

Online Interactions and Privacy Issues

Giovanni Livraga

Security, Privacy, and Data Protection Laboratory

Dipartimento di Informatica

Università degli Studi di Milano

giovanni.livraga@unimi.it

<http://www.di.unimi.it/livraga>

Privacy? – 1

- In 2006, AOL released 20 million search queries for 650,000 users
- Identifying information replaced with [unique identification numbers](#)
- User [4417749](#):
 - “numb fingers”, “60 single men”, “dog that urinates on everything”, “landscapers in Lilburn, Ga”, “Arnold” (several people with this last name)

Privacy? – 1

- In 2006, AOL released 20 million search queries for 650,000 users
- Identifying information replaced with unique identification numbers
- User 4417749:
 - “numb fingers”, “60 single men”, “dog that urinates on everything”, “landscapers in Lilburn, Ga”, “Arnold” (several people with this last name)

A Face Is Exposed for AOL Searcher No. 4417749

By MICHAEL BARBARO and TOM ZELLER Jr.
Published: August 9, 2008

Buried in a list of 20 million Web search queries collected by AOL and recently released on the Internet is user No. 4417749. The number was assigned by the company to protect the searcher's anonymity, but it was not much of a shield.



Erik G. Lessor for The New York Times
Thelma Arnold's identity was betrayed by AOL records of her Web searches, like ones for her dog, Dudley, who clearly has a problem.

No. 4417749 conducted hundreds of searches over a three-month period on topics ranging from “numb fingers” to “60 single men” to “dog that urinates on everything.”

And search by search, click by click, the identity of AOL user No. 4417749 became easier to discern. There are queries for “landscapers in Lilburn, Ga,” several people with the last name Arnold and “homes sold in shadow lake subdivision gwinnett county georgia.”

It did not take much investigating to follow that data trail to Thelma Arnold, a 62-year-old widow who lives in Lilburn, Ga., frequently researches her friends' medical ailments and loves her three dogs. “Those are my searches,” she said, after a reporter read part of the list to her.



Privacy? – 2

What about user 17556639?

- how to kill your wife
- how to kill your wife
- wife killer
- how to kill a wife
- poop
- dead people
- pictures of dead people
- killed people
- dead pictures
- dead pictures
- dead pictures
- murder photo
- steak and cheese
- photo of death
- photo of death
- death
- dead people photos
- photo of dead people
- www.murderdpeople.com
- decapitated photos
- decapitated photos
- car crashes3
- car crashes3
- car crash photo

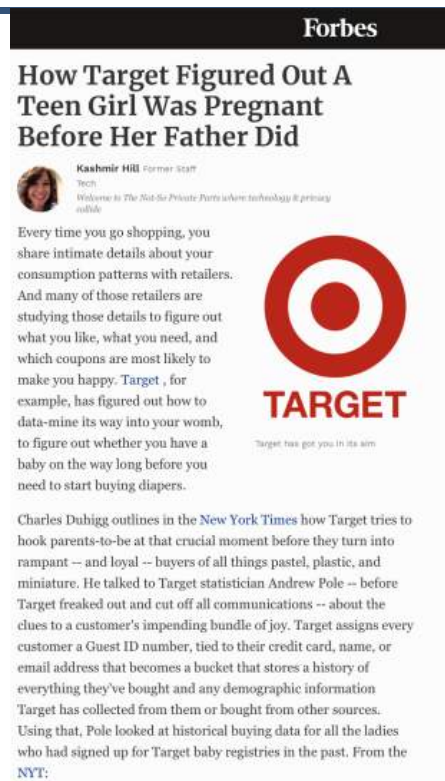
Privacy? – 3

- In 2006, "Netflix Prize" for movie recommendation algorithm
- Netflix provided 100 million records revealing how nearly 500,000 users had rated movies from Oct.'98 to Dec.'05
- Only a sample of the movie ratings database was released
- Some ratings were perturbed (but not much to not alter statistics)
- Identifying information (e.g., username) was removed, but a **unique user identifier** was assigned
- De-identified Netflix data can be re-identified by **linking with external sources** (e.g., user ratings from IMDb users)

