

# 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](http://webtap.princeton.edu)



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

The New York Times - Breaking... <http://www.nytimes.com/> Google

Home Page Computer World Most Popular Times Topics Get Home Delivery-Bay Area Log In Register Times People

**Advertising** **The New York Times** **Advertising**

Tuesday, February 1, 2011 Last Update: 10:16 PM ET

Search [select a foreign language](#) Follow Us [Subscribe to Home Delivery-Bay Area's Personalized Edition](#)

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  
Public Editor  
Sunday Magazine  
T Magazine  
Video

**Mubarak's Week New Term**  
**Opposition Demands That He Leave Sooner**

BY ANTHONY SHADID  
58 minutes ago

**Path to Change in Power Still Unclear**

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)

**Path to Change in Power Still Unclear**

By DAVID D. KIRKPATRICK and MARK LANDLER 1 minute ago

The democracy movement is unfolding so rapidly in Egypt that Washington came close to being left behind.

**Video**

**Quiet Acts of Protest on a Noisy Day**  
By KAREEM FAHIM and ANTHONY SHADID 54 minutes ago

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

King of Jordan Dismisses His Cabinet  
Antiquities Chief Says Sites Are Largely Secure  
New Service Lets Voices From Egypt Be Heard  
The Lede: 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**

**OPINION**  
EGYPT  
Kristof: The White House should support the protesters in Egypt.  
Blog: Mubarak's Speech  
Follow on Twitter  
Brooks: In the democratic wave, a quest for dignity.  
Cohen: From a culture of victimhood to one of self-empowerment.  
Senator John Kerry: A future without Mubarak.

**MORE IN OPINION**  
Editorial: Debt Limit  
Bloggingheads: Abortions

**MARKETS** As 10:14 PM ET

	Japan	Hong Kong	China
Nikkei	15,470.50	23,826.39	2,788.86
	+100.00	+343.44	Closed
	+1.01%	+1.48%	Hong Kong

