Privacy

Privacy?

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

[Peter Steiner,1993]

myth reality

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

[Peter Steiner,1993]

To understand privacy risks

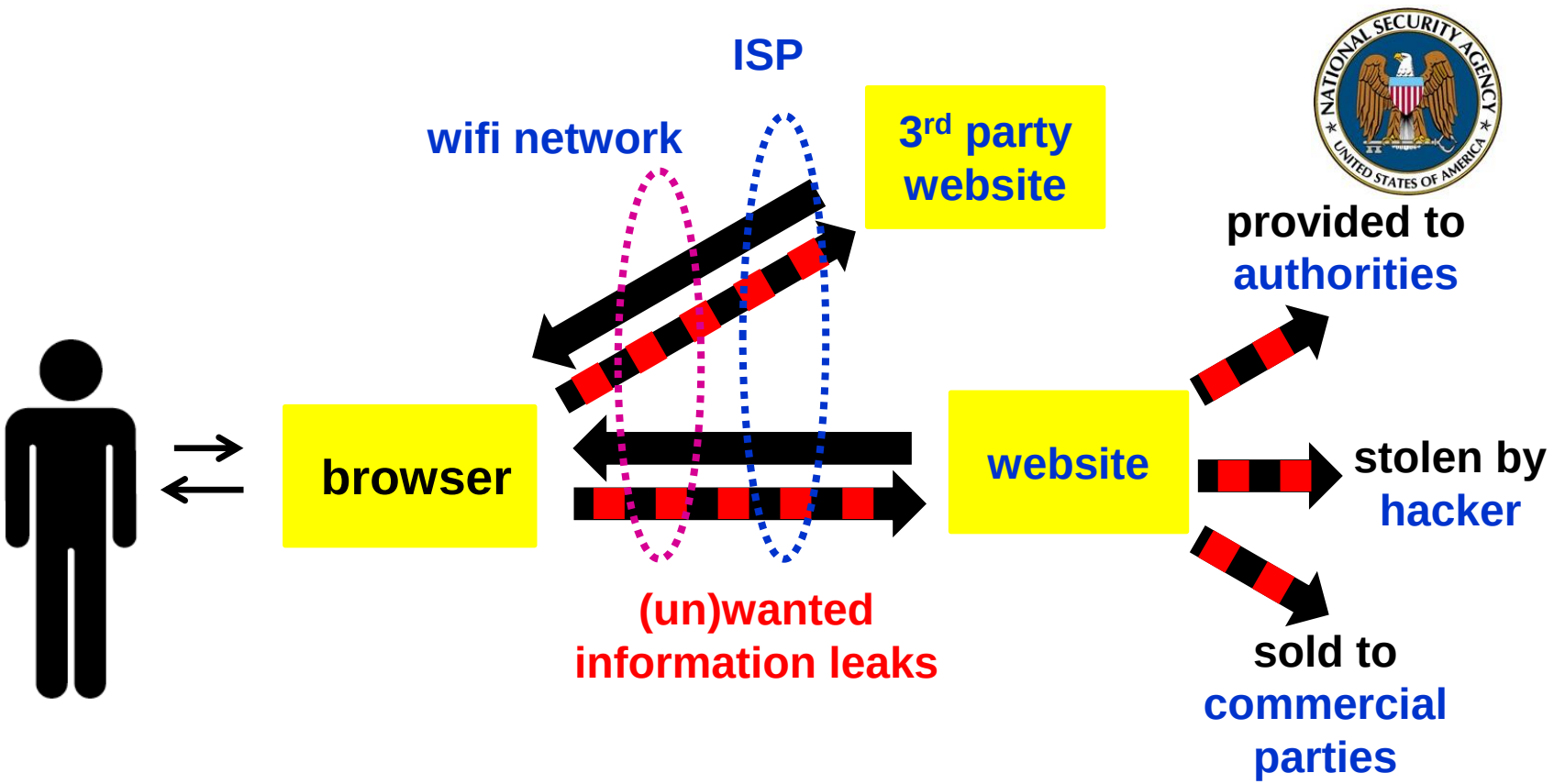
- What information is leaked?
- How is information leaked?
- Who are the parties that might get this information?
  What will they do with it?
  Why *are parties interested in this information?*
- What are the legal rules?

Parties involved

- user
- website visited
- websites providing 3rd party content
- internet service provider (ISP)
- browser
  - producer of the browser, eg Microsoft for IE, Google for Chrome
  - producer of browser plug-ins, eg Adobe for Flash
- public authorities and national security agencies
  - AIVD and MIVD, eg. via CIOT (Centraal Informatiepunt Onderzoek Telecommunicatie)
  - NSA eg. via PRISM
- (organised) criminals, hacktivists, and random hackers

- legislators (national and EU level), government regulators (ACM) and watchdogs (CPB), privacy advocates, scientific researchers....

Privacy


ISP
wifi network
3rd party website
provided to authorities
browser
website
stolen by hacker
(un)wanted information leaks
sold to commercial parties

Beyond the web and the internet

Privacy is just issue for web and internet, but more generally for computing devices and systems storing information, eg

- (mobile) telephones and telephone networks
- other transactions involving *identification*:

  public transport, payment with bank card, customer card at shops,…
- other information *digitally recorded*:

  number plate registration, CCTV security cameras, ..

Issue of growing importance, with the explosion of digital information and the merging of the virtual & physical world into one *cyber-physical world*

What information?

Possible information leaks

- visits to certain web site
- browser history
- “content”, entered certain data at web site
  - search queries
  - look at certain subpages, topics,...
  - email addresses, email content, telephone number
- video & sound via camera and microphone
- geographical location
- ...

- content vs meta-data

What motive?

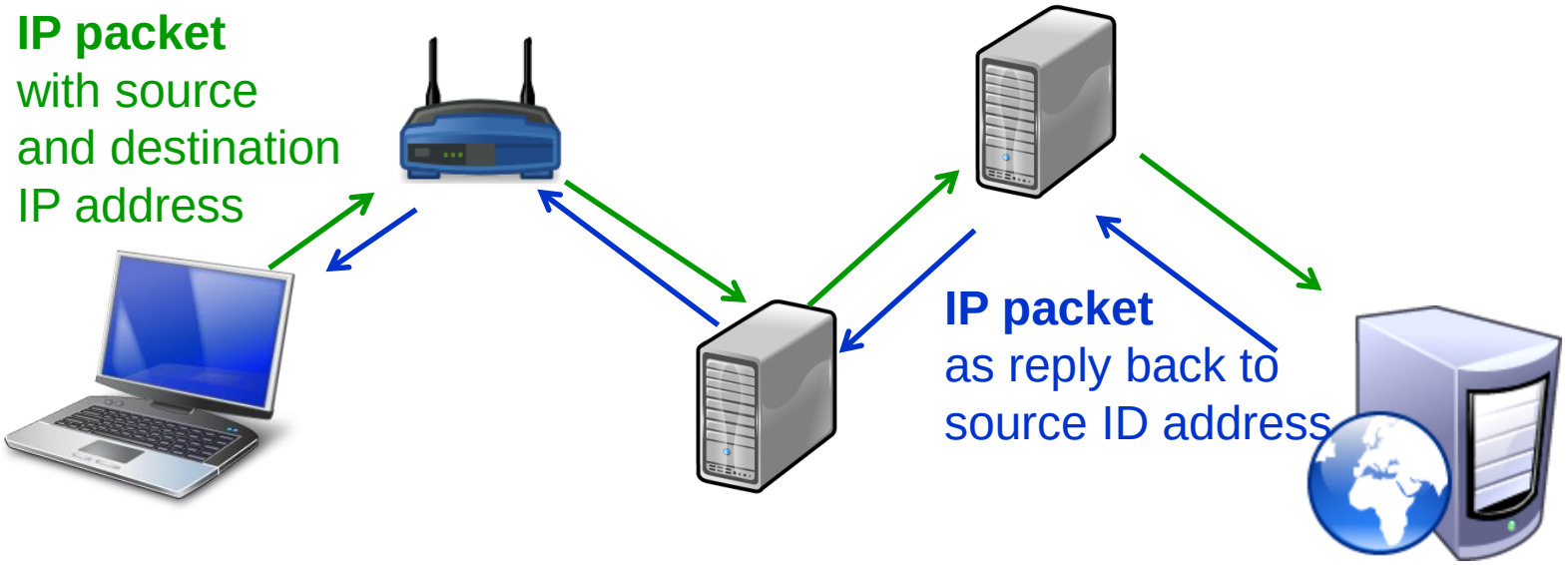
- commercial
  - or `service' to the customer
- law enforcement
- criminal

