

PERSONAL TRACKING DEVICES

A JOURNEY INTO THE TRUE DARK NET

Silvia Puglisi - silviap@ieee.org / [@nopressure](https://twitter.com/nopressure)

HELLO WORLD

My name is Silvia Puglisi. I am a software engineer and Ph.D. candidate at UPC Barcelona Tech.

I research in privacy and web science.

I am here to talk you about the real dark web.

WHAT IS THIS ALL ABOUT?

- Marketing
- Privacy
- User tracking
- Online footprint
- Identity
- Control

WHAT HAS MARKETING TO DO WITH THIS?

Think of the so-called *new economy* without online marketing and advertising..

Aehm.. You probably can't.

Most of the successful *online* companies of the last.. make it 20 years.. use advertising to sustain their business at least in part.

THE SOURCE OF ADVERTISING

Advertising wants you to buy products.

So advertising companies are doing all that they can to know you so they can recommend you products that you are more likely to buy.

The source of advertising is **data**.

Data about you.

HOW DOES ONLINE MARKETING WORK?

eMarketing follows users in their online and sometimes offline activities.

Information about what users do and are interested in are collected by websites, applications, devices.

This information is crawled, analysed and categorised.

WHAT ABOUT ONLINE PRIVACY?



If our actions on the web are constantly collected and analysed.
Do we have online privacy? Have we lost our right to be
anonymous?

In an online context, the right to privacy has commonly been interpreted as a right to “information self-determination”.

Acts typically claimed to breach online privacy concern the collection of personal information without consent, the selling of personal information and the further processing of that information.

IS PRIVACY THE RIGHT TO BE FORGOTTEN?

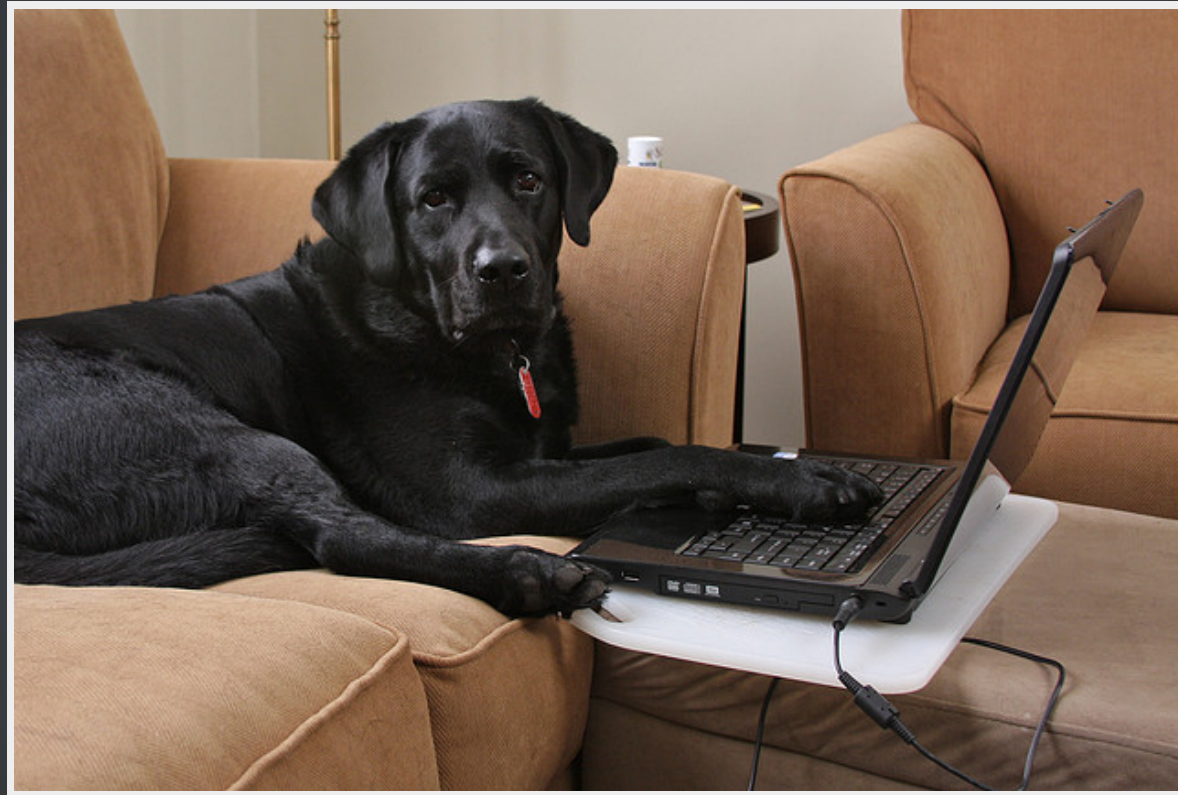
In 2011, the amount of digital information created and replicated globally exceeded 1.8 zettabytes (1.8 trillion gigabytes).

75% of this information is created by individuals through new media fora such as blogs and via social networks.

By the end of 2011, Facebook had 845 million monthly active users, sharing over 30 billion pieces of content.

