

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?

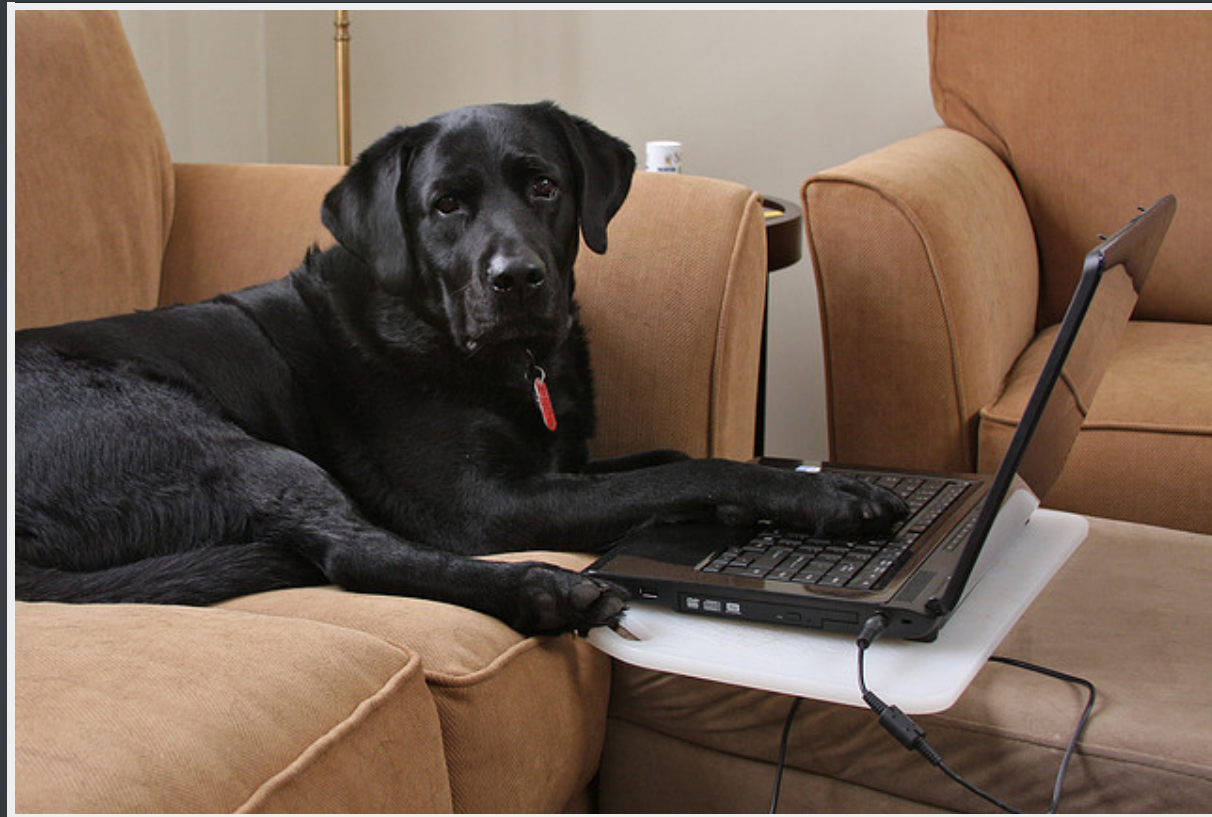
In 2011, the amount of digital information created and replicated globally exceeded 1.8 zettabytes (1.8 trillion gigabytes).

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.

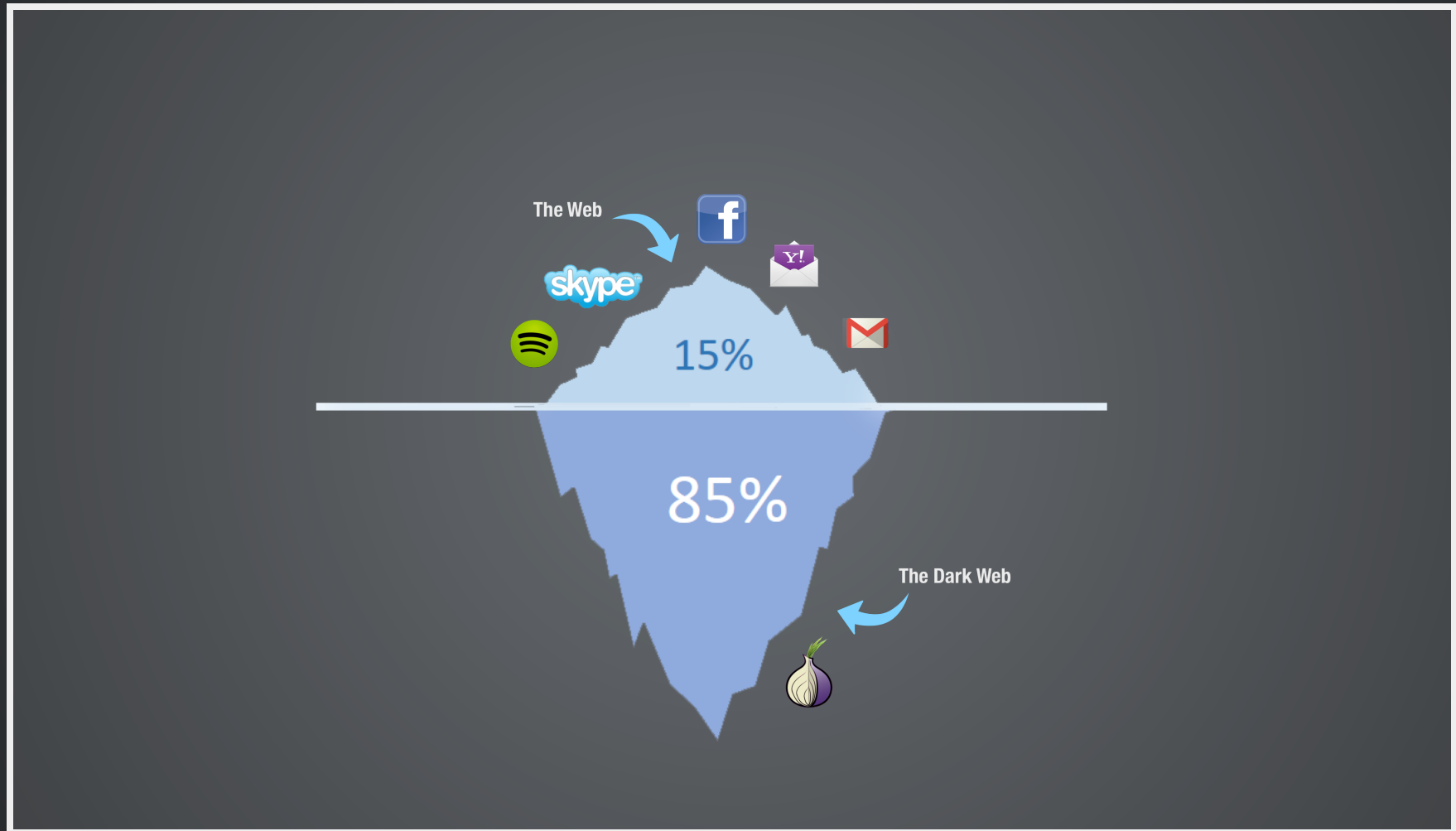
Library Briefing - Library of the European Parliament - 01/03/2012