Delayed at least 15 minutes

**GET QUOTES** My Portfolios  
Stock, ETFs, Funds  Go

**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 Social Media?

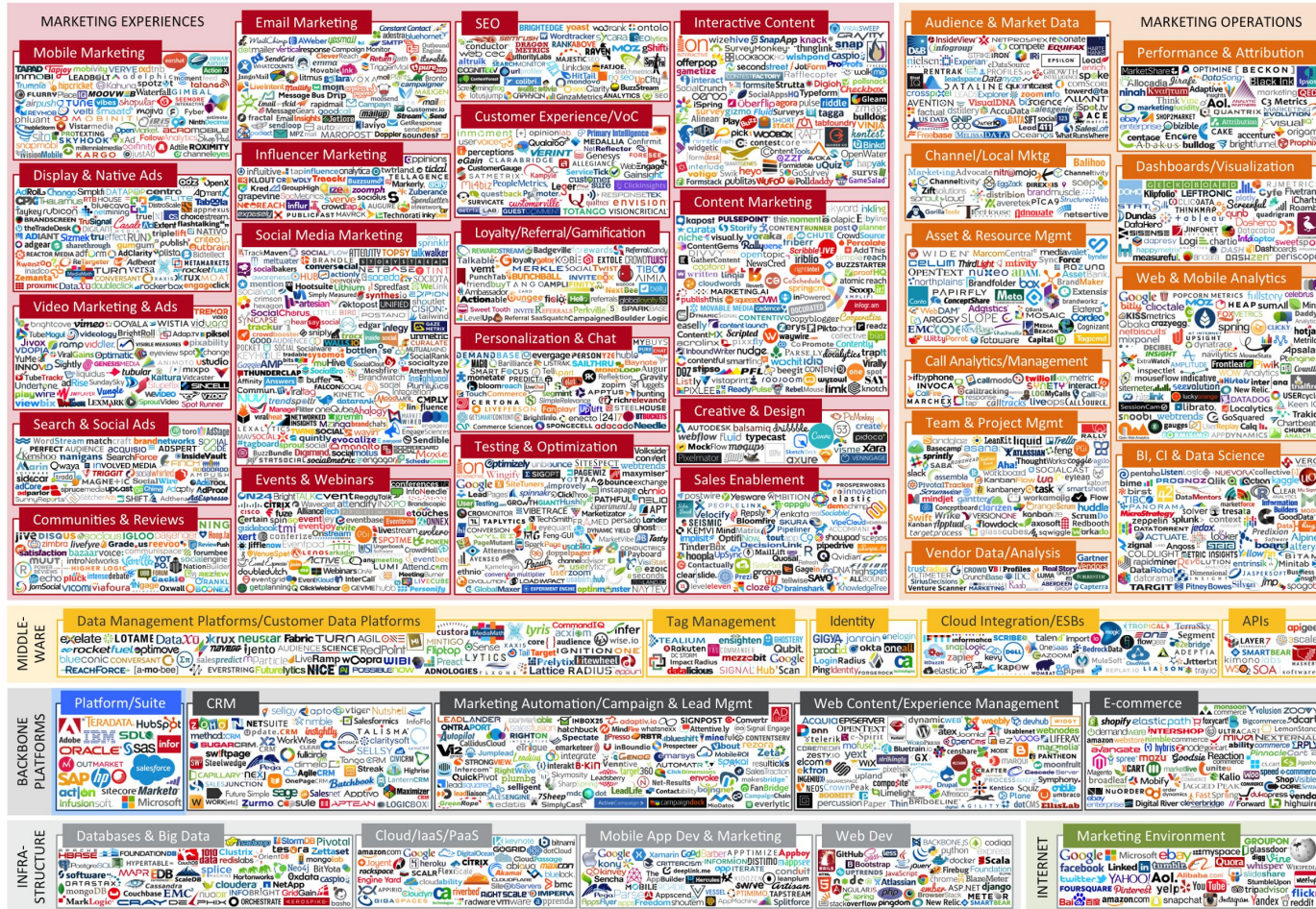
**1.30** NEW FIELD SAFETY ACCOUNT

**Advertising**

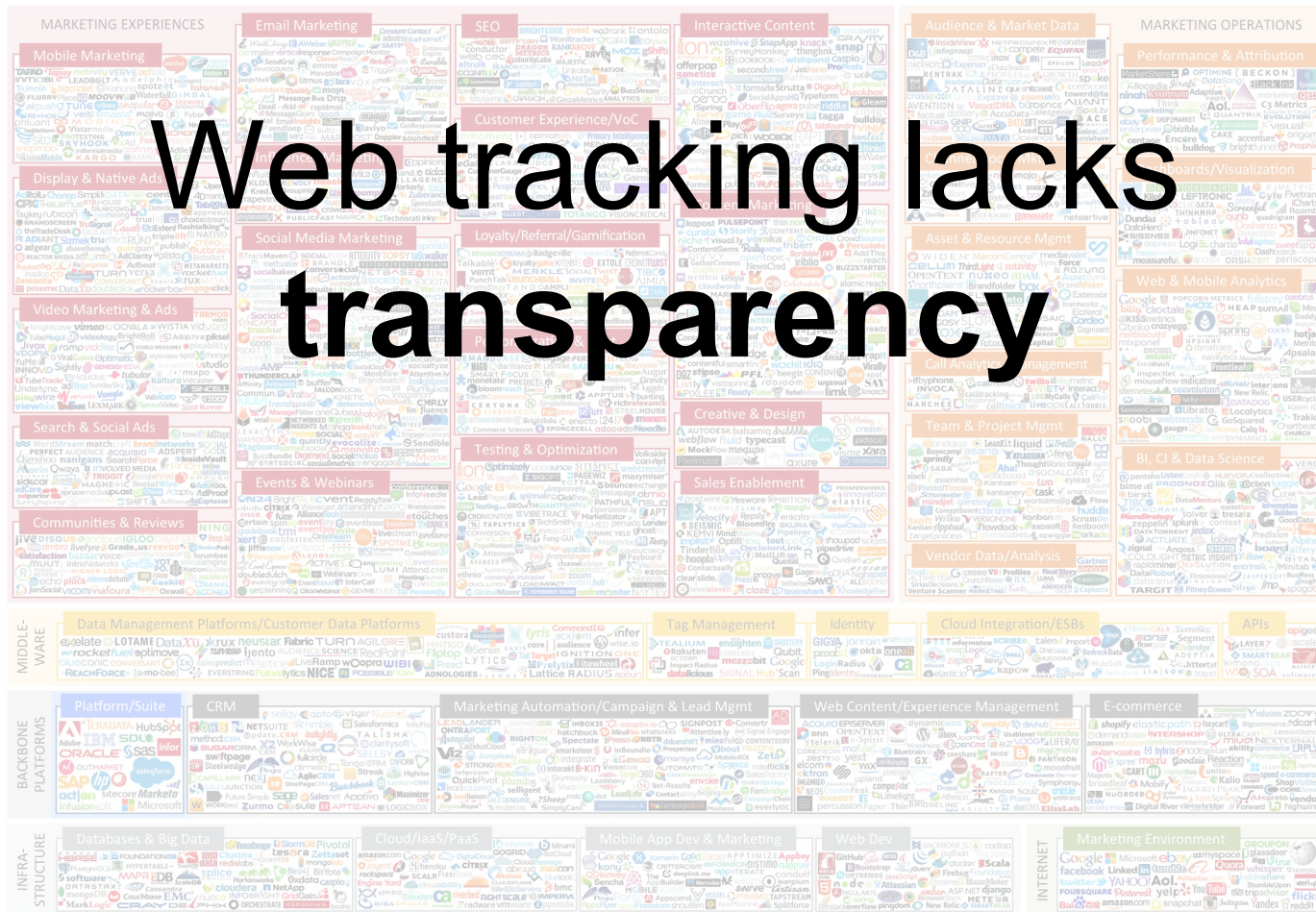
**iMeet built a meeting room just for you.**

