Browser fingerprinting

(how did we get here)

SecAppDev

February 2014

Nick Nikiforakis

www.securitee.org

echo `whoami`

- Postdoctoral researcher at KU Leuven
- Working, mainly, on web security and privacy
- Identify online ecosystems
  - Players
  - Interactions
  - Common patterns
- Search for systematic problems and solutions

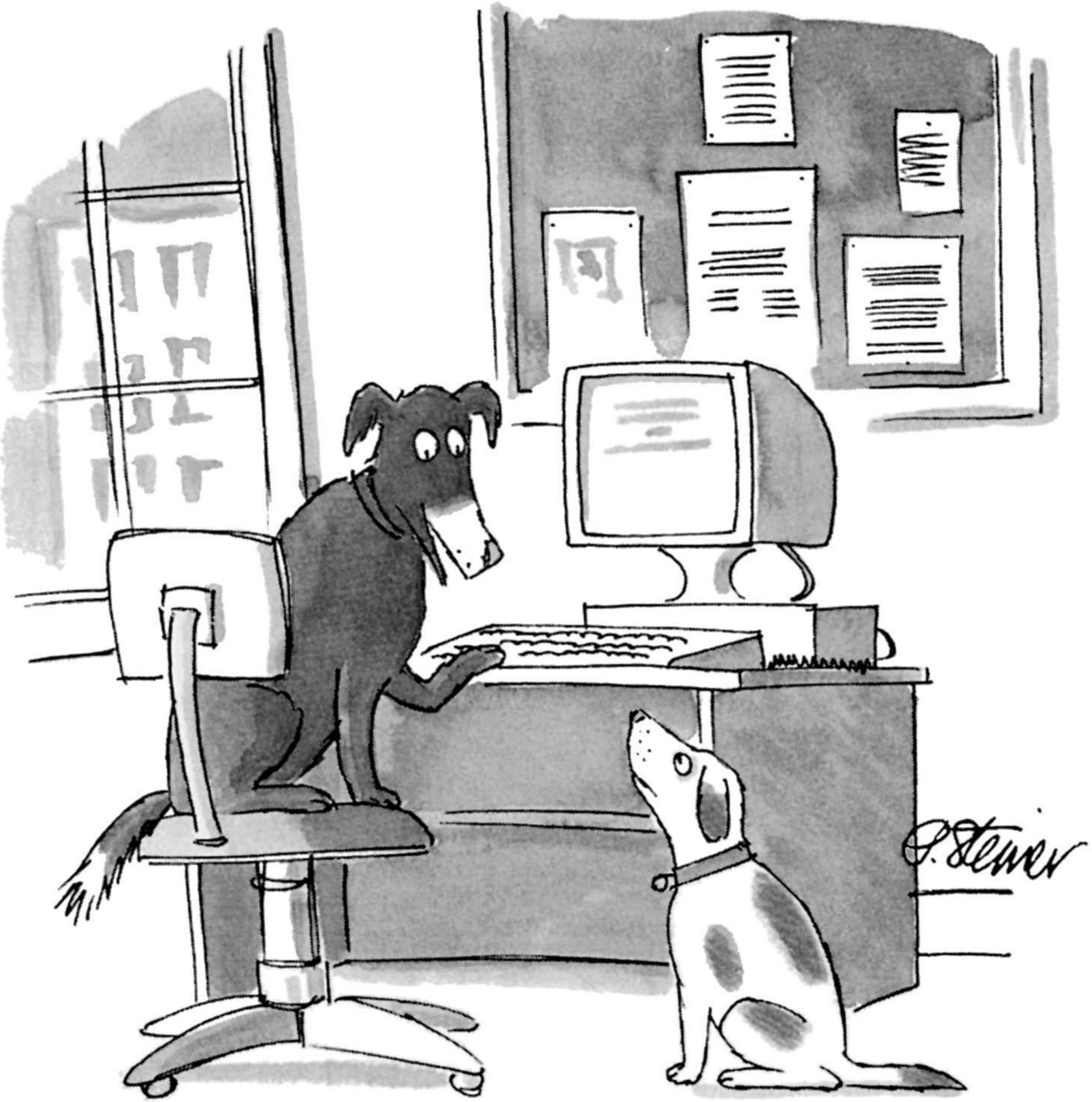
STREWS

DistriNet

iMinds

KU LEUVEN

1993



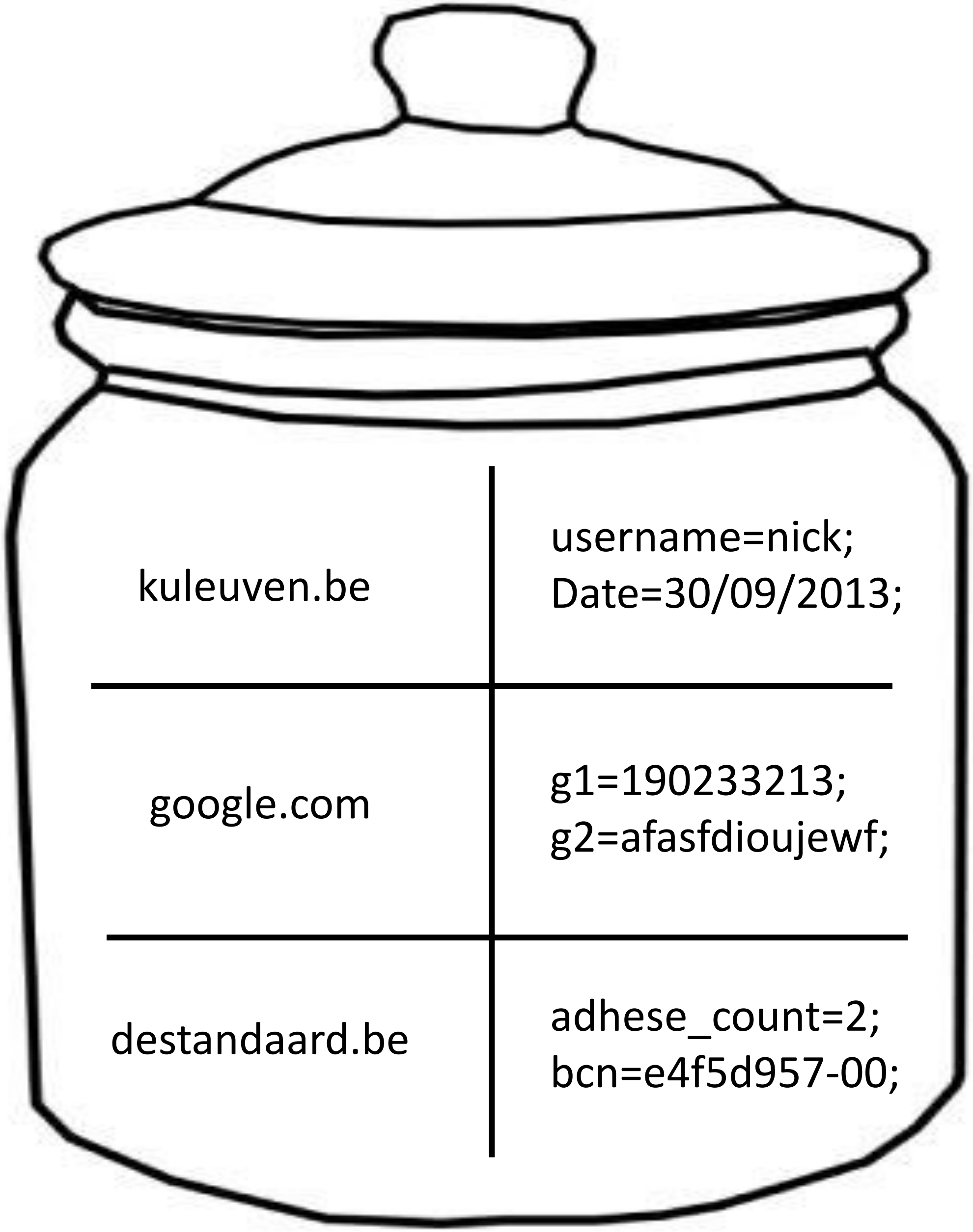
"On the Internet, nobody knows you're a dog."

I need state!

- HTTP is a stateless protocol
  - The server does not know that two or requests originate from the same user

- No state -> No Personalization
  - No e-banking, e-shops, webmail, etc.

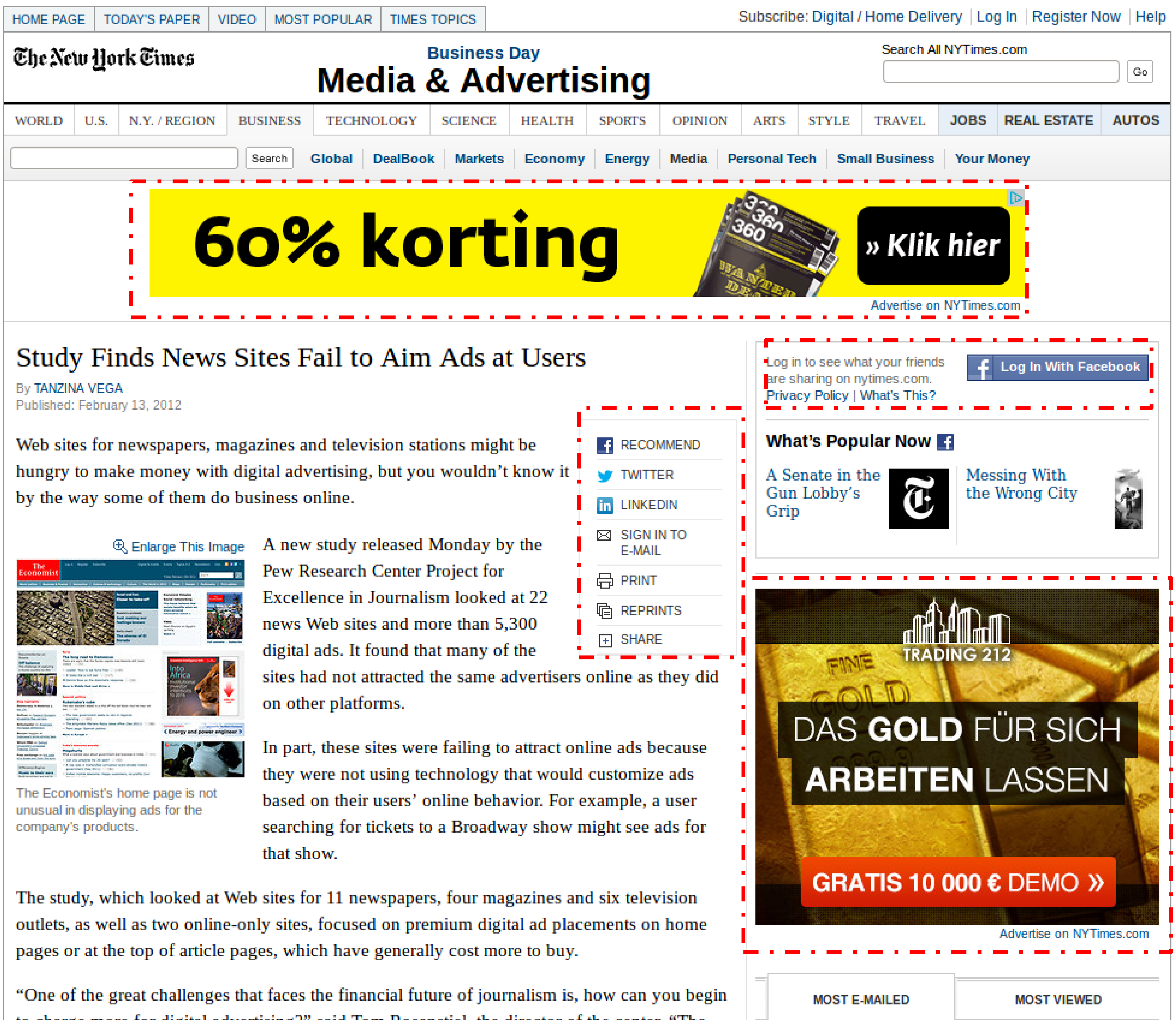
- Solution: Cookies!