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

The screenshot shows the 'Network Log' application interface. At the top, there's a status bar with icons for USB, rotation, Wi-Fi, signal strength, and battery (60%), along with the time 10:00. Below the status bar, the app title 'Network Log' is displayed next to an Android icon, and a 'Logging On' button is on the right. A summary line indicates 'Logfile size: 2.45MB Traffic: 0bps/0bps'. The main content area is a table with two columns: 'Log' and 'Apps'. The table lists several network events, each with a small app icon, a red dashed box around the app name, and a 'TCP/wlan0' protocol indicator. The events include traffic from SRC: 192.168.1.100 to DST: 74.125.227.196 (HTTPS) and traffic from SRC: 64.233.160.188 to DST: 192.168.1.100 (HTTPS). The apps listed are '(10018) Network Location' and '(10159) Facebook'. At the bottom, there are three buttons: 'GRAPH', 'FILTER', and 'CLEAR'.

| Log | Apps |
|--|----------------|
| SRC: 192.168.1.100 DST: 74.125.227.196 LEN: 52 2014-12-23 09:59:24.093 | 38754 HTTPS |
| (10018) Network Location | TCP/wlan0 |
| SRC: 192.168.1.100 DST: 64.233.160.188 LEN: 147 2014-12-23 09:59:24.094 | 54195 HTTPS |
| (10018) Network Location | TCP/wlan0 |
| SRC: 64.233.160.188 DST: 192.168.1.100 LEN: 52 2014-12-23 09:59:24.095 | 54195 HTTPS |
| (10018) Network Location | TCP/wlan0 |
| SRC: 192.168.1.100 DST: 64.233.160.188 LEN: 145 2014-12-23 09:59:24.096 | 54195 HTTPS |
| (10018) Network Location | TCP/wlan0 |
| SRC: 64.233.160.188 DST: 192.168.1.100 LEN: 52 2014-12-23 09:59:24.096 | 54195 HTTPS |
| (10159) Facebook | TCP/wlan0 |
| SRC: 192.168.1.100 | 41343 |

GRAPH | FILTER | CLEAR

WEARABLES FOR WELLBEING



WEARABLES FOR PRODUCTIVITY

The Telegraph

Home Video News World Sport Finance Comment Culture Travel Life Women Fa
Apple iPhone Technology News Technology Companies Technology Reviews Video Games

HOME » TECHNOLOGY » TECHNOLOGY NEWS

Wearables can boost employee productivity by almost 10pc

Fitting workers with wearable technologies can boost their efficiency by up to 8.5 per cent, a new study has found

 25  169  0  19  213  Email



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

WEARABLES FOR ...

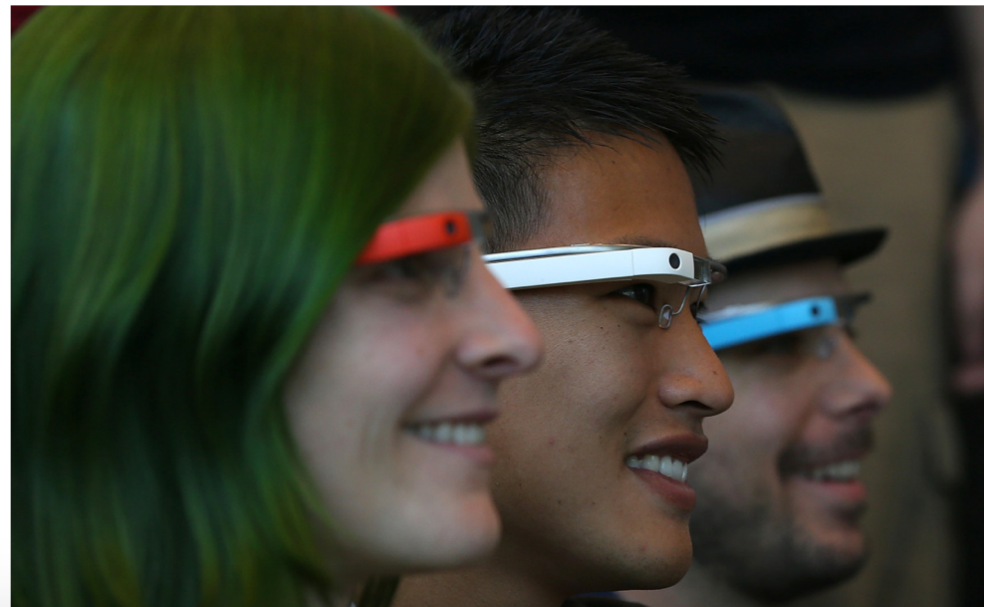
Google Glass Doesn't Have a Privacy Problem. You Do.