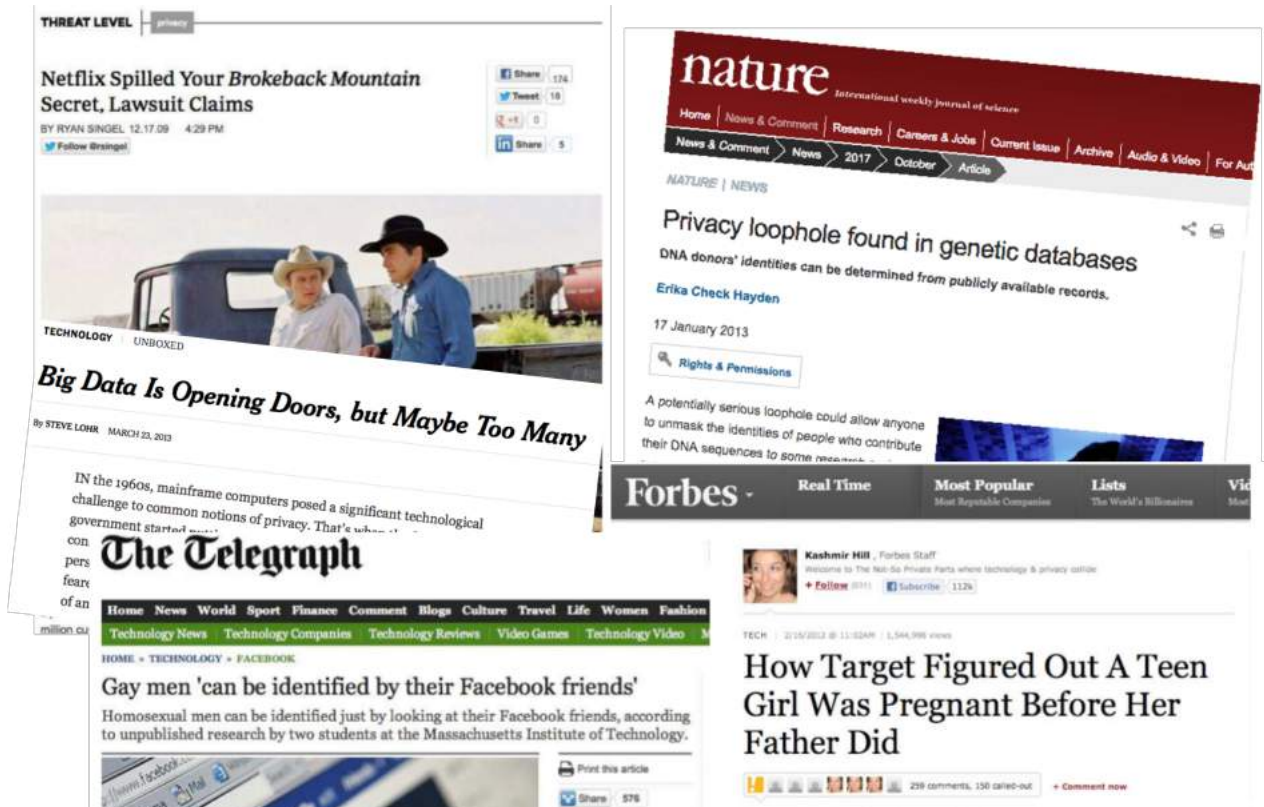


Privacy? – 4

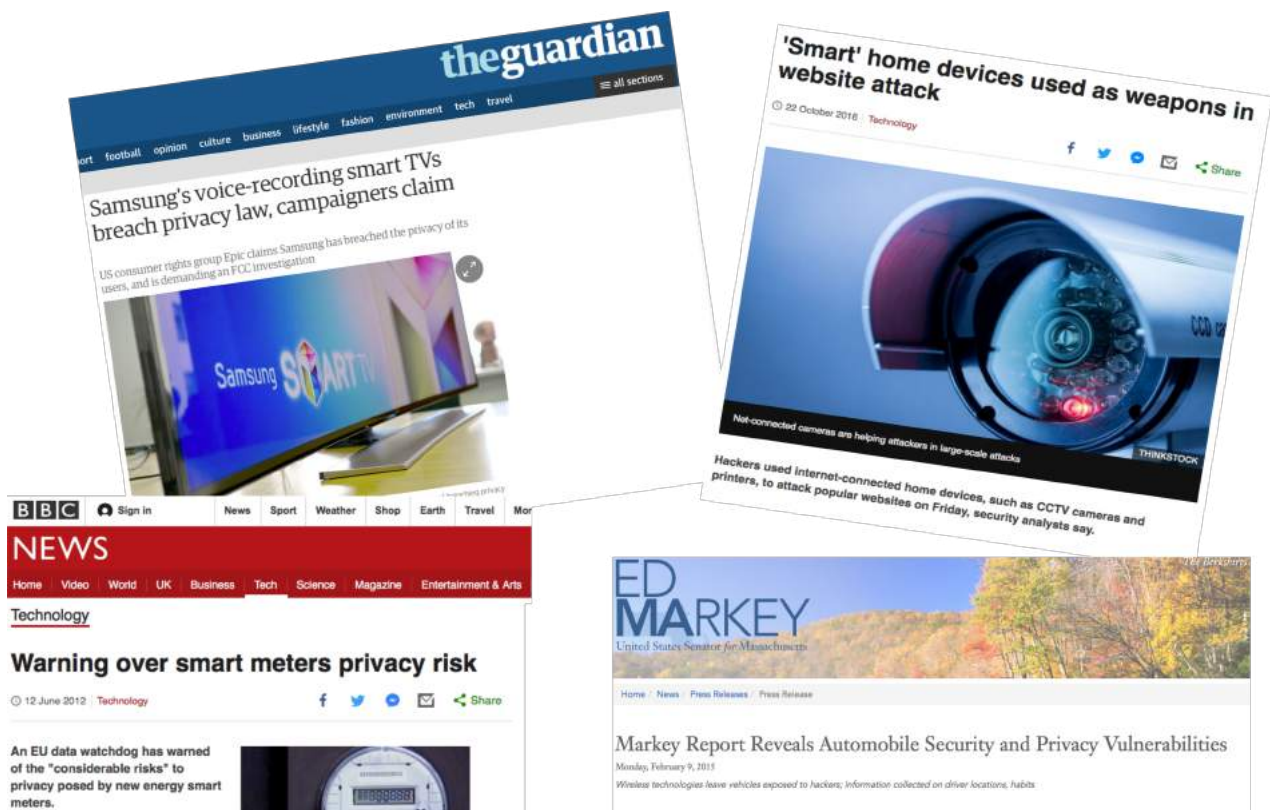
- In 2012, Target found to **mine purchase data** and apply pregnancy prediction model on female customers
- Target assigns every customer a **Guest ID number**
- Analysts at Target identified 25 products that assign each shopper a **pregnancy prediction score**
 - **woman, 23 y.o., buying in March cocoa-butter lotion, a purse large enough to double as a diaper bag, zinc and magnesium supplements and a bright blue rug**
⇒ **87% due late August**



Privacy? – 5



Privacy? – 6



Privacy? – 7



Privacy issues

- Data about users is every day **collected**, analyzed, and possibly **shared** and disseminated
 - with the consent of the user (e.g., application forms)
 - without the consent of the user (e.g., cookies, tracking, big data)
 - made available by the user herself (e.g., social media)
- Data from multiple sources may be **linked**
 - non-identifiable information can become identifiable when linked
- Users **lose control** on:
 - what information is collected
 - how the information is used
 - secondary use

Privacy – Example (1)

Medical confidentiality is one of the basic ethics for a physician (the Hippocratic Oath)

- There is no more a physician ... there is a complex system
 - family doctor
 - specialist
 - hospital
 - medical laboratory
 - chemist
 - employer
 - insurance
 - ...

Privacy – Example (2)

- **Electronic patient's medical record** contains a large amount of both clinical data (e.g., diagnoses, medicines, ...) and personal information
- The record contains therefore information accessible by different users/organizations and for different purposes:
 - administrative tasks
 - clinical practices (diagnostic and therapeutic decisions)
 - research practices
- Medical data are collected and aggregated
 - + research advancements
 - can compromise patients' privacy

Privacy violations – 1

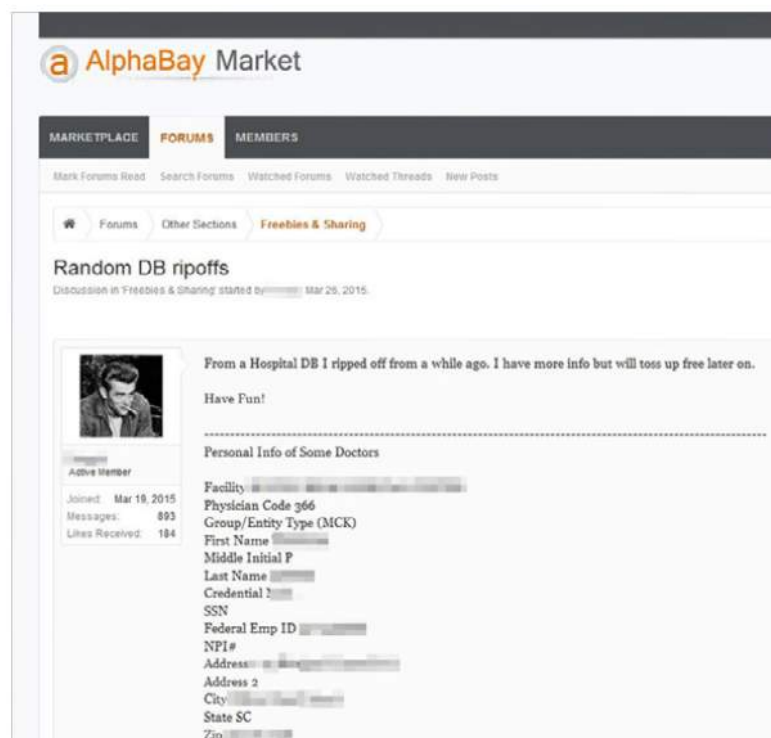


The screenshot shows a web browser displaying the Labio medical laboratory website. The page includes contact information for the laboratory, patient details, and a hematology report. The report is titled 'HEMATOLOGIE' and contains various blood test results.

HEMATOLOGIE			
Valeurs de référence			
Résultats de données relatives le 22/07/2015, validés le 26/07/2015			
NUMERATION GLOBALE			
(Abdel Rely-Optimatus de Bio) (SA)			
* Leucocytes	13.90	10 ⁹ /L	4.810
* Hémoglobine	6.06	10 ¹² /L	45.813
Hémoglobine	16.8	g/dL	132.417,7
Hématocrite	50.9	%	46.54
T.C.M.H	27.30	pg	27.31
C.C.M.H	35.1	%	32.434
V.C.M	84	μd	80.100
FORMULE LEUCOCYTAIRE			
Poly. neutrophiles	50.0	%	7.48
Poly. éosinophiles	0.8	%	0.10
Poly. basophiles	0.6	%	0.08
* Lymphocytes	33.5	%	4.32
Monocytes	7.2	%	0.93
NUMERATION PLAQUETTAIRE			
Plaquettes	272	10 ⁹ /L	130.440
HEMOCTASE			

Source: McAfee Labs, *The Hidden Data Economy*, 2015

Privacy violations – 2



The screenshot shows a forum post on the AlphaBay Market website. The post is titled 'Random DB ripoffs' and is part of a discussion in the 'Freebies & Sharing' section. The post contains a profile picture of a user and a list of personal information for several doctors, including facility names, physician codes, and addresses.