kuleuven.be	username=nick;
	Date=30/09/2013;

google.com	g1=190233213;
	g2=afasfdioujewf;

destandaard.be	adhese_count=2;
	bcn=e4f5d957-00;



HOME PAGE | TODAY'S PAPER | VIDEO | MOST POPULAR | TIMES TOPICS

Subscribe: Digital / Home Delivery | Log In | Register Now | Help

The New York Times

Business Day

Media & Advertising

Search All NYTimes.com

Go

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION | ARTS | STYLE | TRAVEL | JOBS | REAL ESTATE | AUTOS

Search | Global | DealBook | Markets | Economy | Energy | Media | Personal Tech | Small Business | Your Money

Study Finds News Sites Fail to Aim Ads at Users

By TANZINA VEGA

Published: February 13, 2012

RECOMMEND
TWITTER
LINKEDIN
SIGN IN TO E-MAIL
PRINT
REPRINTS
SHARE

Web sites for newspapers, magazines and television stations might be hungry to make money with digital advertising, but you wouldn't know it by the way some of them do business online.

Enlarge This Image

The Economist's home page is not unusual in displaying ads for the company's products.

A new study released Monday by the Pew Research Center Project for Excellence in Journalism looked at 22 news Web sites and more than 5,300 digital ads. It found that many of the sites had not attracted the same advertisers online as they did on other platforms.

In part, these sites were failing to attract online ads because they were not using technology that would customize ads based on their users' online behavior. For example, a user searching for tickets to a Broadway show might see ads for that show.

The study, which looked at Web sites for 11 newspapers, four magazines and six television outlets, as well as two online-only sites, focused on premium digital ad placements on home pages or at the top of article pages, which have generally cost more to buy.

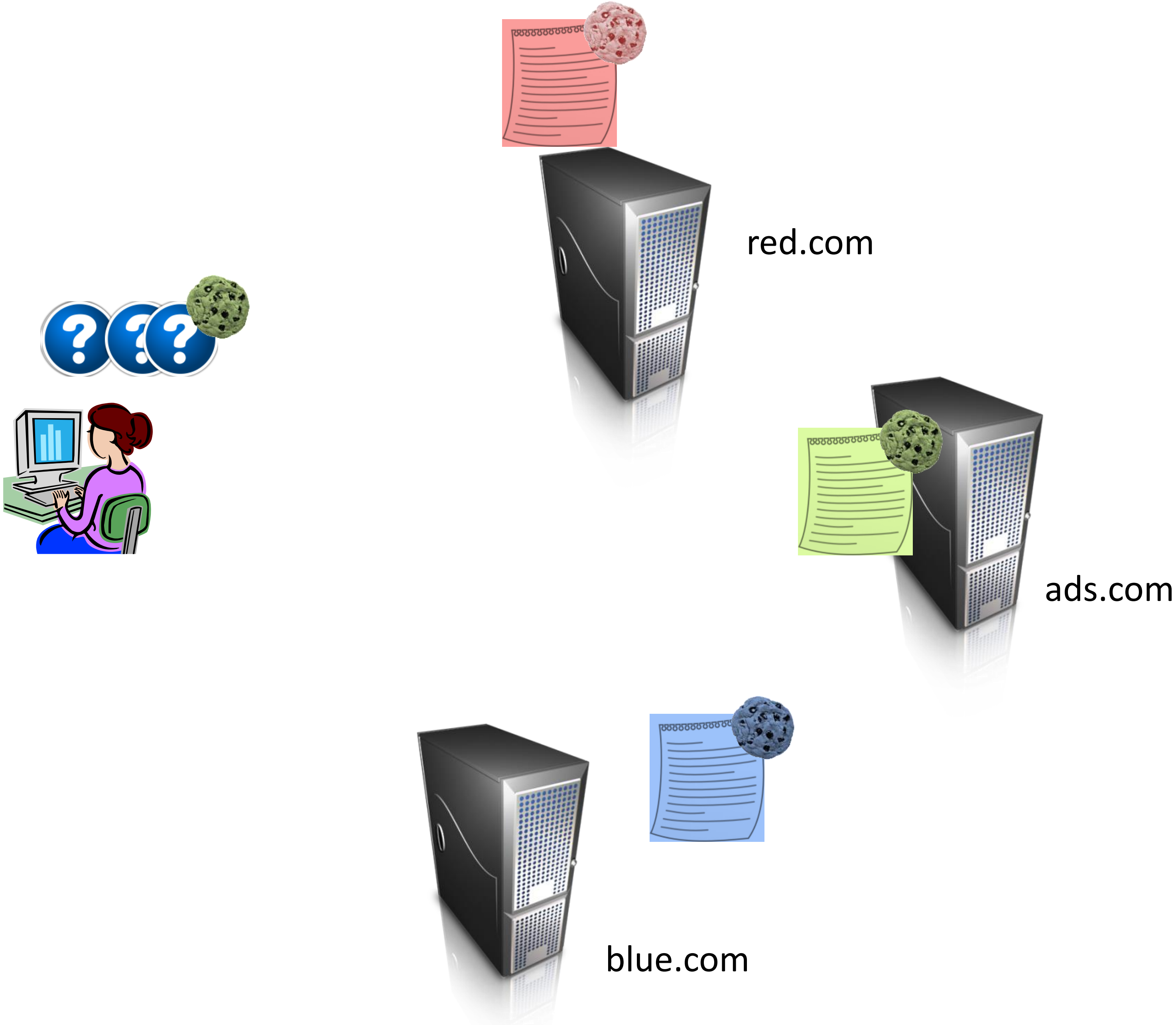
“One of the great challenges that faces the financial future of journalism is, how can you begin to charge more for digital advertising?” said Tom Rosenstiel, the director of the center. “The

Log in to see what your friends are sharing on nytimes.com. Privacy Policy | What's This?
Log In With Facebook

What's Popular Now
A Senate in the Gun Lobby's Grip
Messing With the Wrong City

MOST E-MAILED | MOST VIEWED

A cookie's life



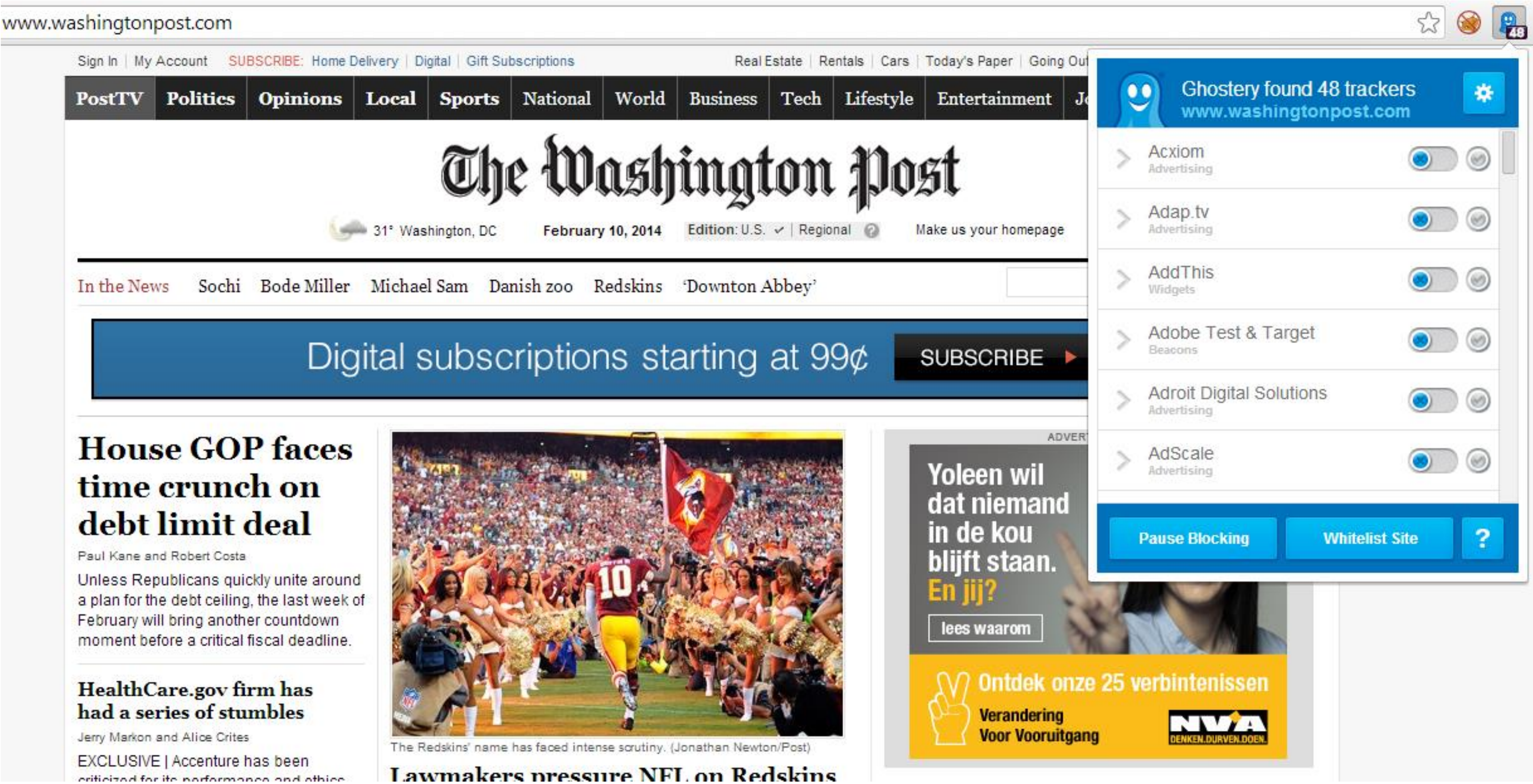
3rd Party Tracking

- “Suddenly” all sorts of websites that you’ve never heard about, can create a browsing profile of you and sell it to advertising companies
  - quantserve.com
  - scorecardresearch.com
  - addthis.com

Users reacted…

- 1/3 of users delete first & third-party cookies within a month after they’ve been setup
- Multiple extensions revealing hidden trackers
  - Ghostery
  - Lightbeam
- Private mode of browsers used to avoid traces of cookies from certain websites

Ghostery


www.washingtonpost.com
Sign In | My Account SUBSCRIBE: Home Delivery | Digital | Gift Subscriptions
Real Estate | Rentals | Cars | Today's Paper | Going Out
PostTV
Politics
Opinions
Local
Sports
National
World
Business
Tech
Lifestyle
Entertainment
The Washington Post
31° Washington, DC
February 10, 2014
Edition: U.S. ✓ | Regional
Make us your homepage
In the News
Sochi
Bode Miller
Michael Sam
Danish zoo
Redskins
'Downton Abbey'
Digital subscriptions starting at 99¢
SUBSCRIBE
ADVERT
Yoleen wil dat niemand in de kou blijft staan. En jij?
lees waarom
Ontdek onze 25 verbintenissen
Verandering Voor Vooruitgang
N-VA
DENKEN.DURVEN.DOEN.