Whitney Erin Boesel @weboesel | May 19, 2014



Blaming a particular device is a lot easier than confronting, for example, America's cultural faultlines and problems of power and privilege.

Brace yourself for more Glassholes. Beginning last week, Google started to make available a limited supply of Google Glass to



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

WEARABLES FOR ...

engadget™

REVIEWS ▾

FEATURES ▾

GUIDES ▾

VIDEOS ▾

GALLERIES ▾

FORUMS ▾

EVENTS ▾

Search Products &

SexFit is a pedometer for your penis

 by [Daniel Cooper](#) | [@danielwcooper](#) | August 7th 2014 at 6:55 am



SEXFIT. UPGRADING YOUR SEX LIFE ...
Brought to you by [bondara.co.uk](#)

PERFORMANCE LIGHT INDICATORS

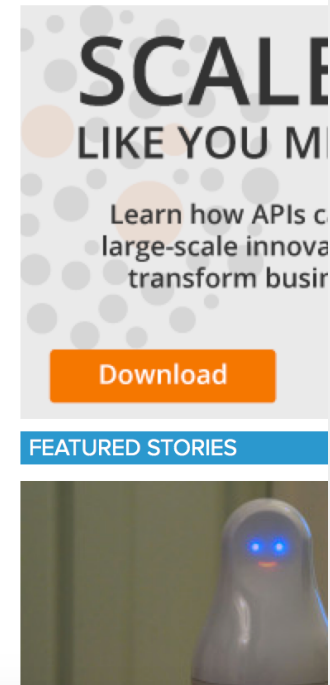
SEXFIT BIT WITH BUILT IN WIFI AND BLUETOOTH CONNECTIVITY

FEATURES:

- * LED performance indicator lights
- * Personal trainer vibration mode
- * SexFit Downloadable App
- * Bluetooth connectivity controls
- * Wi-Fi performance evaluation sync
- * Built in controls for 'on the go' adjustments

COLOURS AVAILABLE:

MULTI PRE-PROGRAMMABLE VIBRATIONS

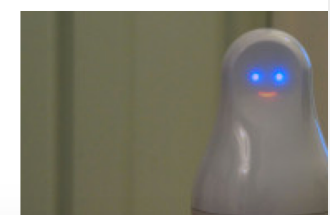


SCALE
LIKE YOU M...