From a Hospital DB I ripped off from a while ago. I have more info but will toss up free later on.

Have Fun!

Personal Info of Some Doctors

Facility	Physician Code	Group/Entity Type (MICK)	First Name	Middle Initial P	Last Name	Credential ?	SSN	Federal Emp ID	NPI#	Address	Address 2	City	State SC	Zip
[REDACTED]	366	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

Source: McAfee Labs, *The Hidden Data Economy*, 2015

Privacy violations – 3

```
Username(s): [REDACTED]
Emails: [REDACTED]@hotmail.co.uk, [REDACTED]zz@gmail.com
First Name: [REDACTED]
Last Name: Bright
Gender: Male
Date of Birth: [REDACTED], 1995
Age: 19
Skype(s): [REDACTED]
City: Norwich, Norfolk
Postal Code: [REDACTED]
State/Province: Norwich, Norfolk
Country: United Kingdom
Street: [REDACTED]
PlayStation Network: [REDACTED]

===== Accounts =====
http://Twitter.com/[REDACTED]
http://Instagram.com/[REDACTED]
http://Twitch.tv/[REDACTED]
http://Youtube.com/user/[REDACTED]
http://Youtube.com/MzBom3rZz - [REDACTED]
http://Youtube.com/user/[REDACTED]
http://Facebook.com/[REDACTED]
http://Plus.Google.com/[REDACTED]
http://Plus.Google.com/[REDACTED]
http://SteamCommunity.com/profile/[REDACTED]
http://ConsoleCrunch.org/index.php?members/[REDACTED]
http://NextGenUpdate.com/forums/member/[REDACTED]
http://XenForo.com/community/members/[REDACTED]

===== Family =====
Mother: [REDACTED]
http://Facebook.com/[REDACTED]
Father: [REDACTED]
```

Source: McAfee Labs, *The Hidden Data Economy*, 2015

Privacy violations – 4

Payment Card Number With CVV2	United States	United Kingdom	Canada	Australia	European Union
Software-generated	\$5–\$8	\$20–\$25	\$20–\$25	\$21–\$25	\$25–\$30
With Bank ID Number	\$15	\$25	\$25	\$25	\$30
With Date of Birth	\$15	\$30	\$30	\$30	\$35
With Fullinfo	\$30	\$35	\$40	\$40	\$45

Source: McAfee Labs, *The Hidden Data Economy*, 2015

Laws and regulations

- Privacy is **not only** a technological issue
- Privacy **laws and regulations** vary widely throughout the world
- Examples of US laws:
 - **HIPAA** (Health Insurance Portability and Accountability Act)
 - it should protect medical records and other individually identifiable health information
 - **COPPA** (Children's Online Privacy Protection Act)
 - web sites that target children must obtain parental consent before collecting personal information from children under the age of 13
 - **GLB** (Gramm-Leach-Bliley-Act)
 - requires privacy policy disclosure and opt-out mechanisms from financial service institutions

History: Italian legislation – 1

Decree 30/06/2003, n. 196 \Rightarrow “Codice in materia di **protezione dei dati personali**”

- art. I: “**Chiunque** ha diritto alla **protezione dei dati personali** che lo riguardano”
- Acknowledges an individuals' rights over her personal data and regulates all data manipulations
 - collection
 - elaboration
 - deletion
 - updates
 - communications
 - disclosure and sharing
- **Autorità Garante per la protezione dei dati personali** is in charge of evaluating fulfilment of the laws

History: Italian legislation – 2

Titolo II of Testo Unico, artt. 8 – 10, regulates:

- rights of individuals
- how personal data can be collected
- requirements over collected data
- obligations to be satisfied by subjects collecting personal data
- obligations to be satisfied by subjects storing or manipulating personal data
- responsibilities and penalties in cases of misbehaviors

EU data protection legislation

January 5, 2012: the European Commission presented its revision proposal for the EU data protection legislation \Rightarrow GDPR (General Data Protection Regulation) adopted Apr 2016, start May 2018

- strengthen users' rights w.r.t. privacy
- uniform the 27 different laws of EU countries
 - Increase in the transparency in privacy policies
 - Explicit consent to usage of personal data (impact on the collection of browsing data!)
 - Right to be forgotten (access to information and possibility of deleting or moving them)
 - Data Protection Officer
 - Penalties and fines

EU GDPR – 1

- The EU GDPR applies to **all parties** (even non-European) that **process data of EU citizens**
- Key definitions:
 - **data controller**: an organisation that collects data from EU residents
 - **data processor**: an organisation that processes data on behalf of a data controller (e.g., cloud provider)
 - **data subject**: a natural person
 - **personal data**: any information related to an identified or (directly or indirectly) identifiable data subject
 - **sensitive personal data**: data revealing racial or ethnic origin, religious or philosophical beliefs, trade-union membership, health status, sex life, sexual orientation, and genetic and biometric data

EU GDPR – 2

- Key definitions (cont.):
 - **anonymous data**: data for which the related subject cannot be identified directly or indirectly, by any means or by any person
 - **pseudonymous data**: data for which the related subject cannot be identified directly or indirectly without a key
- Key requirements:
 - **Lawful, fair and transparent processing**: data processing must be based on legitimate purpose and subjects must be informed
 - **Limitation of purpose, data and storage**: companies must collect only the data strictly needed, data should not be kept after the processing
 - **Data subject rights**: for transparency, correction, objection to processing, data deletion and transfer

- Key requirements (cont.):
 - **Consent**: clear and explicit, must be acquired from subjects
 - **Personal data breaches**: regulator and subject may be timely informed
 - **Privacy by Design and by Default**: should be put in place
 - **Data Protection Impact Assessment**: should be conducted when needed
 - **Data transfers**: data must be kept and processed according to the GDPR even when transferred to third parties
- EU institutions are also working towards the **ePrivacy regulation**, which would be complementary to the GDPR, focusing on confidentiality and integrity of electronic communications

Web and data collection

Information collection

- **Personal information** can be collected in the Web
 - application forms
 - questionnaires and surveys
 - ...
- Users **can know** that their personal information is collected, but lose control on **how** and to **whom** the information is disclosed
- Often users **do not even know** that their personal information is collected
 - referral links
 - log files
 - cookies
 - ...