Ghostery found 48 trackers
www.washingtonpost.com

- Acxiom — Advertising
- Adap.tv — Advertising
- AddThis — Widgets
- Adobe Test & Target — Beacons
- Adroit Digital Solutions — Advertising
- AdScale — Advertising

House GOP faces time crunch on debt limit deal

Paul Kane and Robert Costa

Unless Republicans quickly unite around a plan for the debt ceiling, the last week of February will bring another countdown moment before a critical fiscal deadline.

HealthCare.gov firm has had a series of stumbles

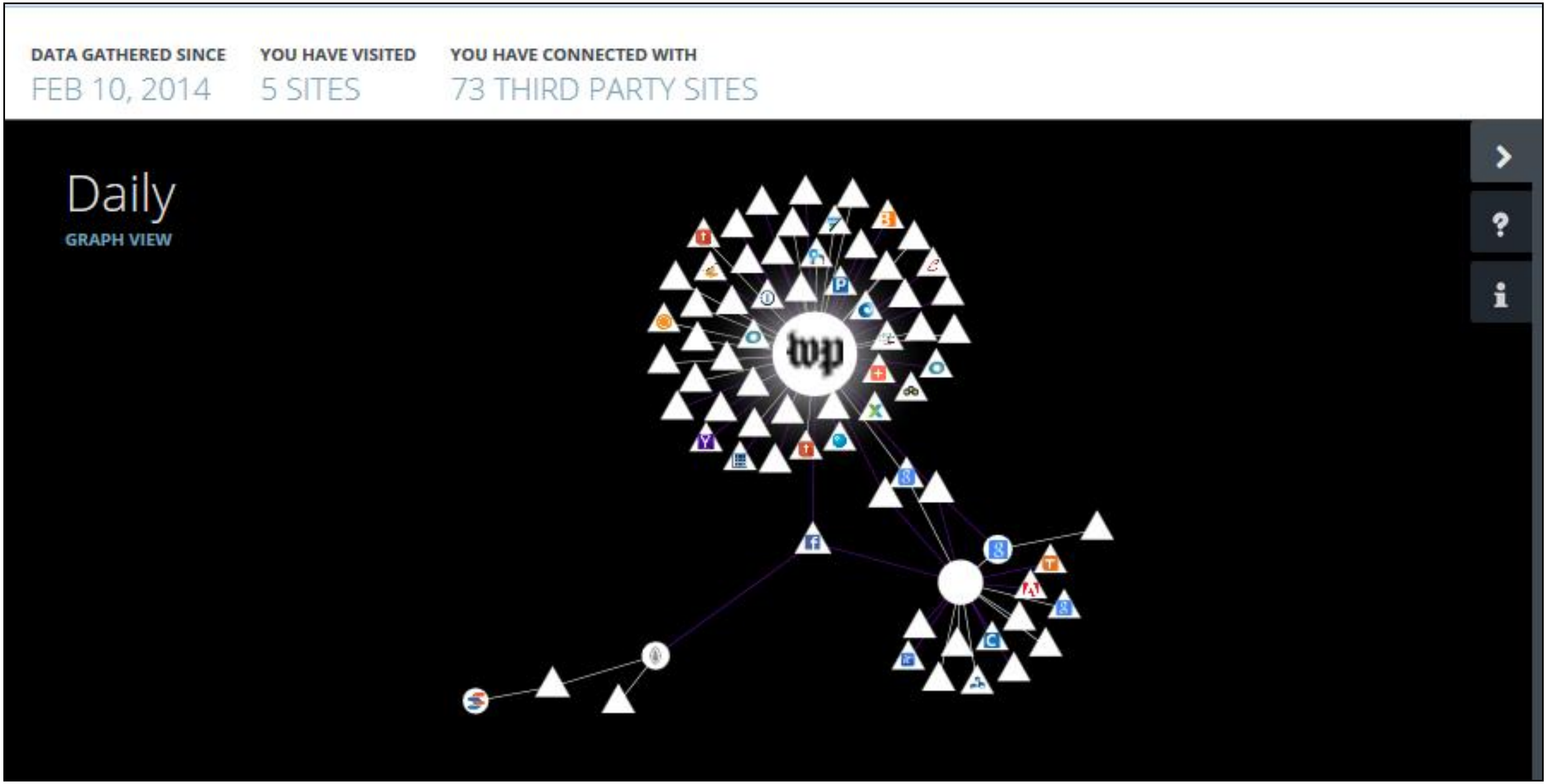
Jerry Markon and Alice Crites

EXCLUSIVE | Accenture has been criticized for its performance and ethics

The Redskins' name has faced intense scrutiny. (Jonathan Newton/Post)

Lawmakers pressure NFL on Redskins

Lightbeam



DATA GATHERED SINCE
FEB 10, 2014

YOU HAVE VISITED
5 SITES

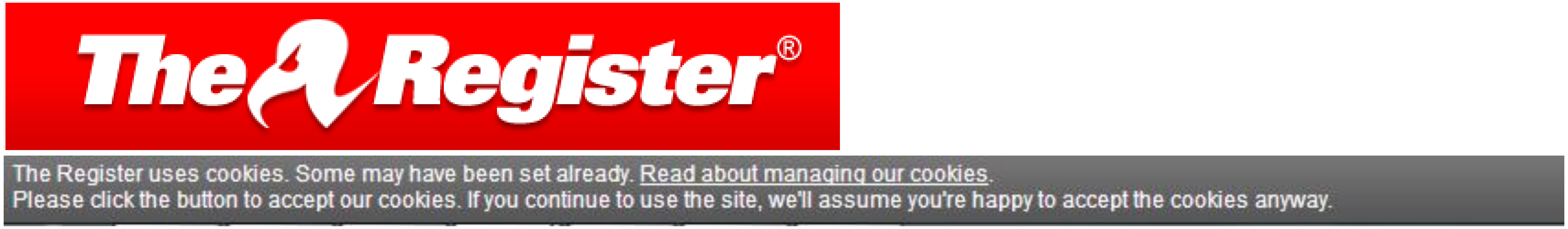
YOU HAVE CONNECTED WITH
73 THIRD PARTY SITES

EU Cookie law

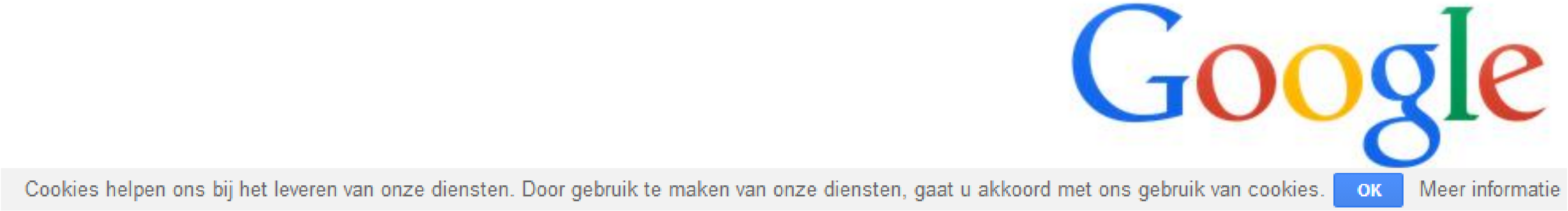
“Member States shall ensure that the storing of information, or the gaining of access to information already stored, in the terminal equipment of a subscriber or user is only allowed on condition that the subscriber or user concerned has given his or her consent, having been provided with clear and comprehensive information, in accordance with Directive 95/46/EC, inter alia, about the purposes of the processing. This shall not prevent any technical storage or access for the sole purpose of carrying out the transmission of a communication over an electronic communications network, or as strictly necessary in order for the provider of an information society service explicitly requested by the subscriber or user to provide the service.”

EU Cookie law

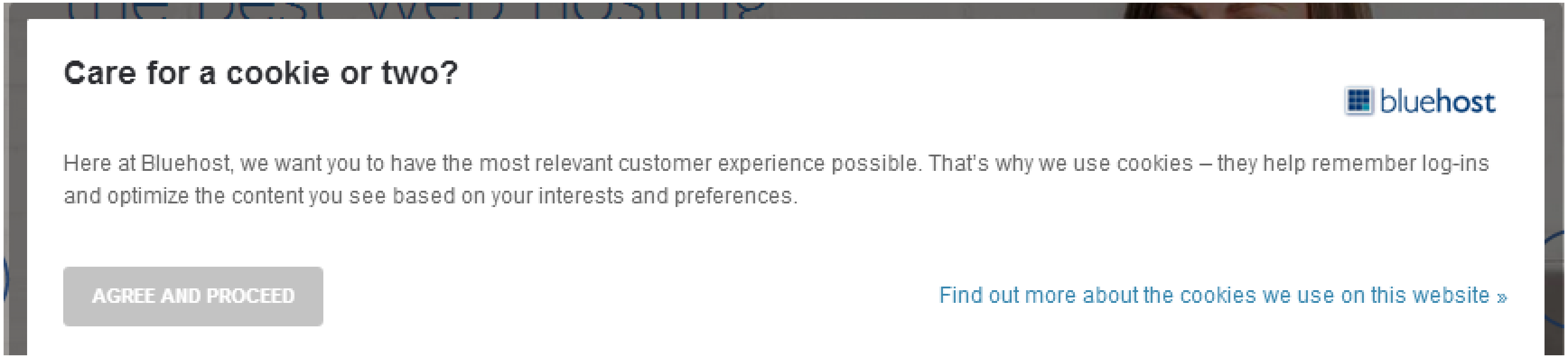
“Member States shall ensure that the storing of information, or the gaining of access to information already stored, in the terminal equipment of a subscriber or user is only allowed on condition that the subscriber or user concerned has given his or her consent, having been provided with clear and comprehensive information, in accordance with Directive 95/46/EC, inter alia, about the purposes of the processing. This shall not prevent any technical storage or access for the sole purpose of carrying out the transmission of a communication over an electronic communications network, or as strictly necessary in order for the provider of an information society service explicitly requested by the subscriber or user to provide the service.”