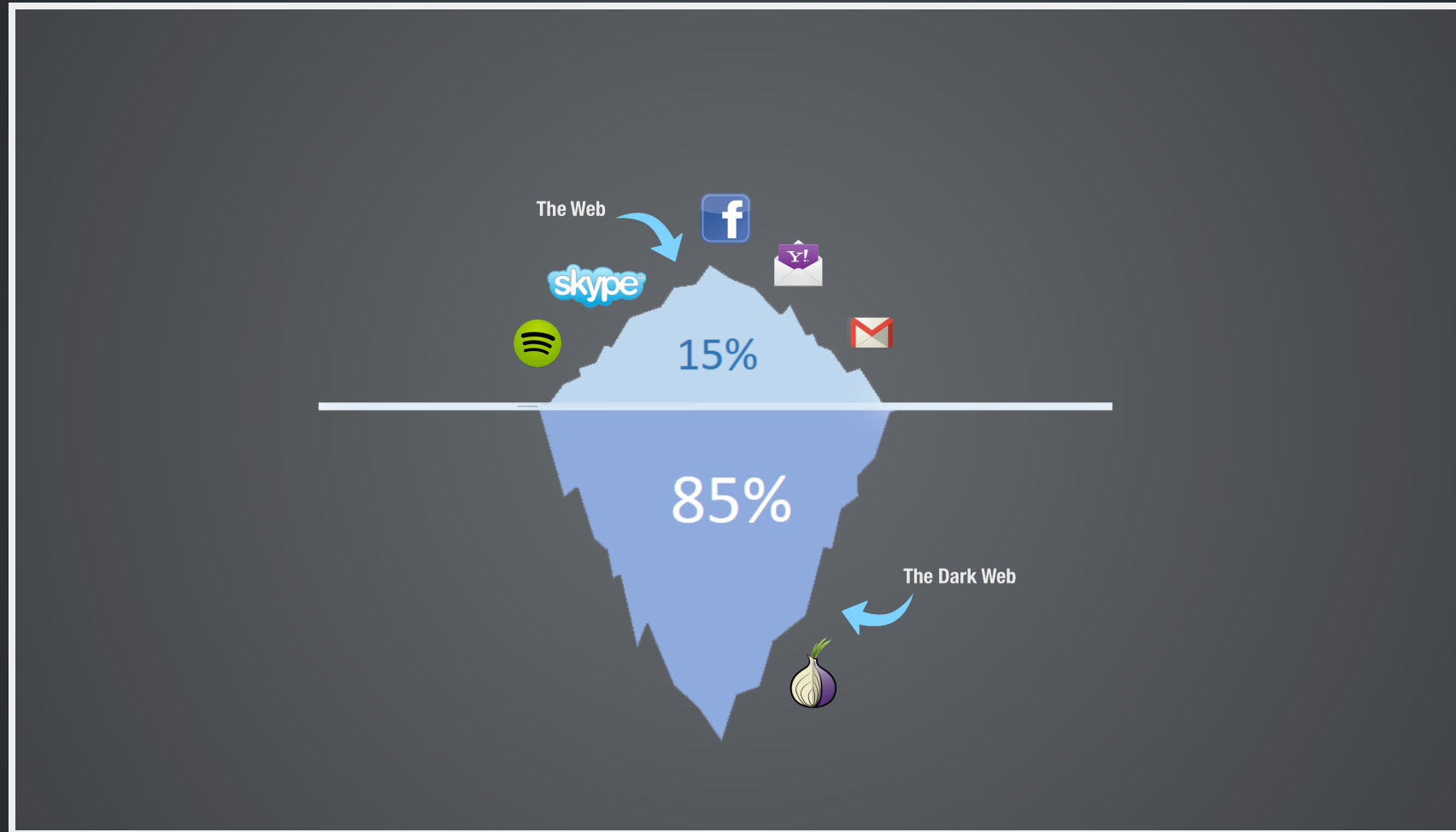
Library Briefing - Library of the European Parliament - 01/03/2012

**A FEW YEARS AGO ON THE INTERNET NOBODY KNEW YOU
WERE A DOG.**



Now they probably also know the colour of your fur.

THIS IS THE DARK WEB



Or at least this is how "they" pictured it.

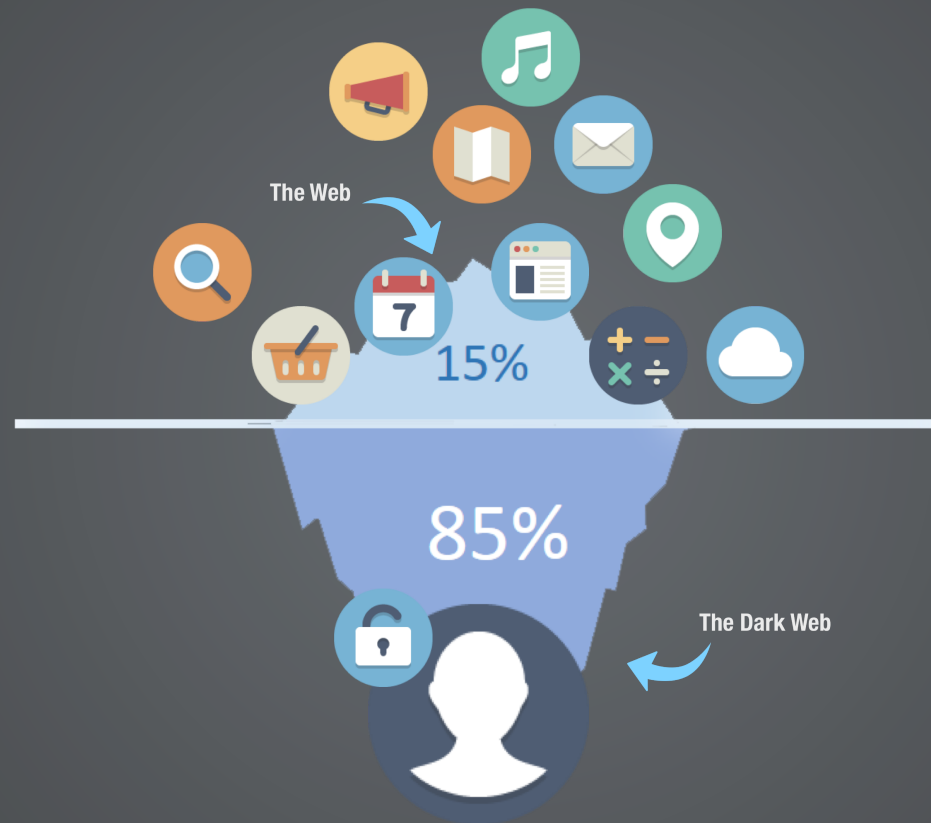
THE DARK WEB IS THE WEB THAT CANNOT BE CRAWLED.