HOW:
internet basics

IP basics

Home PC and website identified by IP address:
unique address of individual computer

Web browers requests webpage, web server returns webpage

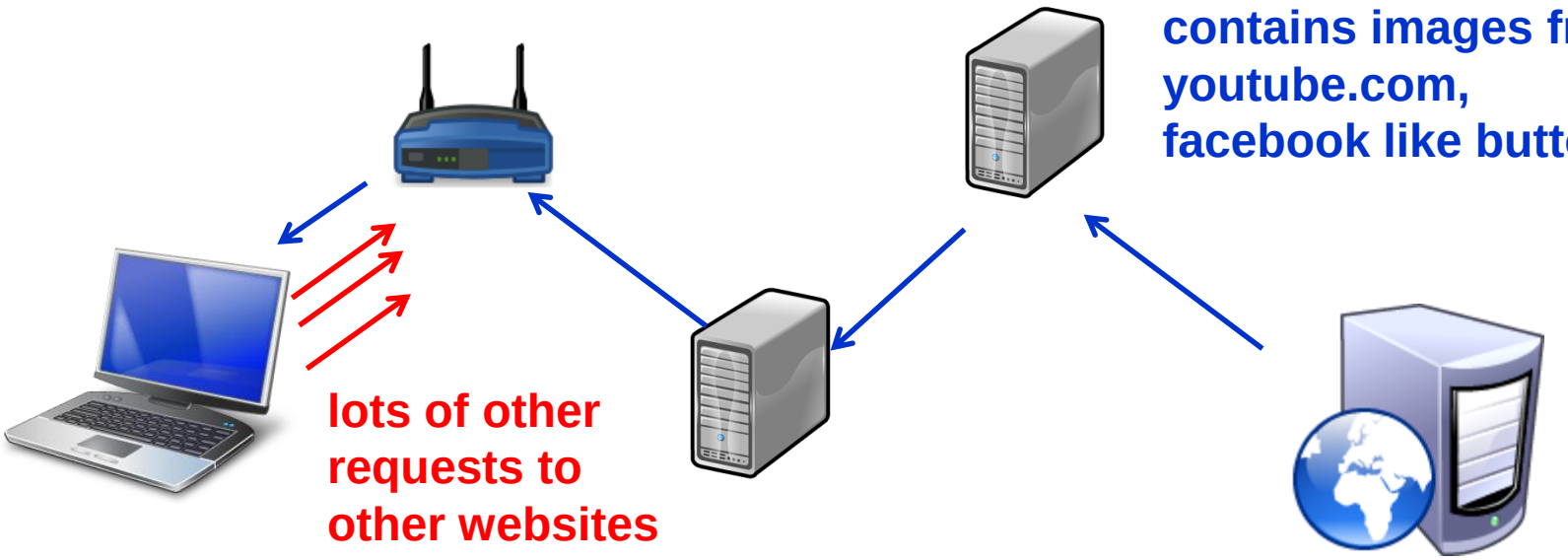


IP address 123.123.123.45

web site IP address
234.234.234.56

Third party content

A web page returned by a website will usually contain content from other website, which the browser will immediately fetch



IP address 123.123.123.45

web site IP address
234.234.234.56

(Lack of) anonymity in normal internet use

- any website you visits knows your IP address
  - as do all websites that provide third-party content to this website

- ISPs and telcos report which person uses which IP address & telephone number to a central point for law enforcement

  In Netherlands: Centraal Informatiepunt Onderzoek Telecommunicatie (CIOT);

  consulted 2.9 million times/year in 2009

  [Source: Bits of Freedom, bof.nl]

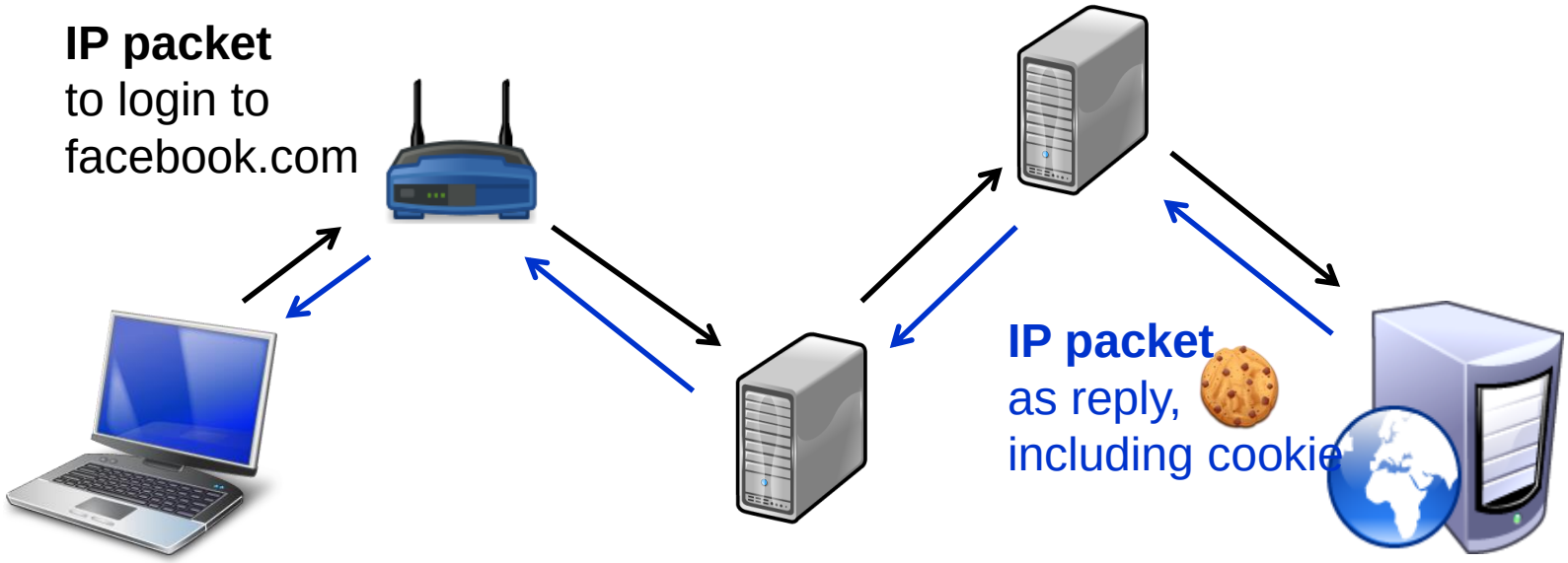
Cookies

Cookies installed by website in browser to

- maintain a session after the user logs in
  - after logging in to gmail or facebook, a cookie stored on your machine to authenticate you, so that you don't have to login for the next N hours
- record user preferences
  - eg information in English or Dutch
- track a user across many websites
  - eg for targetted aka behavourial advertising

