user profiles and a main header with the text "mozilla support". Below the header, a breadcrumb navigation shows "HOME > SUPPORT FORUM > FIREFOX > WHY IS FIREFOX ALLOWING "CANVAS ...". The main title "Support Forum" is displayed prominently. A yellow banner at the top of the main content area states, "This thread was archived. Please ask a new question if you need help." The main post title is "Why is FireFox allowing "Canvas Fingerprinting" to track me?". Below the title, it says "2 REPLIES", "14 HAVE THIS PROBLEM", "1010 VIEWS", "LAST REPLY BY JSCHER2000 1 YEAR AGO". A partial view of a reply from a user named "Xavier" is visible on the right, mentioning "one should it-y...". Another partial reply from "The Sne..." is also visible. At the bottom, a user with a blue profile picture asks, "Since "Canvas Fingerprinting" seems to be the new way of tracking one. How does one TURN IT OFF! ---".

Information asymmetry not just between trackers and users.

Dragnets
Tracking Censorship and Surveillance

Meet the Online Tracking Device That is Virtually Impossible to Block

A new kind of tracking tool, canvas fingerprinting, is being used to follow visitors to thousands of top websites, from WhiteHouse.gov to YouPorn.

by [Julia Angwin](#)
ProPublica, July 21, 2014, 8 a.m.

Update: After this article was published, YouPorn contacted us to say it had removed AddThis technology from its website, saying that the website was "completely unaware that AddThis contained a tracking software that had the potential to jeopardize the privacy of our users." A spokeswoman for the German digital marketer Ligatus also said that is no longer running its test of canvas fingerprinting, and that it has no plans to use it in the future.

This story was co-published with [Mashable](#).

A new, extremely persistent type of online tracking is shadowing visitors to thousands of

- IN THIS ISSUE
- [An Unbelievable Story](#)
 - [Devils, Deals and the D](#)
 - [Why Small Debts Matter](#)
 - [Lives](#)
 - [Out of Options, California Troubled Children Out of](#)
 - ['All of This Because So](#)
 - [Work'](#)
 - [Small-Scale Violations](#)
 - [Cause the Most Harm](#)

"YouPorn contacted us to say...'[the website was] completely unaware that AddThis contained a tracking software..."

Transparency is effective at returning control to users and publishers

Automated, large-scale measurements can provide this transparency

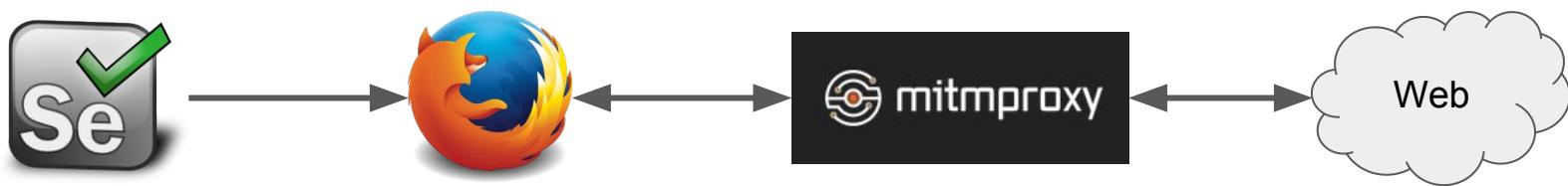
We're doing three things to help:

1. Developing OpenWPM
2. Running monthly, 1 million site measurements
3. Building an analysis layer on top of the data

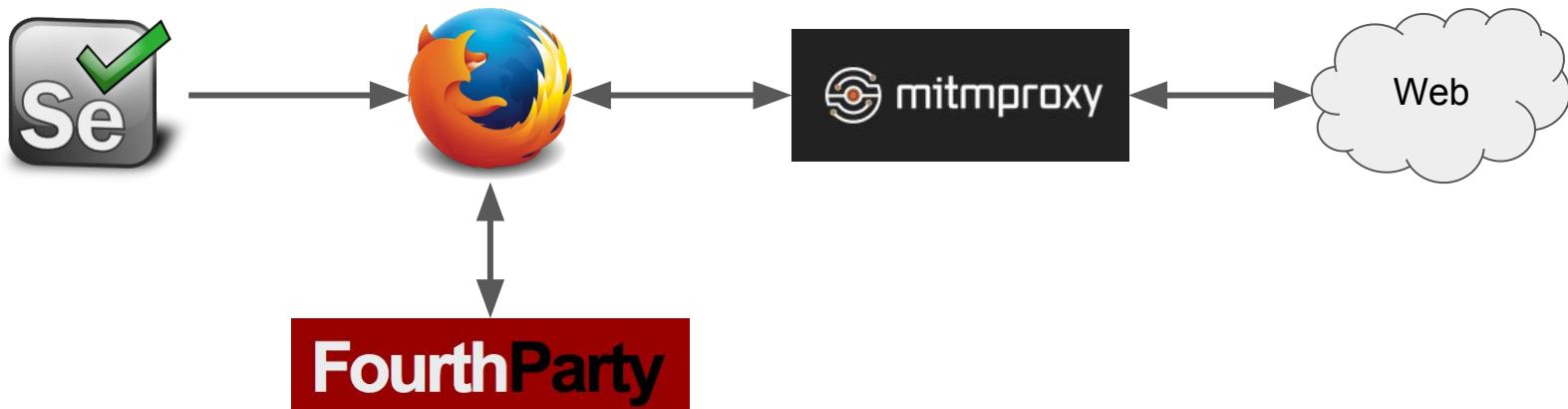
OpenWPM



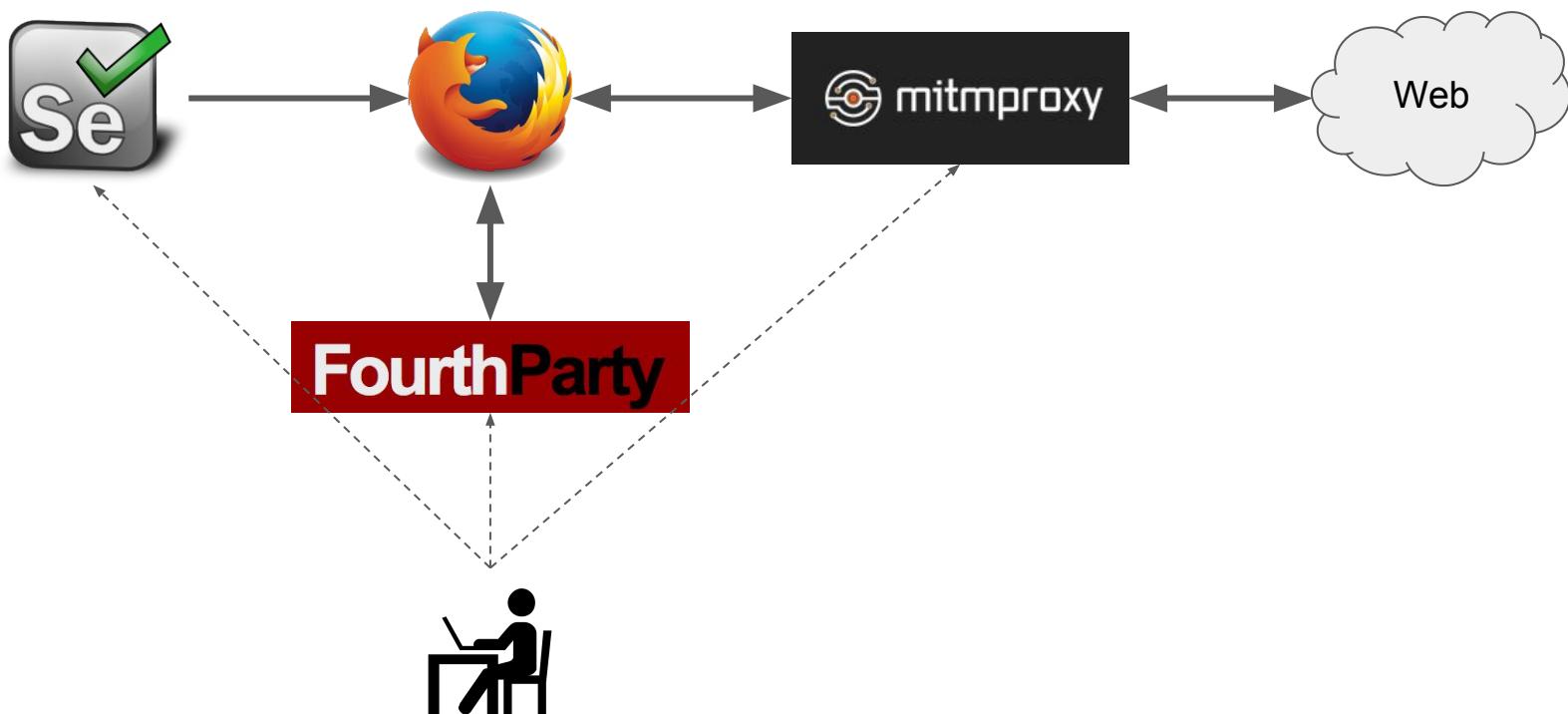
OpenWPM



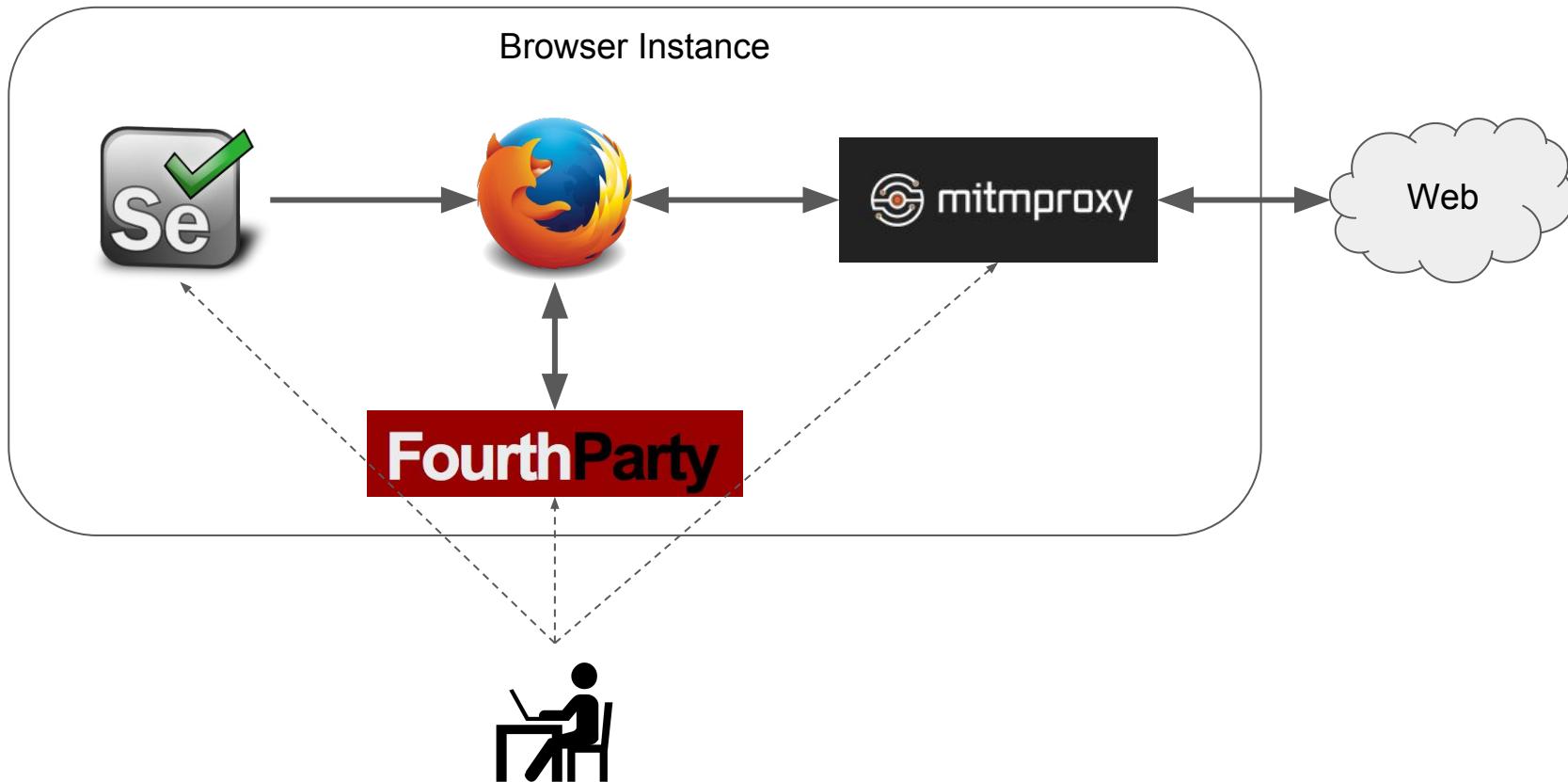
OpenWPM

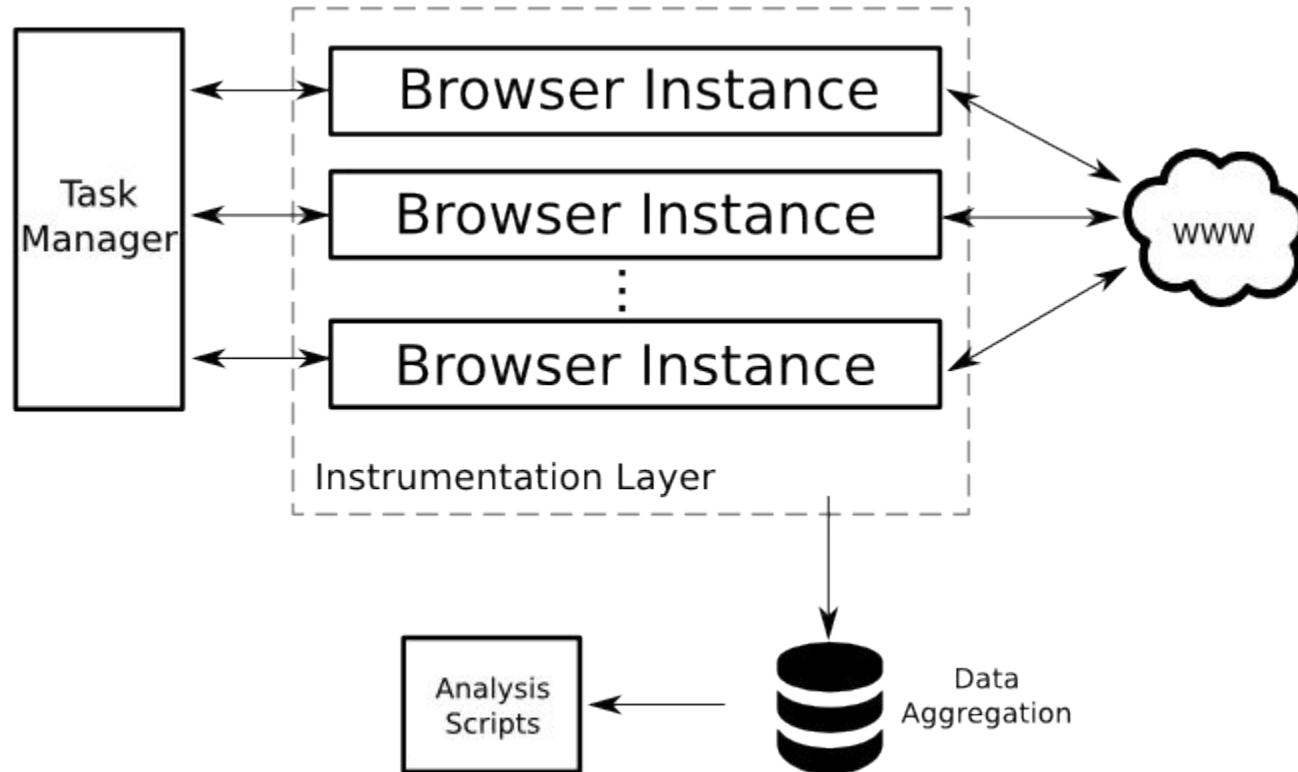


OpenWPM



OpenWPM





OpenWPM supports browsing with persistent state

- Browser can keep profile through crashes and freezes
 - Cookie setting over a session
 - Cookie synchronization (id sharing)
 - Zombie Cookies

OpenWPM uses a real browser

- Extensions
 - AdBlock Plus, Ghostery, ...
- Privacy Features
 - Block third-party cookies, FF tracking protection, ...
- Support for new web technologies
 - WebRTC, Audio, Video, WebGL

OpenWPM is already used by at least 7 research groups