In a way is the web that companies can reach and control.

LET US SWITCH THIS MYTH AROUND

What if things were actually a bit different..

THE DARK WEB OF MARKETING



THE DARK WEB OF MARKETING IS THIS CONCEPT THAT WE USE SOFTWARE AND HARDWARE THAT WE DO NOT CONTROL.

- We do not know how these object are made.
- We cannot make modifications.
- They collect a huge amount of data about ourselves.
- We are perfectly content with it.

THE AGE OF THE METADATA

“meta-data” is collected and stored by public and private organisations about where, when and who created and accessed a particular online content.

Websites have embeded structured data for a few years now.

-

Structured data is used to describe product, services, events, and make user information available already into their HTML pages using markup standards such as Microformats, Microdata and RDFa.

LET'S MAKE A FEW EXAMPLES





Google Conversion Tracking


```
<html>
  <body>
    <!-- Below is sample text link with a phone number.
           You need to replace the number with your own
           phone number and the CALL NOW text with the text
           you want to hyperlink. -->
    <a onclick="goog_report_conversion('tel:949-555-1234')"
       href="#" >CALL NOW</a>
  </body>
</html>
```


HOW DOES CONVERSIONS TRACKING WORK?

Let's say your Google search ads send people to your website where they research and learn more about your business. Website call conversions dynamically inserts a Google forwarding number on your website that measures the calls made by these customers. Whether they click on the number or dial it directly from their phone, you can attribute the call conversion and conversion value back to the keyword and ad that drove the customer. You can learn more about setting up website call conversions [here](#).