The Register uses cookies. Some may have been set already. Read about managing our cookies.
Please click the button to accept our cookies. If you continue to use the site, we'll assume you're happy to accept the cookies anyway.



Cookies helpen ons bij het leveren van onze diensten. Door gebruik te maken van onze diensten, gaat u akkoord met ons gebruik van cookies.

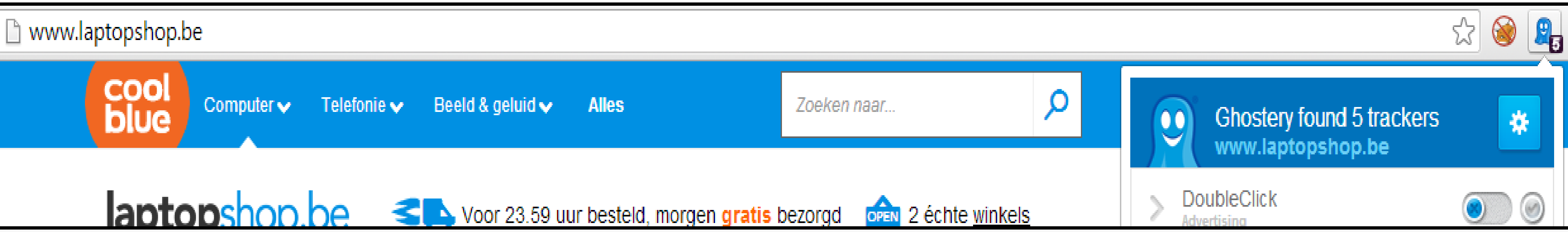


Care for a cookie or two?

Here at Bluehost, we want you to have the most relevant customer experience possible. That's why we use cookies – they help remember log-ins and optimize the content you see based on your interests and preferences.

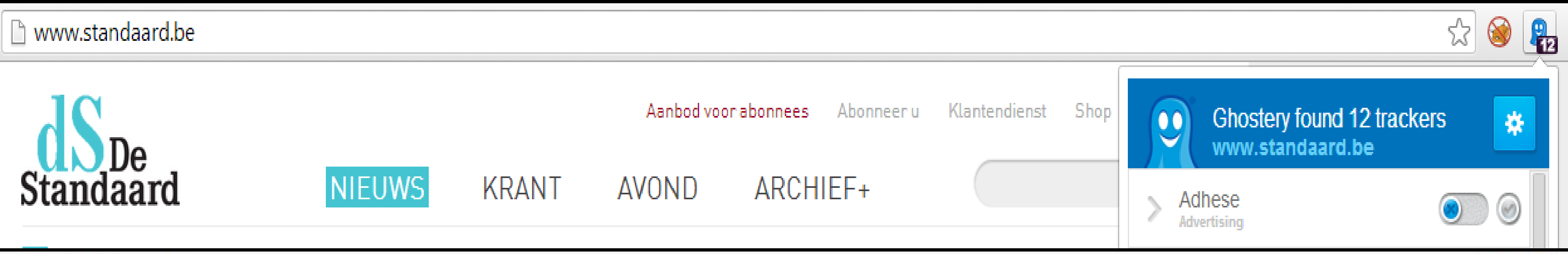
Find out more about the cookies we use on this website »

Belgium?


www.laptopshop.be

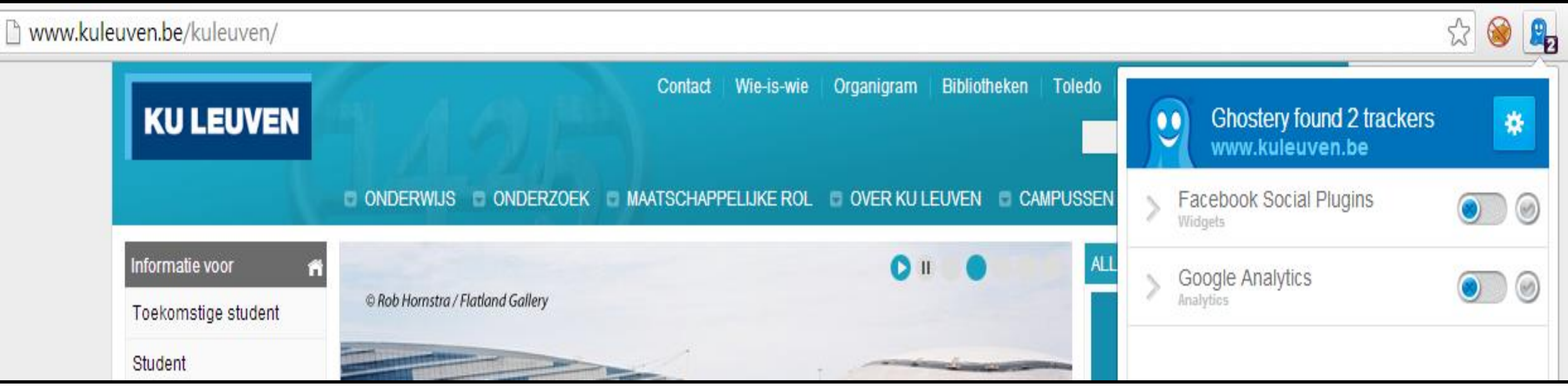
Ghostery found 5 trackers
www.laptopshop.be