Learn how APIs c
large-scale innova
transform busin

Download

FEATURED STORIES



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

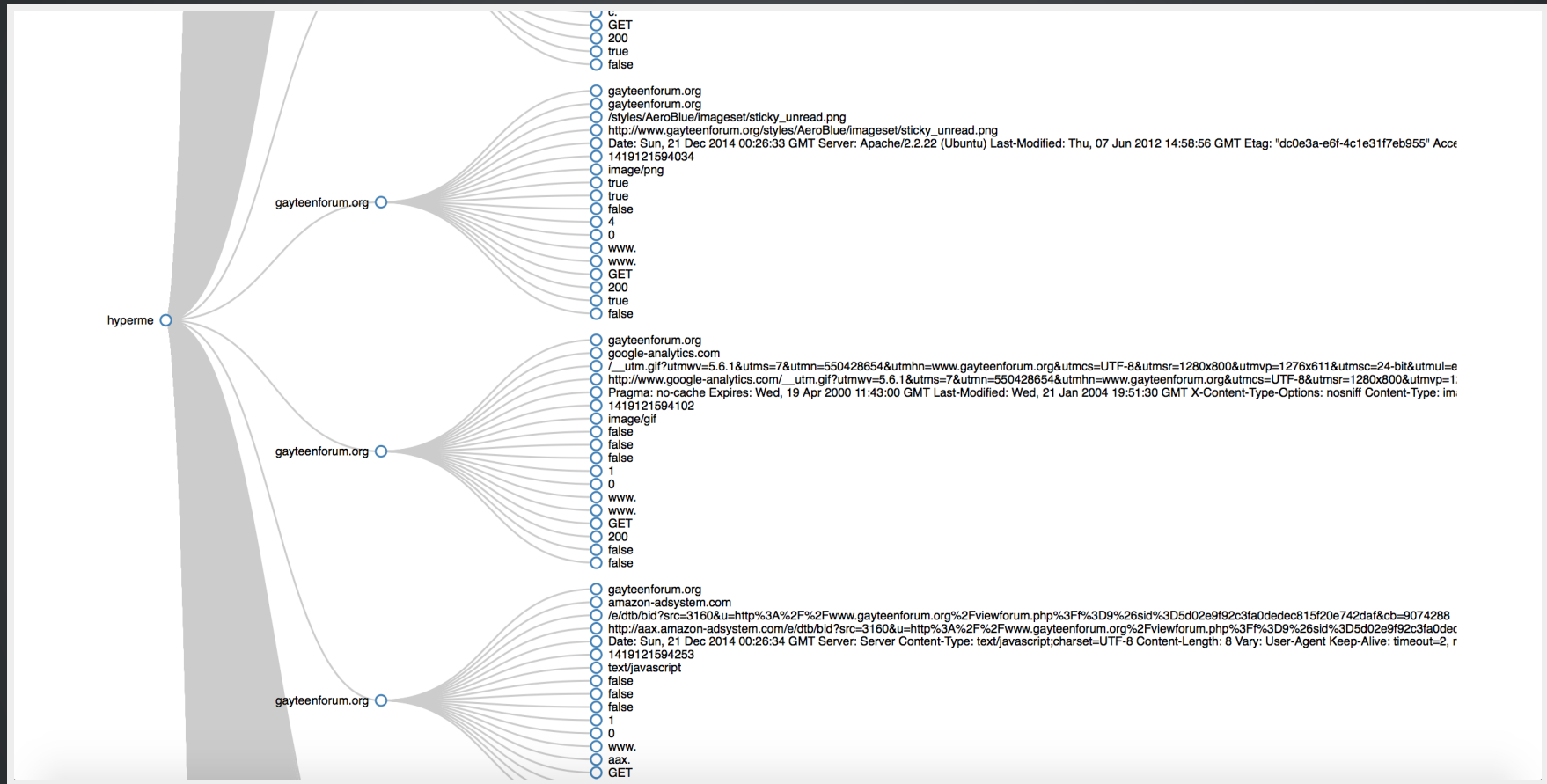
WEBSITE ADVERTISING

```
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

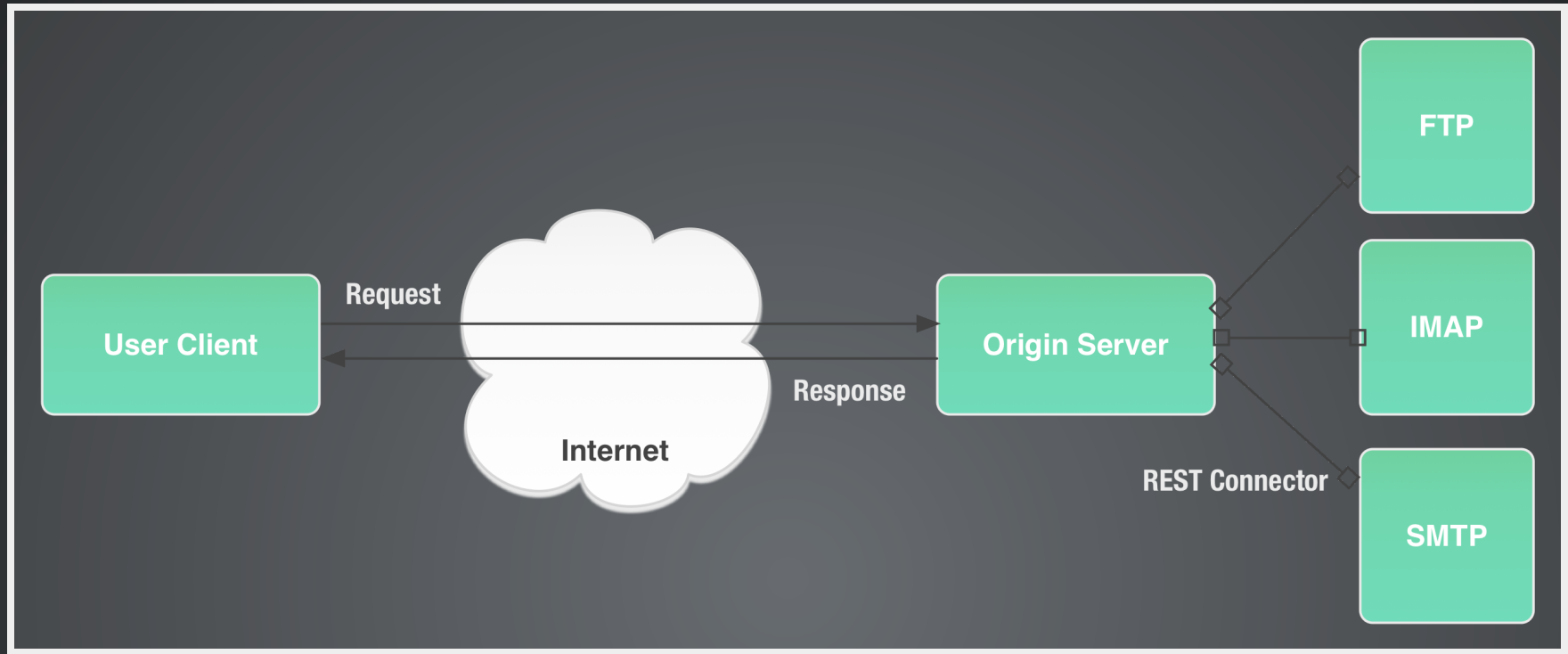
REST, an architectural style introduced by Roy Thomas Fielding in 2000, which has been at the core of the web design and development.

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

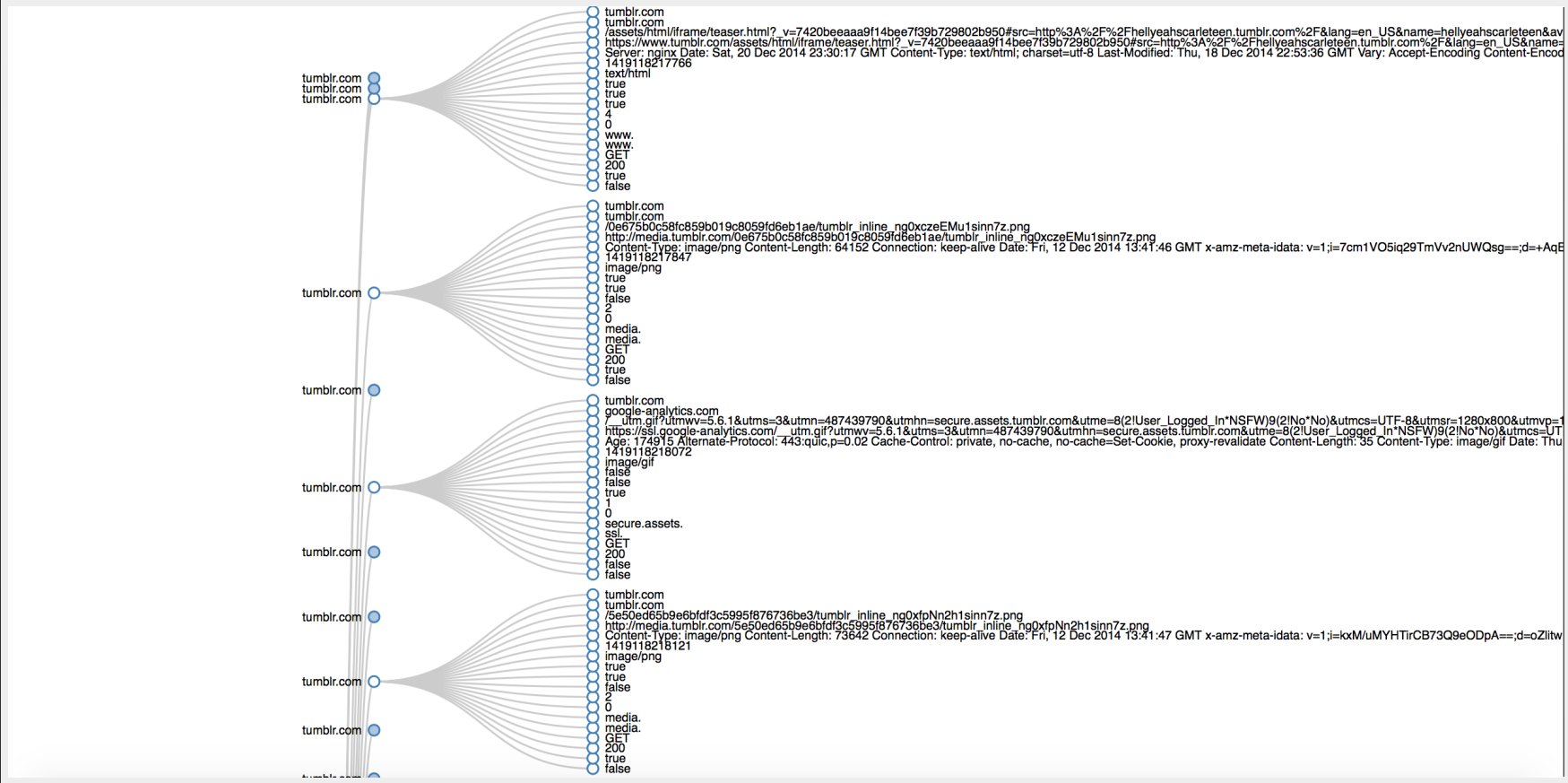
The uniformity of REST interfaces is build upon four guiding principles:

- 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 ads up more information.

THE IDENTITY HYPERGRAPH

In mathematics, a hypergraph is a generalization of a graph in which an edge can connect any number of vertices. Formally, a hypergraph H is a pair $H = (X, E)$ where X is a set of elements called nodes or vertices, and E is a set of non-empty subsets of X called hyperedges or edges. Therefore, E is a subset of $P(X) \setminus \{0\}$, where $P(X)$ is the power set of X .

Google tracking cookie used to pinpoint targets