MOBILE APP DATA



 R 60% 10:00


 Network Log

Logging On

Logfile size: 2.45MB Traffic: 0bps/0bps

Log

Apps




SRC: 192.168.1.100

DST: 74.125.227.196

LEN: 52

2014-12-23 09:59:24.093



(10018) Network Location


TCP/wlan0

SRC: 192.168.1.100

DST: 64.233.160.188

LEN: 147

2014-12-23 09:59:24.094



(10018) Network Location


TCP/wlan0

SRC: 64.233.160.188

DST: 192.168.1.100

LEN: 52

2014-12-23 09:59:24.095



(10018) Network Location


TCP/wlan0

SRC: 192.168.1.100

DST: 64.233.160.188

LEN: 145

2014-12-23 09:59:24.096



(10018) Network Location


TCP/wlan0

SRC: 64.233.160.188

DST: 192.168.1.100

LEN: 52

2014-12-23 09:59:24.096



(10159) Facebook

TCP/wlan0

SRC: 192.168.1.100

41343

GRAPH

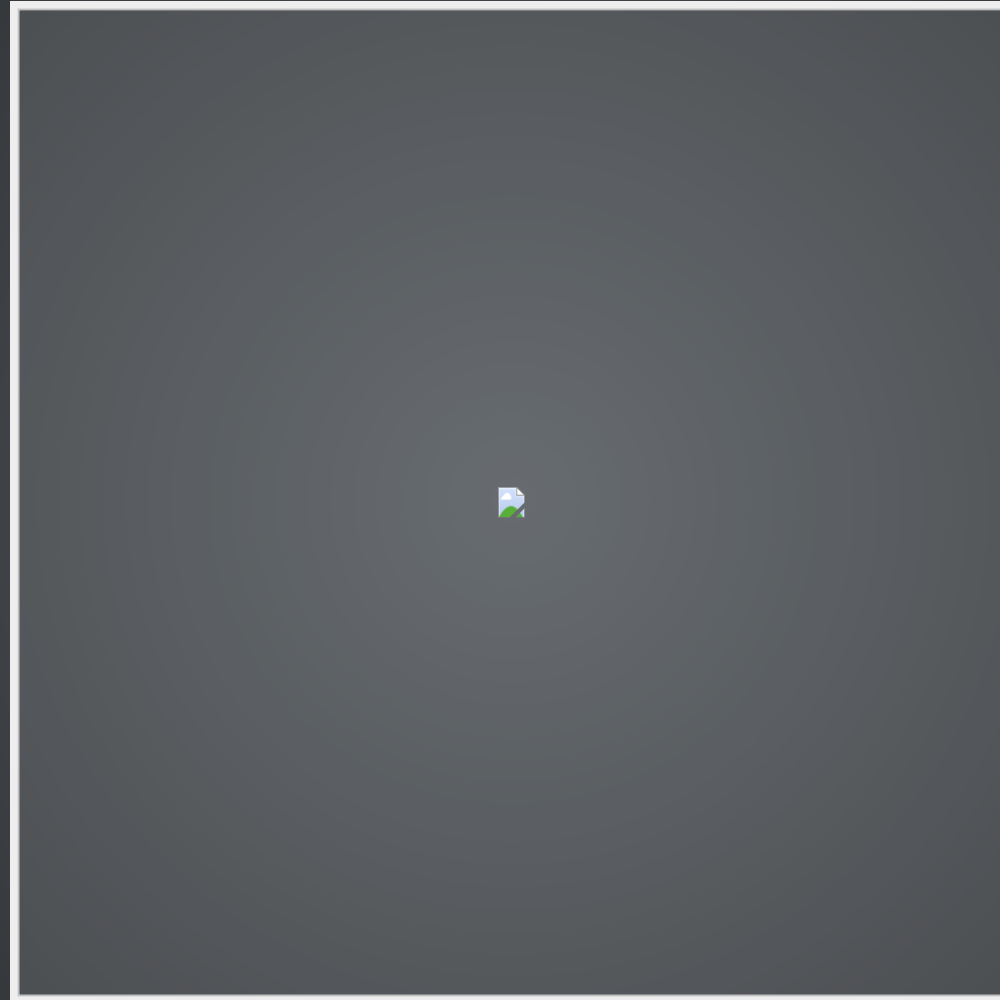
FILTER

CLEAR

WEARABLES FOR WELLBEING

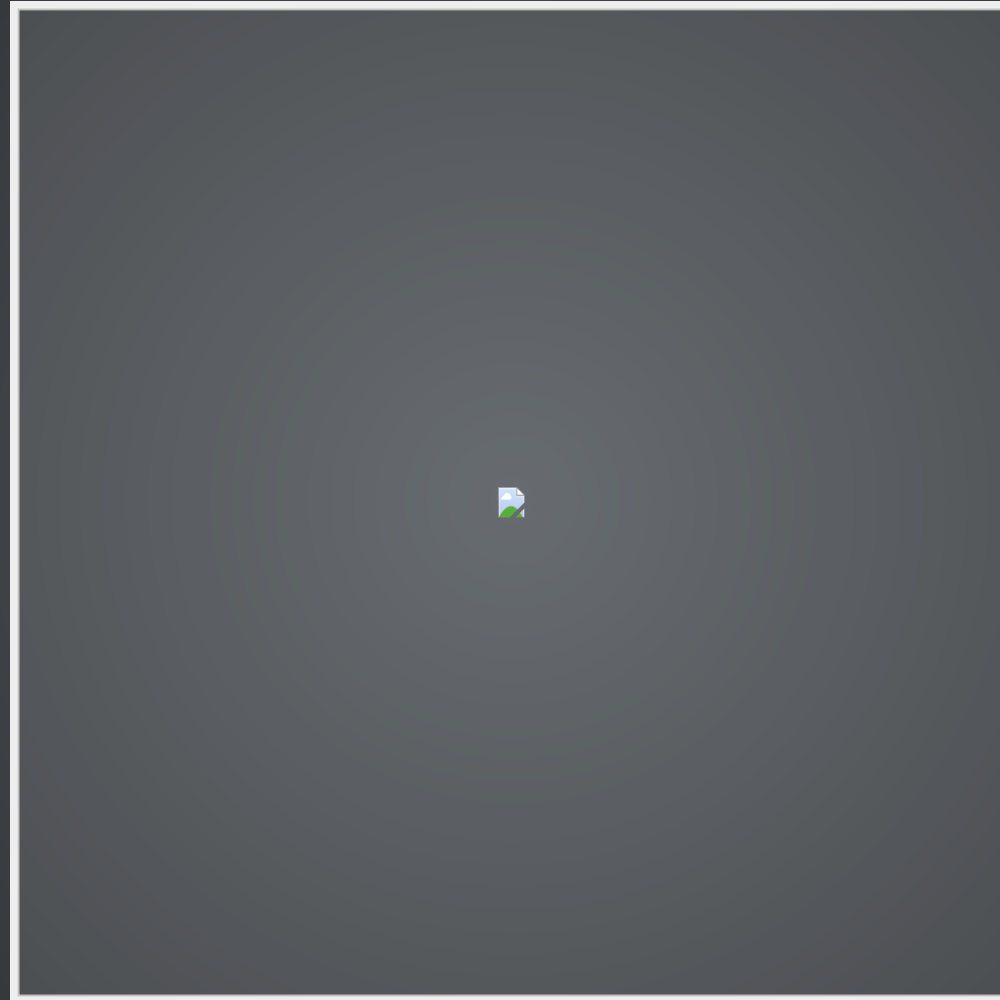


WEARABLES FOR PRODUCTIVITY



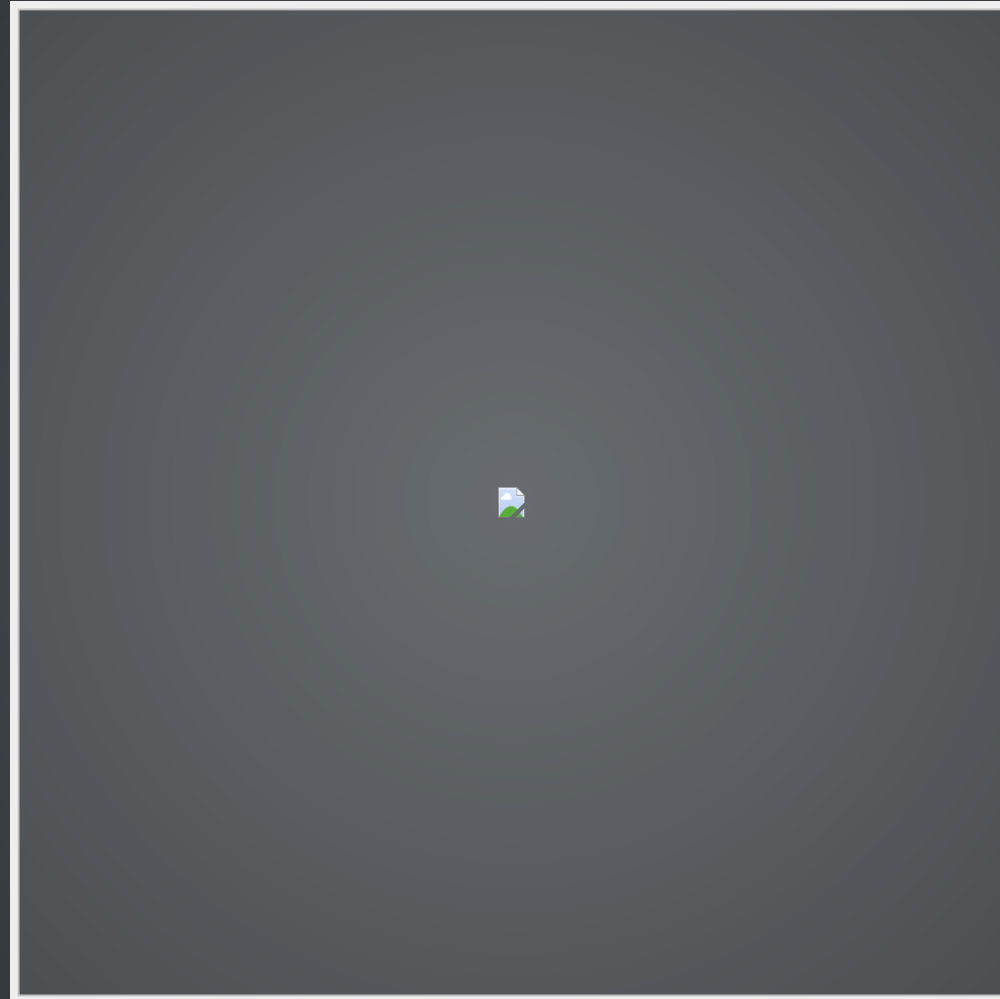
* <http://www.telegraph.co.uk/technology/news/10801687/Wearables-can-boost-employee-productivity-by-almost-10pc.html>

WEARABLES FOR ...



* <http://time.com/103510/google-glass-privacy-foregrounding/>

WEARABLES FOR ...



* <http://www.engadget.com/2014/08/07/bondara-sexfit-i-used-to-be-a-serious-journalist/>

WEBSITE ADVERTISING

```
src="http://pubads.g.doubleclick.net/gampad/ads?gdfp_req=1&
correlator=2213477108582874&output=json_html&
callback=callbackProxy&impl=fif&sfv=1-0-1&
iu=%2F1004060%2FMNT_MREC_Right&sz=300x250&