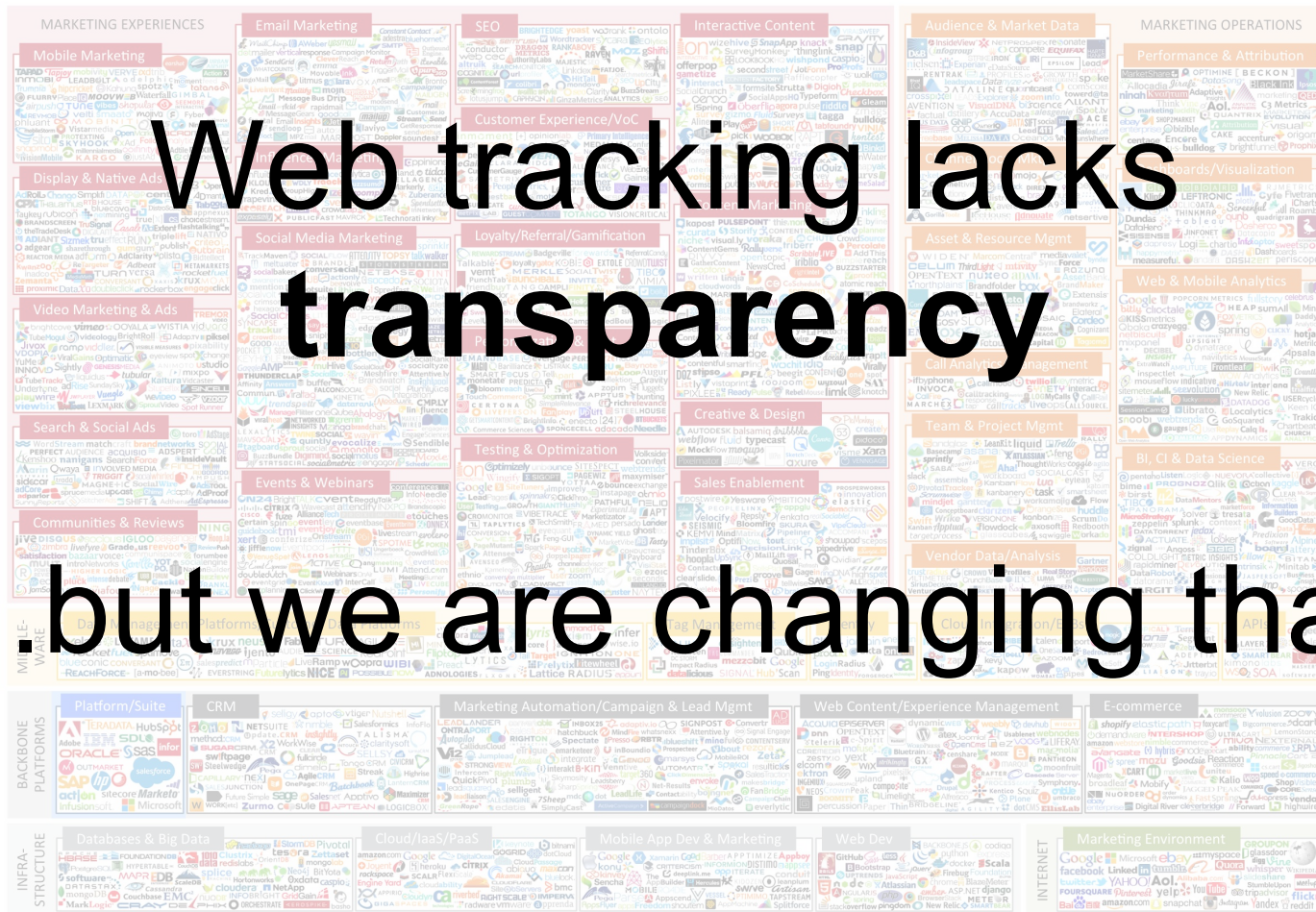
**Advertising** 30-DAY FREE TRIAL

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





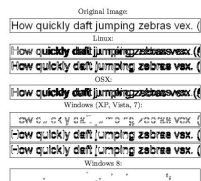








# Transparency encourages best practices



Canvas  
Fingerprinting  
Introduced

May 2012

# Transparency encourages best practices



# Canvas Fingerprinting Introduced



May 2014

May 2012



# Transparency encourages best practices



Canvas  
Fingerprinting  
Introduced

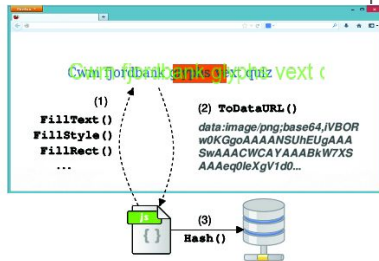
 Add This

 LIGATUS

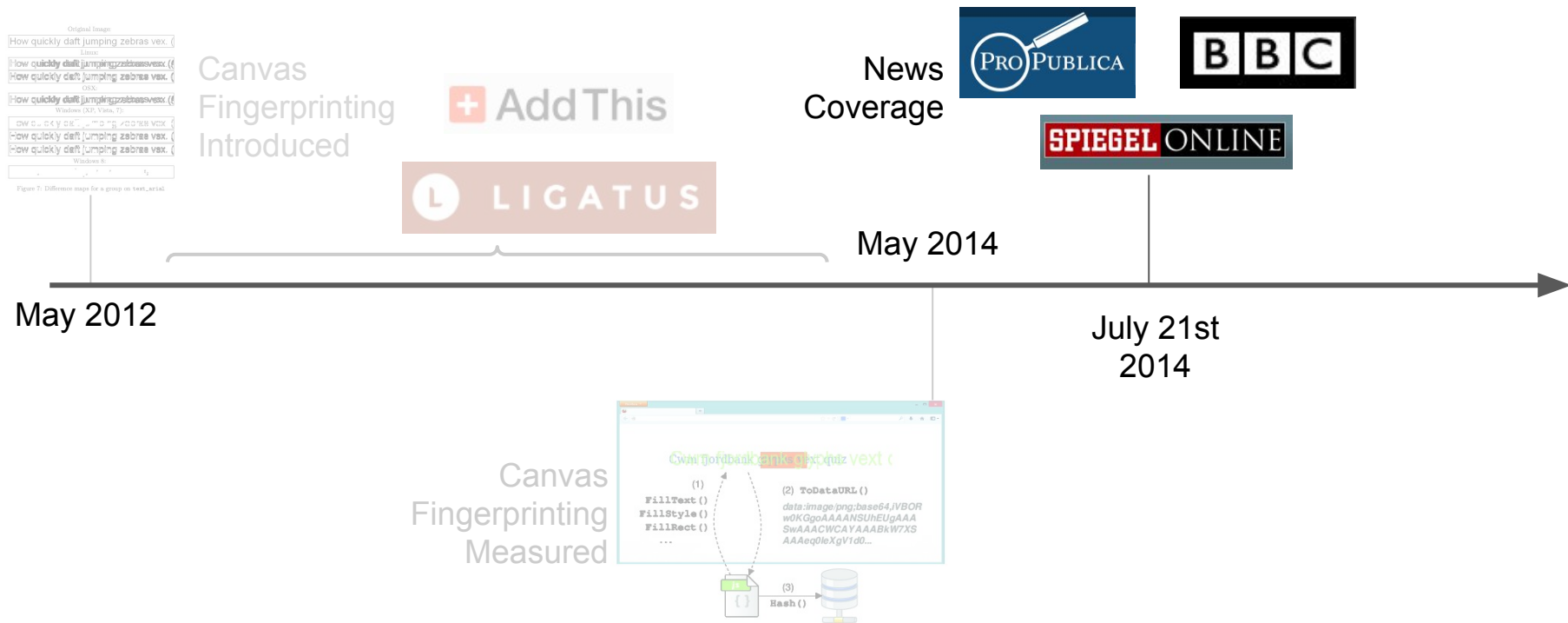
May 2014

May 2012

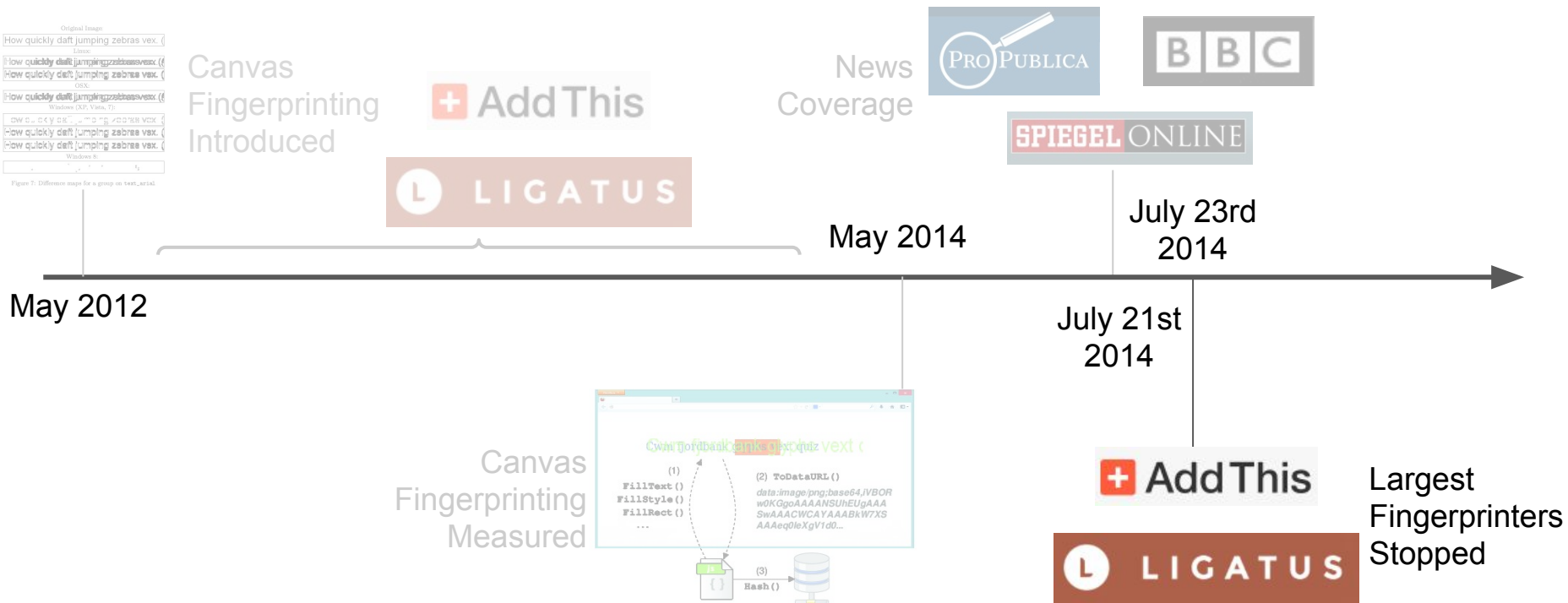
Canvas  
Fingerprinting  
Measured



# 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.



## 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 image is a screenshot of a web browser displaying a BBC News article. The top navigation bar is blue with the 'PRO PUBLICA' logo on the left, which includes a magnifying glass icon. To the right of the logo are links for 'Home' and 'Our Investigations'. Further right is the BBC logo, a 'Sign in' button, and a menu of categories: News, Sport, Weather, Shop, Earth, Travel, and More. Below this is a red banner with the word 'NEWS' in large white letters. Underneath the banner is another navigation bar with links: Home, Video, World, US & Canada, UK, Business, Tech (which is highlighted), Science, Magazine, and Entertainment. The main headline of the article is 'Meet the Virtually Invisible Browser 'fingerprints' help track users'. Below the headline is a sub-headline: 'A new kind of tracking technology has been used to track visitors to the WhiteHouse.gov to YouTube.com'. The byline reads 'by Julia Angwin' and 'ProPublica, July 21, 2014, updated'. A red rectangular placeholder is visible below the byline. At the bottom of the article preview is a photograph of the White House dome under a blue sky with some clouds.

**PRO PUBLICA**

Home | Our Investigations

**BBC** Sign in | News | Sport | Weather | Shop | Earth | Travel | More

**NEWS**

Home | Video | World | US & Canada | UK | Business | **Tech** | Science | Magazine | Entertainment

**Meet the Virtually Invisible Browser 'fingerprints' help track users**

A new kind of tracking technology has been used to track visitors to the WhiteHouse.gov to YouTube.com

by *Julia Angwin*  
ProPublica, July 21, 2014, updated

22 July 2014 | Technology

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/1pKLIhl](http://bit.ly/1pKLIhl)

fingerprinting scripts

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

Search	Find
	Find
	Find
	Find
	Find
	Find
	Find
	Find
	Find
	Find
	Find

**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](http://darknet.org.uk)



1



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.



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** @cymcymys · 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...](https://news.genius.com/Gizmodo-what-y...)