The NSA and its British counterpart GCHQ utilize Web browser tracking files, known as "cookies," to pinpoint targets for government hacking and to bolster surveillance. The intelligence agencies have found particular use for a Google-specific tracking cookie known as the "PREFID." These cookies typically don't contain personal information, but they do contain numeric codes that enable Web sites to uniquely identify a person's browser. In addition to tracking Web visits, this cookie allows the agencies to single out an individual's communications among the sea of Internet data in order to send out software that can hack that person's computer.

NSA partnered with NGA

The NSA partnered with the National Geospatial-Intelligence Agency, which provides mapping and imagery analysis to intelligence agencies to confirm target locations.

What is TAO?

Tailored access operations — one of the NSA teams that helps conduct technical surveillance.

DNR/DNI metadata

This slide describes the NSA's use of Dialed Number Recognition (information from phones, both mobile and land lines) and Digital Network Intelligence (information from the Internet) metadata to identify new surveillance targets.

mission example and result. The TAO/TAO analytic aggregates leaked location-based service / location-aware application data to infer IP address geo-locations. SDS identified 'Public' and 'private' usage of the same IP address that caused HAPPYFOOT to assign [redacted] netblocks to [redacted] geo-locations (the IP address was used in both countries). This private network is now being realmed and properly geo-located. Ongoing work will solve this realming problem for networks affecting other cloud analytics.

Our Approach

- Tracked [redacted] target's converged communications and CNE accesses.
- Monitored passive internet traffic; created automated processes where possible (XKS ANCHORMAN, Workflows, Fingerprints).
- Provided TAO/GCHQ with WLLids/DSL accounts, Cookies, GooglePREFIDs to enable remote exploitation.
- Partnered with NGA and R4 to confirm locations and USRP equipment based on collected photographs.
- Drove CNE collection and partnered with TAO to increase USRP specific endpoint accesses.
- Provided knowledge to interagency partners for potential on the ground survey options and FBI-led intelligence guiding efforts.



(S//SI//REL TO USA, FVEY) **Metadata/Target Discovery:** Analyze DNR/DNI/Convergence metadata for target discovery, identify gaps in collection, processing, and analytic methodologies; Improve metadata collection and processing; Create analytics that automate or improve analytic methodologies; Conduct target discovery through multiple technology thrusts, including endpoint, web-based