cust_params=main%3Danxiety%26sub%3Dpsychology-psychiatry%252Cneurology
&cookie_enabled=1&lmt=1419215950&dt=1419215950968&cc=100
&frm=20&biw=1126&bih=623&oid=3&adx=758&ady=378
&adk=3903700250&color_bg=FFFFFF&gut=v2&ifi=1
&u_tz=60&u_his=22&u_java=true&u_h=800
&u_w=1280&u_ah=725&u_aw=1280&u_cd=24&u_nplug=7
&u_nmime=65&u_sd=2&flash=16.0.0&
url=http%3A%2F%2Fwww.medicalnewstoday.com%2Farticles%2F287291.php&
vrg=55&vrp=55&ga_vid=182404171.1419215951&ga_sid=1419215951
&ga_hid=1886735435"
```


WEBSITE ADVERTISING

```
<meta name="title" content="Advice for ADD Adults Experiencing Holiday  
Overload"><meta name="copyright" content="Copyright (c) 1998 - 2013  
New Hope Media LLC. All rights reserved."><meta name="keywords"  
content="ADD symptoms, ADHD information, help, diagnosis, treatment,  
attention deficit disorder resources, attention deficit hyperactivity  
disorder tools, medication, ADD in adults, ADHD in children, forums,  
community, blog, learning disabilities, dyslexia, bipolar, anxiety,  
Adderall, Concerta, Vyvanse, Ritalin, Strattera, Prozac adult add,  
married to adhd, adhd during the holidays, Travel and Vacation Tips,  
ADHD and Marriage, ADHD and Relationships, ADHD Social Skills">
```

**EVERY HTTP CONNECTION IN THE BROWSER HAS DIFFERENT
SELECTOR ASSOCIATED WITH IT**



HYPERDATA & HYPERMEDIA

Hyperdata indicates data objects linked to other data objects in other places, as hypertext indicates text linked to other text in other places.