Referral link – 1

- When a user looks for a new Web page, the **URL of the page** that the user is **currently** looking at can be sent along
- Benefits for companies, as it permits to
 - track how customers move through sites
 - therefore also check the effectiveness of advertisements they purchase
- The referral link field can also reveal **personal information**
 - the URL of the page the user was looking at before she clicked on a link
 - when a user clicks the link of a web search engine, the refer link sent to the server encodes the query asked by the user

Referral link – 2

- **Referral marketing** is used to promote products to new customers through **referrals** (i.e., people who already know the products) as traditional **word of mouth**
- **Dropbox** in 2009 used a referral program to acquire new users
 - they encouraged their users to tell their friends about Dropbox
 - the person who signs up for Dropbox through a referral link gets more space than through a normal sign up
 - the referrer gets additional space on her account
 - the referral program increased the signups by 60%
- In the Dropbox example, the referral link reveals the identity the users who invited her friend

Log file – 1

Every time a browser downloads a page on the Web, this event may be recorded in the **log files** of the remote Web server, which can keep the control on the log file and register:

- the name and IP address of the computer that downloaded the web page
- the time of the request
- the URL that was requested
- the time it took to download the file
- if HTTP authentication was used, the log file contains the username of the person who downloaded the file
- any errors that occurred
- the previous web page that was downloaded by the web browser (referral link)

Log file – 2

- This information can be combined with other log files (e.g., login/logout information from Internet service providers, or logs from mail servers) to discover the actual **identity** of a user
- Normally this kind of correlation requires the **assistance of another organization**, but that is not always the case

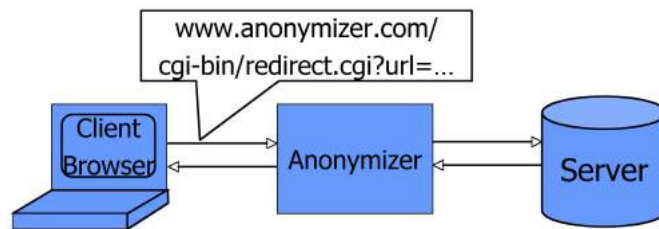
Proxy servers and anonymizers

Privacy can be enhanced by using:

- **Proxy servers**
 - act as **intermediaries** between a client and a server
 - the client connects to the proxy requesting a service available from server X
 - the proxy sends the request to X and sends back the service to the client
 - X logs the request as coming from the proxy
- **Anonymizers**
 - proxies specifically designed to maintain their accesses on the web anonymous (e.g., www.anonymizer.com, www.your-freedom.net, www.the-cloak.com)

Browsing anonymizers

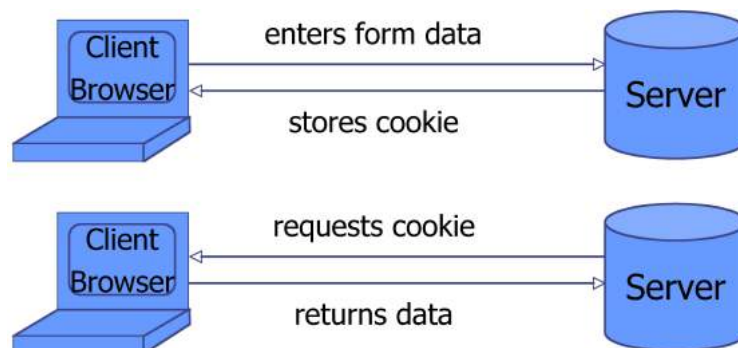
- The IP address is visible to web sites; this may
 - reveal users' employers, ISPs, ...
 - enable linking activities across sites over time
- A web anonymizer hides IP addresses



- The anonymizer sees all web traffic and it can add ads to pages
- Protection is not absolute (HTTP headers can leak some information)

Cookies – 1

- A **cookie** is a **small file** created by an Internet site to **store** information on a user's computer



- Cookies allow websites to **identify a user/device** that **accesses them** and her/its **preferences**, such as **language** and other **identifiable information**

Cookies – 2

- A cookie is part of an HTTP header (request or response)
- A cookie is generally composed of
 - arbitrary text string
 - expiry date
 - pattern to recognize the domain(s) to which the cookie can be sent back
- HTTP is stateless protocol; cookies add state
 - if a server sets a cookie, the cookie will be included in the answer of a request
 - once server X sets a cookie, every time the browser visits X it also sends the cookie back to X

Cookies – 3

- Only the Web site that created the cookie can read it
- Cookies can be classified by
 - their kind
 - temporary/session cookies, stored until the user quits her browser
 - persistent/tracking cookies, remain until deleted or expire
 - their source
 - first-party cookies
 - third-party cookies, originated by or sent to a web site other than the one that provided the current page
 - their purpose
 - technical, analytics, preference, ...

Cookies – 4

- Cookies can be **useful**:
 - used like a staple to attach multiple parts of a form together
 - used to recognize when a user returns to a web site so the user do not have to remember a password
 - used to help web sites understand how people use them
 - used to manage sessions
- Cookies can be **used improperly**:
 - profiling users and tracking their activities, especially across web sites

Cookies issues

- Cookies maintain **records of browsing habits**
 - may include information a web site knows about the user
 - sites track users activity from multiple visits to site
- Cookies can be used to store **virtually unlimited amounts of information**
 - cookies have a **maximum size** and store information as name/value pairs
 - workaround: store simply a **unique identifier**, and store all wanted information at server side **linked to the identifier**

Third-party cookies

- Get a page from mysite.com
 - suppose that it contains an image from [doubleclick.com](#)/
 - the image is fetched from DoubleClick.com, which [creates a \(third-party\) cookie](#)
- What happens if other sites include images from [doubleclick.com](#)?
 - the user's DoubleClick cookie is sent back to DoubleClick.com
 - DoubleClick.com can [track](#) the websites visited by the user
 - basically, the images from DoubleClick.com [identify](#) the user!

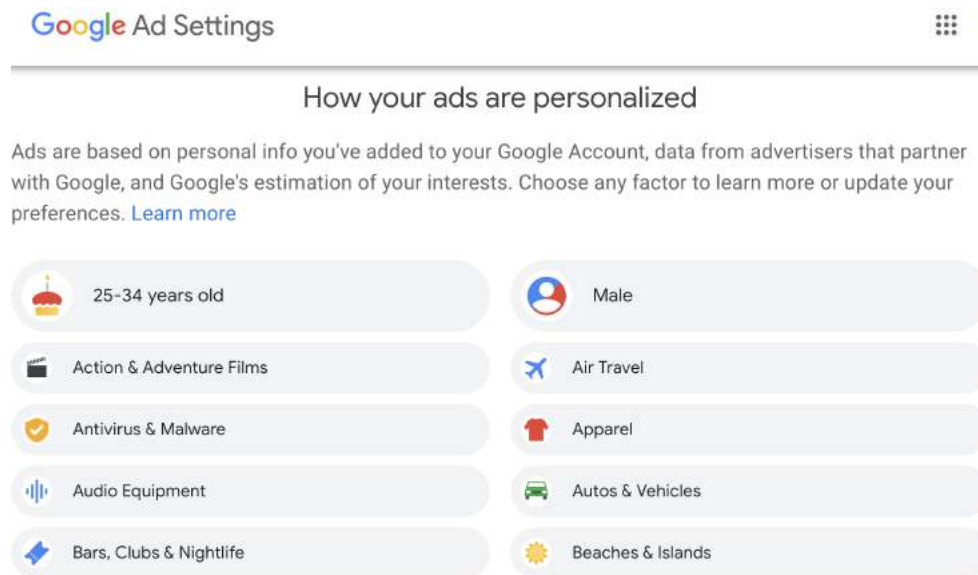


Cookies - Google Example (1)

```
time: 06/Aug/2008 12:01:32
ad_placement_id: 105
ad_id: 1003
userid: 000000000000000001
client_ip: 123.45.67.89
referral_url: "http://youtube.com/categories"
[Example of data stored from a Doubleclick cookie]
```

- It seems [little](#) information, [but...](#)
 - IP address \implies good guess of country and town/city
 - if the website is part of, e.g., Google AdSense, other visits to other sites in AdSense will be recorded and pooled
 - over time, guesses can be made about the interests of the person using that browser

Cookies - Google Example (2)

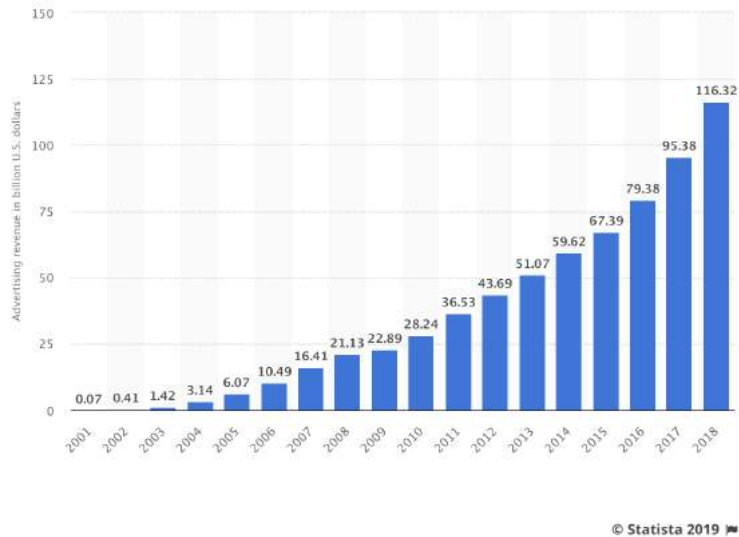


Cookies - Google Example (3)

- Google bought DoubleClick for \$3.1 billion in 2008
- ... was it worth?