Cookies

After first visit to facebook.com to login you receive a cookie

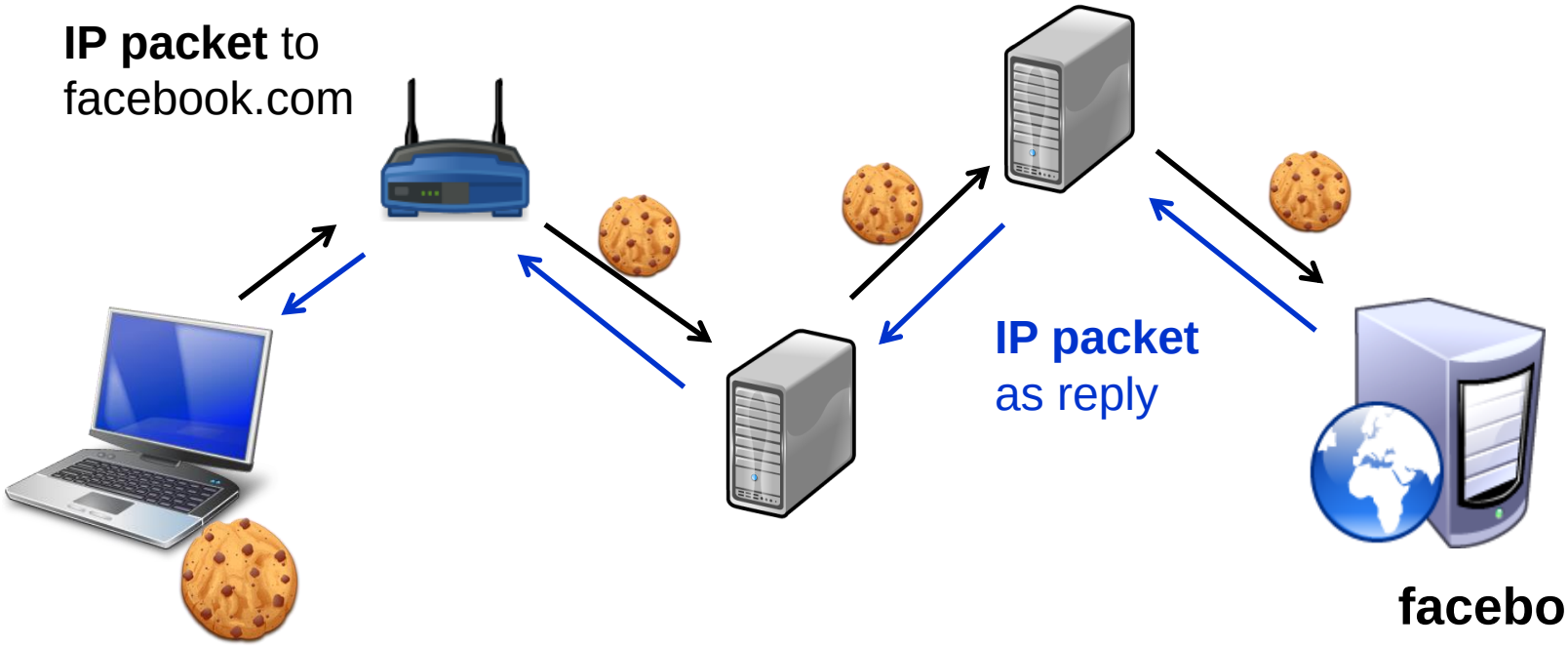


will store the facebook cookie

Cookies

Cookie is sent along to every IP request to facebook.com.
Also when you visit any page with a facebook like button

- Viewing one website you get & send cookies from & to many others!



Cookies vs IP addresses

Why use cookies instead of IP addresses to track users?

- Cookies allow sites to track users across different IP addresses
  - connecting to different Wifi points with your smartphone or laptop will result in different IP addresses
- Legally, an IP address is **personal information**, and there are legal restrictions on what you can do with this
  - personal information = information that can be related to one human individual

privacy threats on the internet

IP addresses

- Any eavesdropper on the network will also see source and destination IP addresses of internet communication
- Server logs will at least record the IP information

- IP address usually gives accurate country & town information

- In Dutch law, IP address counts as persoonsgegeven (personal information), so processing it is subject to Wet bescherming persoonsgegevens (WBP)

- Using HTTPS does not help; this hides the content, but not the source & destination

Leaking your IP address

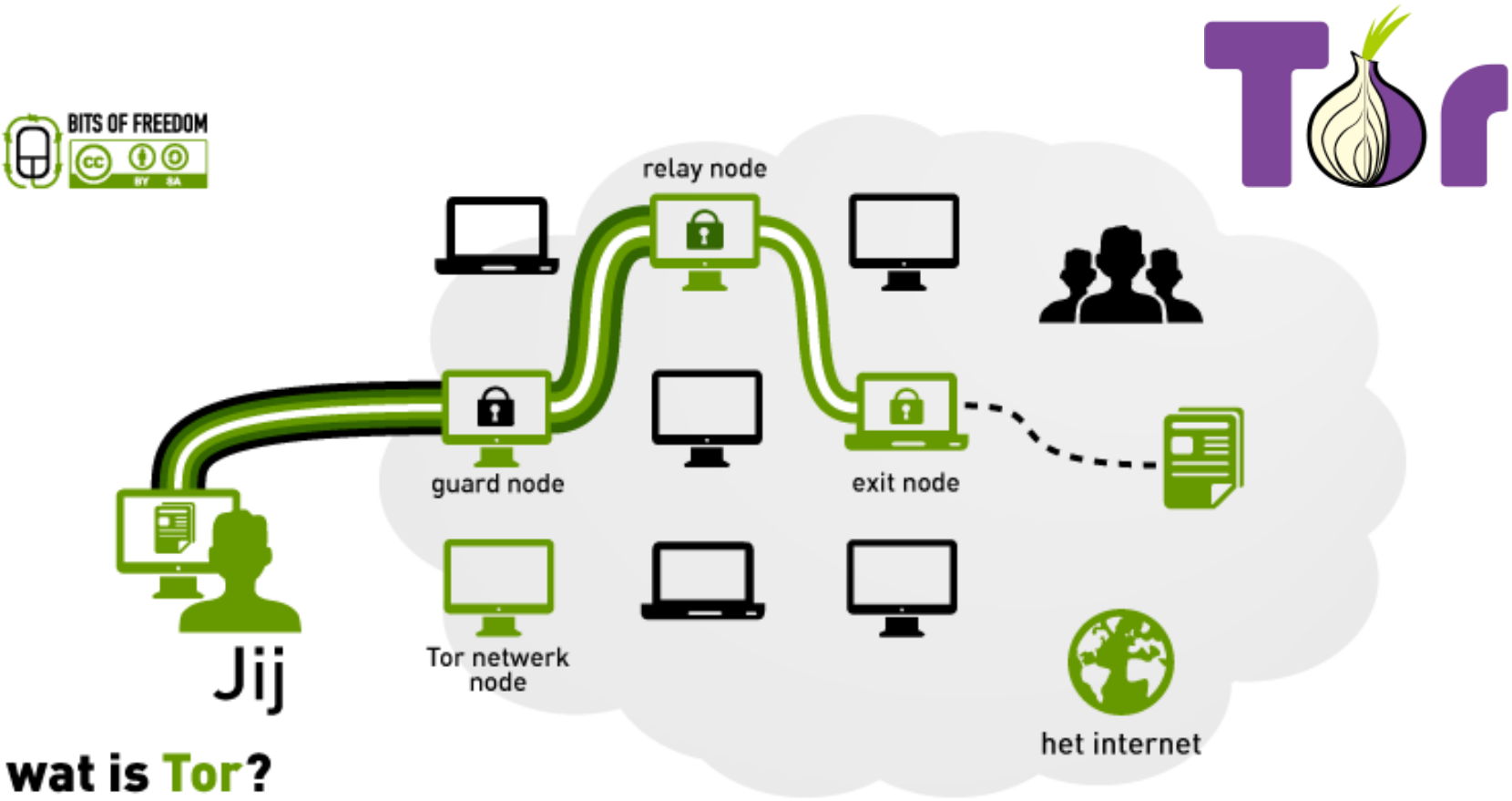
Senate Staffer Tried To Scrub 'Torture' From Torture Report's Wikipedia Entry