Hyperdata enables formation of a web of data, evolving from the "data on the Web" that is not inter-related (or at least, not linked).

Hypermedia, an extension of the term hypertext, is a nonlinear medium of information which includes graphics, audio, video, plain text and hyperlinks.

- Source: Wikipedia

RESTFUL ARCHITECTURES

REST, an architectural style introduced by Roy Thomas Fielding in 2000, which has been at the core of the web design and development.

REST represents an abstraction over the actual architecture of the web.

In REST identification, representation and format are independent concepts.

SPECIFICALLY:

- An URI can identify a resource without knowing what formats the resource uses to exchange representations.
- Likewise the protocols and representations used by the resource to communicate can be modified independently from the URI identifying the resource.



REST INTERFACES

The uniformity of REST interfaces is build upon four guiding principles:

- The identification of resources through the URI mechanism.
- The manipulation of resources through their representations.
- The use of self-descriptive messages.
- Implementing hypermedia as engine of the application state (HATEOAS)

WHY HYPERMEDIA MATTERS FOR PRIVACY PROTECTION

- Information self-determination is not even possible if users have no control on their online footprint.
- Hypermedia provides context over unstructured footprint information.
- Users and applications use REST interfaces to interact with one another exchanging resource representations.
- The web follows REST principles and so do users' online traces.

EACH SELECTOR CONTRIBUTE TO THE BUILD A USER PROFILE.



Each selector makes your profile more unique, because it adds up more information.



* <http://apps.washingtonpost.com/g/page/national/nsa-signal-surveillance-success-stories/647>

UNIQUE FOOTPRINTS

While you surf the web you carry a unique footprint:

- HTTP cookies, often set by with 3rd party analytics and advertising domains
- sessions and storage.
- Browser or device personalised settings. Resolution, Fonts, Plugins.
- Accounts ID.
- Devices ID.

* <https://panoptickick.eff.org/browser-uniqueness.pdf>.

HOW UNIQUE IS A FOOTPRINT?

To answer this question we can do a set of different things.

- We can profile our activity within a set of categories
- We can calculate how many bits of information are introduced everytime we add a feature to our profile.
- We can analyse how many unique feature we are sharing across the network.