Cookies - Google Example (3)

- Google bought DoubleClick for \$3.1 billion in 2008
- ... was it worth?



[Advertising revenue of Google from 2001 to 2018 (in billion U.S. dollars)]

Disabling cookies

- Usually browsers have options that allow you to
 - be notified when a cookie is received
 - accept all cookies, reject all cookies, or be prompted for each cookie and explicitly accept it or not
 - control cookie acceptance on a site-by-site basis
- but...
 - functionalities that do not work if cookies are disabled (e.g., “basket” in on-line purchases)
 - zombie cookies
 - ...

Cookies issues – 1

- **13 Sept 2012:** Google Chrome 18 on Android allows remote attackers to obtain cookie information via a crafted application
- **12 March 2012:** Apple Safari 5.1.4 does not properly implement “From third parties and advertisers” cookie blocking \implies easier for web servers to track users via a cookie
- **25 Oct 2011:** Google Chrome 15 does not properly handle javascript: URLs, which allows remote attackers to bypass intended access restrictions and read cookies via unspecified vectors
- Some shopping cart software has been found to enable web users to view order information of other shoppers

Cookies issues – 2

THE WALL STREET JOURNAL

Home World U.S. Politics Economy Business Tech Markets Opinion Life & Arts Real Estate WSJ. Mail

TECH

Yahoo Triples Estimate of Breached Accounts to 3 Billion

Company disclosed late last year that 2013 hack exposed private information of over 1 billion users

By *Robert McMillan* and *Ryan Knutson*
Updated Oct. 3, 2017 9:23 p.m. ET

A massive data breach at Yahoo in 2013 was far more extensive than previously disclosed, affecting all of its 3 billion user accounts, new parent company Verizon Communications Inc. said on Tuesday.

The figure, which Verizon said was based on new information, is three times the 1 billion accounts Yahoo said were affected when it first disclosed the breach in December 2016. The new disclosure, four months after Verizon completed its acquisition of Yahoo, shows that



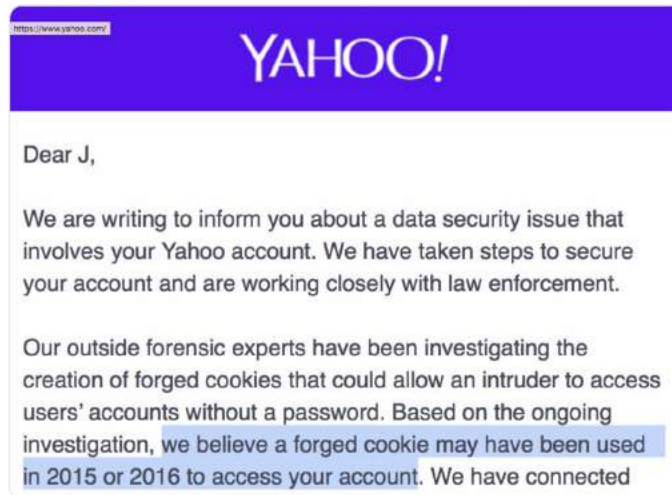
Cookies issues – 3



Joshua B. Plotkin
@jplotkin

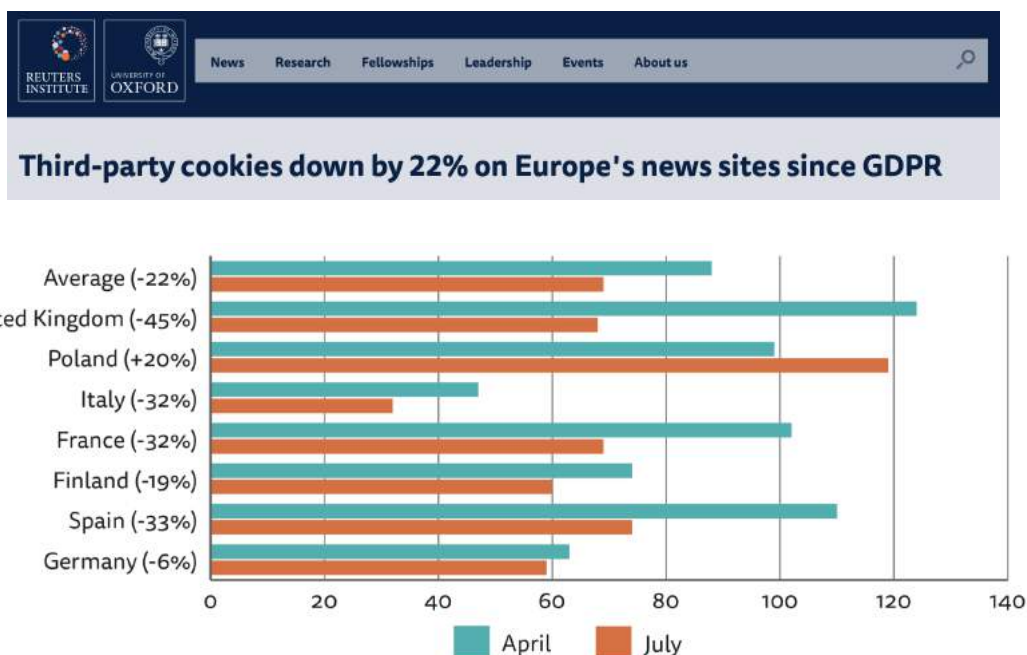
Follow

Hopefully the cookie was forged by a state known for such delicacies. #yahoo #security #baking



8:51 am - 15 Feb 2017

Cookies and EU regulations

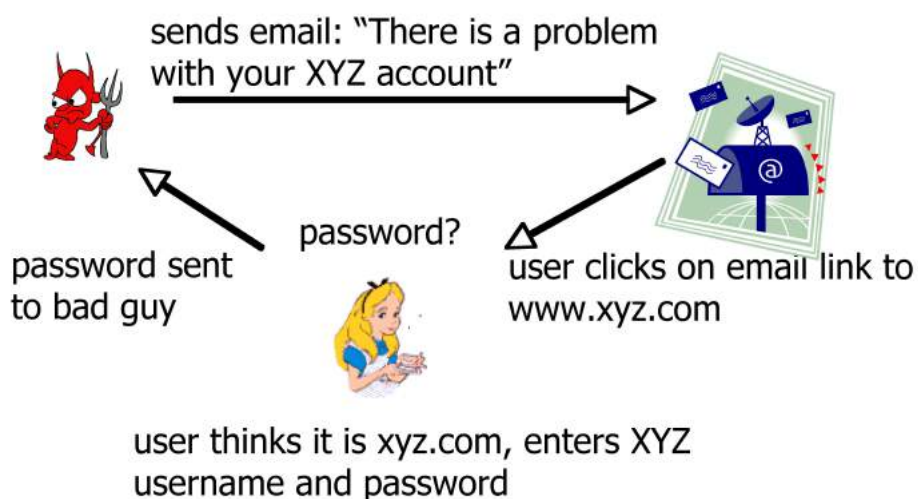


[Third-party cookies per page by country (April-July 2018 change)]

Seal programs

- They help to protect online consumers by identifying safe and trustworthy Web sites
- In general, a seal program provides a trusted third party verification service that should certify compliance with stated policy
 - limited ability to detect non-compliance
- Example:
 - TRUSTArc (formerly TRUSTe)
 - ...