DoubleClick
Advertising


www.standaard.be

Ghostery found 12 trackers
www.standaard.be

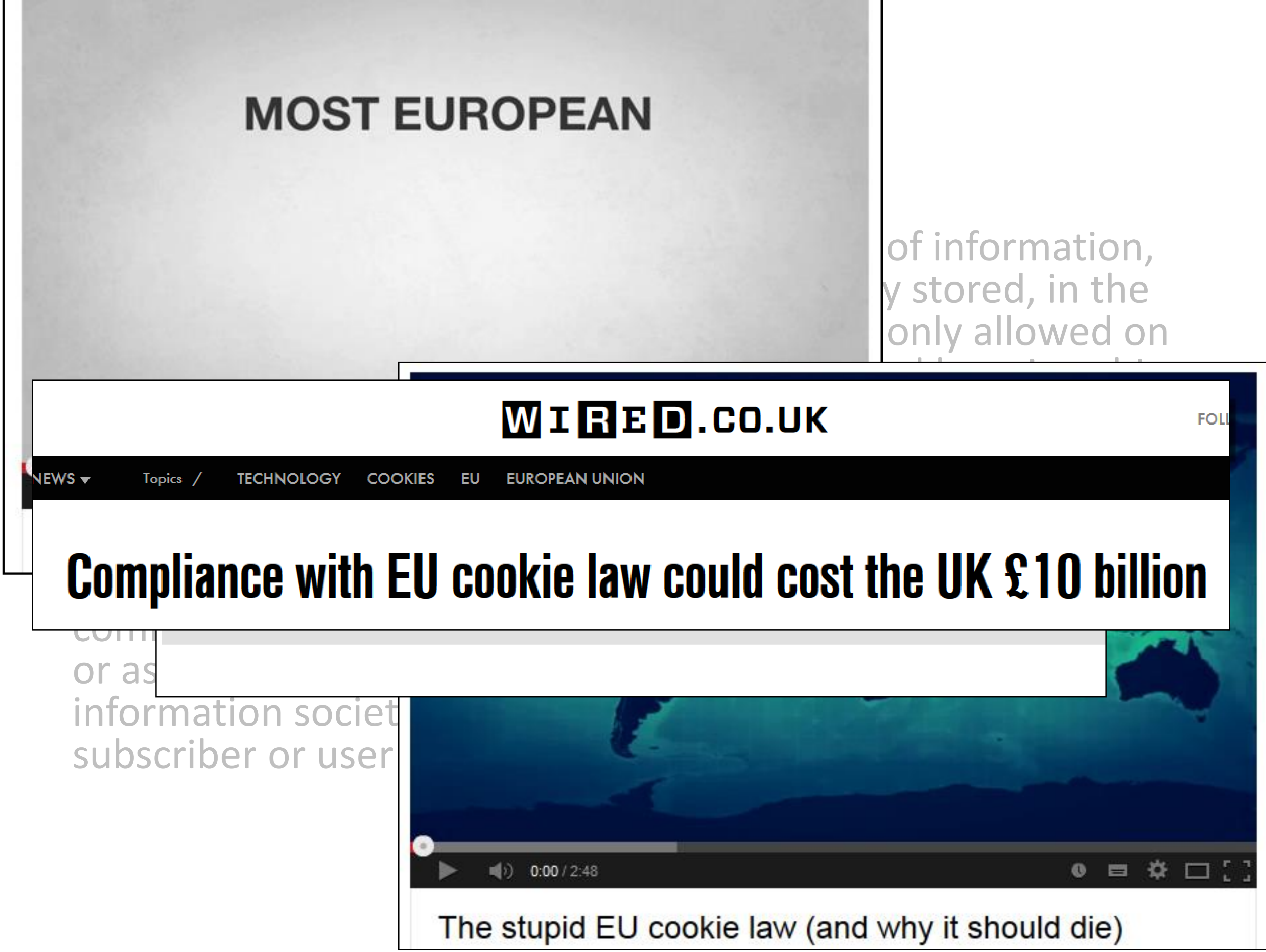
Adhese
Advertising


www.kuleuven.be/kuleuven/

Ghostery found 2 trackers
www.kuleuven.be

Facebook Social Plugins
Widgets

Google Analytics
Analytics



MOST EUROPEAN

of information,
y stored, in the
only allowed on

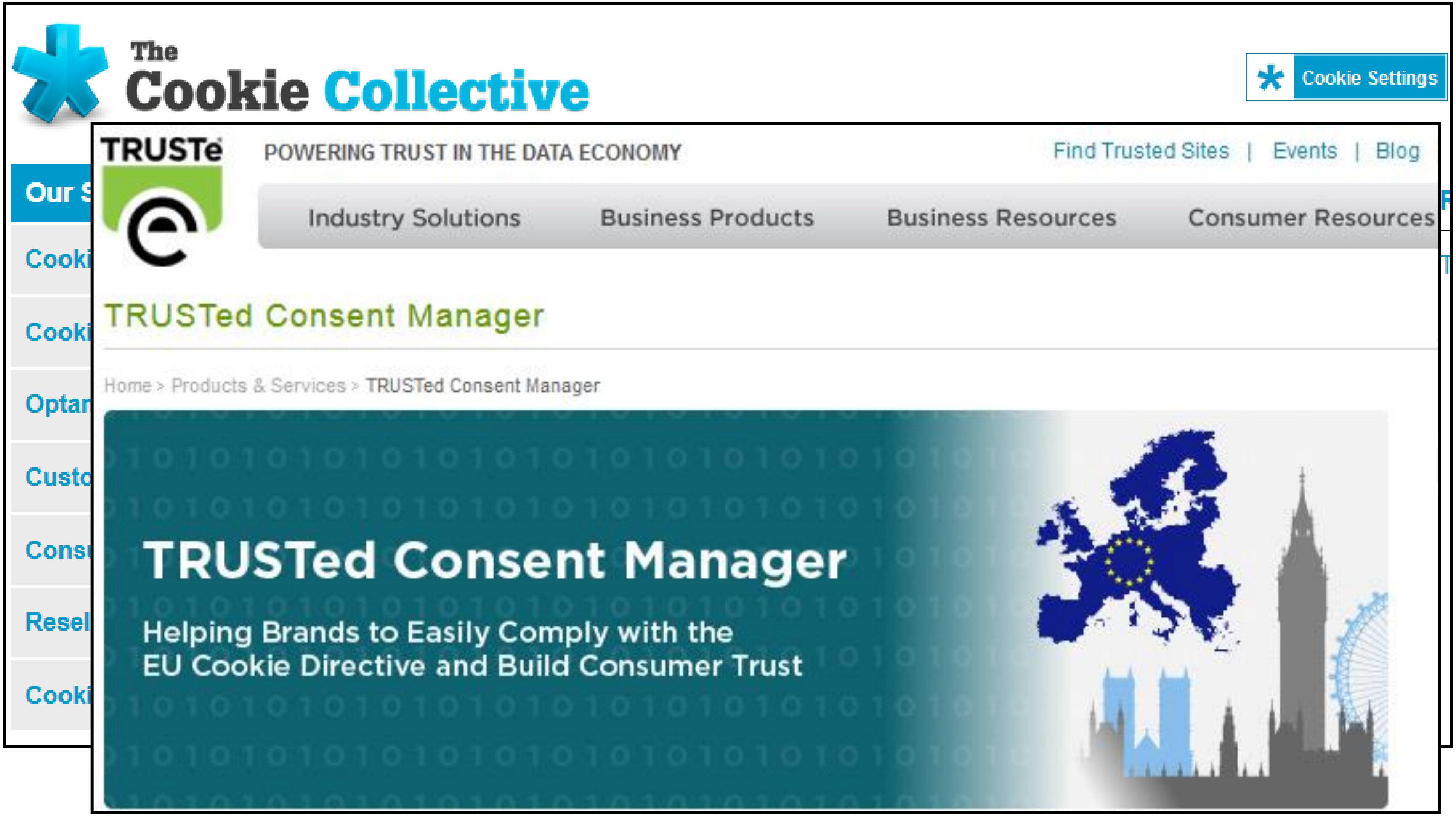
or as
information societ
subscriber or user

WIRED.CO.UK

Compliance with EU cookie law could cost the UK £10 billion

The stupid EU cookie law (and why it should die)

Money making



TRUSTed Consent Manager

Home > Products & Services > TRUSTed Consent Manager

TRUSTed Consent Manager

Helping Brands to Easily Comply with the EU Cookie Directive and Build Consumer Trust

= PROBLEM

=

Right?

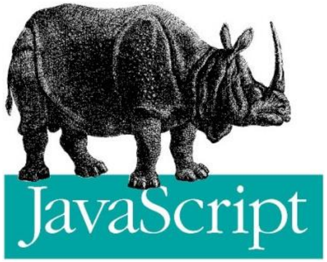
However…

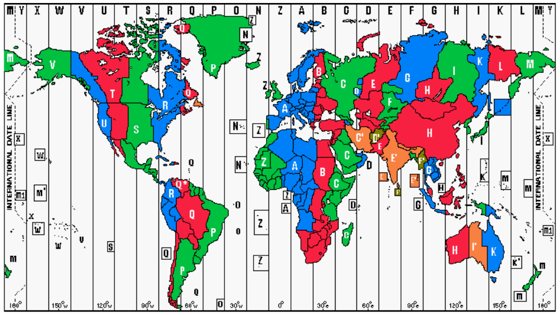
- What if you could track users without the need of cookies or any other stateful client-side identifier?
  - Hidden from users
  - Hard to avoid it / opt-out

Web-based device fingerprinting

- Eckersley showed in 2010 that certain attributes of your browsing environment can be used to accurately track you
- These attributes, when combined, created a quite unique fingerprint of your system?
  - How?

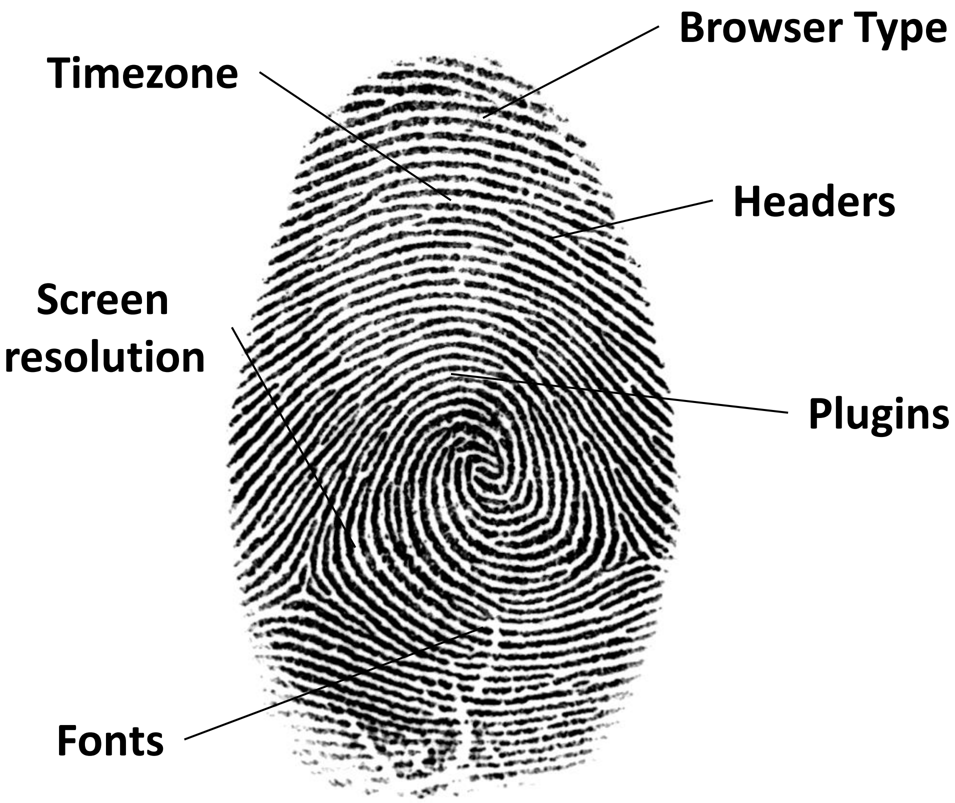
Properties fingerprinted by Panopticlick





Maverick
Ocean Front Villas
mandarin tea
Regency
Sassafras & Ginger
Dollhouse
Athletics Dept.

Resulting fingerprints



- 94.2% of the users with Flash/Java could be uniquely identified
- Simple heuristic algorithms could track updates of the same browser

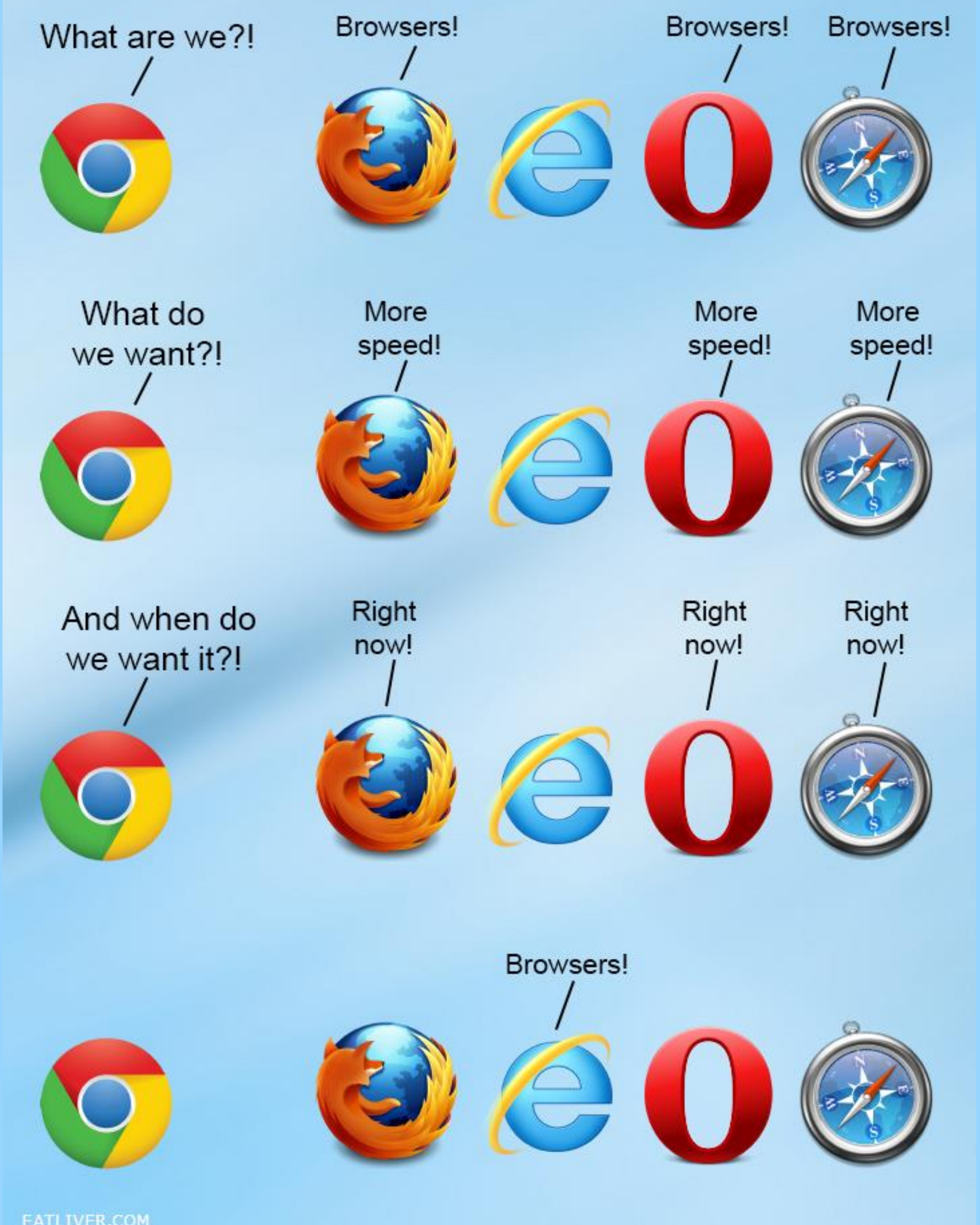
Other proposed ways

- Eckersley paved the way of stateless tracking through fingerprinting
- After Eckersley, other researchers proposed ways of fingerprinting browsers, based on:
  - Speed
  - Implementation coverage
  - Rendering of elements