PROFILING OVER A SET OF CATEGORIES

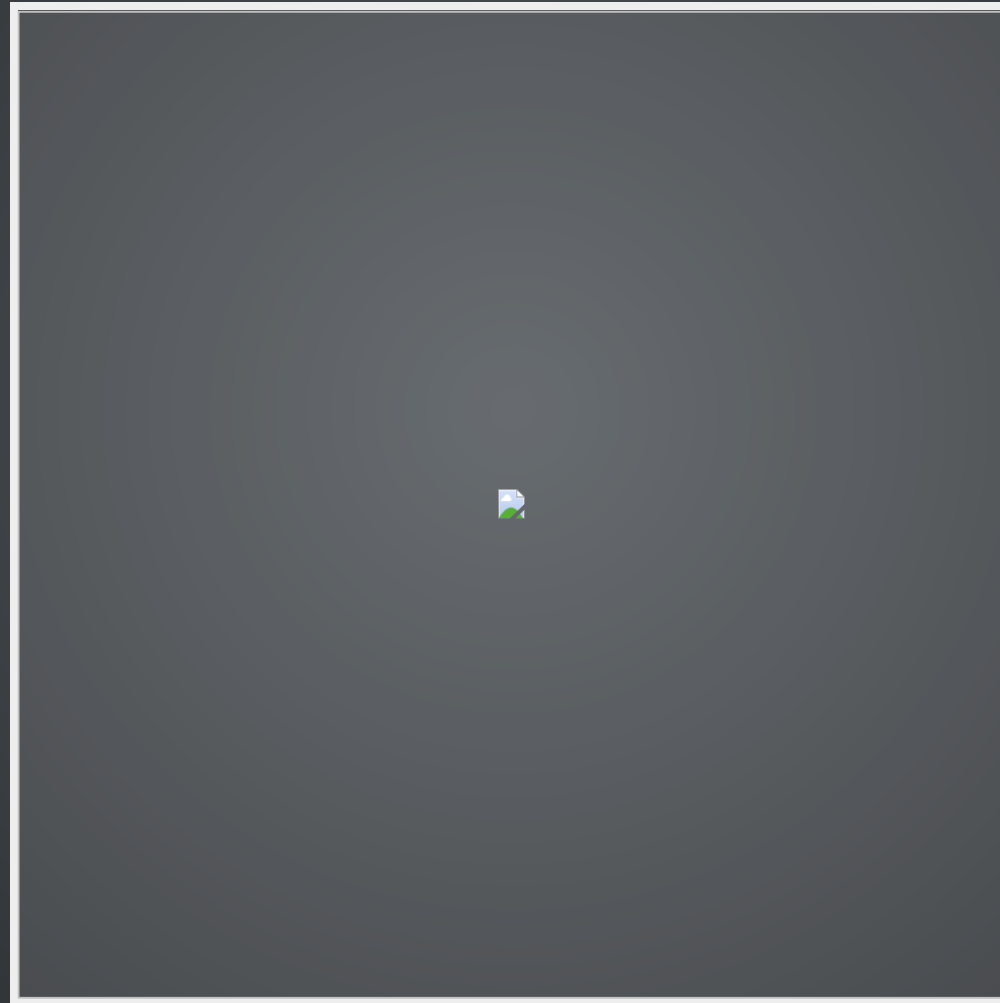


CALCULATING HOW MANY BITS OF INFORMATION ARE INTRODUCED BY UNIQUE FEATURES

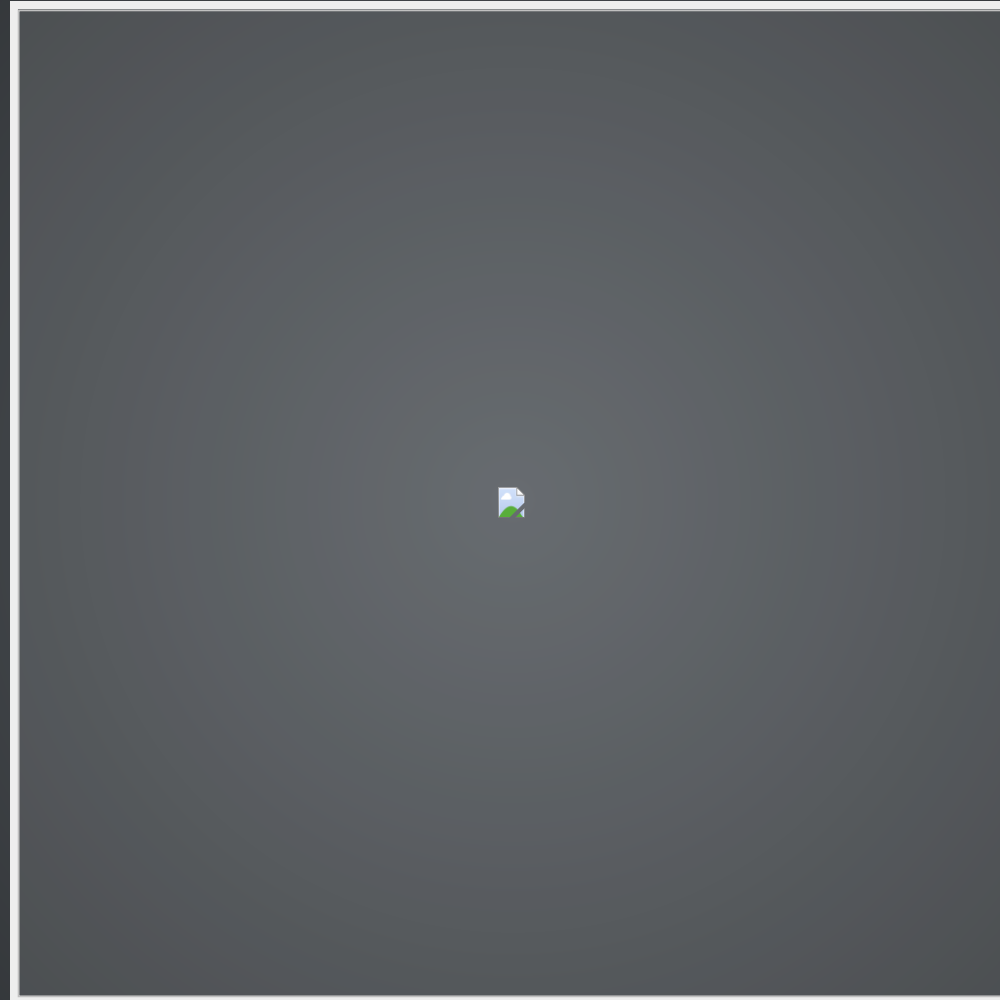
* <https://panopticlick.eff.org>



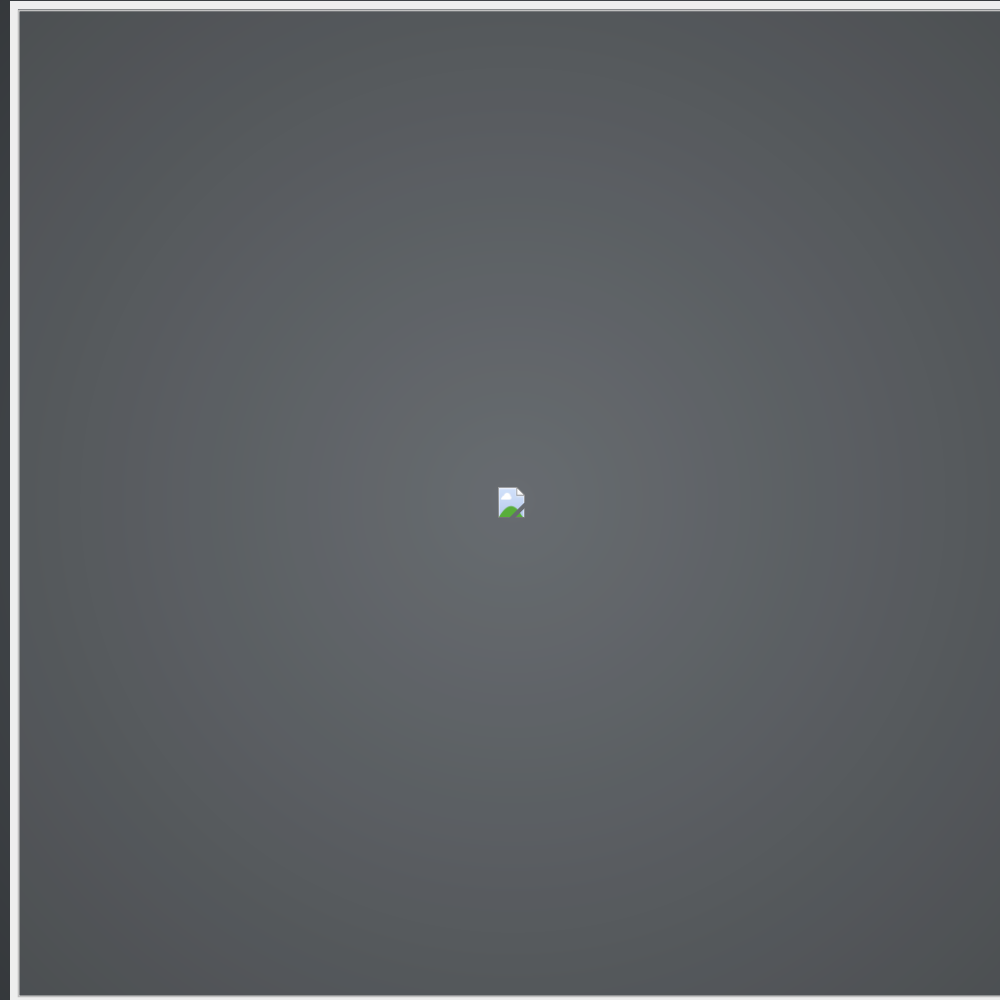
ANALYSING HOW MANY UNIQUE TRACES WE LEAVE AROUND THE WEB