Phishing



Web bugs

Web bugs (also called web beacons, clear gifs, tracker gifs, ...)

- Invisible images (1x1 pixel, transparent) embedded in web pages
- Work like banner ads, but only visible looking at the source code behind the page
- Also embedded in HTML formatted emails, Word documents, ...
- Allow the server to discover that the user opened the page including the web beacon
- Can transfer simple information (e.g., IP address, browser type) and the existence of cookies previously set by the server
- Example: used by Facebook for their advertisement, shut down in 2009 after a long controversy

Spyware

Spyware

- Malware (malicious software) that collects information about users without their knowledge or explicit permission
- Are typically used to monitor users
- Can collect almost any type of data (e.g., personal information, user logins, bank or credit card information)
- Can install additional software or redirecting Web browsers or change computer settings
- Example: keyloggers

Privacy and Social Media

Social Media – 1

In recent years, users have played an **active role** in the Information Society, making their information available themselves

- They can **stay in contact** with colleagues, friends, relatives
- They can access countless **online services** with a single mouse click
- They can **share** their own ideas, user-generated contents, and preferences
- ...

Social Media – 2

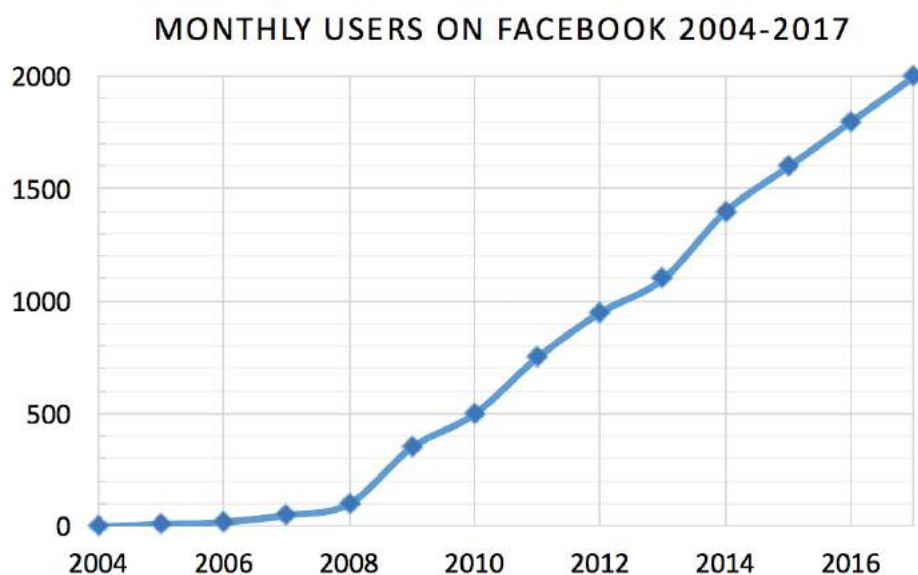
Social Media offer:

- definition of a personal profile, management of friendships, “wall”, sharing of preferences and interests ...
⇒ **digital identity** and related social relationships
- novel communication channels: one-to-one, one-to-many, many-to-many, in synchronous and asynchronous fashions
- sharing of contents (images, videos, texts and links) as an alternative or coupled with dedicated sharing services such as YouTube (videos) and Flickr (images)



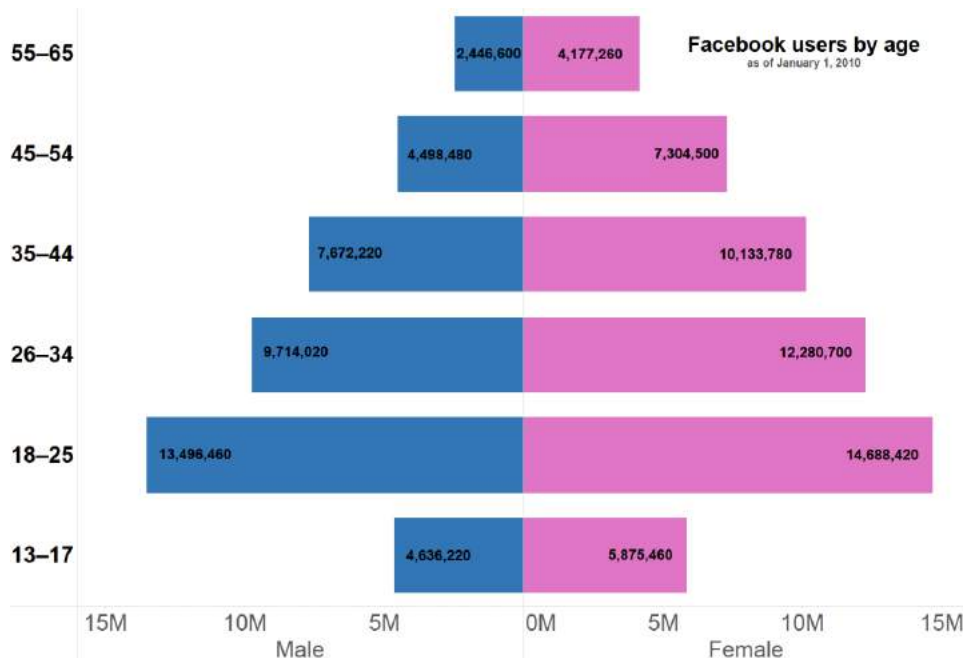
The explosion of Facebook

- Facebook counts more than 2.9 billion monthly active users in the world as of March 2022



The explosion of Facebook

- Facebook counts more than 2.9 billion monthly active users in the world as of March 2022



The content of Facebook – 1

- Success of Social Media \implies hundreds of millions of individuals have spontaneously chosen to [make their personal information available online](#) (to an audience ranging from some friends to the entire cyberspace)
- Users publish on Facebook a tremendous amount of information:
 - personal
 - sensitive
 - not sensitive but exploitable to infer sensitive information
- Some figures:
 - trading on NASDAQ: original peak market capitalization of [USD 104 billion](#)
 - fastest company in the S&P 500 Index to reach a market cap of [USD 250 billion](#)
 - [USD 500+ billion](#) market cap on May 02, 2019

The content of Facebook – 2

- April 2012: Facebook acquires Instagram (USD 1 billion)
- February 2014: Facebook acquires Whatsapp (USD 19.3 billion)
- ...

⇒ The value of Facebook is strictly linked to the amount of information it holds!

The content of Facebook – 3

- March 2012: press news reporting how U.S. interviewers increasingly ask future employees to access their Facebook profile
- In some cases, credentials to access Facebook profile are required
 - violation of individuals' privacy
 - novel professional figure which, upon payment, helps users to clean their Facebook profile

⇒ Facebook contains important information to profile an individual!

The content of Facebook – 4



Some cases of Facebook information usage – 1

The Telegraph

Home Video News World Sport Business Money Comment Culture Travel Life W
Apple iPhone Technology News Technology Companies Technology Reviews Video Games T

Office worker sacked for branding work boring on Facebook

Teenage office worker Kimberley Swann was sacked from her job after branding it 'boring' on Facebook.