technologies/services, mobile applications and networks, geo-location analysis, correlations/identity Analysis, Social Network Analysis; Collaboration/facilitation with TAO, S3, CIA, ODNI, SSO, CES, and SSG centers.



Page

1

of 4



* <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.

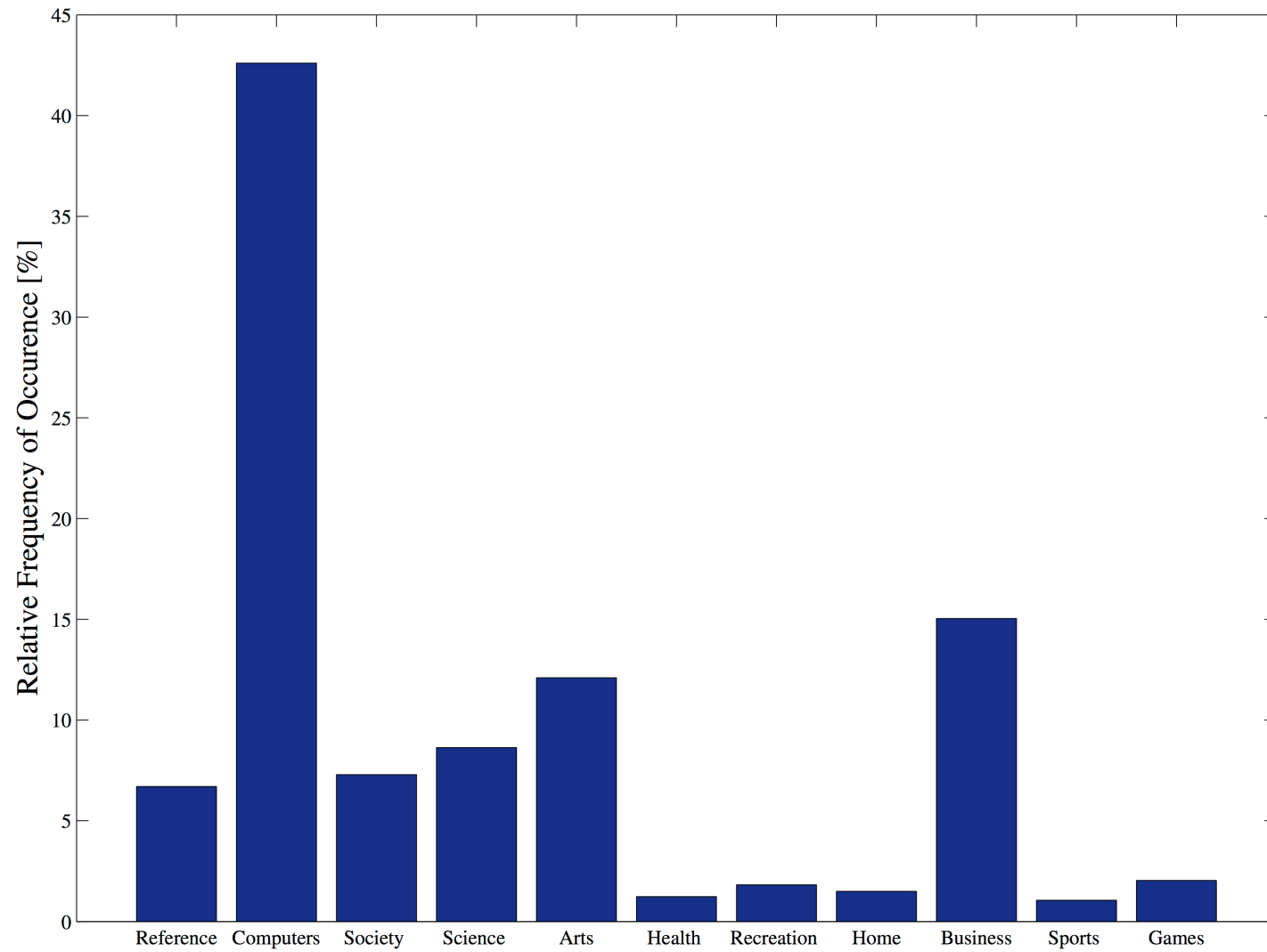
* <https://panoptickick.eff.org/browser-uniqueness.pdf>.

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

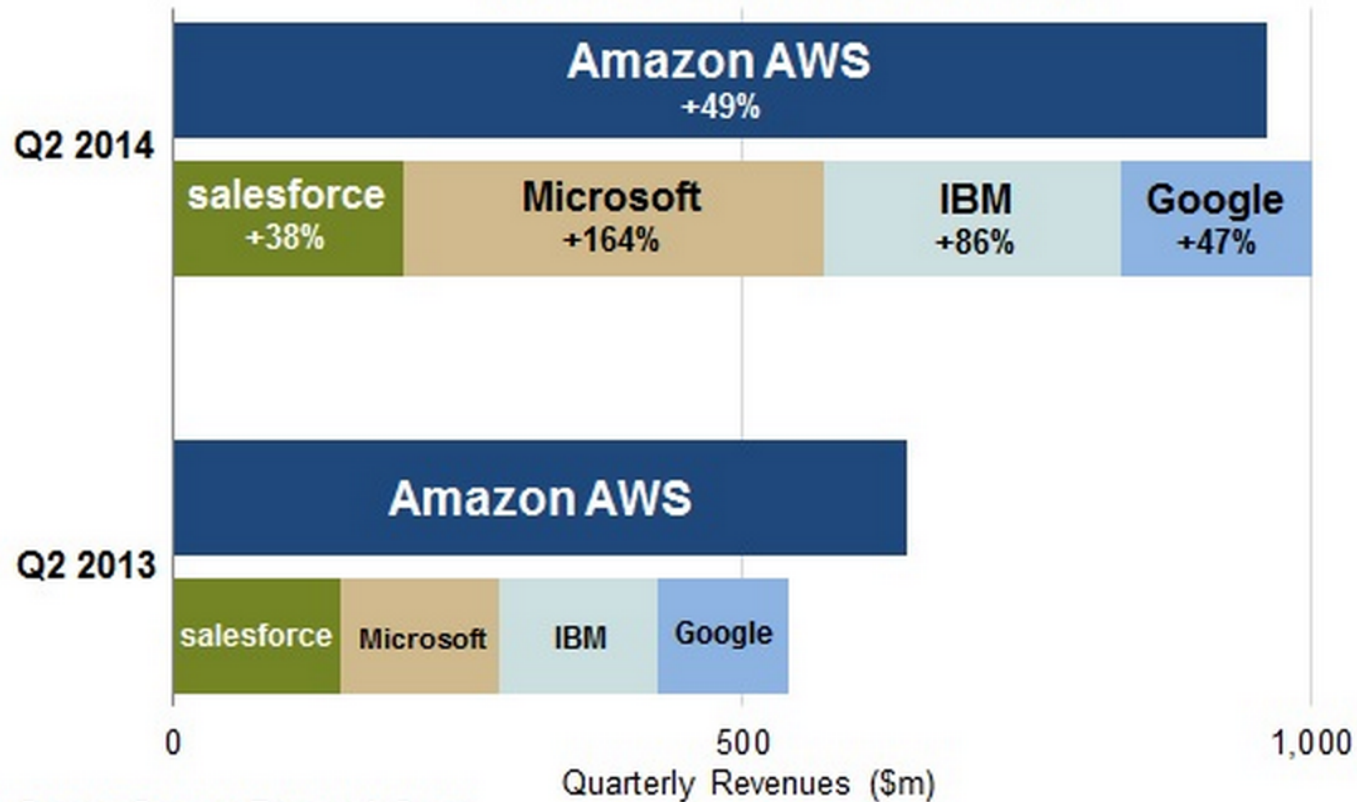


WHAT IS THE CLOUD REALLY?



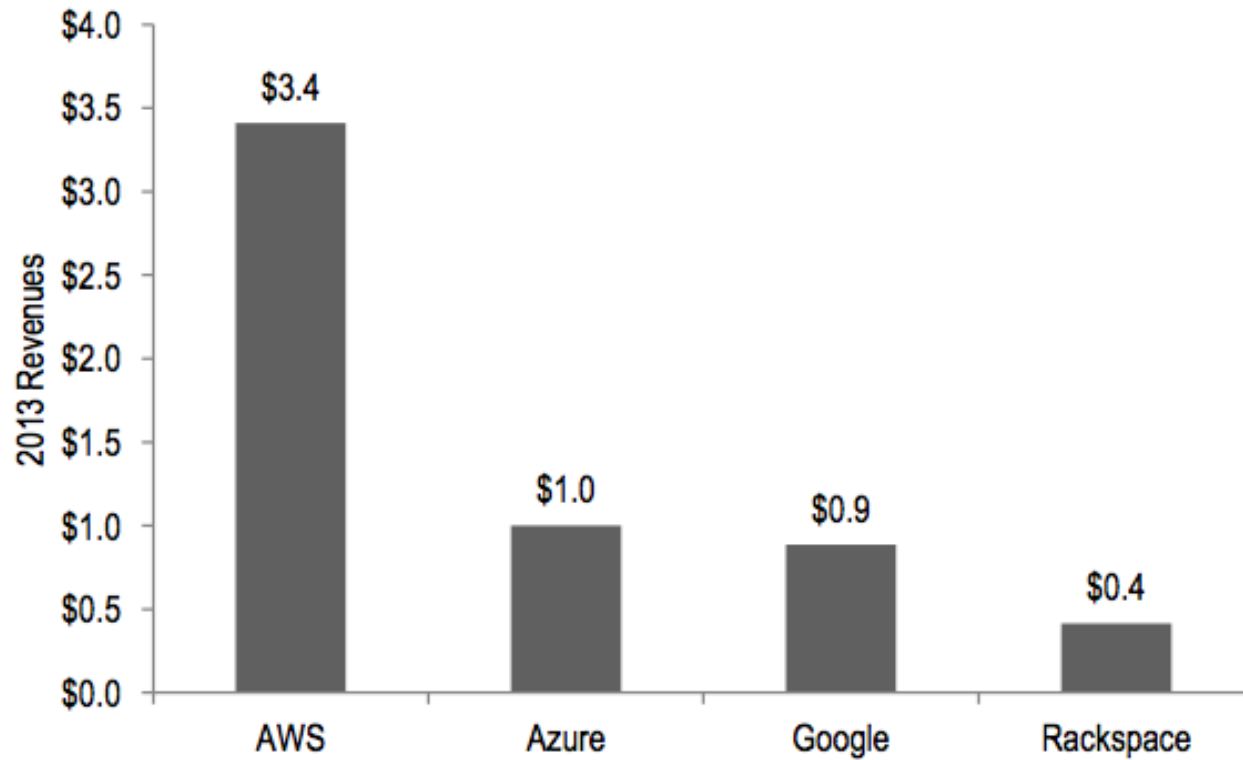
WHO OWNS THE CLOUD?

Cloud Infrastructure Services Revenue Growth
(IaaS, PaaS, Private & Hybrid combined)



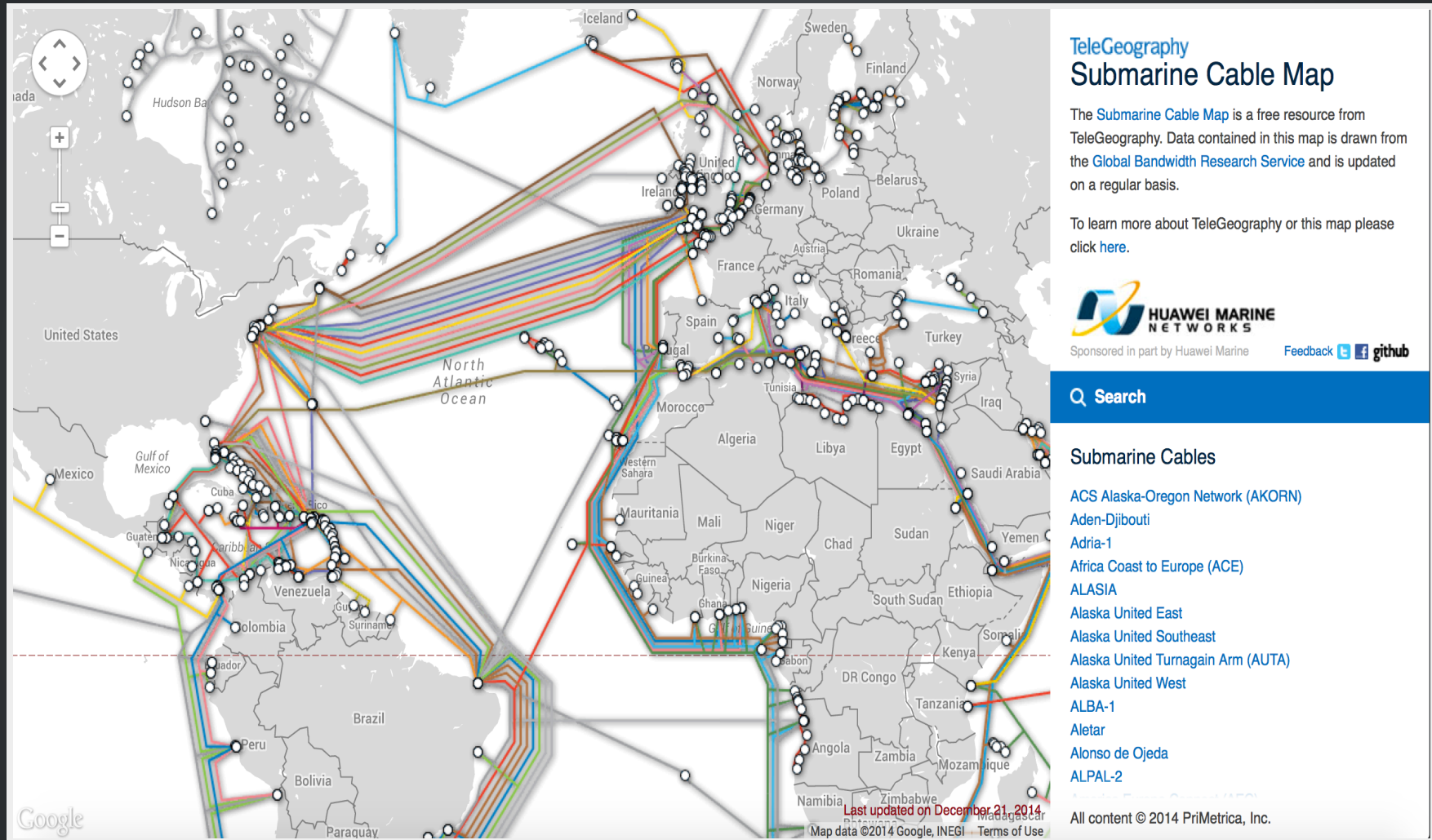
Source: Synergy Research Group

Figure 4. AWS vs. Competition on 2013 Revenue



Source: Company data, Evercore Group L.L.C. Research. Azure based on MSFT comments about a \$1bn rev run rate in May2013. Google based on estimate by TBR (Technology Business Research).

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

Huawei, ZTE and Ericsson to Dominate Telecom Infrastructure Equipment Market - or not?

Mon, 02/14/2011 - 14:52 — Alan Weissberger

The Deah of Network Infrastructure Companies- Will Only 3 remain standing?

Most network infrastructure equipment companies, with the exception of Chinese based Huawei and ZTE, seem to be either losing money or exiting the business. This includes the big three of Alcatel-Lucent, Nokia-Siemens and Ericsson, but also a whole lot of niche market players that have been around for a very long time, e.g. Fujitsu, NEC, Hitachi, ADC Telecom, ADTRAN, etc.

The telecom infrastructure market seems to be dominated now by tje Chinese dynamic duo of Huawei and ZTE which are breathing down LM Ericsson's neck to be the leader in wireless infrastructure gear. Chinese upstart UT Starcom is not too far behind them.

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

I grew up with the understanding that the world I lived in was one where people enjoyed a sort of freedom to communicate with each other in privacy, without it being monitored, without it being measured or analyzed or sort of judged by these shadowy figures or systems, any time they mention anything that travels across public lines.

- Edward Snowden

- Research in open infrastructure
- Collaborate with researchers outside of computer/telecom science/eng
- Be mindful about your online footprint

THE END

This presentation is available on: <http://www.nopressure.co.uk/31C3/>

Github: <https://github.com/nopressurelabs/31C3>

<http://hiromipaw.github.io/>