**Tor Works: 2**

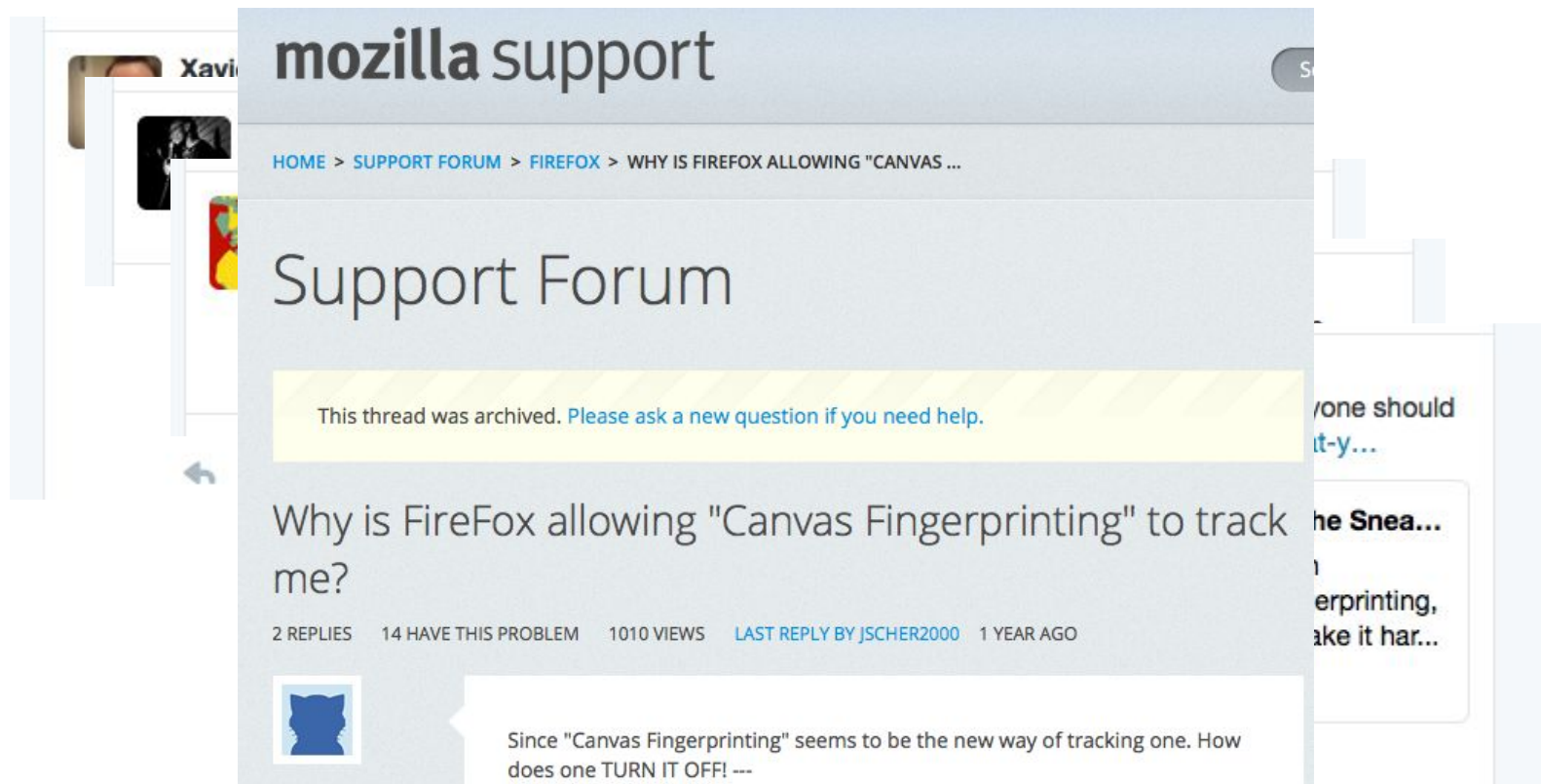


**Gizmodo – What You Need to Know About the Sne...**

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](https://news.genius.com/Gizmodo-what-y...)

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



# 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 it 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

### Must read

An Unbelievable Story of

Devils, Deals and the D

Why Small Debts Matter  
Lives

Out of Options, California  
Troubled Children Out of

'All of This Because Society  
Work'

Small-Scale Violations of  
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

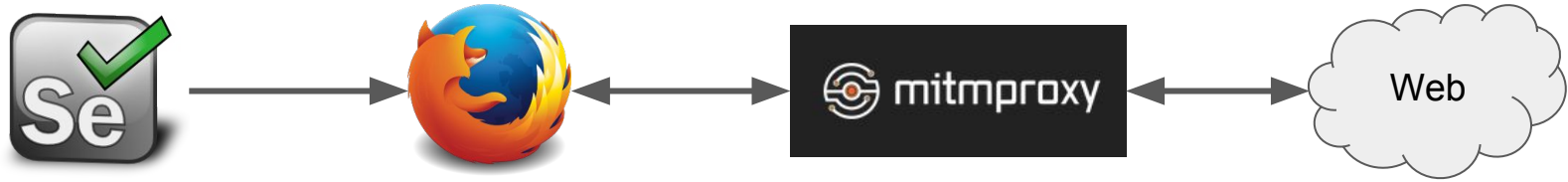
# We're doing three things to help:

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

# OpenWPM

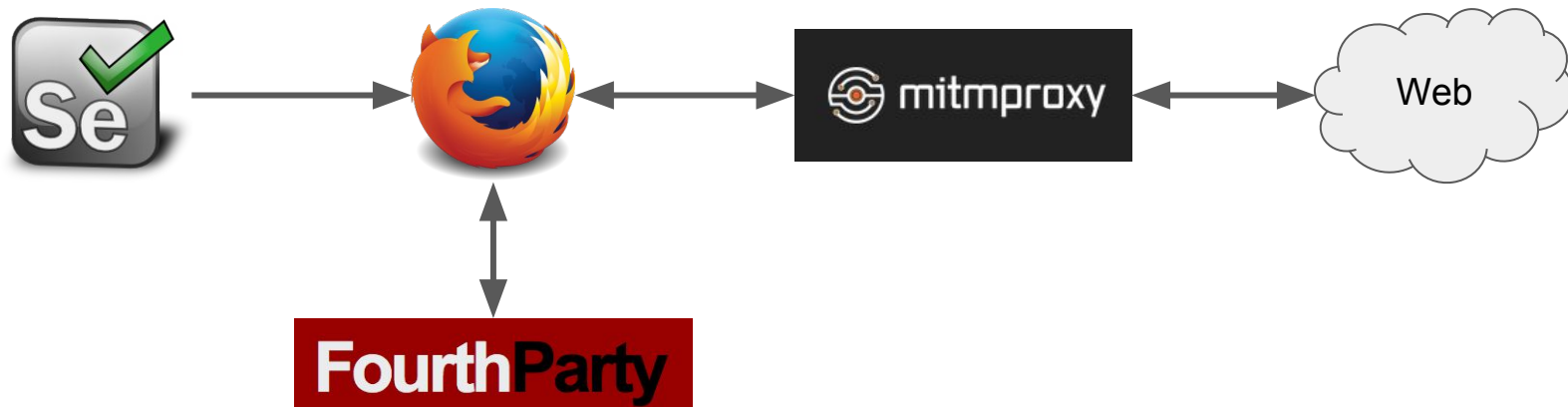


# OpenWPM

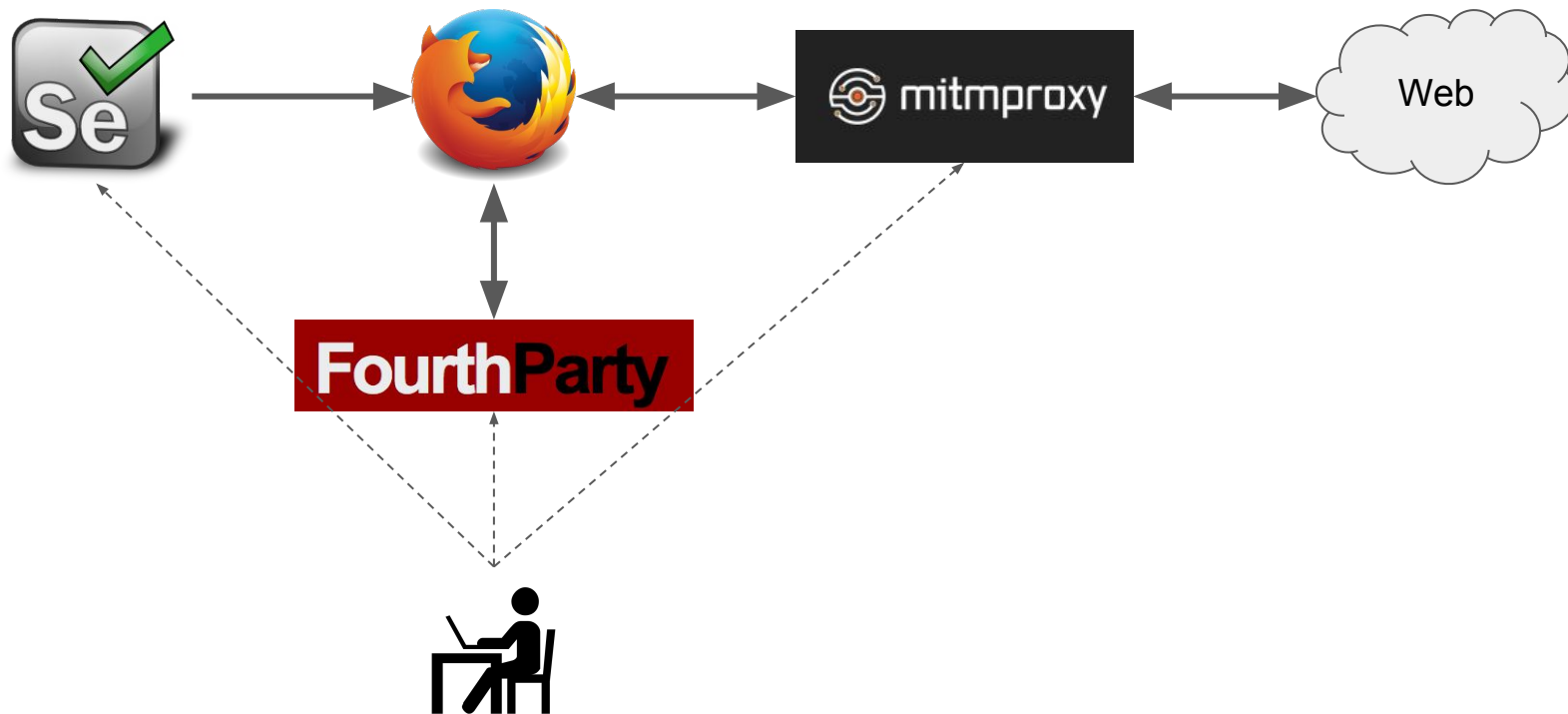




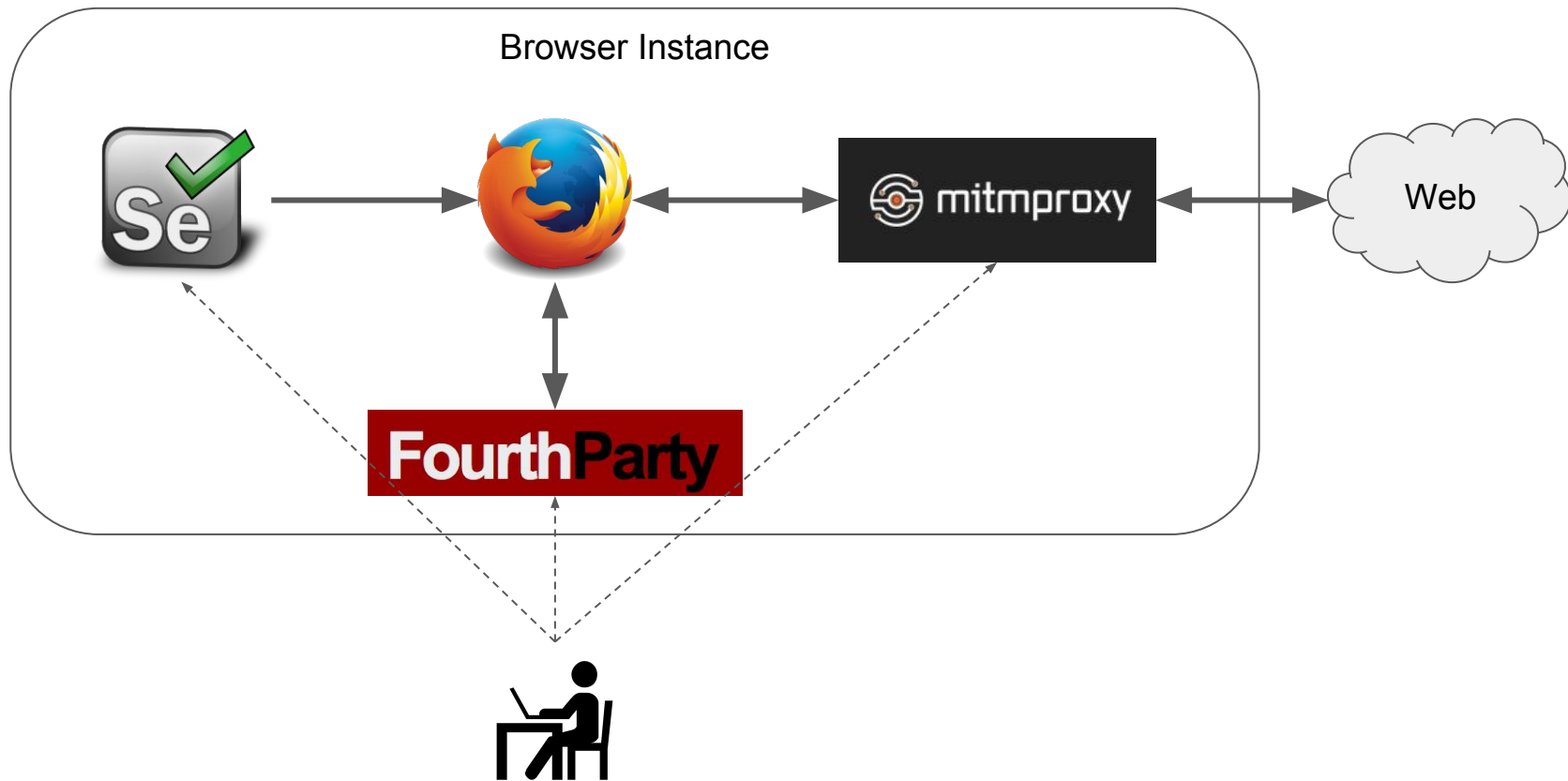
# OpenWPM

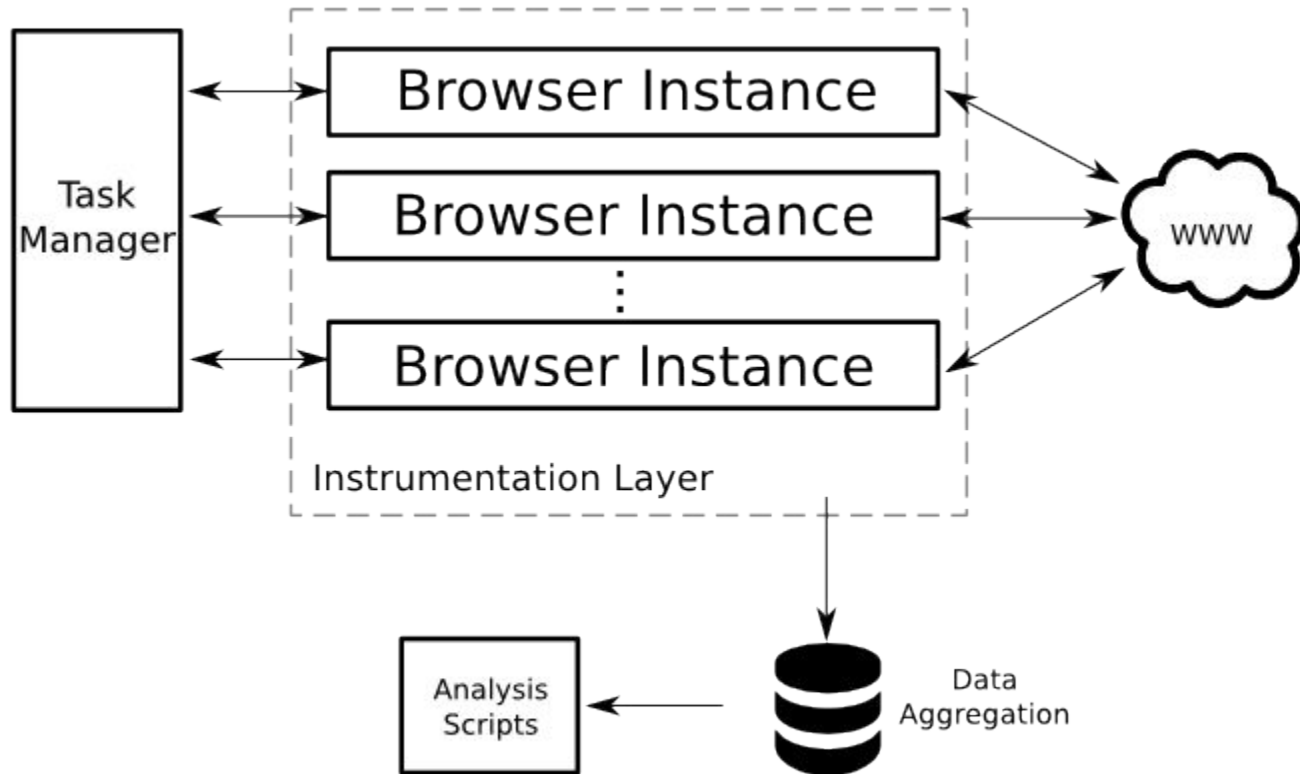


# OpenWPM



# 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, pur	
How quickly daft jumping zebras vex. (Also, pur	
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	
<b>How quickly daft jumping zebras vex. (Also, pu</b>	
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