A 16 y.o. woman from Essex (U.K.), has been fired because she had defined on her Facebook profile her job as boring and complained about her filing, stapling and hole-punching duties

Some cases of Facebook information usage – 2



A Canadian 29 y.o. woman lost sick-leave benefits for depression after she posted on her Facebook profile pictures of her in vacation to the sea

Some cases of Facebook information usage – 3



Insurance companies have increased the premium for customers with a social media profile as they are claimed to be easier targets for robberies

Some cases of Facebook information usage – 4

MailOnline

Home | News | U.S. | Sport | TV&Showbiz | Australia | Femail | Health | Science | Money | Latest Headlines | Royal Family | News | World News | Arts | Headlines | France | Most read | Wires | Discounts

MI6 chief faces probe after wife exposes their life on Net

By KIRSTY WALKER
UPDATED: 12:34 BST, 6 July 2009

MI6 faced calls for an inquiry last night after an extraordinary lapse of judgment led to the new head of MI6's personal details being plastered over Facebook.

Millions of people could have gained access to compromising photographs of Sir John Sawers and his



Share | Twitter | Pinterest | Facebook | Email | Print | View comments

Sir John Sawers, head of MI6 U.K. Secret Service, appeared in July 2009 in some pictures posted by his wife on Facebook, revealing details about his family life and movements

Some cases of Facebook information usage – 5

CORRIERE DELLA SERA
Cronache

Home | Opinioni | Economia | Cultura | Spettacoli | Cinema | Sport | Salute | Tecnica

< CRONACHE

Corriere della Sera - Cronache - Grazie a Facebook arrestato a Bangkok il boss siciliano Vito Roberto Palazzolo

MAFIA E AFFARI

Grazie a Facebook arrestato a Bangkok il boss siciliano Vito Roberto Palazzolo

Considerato il banchiere di Riina, è stato condannato a 9 anni



MILANO - E' stato rintracciato mettendo sotto controllo il suo profilo Facebook e quello di alcuni familiari. Oltre alle indagini tradizionali i carabinieri del reparto operativo di Palermo, in collaborazione con l'Interpol, hanno battuto anche le piste telematiche per rintracciare e arrestare il boss Vito Roberto Palazzolo, considerato il tesoriere del «capo dei capi» Totò Riina. Palazzolo è stato fermato all'aeroporto di Bangkok da dove probabilmente si preparava a lasciare il paese. Dopo un lungo braccio di ferro la Thailandia ha consentito la «traduzione in Italia per la consegna alle autorità competenti» anche se sono ancora in corso tentativi da parte dei legali che stanno facendo pressioni sulle autorità dei due paesi. Ufficialmente Palazzolo ha infatti la cittadinanza del Sudafrica,

178 Like | Tweet

OGGI IN cronache >

«Benedetto XVI non è una statua Partecipa alla vita della Chiesa»

Esplode palazzina nell'Aquilano: un morto

Il fratello Beppe: «Fiorello sta molto male»

A Sicilian mafia boss was found and arrested in 2012 in Bangkok thanks to tracking his Facebook profile and those of his relatives (note: good example of “privacy” that we do not want to protect!)

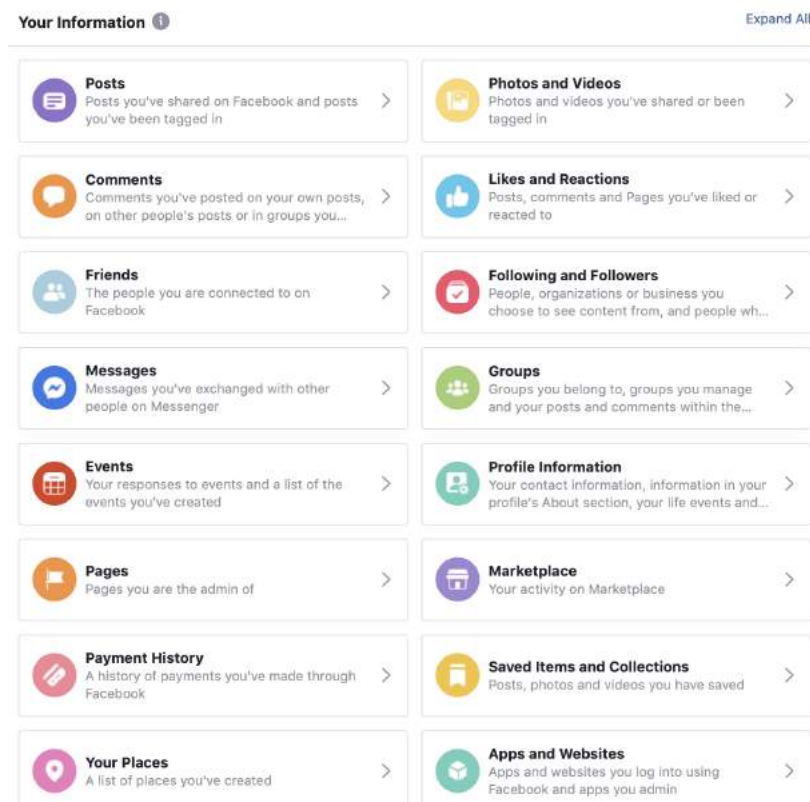
Some cases of Facebook information usage – 6



Girls-around-me: (now removed from store) app for iPhone that analyses the position of Foursquare users (location-based service) and aggregates it with information posted on Facebook by the same users.

Revealing to users information about the female users currently in the nearby area

What Facebook knows



What Facebook knows

Information About You ⓘ

Expand All



Ads

Your interests, interactions, and existing relationships you have with advertisers that...



About You

Information associated with your Facebook account



Search History

A history of the words, phrases and names you've searched for



Location

Information related to your location



Calls and Messages

Logs of your calls and messages that you've chosen to share in your device settings

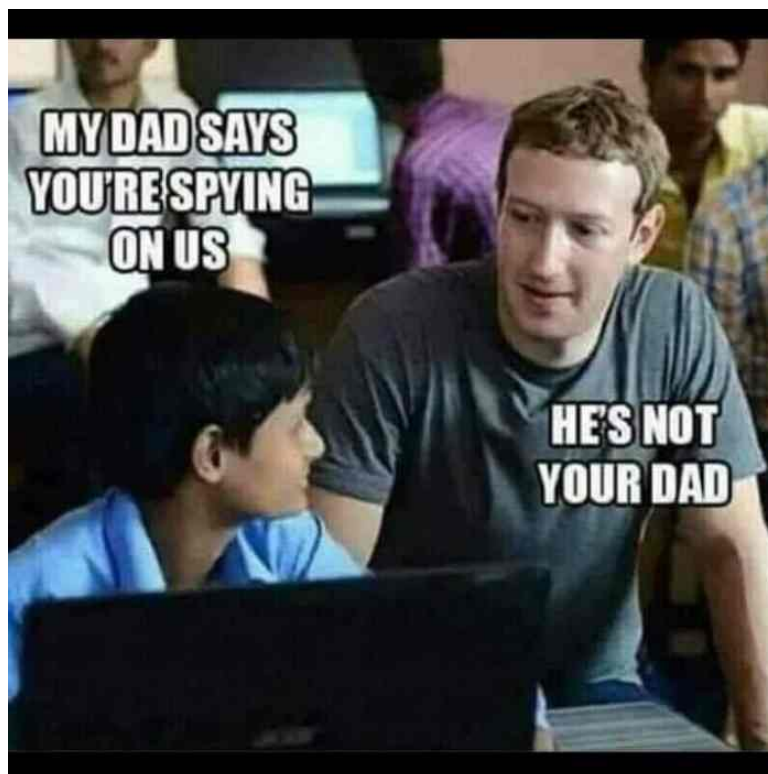


Security and Login Information

Your login history, session length, and other security-related inform... security information