The Huffington Post | By Igor Bobic

Posted: 12/11/2014 11:08 am EST | Updated: 12/12/2014 12:59 am EST

Countermeasure: Tor

Tor works with layered encryption, which traffic relayed via multiple nodes, with each node `peeling off' one layer of encryption



Tor

- Tor (The Onion Router) networks aims to provide anonymity on the internet:

  No single node knows both source & destination IP address
- Started by US Naval Research Laboratory, and still partly US funded
- Has both legitimate and illegitimate use
  - eg used by Edward Snowden to leak information
- Not immune to all attacks

cookies & 3rd party cookies

Most websites will include 3rd party content from eg

- social networks
- advertising networks
- web analytic services (eg google-analytics)
- ...

Of course, borders between categories above are vague/non-existent.

Very little 3rd party content is actually useful to users,

apart from google-maps?

Using cookies, these 3rd party websites can track users across web

Browser plugins such as Lightbeam or Ghostery provide insight in the large numbers of 3rd parties that are following your browsing!

Example 3rd party content: Facebook Like button

- Facebook tracks *members* across sites that have Like or Share buttons
  - because the Facebook cookie that identifies user is included with all requests to facebook.com
  - Note: this happens *before* the user clicks the Like button.

- Facebook even tracked *non-members*
  - the Connect button *installed* a cookie, with a life time of 2 years
    - when button is shown, not only after it is clicked
    - the Like button did not install cookie; for both Facebook receive any cookies already set
  - if non-member joins facebook later, histories can be linked
  - similary, if a facebook member surfs anonymously (for Facebook), because he's not logged in, his browsing can be linked as soon as he does log in

Example 3rd party content: Facebook Like button

- Initiative for a privacy-friendly two-click Like button:
  1st click downloads real like button; 2nd click clicked it

- Facebook claimed this violated their copyright to Facebook logos

Why: behavioural advertising & profiling

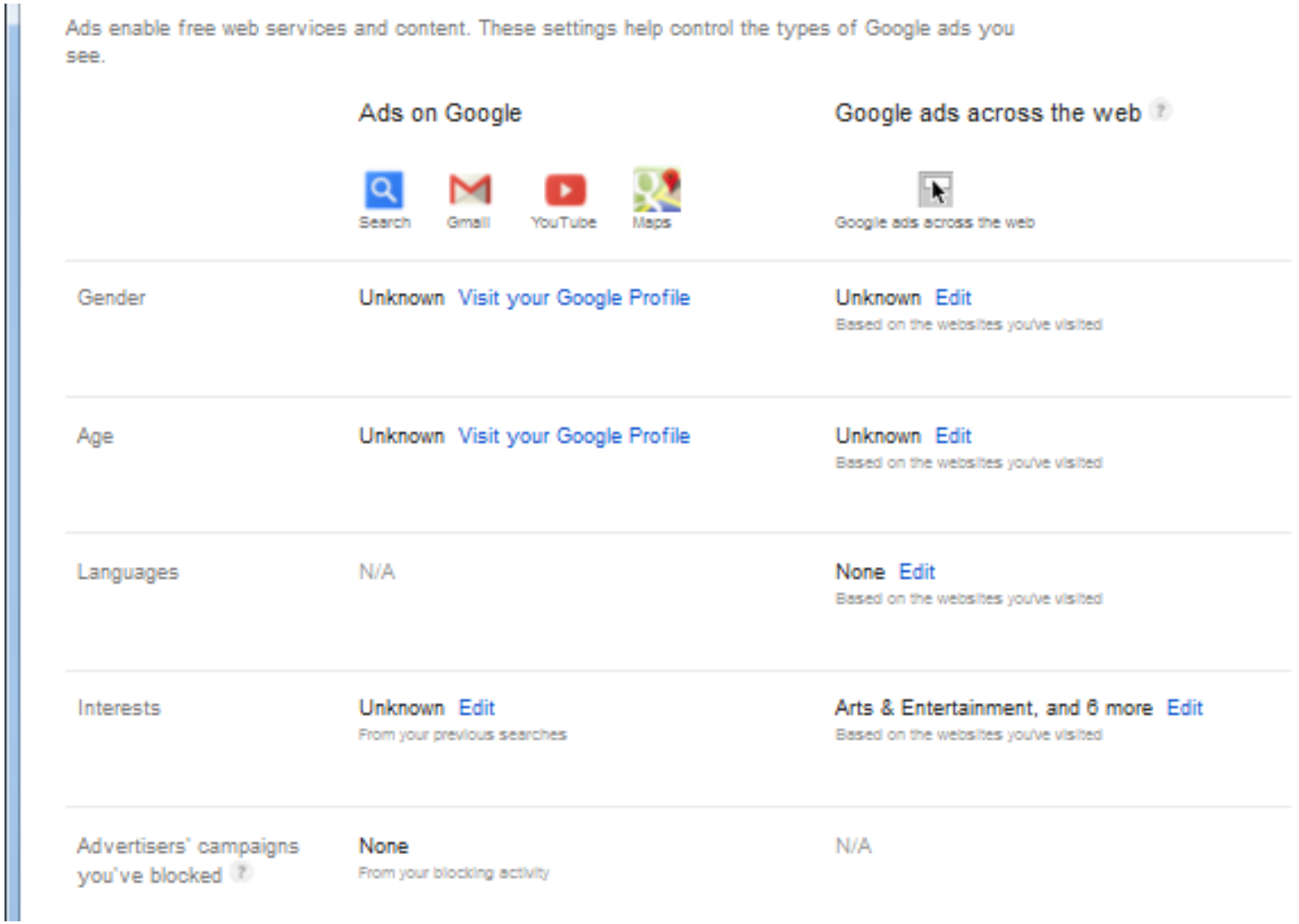
Data can be used for

- targetted aka behavioural advertising
- targetted pricing
  - eg online shop asking higher prices from rich people

    or slowly in/decreasing price to see how customers react
- targetted offering of products and services
  - eg online shops *not* offering insurance to people in certain profiles ...