They will know you by your speed…

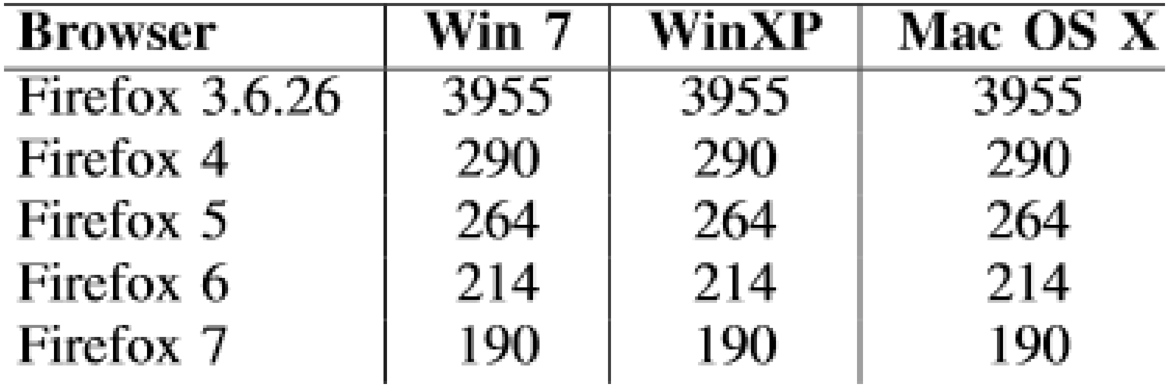
- Mowery et al. (W2SP 2011) proposed the use of performance benchmarks to tell different JavaScript engines apart
  - Different JavaScript engine -> Different browser

- Collected performance signatures (39 tests) from approx. 1000 users
  - 98.2% correct browser family detection
  - Overall accuracy (versions included): 79.8%


What are we?!
Browsers!
Browsers!
Browsers!
What do we want?!
More speed!
More speed!
More speed!
And when do we want it?!
Right now!
Right now!
Right now!
Browsers!

As well as your features…

- Mulazzani et al. (W2SP 2013) proposed the use of missing functionality in JavaScript engines
  - Different browsers, implement JavaScript standards, at a different rate



Browser	Win 7	WinXP	Mac OS X
Firefox 3.6.26	3955	3955	3955
Firefox 4	290	290	290
Firefox 5	264	264	264
Firefox 6	214	214	214
Firefox 7	190	190	190

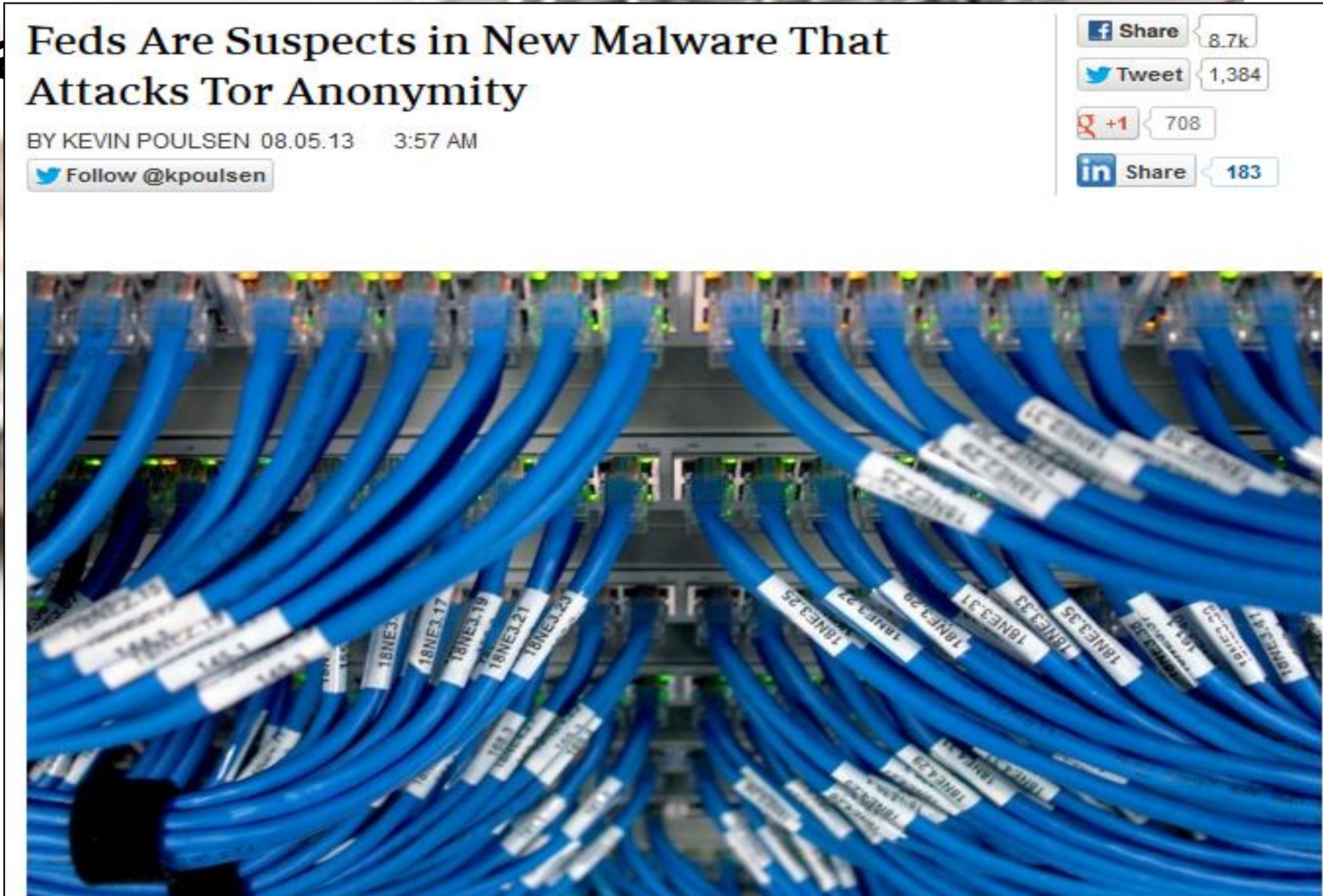
cHRome!
FireFox!
GLUE

As well as your artistic talent

- Mowery et al. (W2SP) proposed the use of the HTML5 canvas to detect browser-specific renderings of the same string
  - Write some text in canvas, read it out as an image
  - Different browsers/hardware combinations will create slightly different images
  - http://jsbin.com/ePAheCi/2/edit

ADS

Pa



Feds Are Suspects in New Malware That Attacks Tor Anonymity

BY KEVIN POULSEN 08.05.13 3:57 AM

EU Cookie law

“Member States shall ensure that the storing of information, or the gaining of access to information already stored, in the terminal equipment of a subscriber or user is only allowed on condition that the subscriber or user concerned has given his or her consent, having been provided with clear and comprehensive information, in accordance with Directive 95/46/EC, inter alia, about the purposes of the processing. This shall not prevent any technical storage or access for the sole purpose of carrying out the transmission of a communication over an electronic communications network, or as strictly necessary in order for the provider of an information society service explicitly requested by the subscriber or user to provide the service.”

Does it apply to fingerprinting?

What’s happening out there?

- In mid 2012, all we knew is that fingerprinting is possible and that a small number of companies offer it as a service

- Questions that begged answering:
  - How are they doing it?
  - Could they do more?
  - Who is using them?
  - How are users trying to hide?
    - Is it working?

Manual analysis of 3 fingerprinting companies

bluecava
where devices rule.

ThreatMetrix™

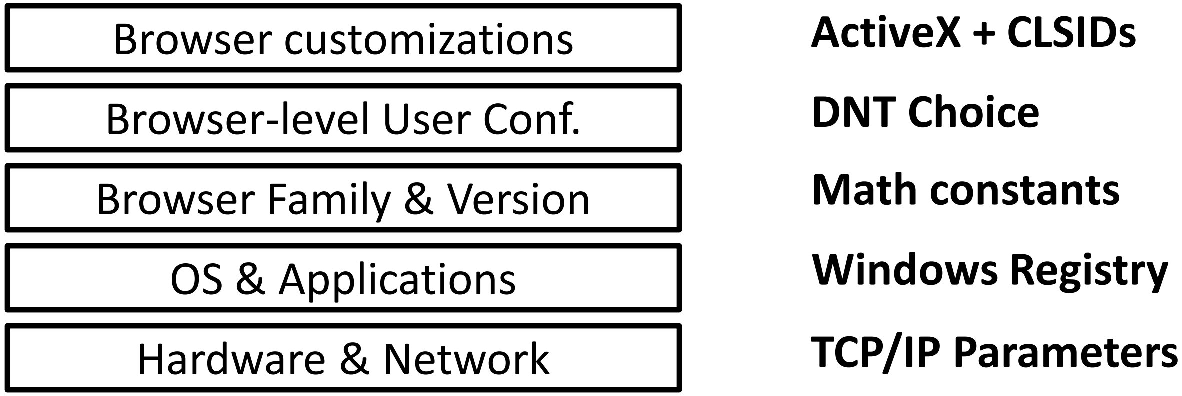
iovation®

1. Find the domains that they use to serve their fingerprinting scripts
2. Find some websites that use them and extract the code
3. De-obfuscate and analyze
4. Compare and classify

```
return;}var _i_b=_i_aa.getElementById(window.io_bbout_element_id);_i_b["value"]=_if_fa;}func
ndow.io_bb_callback:__if_d;_i_c(_if_fa,_if_fb);}var _i_d={__if_p:function(_if_fc){return _if_
l(_if_fc.getUTCDate(),2)+" "+this.__if_ad(_if_fc.getUTCHours(),2)+":"+this.__if_ad(_if_fc.get
r _i_e=_if_fd.toString(16);return(_i_m)?this.__if_ad(_i_e,_i_m):_i_e;},__if_u:function(_i_bz)
odeAt(_i_g);if(_i_h>=56320&&_i_h<57344)continue;if(_i_h>=55296&&_i_h<56320){if(_i_g+1>=_i_bz.
nue;_i_h=((_i_h-55296)<<10)+(s-56320)+65536;}if(_i_h<128)_i_f+=String.fromCharCode(_i_h);els
_f+=String.fromCharCode(224+(_i_h>>12),128+((_i_h>>6)&63),128+(_i_h&63));else _i_f+=String.fr
rn _i_f;},__if_y:function(_if_fe){if(typeof(encodeURIComponent)=="function")return encodeURI
length;_i_g++){var _i_k=_i_j.charAt(_i_g);var _i_l=new RegExp("[a-zA-Z0-9-_.!~*'()]");_i_f+=
nction(_i_bz,_if_ff){var _i_m="";var _i_n=_if_ff-_i_bz.length;while(_i_m.length<_i_n)_i_m+="
IJKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz0123456789+/=",__if_aj:function(_i_bz){var _i_e="
_i_bz.charCodeAt(_i_g+1);var _i_r=_i_bz.charCodeAt(_i_g+2);var _i_s=_i_p>>2;var _i_t=((_i_p&3
v=64;}else if(isNaN(_i_r)){_i_v=64;}_i_e=_i_e+this._i_ej.charAt(_i_s)+this._i_ej.charAt(_i_t)
nction(_i_bz){var _i_w="";var _i_x,chr2,chr3="";var _i_s,_i_t,_i_u,_i_v="";var _i_g=0;var _i
.indexOf(_i_bz.charAt(_i_g++));_i_t=this._i_ej.indexOf(_i_bz.charAt(_i_g++));_i_u=this._i_ej.
2)|(_i_t>>4);chr2=((_i_t&15)<<4)|(_i_u>>2);chr3=((_i_u&3)<<6)|_i_v;_i_w=_i_w+String.fromCharC
ring.fromCharCode(chr3);_i_x=chr2=chr3="";_i_s=_i_t=_i_u=_i_v="";}while(_i_g<_i_bz.length);re
el:12,_i_em:false,_i_en:"",_i_eo:"",_i_ep:true};if(typeof(window.io_install_stm)!="boolean")w
.io_install_flash=_i_z._i_em;if(typeof(window.io_exclude_stm)!="number")window.io_exclude_st
ab_url===undefined)window.io_stm_cab_url=_i_o.__if_aq("aHR0cHM6Ly9tcHNuYXJlLmllc25hcmUuY29t")
l_stm_error_handler===undefined)window.io_install_stm_error_handler=_i_z._i_en;if
_needs_update_handler===undefined)window.io_flash_needs_update_handler=_i_z._i_eo;if(typeof(w
function(_if_fg){if(_if_fg===undefined)return null;if(typeof(_if_fg)=="object"&&_if_fg.tagNam
etElementsByName(_if_fg);for(var _i_g=0;_i_g<_i_ab.length;_i_g++)if(_i_ab[_i_g]._i_dc&&_i_ab[
```