# 2012: Canvas Fingerprinting Introduced

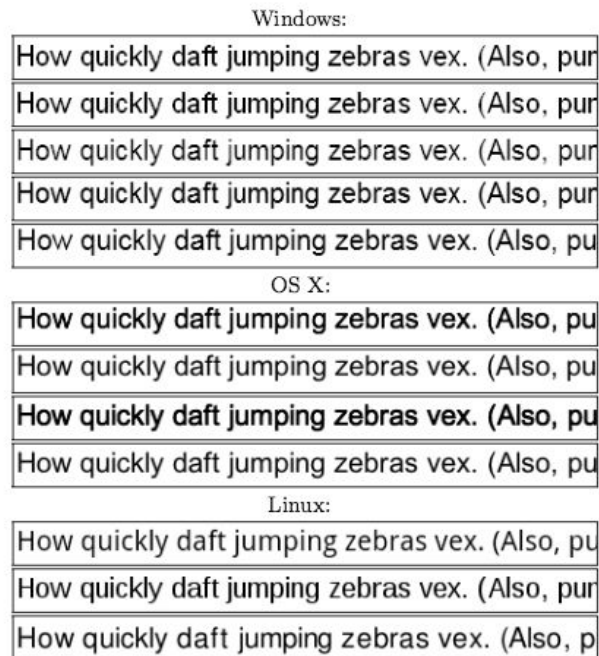


Figure 6: 13 ways to render 20px Arial

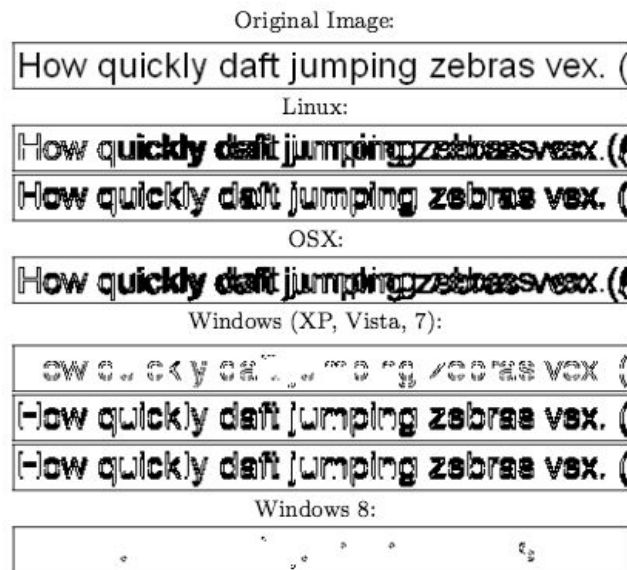


Figure 7: Difference maps for a group on text\_arial

# 2014: Canvas Fingerprinting Measured



## The Web Never Forgets: Persistent Tracking Mechanisms in the Wild

Gunes Acar<sup>1</sup>, Christian Eubank<sup>2</sup>, Steven Englehardt<sup>2</sup>, Marc Juarez<sup>1</sup>  
Arvind Narayanan<sup>2</sup>, Claudia Diaz<sup>1</sup>