What profiles are being used to categorise people?

Google Ads settings



Ads enable free web services and content. These settings help control the types of Google ads you see.

	Ads on Google Search Gmail YouTube Maps	Google ads across the web Google ads across the web
Gender	Unknown Visit your Google Profile	Unknown Edit Based on the websites you've visited
Age	Unknown Visit your Google Profile	Unknown Edit Based on the websites you've visited
Languages	N/A	None Edit Based on the websites you've visited
Interests	Unknown Edit From your previous searches	Arts & Entertainment, and 6 more Edit Based on the websites you've visited
Advertisers' campaigns you've blocked	None From your blocking activity	N/A


Candy Free
Checkout
Lane

3rd parties & their cookies: countermeasures

- Deleting cookies regularly
- Using private browsing modes
- Blocking (all) 3rd party cookies
  - or some plugin for finer-grained cookie control
- Block (some) 3rd party content
  - eg by an AdBlocker
- Browser plugins to reduce tracking, such as Ghostery

- Some browser support for controlling tracking and opt-out initiatives like http://donottrack.us/

if you are not paying for it,

then <u>you</u> are the product being sold

All ‘free’ services (gmail, facebook, twitter, WhatsApp..) are paid with ads and collecting personal information for marketing

Flash cookies

- aka LSO (Locally Shared Objects) or supercookies
- information stored & used by Adobe Flash Player
- Characteristics
  - stored in hidden folder on the OS file system
  - no expiry date
  - up to 100 Kbyte
  - work across multiple browsers
- In 2009, 50% of common websites used Flash cookies
- Flash cookies have been used to restore deleted HTTP cookies, so-called zombie cookies

- Flash cookies can be controlled in Adobe Website Storage Settings Panel
  https://www.adobe.com/support/documentation/en/flashplayer/help/settings_manager07.html
  but nowadays also from most browsers

Web beacons

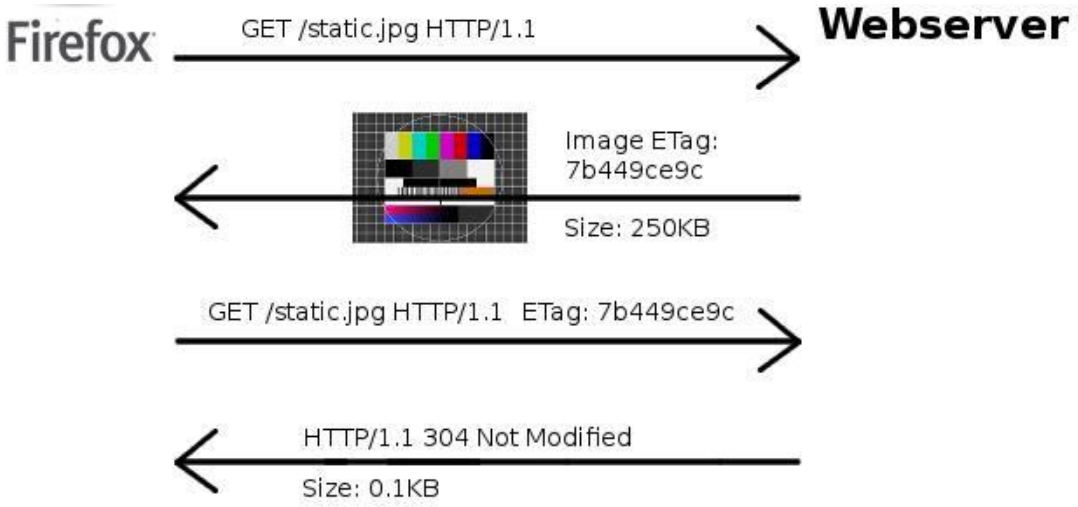
- aka web bugs aka tracking bugs aka pixel tags
  aka JavaScript tags if they use JavaScript
- invisible 1x1 pixel image included in document (eg web page or email) via a link to remote server
  - image will be downloaded from server when document is read

- used in emails
  - to see when an email is being read, from which IP address, ...
  - used by spammers to see if spam is read, meaning that email address is real and email gets past the spam filter
- used in web pages
  - to gather web statistics
  - if 3rd party cookies are blocked, then web beacons cannot directly be used to track visitors across website

Cookieless cookies using ETags

ETags are identifiers added to resources to enable caching

- When browser ask for a resource, it can say which version of that resource it already has in its cache, by giving the ETag



This allows a server to identify the browser...

See http://lucb1e.com/rp/cookielesscookies/

Cookieless tracing via URLs

Of course, the simplest form to trace someone a cross websites is by including a unique ID paramters in the URL

Eg

http://www.google.com.br/settings/ads?hl=pt–BR&sig=ACi0TCjkFq1TS2cz_RFuc_KqcHfC7mp_iJR_uRA_G6UvcHcoFt4d6IEFTU4xuggdkY4DlBz7pr0ToFe8S9vYXrVKVeFVnrdzrYcb84KJZpi0Ffsl2ppZWGGthblqoxeLnab5YDaUHC0rxMzVXp8nxvHnlL0YGjBIY8iKzRIUsPT8iBF4uEzVl_YmichoYgV3vBEXza3

Browser fingerprinting

- Browsers are complex pieces of software that have with many characteristics
  - versions, language, OS, screen size, fonts, plugins,...

- These characteristics leak lots of information, and may even uniquely identify a browser.

  Eg see https://panopticlick.eff.org/

spying on browsing history

- A largely historic attack, as modern browsers have good mechanisms to prevent this, but nice illustration of unexpected power of complex content

- Using executable content in a webpage, the page can reveal the browser history
  - ie which sites have been visited

- This was possible using JavaScript, or just CSS

- This could be used for good purposes (eg checking which social network someone is active on, and then presenting right links for that visitor), but it can also be a privacy threat.

spying on browser history: HTML vs CSS

- CSS (Cascading Style Sheets) are used to improve HTML by separating presentation & layout from the content
  - HTML specifies the content of a web page
  - CSS specifies style, ie how that content is displayed

Example CSS

To underline links, and give visited links a different colour from unvisited links:

```
:link, :visited {  /* for all links */
   text-decoration: underline;
}
:link { /* for unvisited links */
    color: blue;
}
:visited {/* for visited links */
    color: purple;
}
```

Using JavaSacript and the DOM we can now see if a link is visited.

How? JavaScript code can check the color of links!