- At Princeton
 - 4 published studies and several ongoing
- Ongoing Research
 - Columbia University
- In published studies:
 - The Web Privacy Census (UC Berkeley / Berkeley Law)
 - Variations in Tracking in Relation to Geographic Location (CMU / RAND)
 - Forthcoming WWW'16 study by Nick Nikiforakis (Stony Brook)
- By journalists
- By regulators

The Web Transparency Census

Monthly
1 Million Site Crawl

The Web Transparency Census

Monthly
1 Million Site Crawl

Collecting:

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

Supporting a variety of measurements

1. Effectiveness of Privacy Tools

- Ghostery
- AdBlock Plus
- HTTPS Everywhere

Supporting a variety of measurements

1. Effectiveness of Privacy Tools

- Ghostery
- AdBlock Plus
- HTTPS Everywhere

2. Effectiveness Browser Protections

- DNT
- Third-party cookie Blocking
- Firefox Tracking Protection

Supporting a variety of measurements

1. Effectiveness of Privacy Tools

- Ghostery
- AdBlock Plus
- HTTPS Everywhere

2. Effectiveness Browser Protections

- DNT
- Third-party cookie Blocking
- Firefox Tracking Protection

3. Use of javascript for tracking

- Canvas Fingerprinting
- Property Enumeration
- WebRTC Local IP Sniffing

Supporting a variety of measurements

1. Effectiveness of Privacy Tools

- Ghostery
- AdBlock Plus
- HTTPS Everywhere

2. Effectiveness Browser Protections

- DNT
- Third-party cookie Blocking
- Firefox Tracking Protection

3. Use of javascript for tracking

- Canvas Fingerprinting
- Property Enumeration
- WebRTC Local IP Sniffing

4. Tracking Practices

- Cookie Syncing
- Cookie Respawning
- Setting ID cookies

Case Study 1: Canvas Fingerprinting

Case Study 2: WebRTC Local IP Sniffing

2012: Canvas Fingerprinting Introduced

Windows:

How quickly daft jumping zebras vex. (Also, pur
How quickly daft jumping zebras vex. (Also, pu

OS X:

How quickly daft jumping zebras vex. (Also, pu
How quickly daft jumping zebras vex. (Also, pu
How quickly daft jumping zebras vex. (Also, pu
How quickly daft jumping zebras vex. (Also, pu

Linux:

How quickly daft jumping zebras vex. (Also, pu
How quickly daft jumping zebras vex. (Also, pur
How quickly daft jumping zebras vex. (Also, p

Figure 6: 13 ways to render 20px Arial

Source: Mowery and Shacham; Pixel Perfect: Fingerprinting Canvas in HTML5

2012: Canvas Fingerprinting Introduced

Windows:

How quickly daft jumping zebras vex. (Also, pur
How quickly daft jumping zebras vex. (Also, pu

OS X:

How quickly daft jumping zebras vex. (Also, pu
How quickly daft jumping zebras vex. (Also, pu
How quickly daft jumping zebras vex. (Also, pu
How quickly daft jumping zebras vex. (Also, pu

Linux:

How quickly daft jumping zebras vex. (Also, pu
How quickly daft jumping zebras vex. (Also, pur
How quickly daft jumping zebras vex. (Also, p

Figure 6: 13 ways to render 20px Arial

Original Image:

How quickly daft jumping zebras vex. (

Linux:

How quickly daft jumping zebras vex. (

How quickly daft jumping zebras vex. (

OSX:

How quickly daft jumping zebras vex. (

Windows (XP, Vista, 7):

CW Q .. C\< Y ca T .. m P pg /zebras vex. (

(-How quickly daft jumping zebras vex. (

(-How quickly daft jumping zebras vex. (

Windows 8:

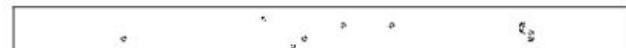


Figure 7: Difference maps for a group on `text_arial`

Source: Mowery and Shacham; Pixel Perfect: Fingerprinting Canvas in HTML5

2014: Canvas Fingerprinting Measured



The Web Never Forgets: Persistent Tracking Mechanisms in the Wild

Gunes Acar¹, Christian Eubank², Steven Englehardt², Marc Juarez¹
Arvind Narayanan², Claudia Diaz¹

¹KU Leuven, ESAT/COSIC and iMinds, Leuven, Belgium
{name.surname}@esat.kuleuven.be