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

<sup>2</sup>Princeton University  
{cge,ste,arvindn}@cs.princeton.edu

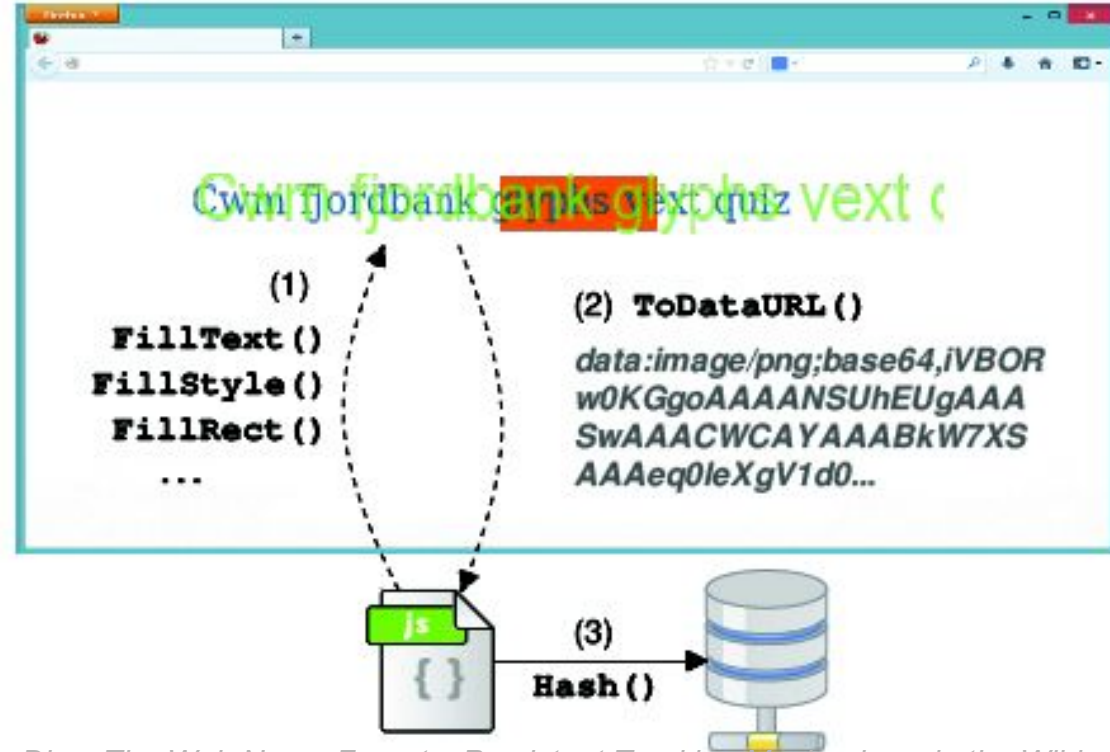
### ABSTRACT

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

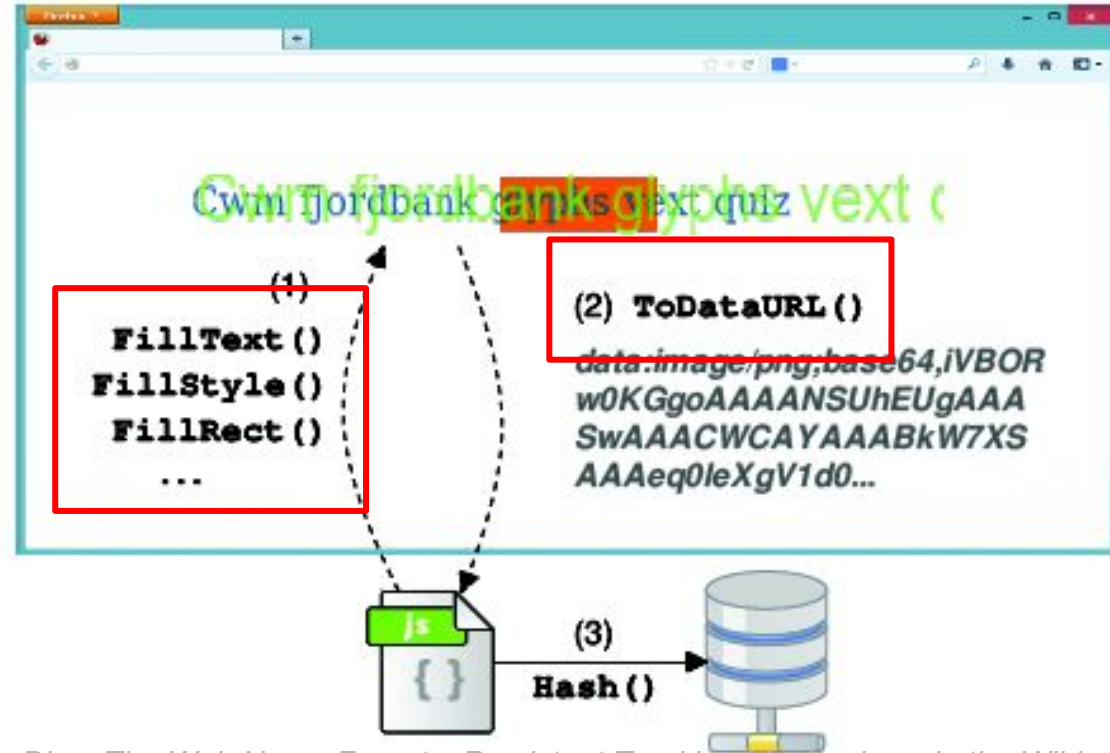
# 2014: Canvas Fingerprinting Measured





# 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](https://www.nytimes.com) to report visitors' local IP addresses.

```
Debugger - http://www.nytimes.com/?WT.z_jog=1&hf
Vector Console Debugger Style Editor Performance Network DOM
8.nytimes.com
kats.on.nytime...
red2.google...
m3
m3.googleads...
4.63/j
ads.g.doublecl...
parvix.com
143871988967000
[...]
```

Time	IP Address	IP Address	IP Address	Operation
148	0.0062285	192.168.1.203	192.168.1.254	DNS Query Operation,
143		192.168.1.119	239.255.255.250	SSDP Request, Method:
571	0.00000005	192.168.1.203	192.168.1.254	DNS Request, Query I
573	0.0148223	192.168.1.203	192.168.1.254	DNS Query Operation,
712		192.168.1.254	192.168.1.203	DNS Response, #Code:
210		192.168.1.107	239.255.255.250	SSDP Request, Method:
359		192.168.1.107	239.255.255.250	SSDP Request, Method:
406		192.168.1.107	239.255.255.250	SSDP Request, Method:
529	0.0268478	192.168.1.203	192.168.1.254	DNS Query Operation,
808	0.00000005	192.168.1.203	192.168.1.254	DNS Request, Query I
906	0.0053289	192.168.1.203	192.168.1.254	DNS Query Operation,
210		192.168.1.254	192.168.1.203	DNS Response, #Code:
336	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,
480	0.0118086	192.168.1.203	192.168.1.254	DNS Query Operation,

## 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**

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

## **WebRTC Local IP Sniffing**

# 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



# 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](https://github.com/citp/OpenWPM)
- **Collaborate:**
  - [webtap.princeton.edu](http://webtap.princeton.edu)

**Email:** [ste@cs.princeton.edu](mailto:ste@cs.princeton.edu)

**Twitter:** [@s\\_englehardt](https://twitter.com/s_englehardt)