Example: JavaScript to spy browser history

```
var links = document.links;
for (var i = 0; i < links.length; ++i) {
  var link = links[i];
  /* exact strings to match actually need to be
    auto-detected using reference elements */
  if (getComputedStyle(link, "").color == "rgb(0,0,128)")
   {
    // we know link.href has not been visited
   } else {
   // we know link.href has been visited
   }
}
```

Modern browsers no longer allow this sort.

privacy threats in the physical world

Google admits collecting Wi-Fi data through Street View cars

German request for data audit reveals the web giant 'accidentally' stored payload information from open networks

Jemima Kiss

Follow @jemimakiss Follow @guardiantech

The Guardian, Saturday 15 May 2010

Technology
Google Street View · Google · Internet

UK news

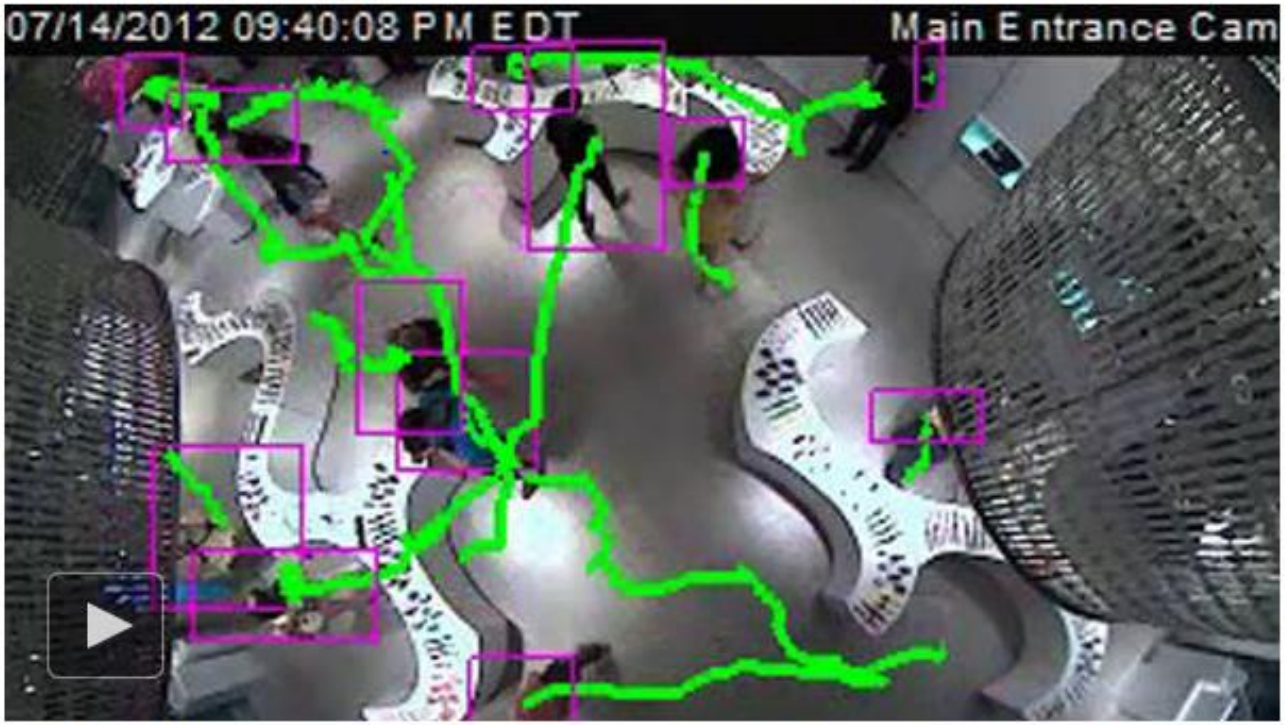
World news
Germany · Europe

More news

More on this story

wifi tracking

Attention, Shoppers: Store Is Tracking Your Cell



Big Data Hits Real Life: Brick-and-mortar stores are looking for a chance to catch up with their online competitors by using software that allows them to watch customers as they shop, and gather data about their behavior.

By STEPHANIE CLIFFORD and QUENTIN HARDY
Published: July 14, 2013 | 410 Comments

Like dozens of other brick-and-mortar retailers, Nordstrom wanted to learn more about its customers — how many came through the doors, how many were repeat visitors — the kind of information that

FACEBOOK
TWITTER
GOOGLE+

Legal context

Legal context (1): Data *Protection* law

Dutch/EU Data protection laws governs the collection
and use of personal data by data controllers.

Three basic ingredients:

1. citizen should consent to personal information be collected & used
2. citizen should be informed
   - that data is collected, and what data is being collected
   - for what purpose
   - if it is shared with third parties
3. citizen has right to see which personal data is collected about them, and the right to have this corrected in case of errors

CBP (College Bescherming Persoonsgegevens) supervises compliance with law

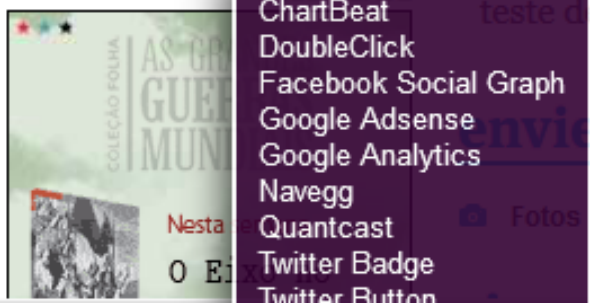
Google pode ser multado em € 15 mi por violar privacidade na Holanda

DA REUTERS, EM AMSTERDÃ

15/12/2014 ⊙ 18h29

O Google pode ser multado em mais de 15 milhões de euros (cerca de R$ 50 milhões) se não interromper a violação de privacidade de usuários de internet na Holanda, disse a agência de proteção de dados holandesa nesta segunda-feira (15).

A companhia norte-americana está violando o ato de



Legal context (2): Data *Retention* law

Dutch data retention act (Wet bewaarplicht telecommunicatiegegevens) governs the collection of telecom and internet data by telco's and ISPs.
Motivation: law enforcement and anti-terrorism

What information is kept?

- For telephone: who is phoning or SMS-ing who, where, when, for how long
  Not the content of call or text message.
- For email: who is emailing who, when
  Not the content of emails
- For internet: time of logging on/off and IP address of client
  Not the IP addresses visited or IP traffic

Note: email sent via gmail and text messages via WhatsApp not recorded

How long?