²Princeton University
{cge,ste,arvindn}@cs.princeton.edu

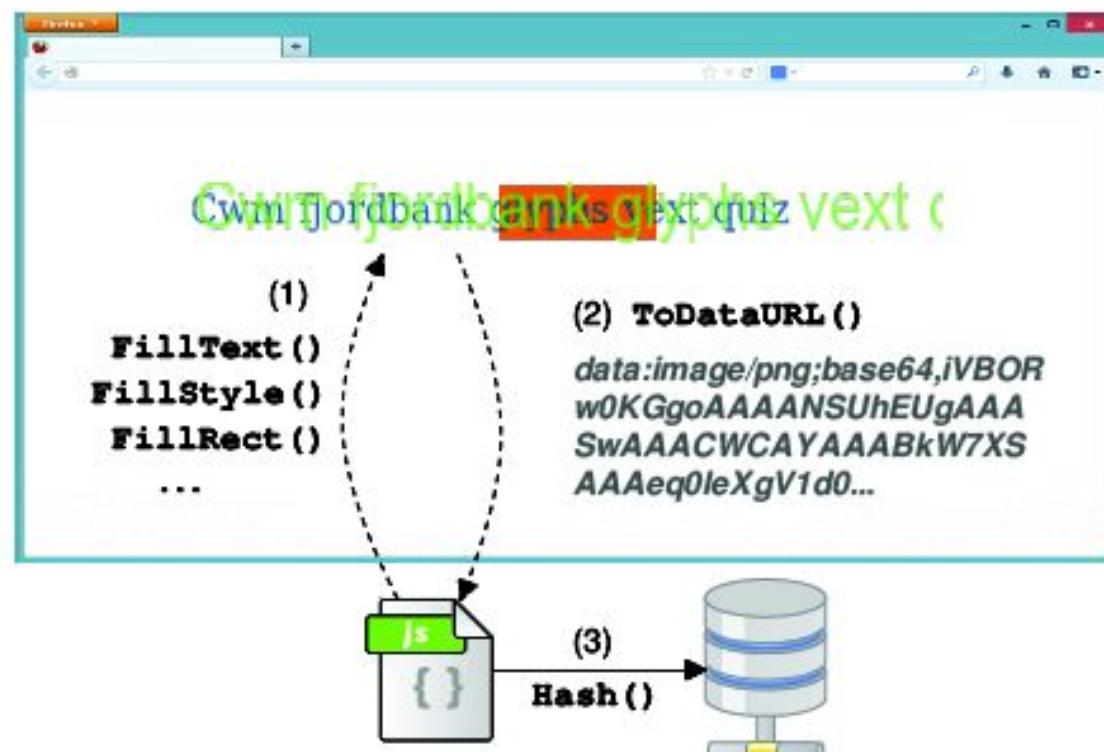
ABSTRACT

We present the first large-scale studies of three advanced web tracking mechanisms — canvas fingerprinting, evercookies and use of “cookie syncing” in conjunction with evercookies.

1. INTRODUCTION

A 1999 New York Times article called cookies comprehensive privacy invaders and described them as “surveillance

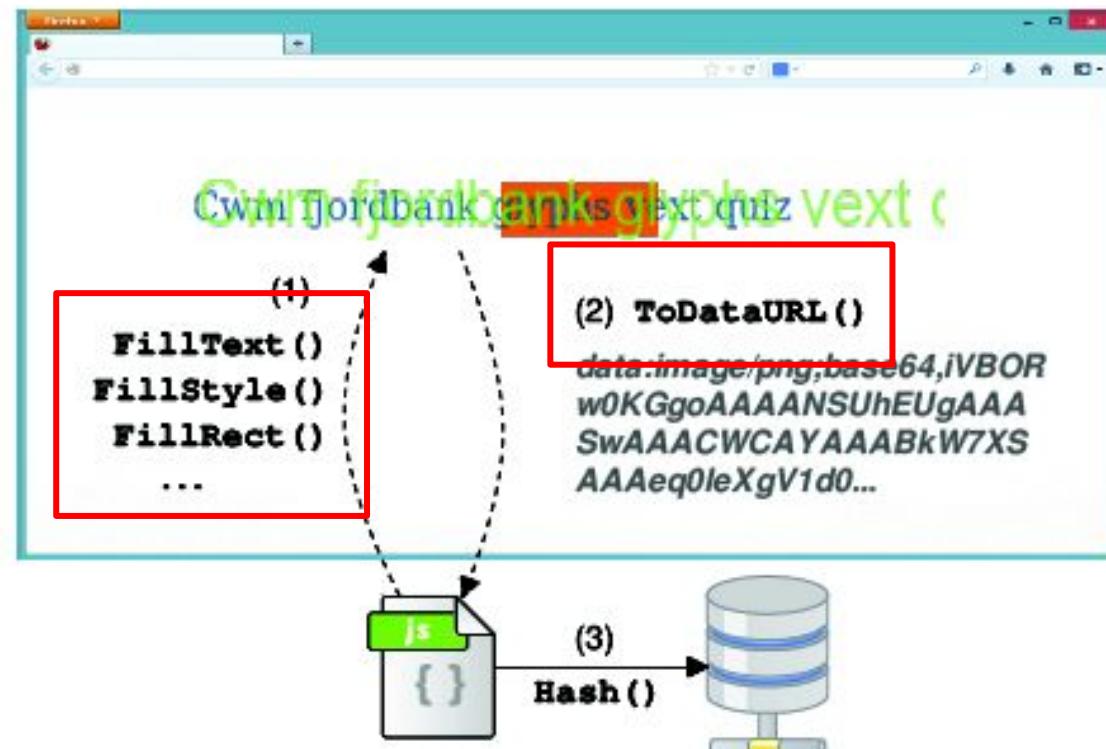
2014: Canvas Fingerprinting Measured



Source: Acar, Eubank, Englehardt, Juarez, Narayanan, Diaz; *The Web Never Forgets: Persistent Tracking Mechanisms in the Wild*

2014: Canvas Fingerprinting Measured

1. Write a Firefox patch
2. Write automation with Selenium
3. Write analysis code



Case Study 1: Canvas Fingerprinting

Case Study 2: WebRTC Local IP Sniffing

1. I saw a tweet that nytimes.com is IP sniffing



Mike O'Neill
@incloud



Follow

WebRTC being used now by embedded 3rd party on nytimes.com to report visitors' local IP addresses.