Results

- After extracting all features, we created a taxonomy of all fingerprinted features, and compared each company to Panopticlick
- Collectively, Panopticlick was fully covered



Browser customizations — **ActiveX + CLSIDs**

Browser-level User Conf. — **DNT Choice**

Browser Family & Version — **Math constants**

OS & Applications — **Windows Registry**

Hardware & Network — **TCP/IP Parameters**

Non-trivial extras

- Non-plugin font detection
  - Comparison of text’s width & height
- Native Fingerprinting plugins
  - Accessing highly-specific registry value
- Fingerprint delivery mechanisms
- Proxy detection

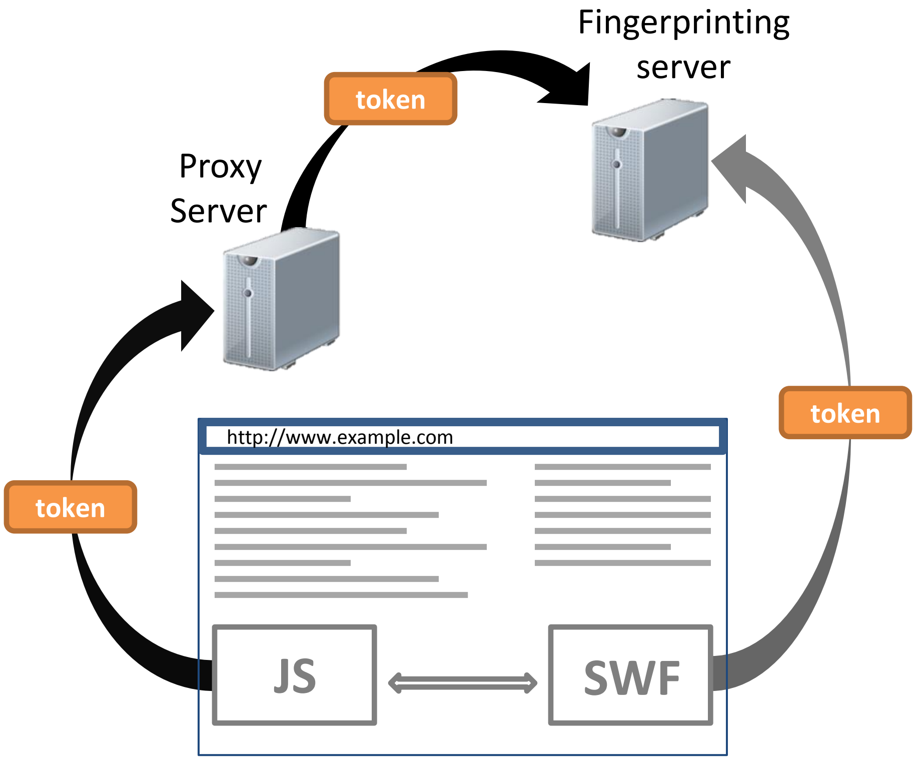
Font Detection through JavaScript

String	**Dimensions**
I_DO_NOT_NEED_FLASH	500 x 84
I_DO_NOT_NEED_FLASH	520 x 84
I_DO_NOT_NEED_FLASH	580 x 87
I_DO_NOT_NEED_FLASH	399 x 82

Non-trivial extras

- Non-plugin font detection
  - Comparison of text’s width & height
- Native Fingerprinting plugins
  - Accessing highly-specific registry values
- Fingerprint delivery mechanisms
- Proxy detection

Proxy-detection



Demo

http://www.orbitz.com

Adoption

Dataset A

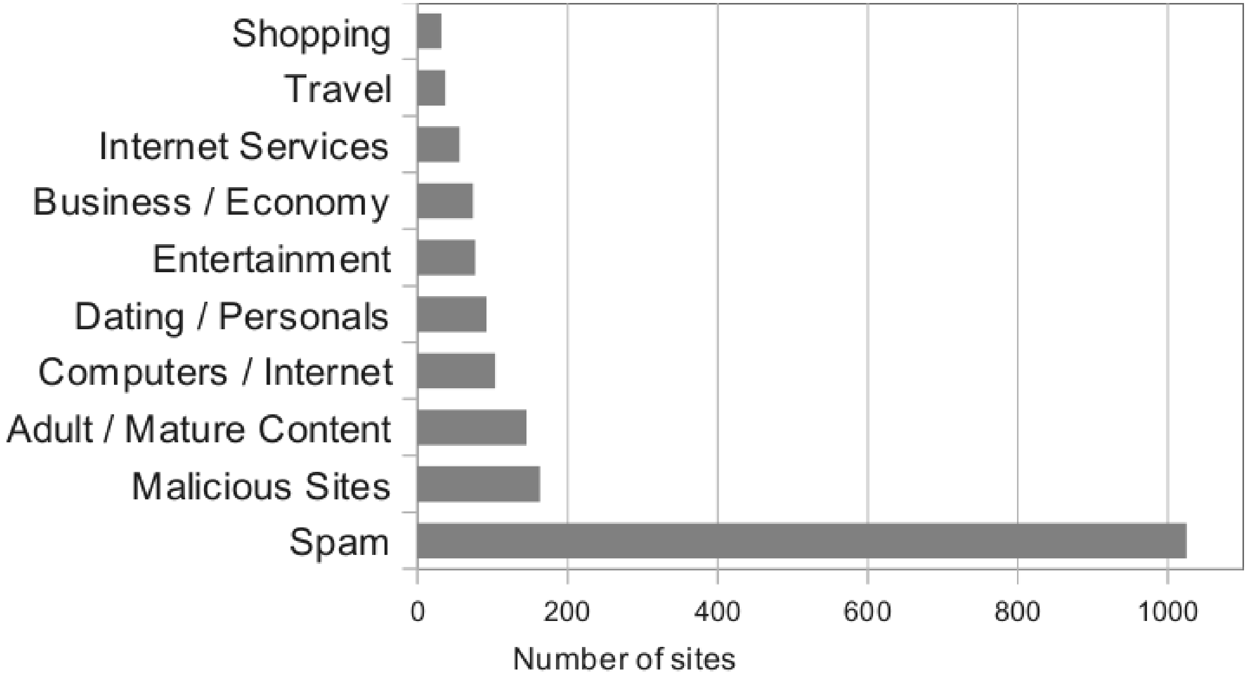
- Crawled top 10,000 sites, searching for inclusions from the 3 fingerprint providers

- 40 sites discovered
  - Porn & dating sites most prominent
    - Shared credentials & Sybil attacks
  - skype.com the highest ranking one

Adoption

Dataset B

- 3,804 domains from Wepawet



But wait… there's more!

- Can we find unknown fingerprinting parties?
  - How do we separate a fingerprinting script from a generic analytics script?

- Fonts!
  - Separating feature between normal analytics and fingerprinting
  - Second most identifying feature according to Eckersley

FPDetective

- Fingerprinting-sensitive crawler
  - If fonts are touched, record site

- Detection of font snooping
  - JS-based font probing (Modified browser)
  - Flash-based font probing (decompilation of Flash)

Adoption (revisited)

Dataset A

- Crawled top 10,000 sites, searching for inclusions from the 3 fingerprint providers

- 40 sites discovered
  - Porn & dating sites most prominent
    - Shared credentials & Sybil attacks

  - skype.com the highest ranking one

Adoption (revisited)

Dataset A

– Cra sions

fro

– 40

- F

– Shared credentials & Sybil attacks

- skype.com the highest ranking one

- **145 fingerprinting sites in the top Alexa 10K**

- **DNT does not matter**

Status

- Fingerprinting is out there
  - Quite a number of new techniques over Panopticlick
- Large and popular sites are using them
- Could they be doing more?
  - How do the browser internals relate to a browser’s identity?

DIY Fingerprinting

- We decided to try some fingerprinting of our own
- Focus on the two special JS objects that fingerprinters probe the most:
  - navigator
  - screen
- Perform a series of everyday operations and search for differences across browsers
  - Add properties
  - Remove properties
  - Modify properties

Novel methods discovered

- E.g., Natural ordering of properties can give away a browser family, and occasionally, a browser version

navigator.geolocation	navigator.appCodeName	⟷	navigator.appCodeName
navigator.onLine	navigator.appName	⟷	navigator.appName
navigator.cookieEnabled	navigator.appVersion		navigator.appMinorVersion
navigator.vendorSub	navigator.language		navigator.cpuClass
navigator.vendor	navigator.mimeTypes		navigator.platform

Status

- Fingerprinting is out there
  - Quite a number of new techniques over Panopticlick
- Large and popular sites are using them
- There could be more fingerprinting done by the companies
- How could a user react?