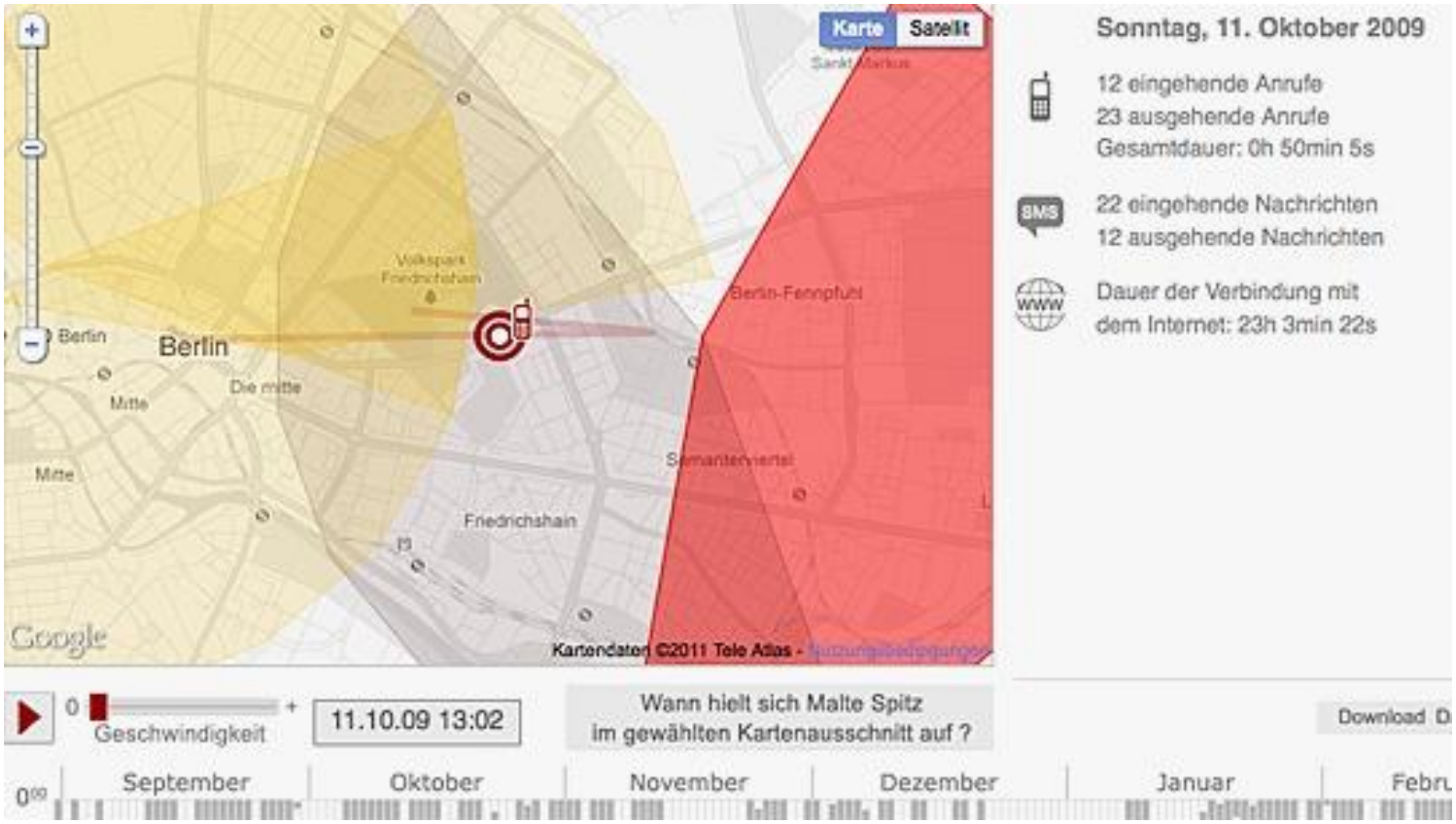
- 12 months, but reduced to 6 months for email & internet

Additionally, public transport card data is kept for 2 years (original plan: 7 years)

Data protection in action

Malte Spitz obtained all the data T-mobile had on him, after long legal battle

See http://www.zeit.de/datenschutz/malte-spitz-vorratsdaten



Data retention in action. Oops...

Some telcos gave the Dutch authorities also the ***content*** of all SMSs
- by accident

Right to be forgotten?

The 'right to be forgotten' online is really a right to be forgiven

Why Europe's 'right to be forgotten' rule means trouble

by Fortune Video

DECEMBER 11, 2014, 2:38 PM EST

General observations on privacy & anonymity

Privacy threats

On the web & internet

- IP addresses, cookies, esp. 3rd party cookies, Flash cookies, Web beacons, Etags, …

In physical world

- mobile phone tracking in shops, transport cards, electronic payments, ..

Growing issue in general, with ever more Big Data

lots of data, and lots of computing power to use it

Future issues:

- Google glasses
- growing power of social networks
- online image search using upload picture of someone face,
- ...

Privacy & Function creep

Privacy is an obvious first casualty in function creep.

Once people have data, they will use it!

Examples:

- *Dutch car navigation system TomTom sold customer data to police for optimal placement of speed cameras..*.
  - *So even if you do pay, you may still be one of the products …*

• 10 maart 2014, 08:36

ING start proef met delen betalingsgedrag klanten

Een bankpas van de ING voor de website van de bank. Foto ANP / Lex van Lieshout

door **Laura Klompenhouwer**

BINNENLAND ECONOMIE **ING wil informatie over betalingsgedrag van klanten gaan delen met bedrijven. Met de informatie zouden bijvoorbeeld retailers gerichter kunnen adverteren,** zegt directeur Particulieren bij ING Hans Hagenaars vandaag in het *Het Financieele Dagblad*. **Het gaat om een proef, die dit najaar begint met enkele duizenden vrijwillige proefpersonen.**

Anonymisation is hard!

It may be harder to anonymise data then you think!

Classic example:

- In 2006, AOL released 2 Gbyte of anonymised search data for research purposes
  - twenty million search queries for over 650,000 users over a 3-month period

- Research then quickly could identify some users, because the search queries contained personally identifying information.

- It also revealed some amusing, sad, and highly disturbing search histories of individuals.

Oops, meta-data…

The file on Iraq of UK government, produced by UK intelligence services prior to the 2nd Gulf War, was distributed as .doc file.

Meta-data in this document included

```
Rev. #1: "cic22" edited file
"C:\DOCUME~1\phamill\Temp\AutoRecovery save of Iraq -
security.asd" ..
 ..
 Rev. #6: "ablackshaw" edited file
"C:\ABlackshaw\Iraq - security.doc" ..
 Rev. #10: "MKhan" edited file
"C:\WINNT\Profiles\mkhan\Desktop\Iraq.doc" ..
```

leaking some of the political people, not experts, who edited it

- Paul Hamill - Foreign Office official
- Alison Blackshaw - personal assistant of the Prime Minister's press secretary
- Murtaza Khan - junior press officer for the Prime Minister

Questions for the future

- Battle *of* and *in* the browsers:

  What will be the *default* policies & configurations of web-browsers and apps?

  – eg wrt. 3rd party cookies
- What parties are controlling this, and what are their motives & business models?
  – The evolution of Google Chrome is steered by different (market) incentives than Mozilla Firefox
- Will web-sites have unique identifiers, even if you block or frequently delete cookies?
  – eg IP address
  – note that web sites are keen to collect unique identifiers, eg phone number (in WhatsApp, or for Google account recovery) or credit card number

Risks


EDITION: INTERNATIONAL
U.S.
MÉXICO
ARABIC
TV:
CNNi
CNN en Español
Set edition preference
CNN
Home
Video
World
U.S.
Africa
Asia
Europe
Latin America
Middle East
Money
World Sport
UK police to investigate possible murder of three bo
claims abuse took place at "military premises."
5 Things to know about the celebrity nude photo hacking scandal
By Alan Duke, CNN
October 13, 2014 -- Updated 0040 GMT (0840 HKT)
SHARE TH
Recomm
ChartBeat
Disqus
DoubleClick DART
Facebook Connect
Facebook Social Plugins
Krux Digital
NetRatings SiteCensus
Omniture (Adobe Analytics)

Privacy risks? **profiling**

Some people will claim they have nothing to hide.
They may even claim that they like targetted advertising.