WHAT IS THE CLOUD REALLY?

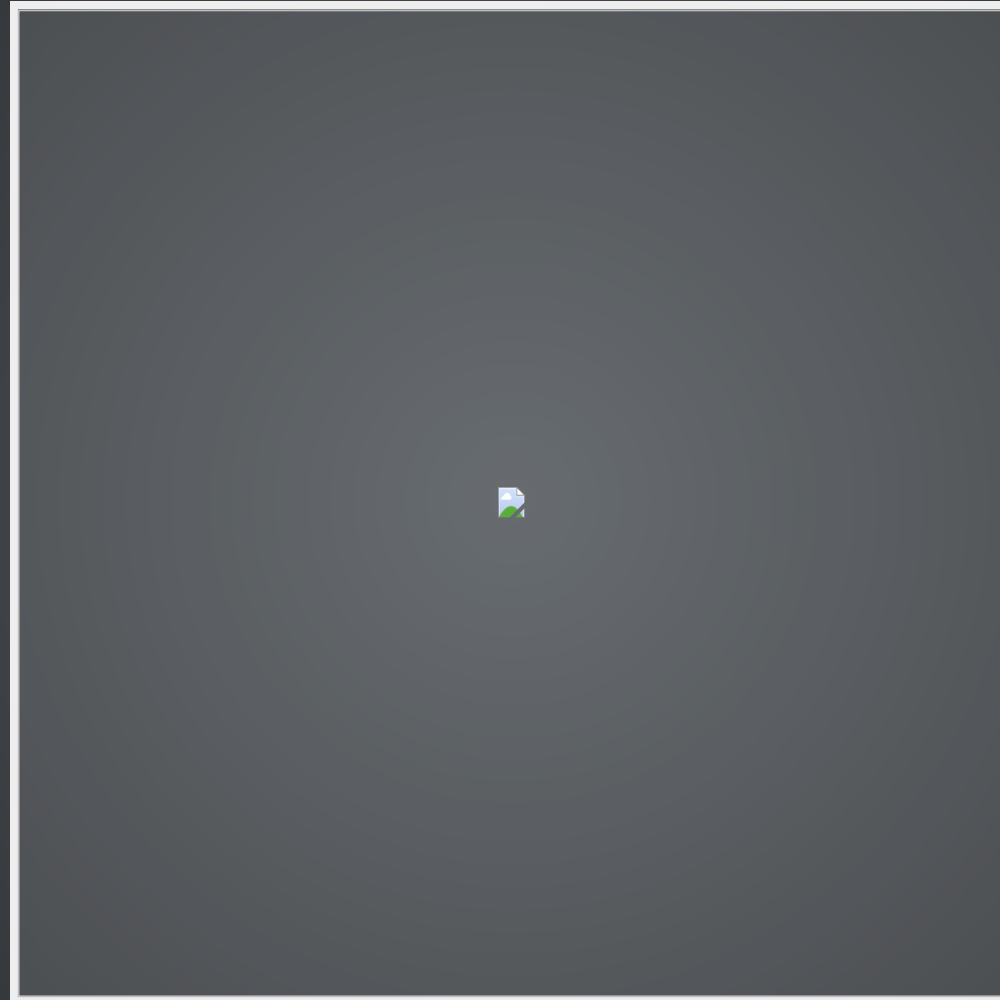


WHO OWNS THE CLOUD?





WHAT IF WE GO BEYOND CLOUD PROVIDERS?



* <http://www.submarinecablemap.com/>

WHAT ABOUT MOBILE COMMUNICATION PROVIDER?

Due to economy of scale property of telecommunication industry, sharing of telecom infrastructure among telecom service providers is becoming the requirement and process of business in the telecom industry where competitors are becoming partners in order to lower their increasing investments.

* http://en.wikipedia.org/wiki/Telecom_infrastructure_sharing



* <http://community.comsoc.org/blogs/alanweissberger/huawei-zte-and-ericsson-dominate-telecom-infrastructure-equipment-market-or-no>