Browser extensions

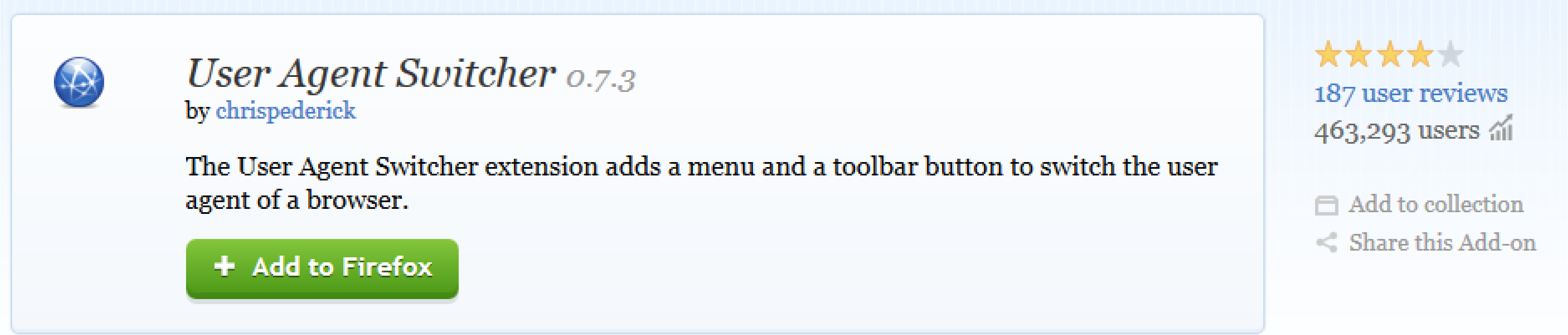
- Reviewed 11 different browser extensions that spoof a browser’s user-agent
  - 8 Firefox + 3 Chrome
  - More than 800,000 users
- Advice to use such extensions:
  - Previous research in web tracking
  - Underground hacking guides
- How do they stand-up against fingerprinting?

Worse than nothing…

- All of them had one or more of the following:
  - Incomplete coverage of the navigator object
  - Impossible configurations
  - Mismatch between UA header and UA property

- Iatrogenic problem:
  - When installing these, a user becomes more visible and more fingerprintable than before

Case Study



User Agent Switcher 0.7.3
by chrispederick

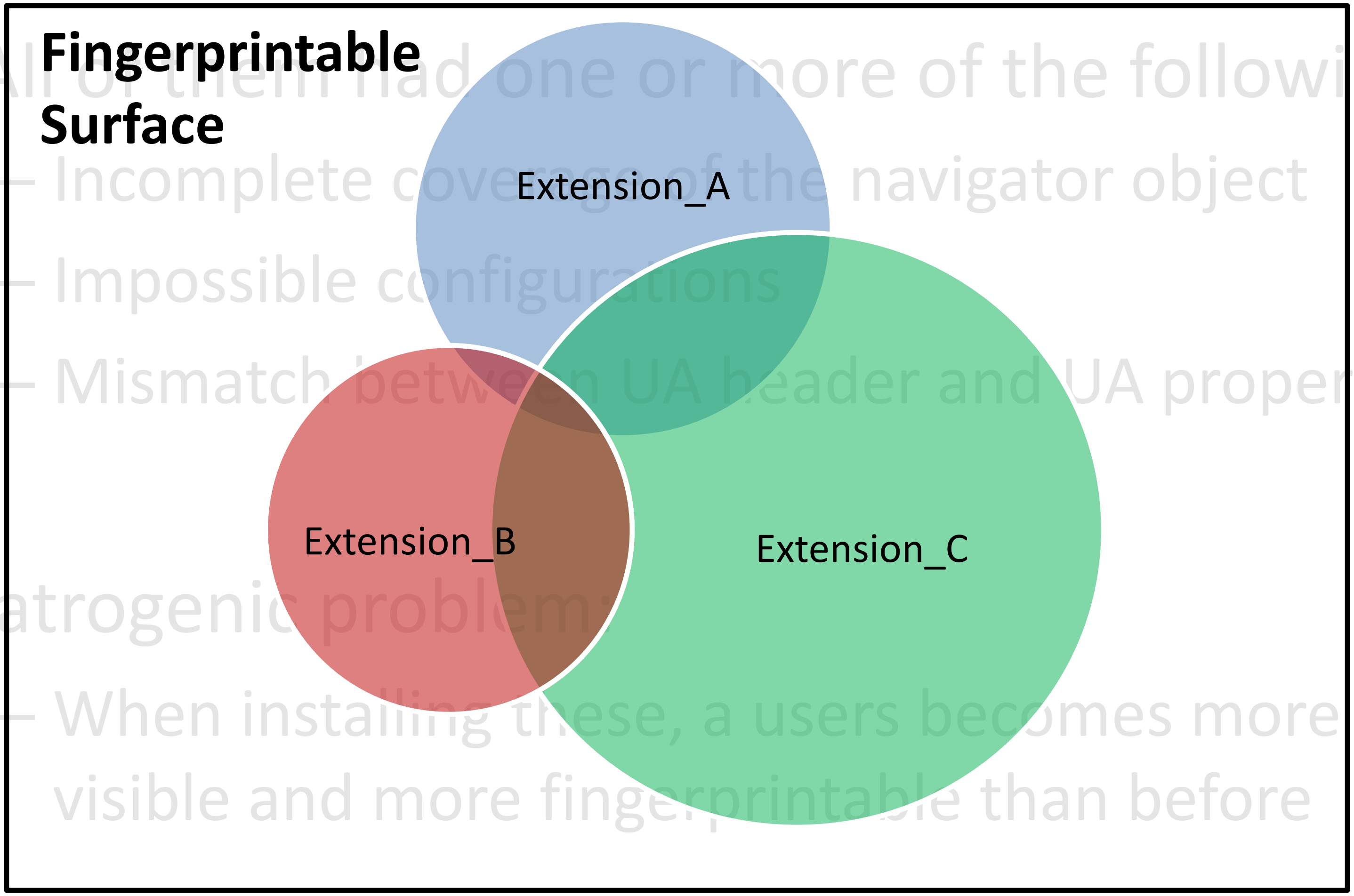
The User Agent Switcher extension adds a menu and a toolbar button to switch the user agent of a browser.

Stats

- 463,293 users
- 187 user reviews
- 4/5 starts

Worse than nothing…

- All of them had one or more of the following:
  - Incomplete coverage of the navigator object
  - Impossible configurations
  - Mismatch between UA header and UA property
- Iatrogenic problems:
  - When installing these, a users becomes more visible and more fingerprintable than before

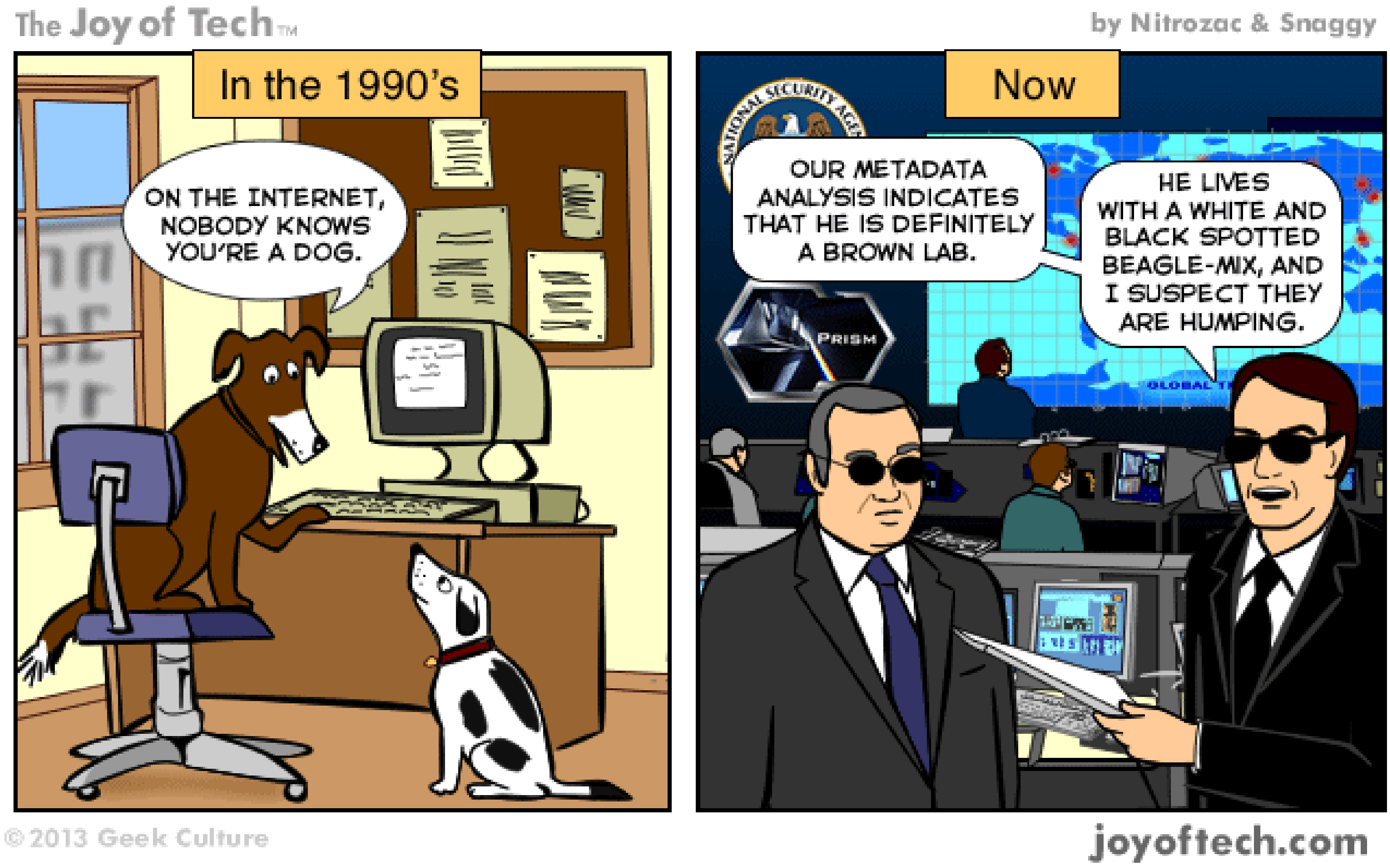


Defenses (today)

- The more generic your system is, the better
  - The more exotic plugins and extensions you have installed, the more chances of being singled out
- Fingerprinters can be black-listed
- Disabling Flash and Java will definitely help
  - No explicit font collection
- Virtual machines? Browsers from a stick?
  - Depends on your balance between hassle and privacy

Conclusion

- Web tracking is so much more than cookies
- Fingerprinting is a real problem
- Browsers are so complex that it is really hard to make them seem identical
- Current browser extensions should not be used for privacy reasons
- Long term solutions will most-likely not be pure technical ones
  - Legislation required, like in stateful tracking


The Joy of Tech™
by Nitrozac & Snaggy
In the 1990's
ON THE INTERNET, NOBODY KNOWS YOU'RE A DOG.
Now
NATIONAL SECURITY AGENCY
OUR METADATA ANALYSIS INDICATES THAT HE IS DEFINITELY A BROWN LAB.
HE LIVES WITH A WHITE AND BLACK SPOTTED BEAGLE-MIX, AND I SUSPECT THEY ARE HUMPING.
PRISM
GLOBAL T
© 2013 Geek Culture
joyoftech.com

nick.nikiforakis@cs.kuleuven.be
http://www.securitee.org