But the risks of profiling go further:

- you may pay a higher price than someone else
  because your profile shows that you are richer
- you may not be offered the same product as someone else
  eg no insurance because your profile indicates higher risks
- you may not get the same information as someone else,
  eg your online newspaper will filter contents you don't like to read
  so getting objective information gets harder

Privacy risks?

Apart from the fundamental loss of privacy & profiling
other risks include

- stolen personal information used for other attacks , eg
  - attacks on your friends, colleagues
  - blackmailing, eg of children for webcam sex
  - ...

  especially useful for social engineering attacks.

- identity theft
  - reported to be the fasted growing form of crime

Privacy risk: more hacking

February 21, 2005 3:35 PM PST

Paris Hilton's cell phone hacked?

By Steven Musil
Staff Writer, CNET News

Related Stories

T-Mobile: Hacker had limited access

January 12, 2005

Hackers steal ID info from Virginia

Paris Hilton seems to be having more trouble keeping her personal life personal, and this time the socialite apparently exposed several A-list celebrities after the contents of her cell phone were published on the Internet.

The content included the phone numbers of the hotel heiress' friends.

Privacy risk: identity theft

Identity theft victims face months of hassle

Story | Comments | Image (3) | Print | Font Size:

Recommend 2 | Tweet | g+1 1 | Pin it 0 | Share 0

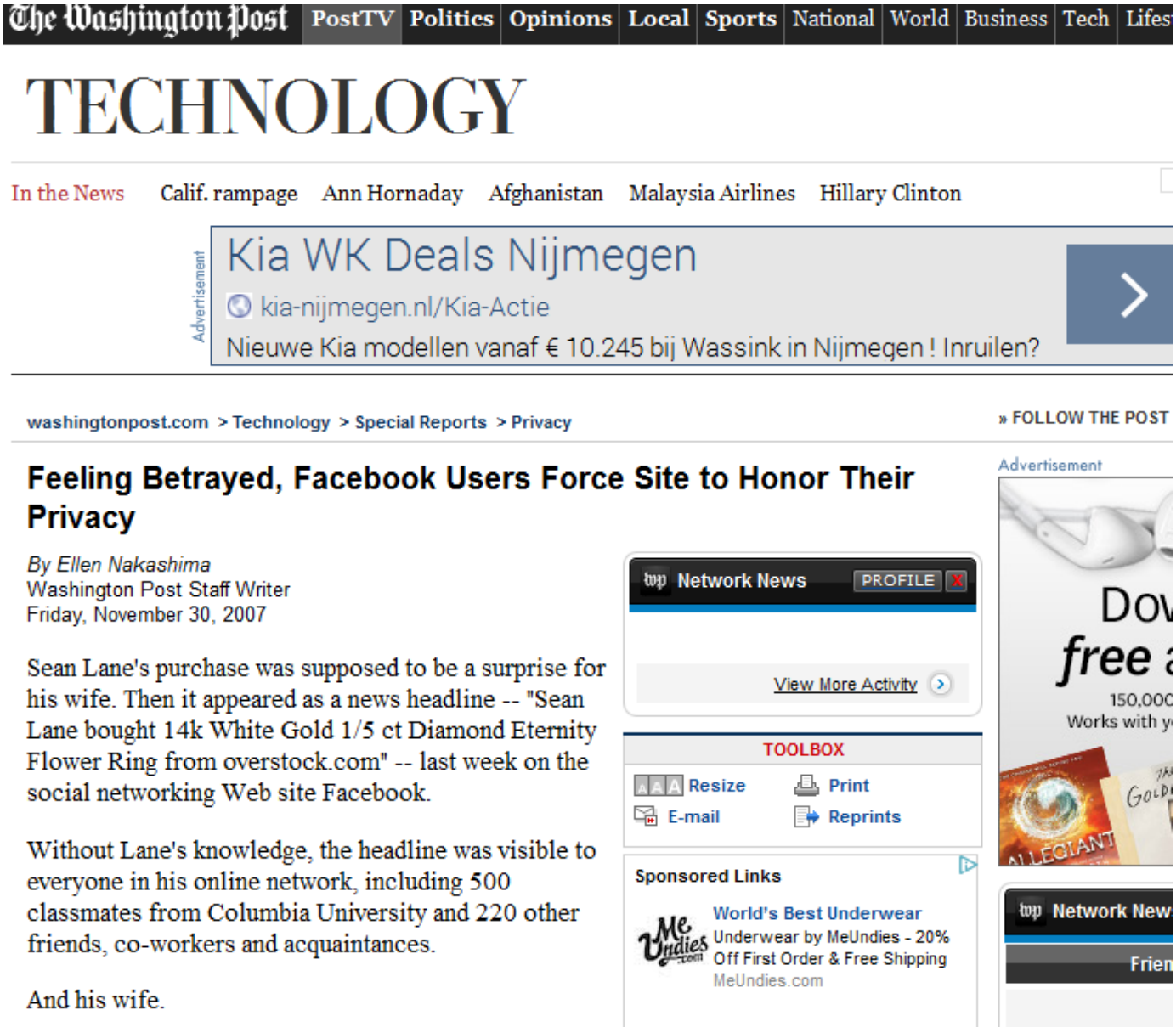
Posted: Sunday, December 14, 2014 9:35 am | *Updated: 9:46 am, Mon Dec 15, 2014.*

Previous | Next

Associated Press | 0 comments

SAN FRANCISCO (AP) — As soon as Mark Kim found out his personal information was compromised in a data breach at Target last year, the 36-year-old tech worker signed up for the retailer's free credit monitoring offer so he would be notified if someone used his identity to commit fraud.

Facebook's Beacon ruining Christmas



The Washington Post | PostTV | Politics | Opinions | Local | Sports | National | World | Business | Tech | Lifes

TECHNOLOGY

In the News Calif. rampage Ann Hornaday Afghanistan Malaysia Airlines Hillary Clinton

washingtonpost.com > Technology > Special Reports > Privacy

» FOLLOW THE POST

Feeling Betrayed, Facebook Users Force Site to Honor Their Privacy

By Ellen Nakashima
Washington Post Staff Writer
Friday, November 30, 2007

Sean Lane's purchase was supposed to be a surprise for his wife. Then it appeared as a news headline -- "Sean Lane bought 14k White Gold 1/5 ct Diamond Eternity Flower Ring from overstock.com" -- last week on the social networking Web site Facebook.

Without Lane's knowledge, the headline was visible to everyone in his online network, including 500 classmates from Columbia University and 220 other friends, co-workers and acquaintances.

And his wife.

Big Brother Pizza Shop

Browser plugins to try out!

- lightbeam Firefox
- ghostery Firefox, Chrome, Safari, Opera en IE
- DNTM (DoNotTrackMe) Chrome, Firefox, IE en Safari

Trust

In Ken Thompson's Turing award acceptance speech
he revealed a backdoor in UNIX and Trojan in C-compiler.

1. backdoor in **`login.c:`**

```
if (name == "ken") {don't check password;
                    log in as root}
```

2. code in C compiler to add backdoor
   when recompiling login.c
3. code in C compiler to add code (2 & 3!)
   when (re)compiling a compiler

Moral of the story: you are trusting more than you expect!