Request	Console	Debugger	Style Editor	Performance	Network	DOM	#	IP	Port	Method	Type	
							570	192.168.1.203	192.168.1.426	DNS	query operation,	
							548	0.00062285	192.168.1.203	192.168.1.254	DNS	Query Operation,
							543		192.168.1.119	239.255.255.258	SSDP	Request, Method:
							571	0.00000005	192.168.1.203	192.168.1.254	DNS	Request, Query I
							537	0.0148223	192.168.1.203	192.168.1.254	DNS	Query Operation,
							712		192.168.1.254	192.168.1.203	DNS	Response, ErrorCode:
							238		192.168.1.187	239.255.255.258	SSDP	Request, Method:
							859		192.168.1.187	239.255.255.258	SSDP	Request, Method:
							486		192.168.1.187	239.255.255.258	SSDP	Request, Method:
							529	0.0268478	192.168.1.203	192.168.1.254	DNS	Query Operation,
							908	0.00000005	192.168.1.203	192.168.1.254	DNS	Request, Query I
							366	0.0053289	192.168.1.203	192.168.1.254	DNS	Query Operation,
							238		192.168.1.254	192.168.1.203	DNS	Response, ErrorCode:
							836	0.0282345	192.168.1.203	192.168.1.254	DNS	Query Operation,
							420	0.0137929	192.168.1.203	192.168.1.254	DNS	Query Operation,
							680	0.0110006	192.168.1.203	192.168.1.254	DNS	Query Operation,
							545					

2. I added code to JS Instrumentation for next crawl

```
// Access to webRTC  
instrumentObject(window.mozRTCPeerConnection.prototype  
    "mozRTCPeerConnection",  
    prototype=true);
```

3. I wrote some analysis code

- Grab all urls that execute
 - mozRTCPeerConnection.onicecandidate
 - mozRTCPeerConnection.createDataChannel
 - mozRTCPeerConnection.createOffer
- Check JS Files to confirm

4. I found several third-parties sniffing local IP

- 121 first-party sites (October 2015)
 - 29 in the top 10k
- 24 unique scripts
- Only 1 of which is blocked by
EasyList/EasyPrivacy

Measurement with OpenWPM is much easier

Canvas Fingerprinting

WebRTC Local IP Sniffing

1. Write a Firefox patch
2. Write automation
with Selenium
3. Write analysis code

Measurement with OpenWPM is much easier

Canvas Fingerprinting

WebRTC Local IP Sniffing

1. ~~Write a Firefox patch~~ → 1. Write 1 line of JavaScript
2. Write automation
with Selenium
3. Write analysis code

Measurement with OpenWPM is much easier

Canvas Fingerprinting

1. Write a Firefox patch

2. Write automation
with Selenium

3. Write analysis code

WebRTC Local IP Sniffing

1. Write 1 line of JavaScript

2. Use OpenWPM

Measurement with OpenWPM is much easier

Canvas Fingerprinting

1. Write a Firefox patch

2. Write automation
with Selenium

3. Write analysis code

WebRTC Local IP Sniffing

1. Write 1 line of JavaScript

2. Use OpenWPM

3. Write analysis code

Where to go from here:

1. Inform the public
2. Provide data for privacy tools
3. Make data more accessible to less technical investigators

We'd like to collaborate with you

1. Submit pull requests for OpenWPM
2. Use OpenWPM to run measurements and release the data
3. Download our data and build analysis on top of it
 - a. (Coming soon!)

Help us make the web more transparent!

- Contribute:
 - github.com/citp/OpenWPM
- Collaborate:
 - webtap.princeton.edu

Email: ste@cs.princeton.edu

Twitter: [@s_englehardt](https://twitter.com/s_englehardt)