What Facebook knows



Privacy controls – 1

General

Security and Login

Your Facebook Information

Privacy

Timeline and Tagging

Location

Blocking

Language

Face Recognition

Notifications

Mobile

Public Posts

Apps and Websites

Instant Games

Business Integrations

Ads

Payments

Support Inbox

Videos

Privacy Settings and Tools

Your Activity	Who can see your future posts?	Friends	Edit
	Review all your posts and things you're tagged in		Use Activity Log
	Limit the audience for posts you've shared with friends of friends or Public?		Limit Past Posts
How People Find and Contact You	Who can send you friend requests?	Everyone	Edit
	Who can see your friends list?	Only me	Edit
	Remember, your friends control who can see their friendships on their own Timelines. If people can see your friendship on another timeline, they'll be able to see it in News Feed, search and other places on Facebook. If you set this to Only me, only you will be able to see your full friends list on your timeline. Other people will see only mutual friends.		
	Who can look you up using the email address you provided?	Friends	Edit
	Who can look you up using the phone number you provided?	Friends	Edit
	Do you want search engines outside of Facebook to link to your profile?	No	Edit

Privacy controls – 2

Limit The Audience for Old Posts on Your Timeline

Close

If you choose to limit your past posts, posts on your timeline that you've shared with Friends of friends, and Public posts, will now be shared only with Friends. Anyone tagged in these posts, and their friends, may also still see these posts.

If you want to change who can see a specific post, you can go to that post and choose a different audience. [Learn about changing old posts](#)

“Anyone tagged in these posts, and their friends, **may also see these posts**”...

Privacy controls – 3

- General
- Security and Login
- Your Facebook Information
- Privacy
- Timeline and Tagging**
- Location
- Blocking
- Language
- Face Recognition
- Notifications
- Mobile
- Public Posts
- Apps and Websites
- Instant Games
- Business Integrations
- Ads
- Payments
- Support Inbox
- Videos

Timeline and Tagging Settings

Timeline	Who can post on your timeline?	Friends	Edit
	Who can see what others post on your timeline?	Friends	Edit
	Allow others to share your posts to their stories?	On	Edit
	Hide comments containing certain words from your timeline	Off	Edit
Tagging	Who can see posts you're tagged in on your timeline?	Only me	Edit
	When you're tagged in a post, who do you want to add to the audience of the post if they can't already see it?	Only me	Edit
Review	Review posts you're tagged in before the post appears on your timeline?	Off	Edit
	Review tags people add to your posts before the tags appear on Facebook?	On	Edit

Privacy controls – 4

When you're tagged in a post, who do you want to add to the audience of the post if they can't already see it?

Close

They'll be able to see these posts in places like News Feed and search.

“Who do you want to add to the audience”...

Friends on Facebook? – 1

“My information is visible **only** to my friends...”

Friends on Facebook? – 2

- In 2011, the University of British Columbia (Canada) conducted an experiment to study how **friendships are created** on Facebook
- The research team implemented a **socialbot**
 - software agent simulating human behaviors
 - used to simulate a non-existing user
- The socialbot sent **friendship requests** to random users
 - who **did not know** the user simulated by the socialbot
- Two-step process: no friends in common, and friends of friends

Friends on Facebook? – 3

- The number of accepted requests is astonishing
 - ⇒ two users out of ten accepted the request from a totally unknown user
 - ⇒ if there were friends in common, six users out of ten accept
- Results in three weeks activity with 102 bots:
 - 3000 friends
 - 46500 e-mail addresses
 - 14500 physical addresses



... not only friends



Is it always the user's fault?

Ultimo aggiornamento: 23 settembre 2011



Se hai domande o reclami riguardo alla nostra politica sulla privacy, contattaci per posta all'indirizzo 1601 S. California Avenue, Palo Alto, CA 94304 (USA) o tramite questa [pagina del Centro assistenza](#).

- Privacy policies of Social Media are oftentimes ambiguous and not easy to understand
- Data protection is often **illusory** and easy to **circumvent**
- **The provider needs to be trusted.** The privacy policy can be changed overnight \implies users' data are no more under their control
 - December 2009: Facebook privacy policy changed overnight and default sharing option was **public**!

Facebook privacy policy evolution – 1

How will we notify you of changes to this policy?

We'll notify you before we make changes to this policy and give you the opportunity to review the revised policy before you choose to continue using our Products.

[Return to top](#)

How to contact Facebook with questions

You can learn more about how privacy works [on Facebook](#) and on [Instagram](#). If you have questions about this policy, you can contact us as described below.

Contact Us

The data controller responsible for your information is Facebook Ireland, which you can [contact online](#), or by mail at:

Facebook Ireland Ltd.
4 Grand Canal Square
Grand Canal Harbour
Dublin 2 Ireland

[Contact the Data Protection Officer for Facebook Ireland Ltd.](#)

You also have the right to lodge a complaint with Facebook Ireland's lead supervisory authority, the Irish Data Protection Commissioner, or your local supervisory authority.

Date of Last Revision: April 19, 2018

Facebook privacy policy evolution – 2

- 2005

No personal information that you submit to the Facebook will be available to any user of the Web Site who does not belong to at least one of the groups specified by you in your privacy settings.

- 2006

We understand you may not want everyone in the world to have the information you share on Facebook; that is why we give you control of your information. Our default privacy settings limit the information displayed in your profile to your school, your specified local area, and other reasonable community limitations that we tell you about.

- 2007

Profile information you submit to Facebook will be available to users of Facebook who belong to at least one of the networks you allow to access the information through your privacy settings (e.g., school, geography, friends of friends). Your name, school name, and profile picture thumbnail will be available in search results across the Facebook network unless you alter your privacy settings.

Facebook privacy policy evolution – 3

- November 2009

Facebook is designed to make it easy for you to share your information with anyone you want. You decide how much information you feel comfortable sharing on Facebook and you control how it is distributed through your privacy settings. You should review the default privacy settings and change them if necessary to reflect your preferences. You should also consider your settings whenever you share information. ...

Information set to “everyone” is publicly available information, may be accessed by everyone on the Internet (including people not logged into Facebook), is subject to indexing by third party search engines, may be associated with you outside of Facebook (such as when you visit other sites on the internet), and may be imported and exported by us and others without privacy limitations. The default privacy setting for certain types of information you post on Facebook is set to “everyone.” You can review and change the default settings in your privacy settings.

Facebook privacy policy evolution – 4

- December 2009

Certain categories of information such as your name, profile photo, list of friends and pages you are a fan of, gender, geographic region, and networks you belong to are considered publicly available to everyone, including Facebook-enhanced applications, and therefore do not have privacy settings. You can, however, limit the ability of others to find this information through search using your search privacy settings.

- April 2010

When you connect with an application or website it will have access to General Information about you. The term General Information includes your and your friends' names, profile pictures, gender, user IDs, connections, and any content shared using the Everyone privacy setting. ... The default privacy setting for certain types of information you post on Facebook is set to "everyone." ... Because it takes two to connect, your privacy settings only control who can see the connection on your profile page. If you are uncomfortable with the connection being publicly available, you should consider removing (or not making) the connection.

Facebook privacy policy evolution – 5

- December 2010

If your friend connects with an application or website, it will be able to access your name, profile picture, gender, user ID, and information you have shared with "everyone." It will also be able to access your connections, except it will not be able to access your friend list. If you have already connected with (or have a separate account with) that website or application, it may also be able to connect you with your friend on that application or website. If the application or website wants to access any of your other content or information (including your friend list), it will have to obtain specific permission from your friend. If your friend grants specific permission to the application or website, it will generally only be able to access content and information about you that your friend can access. In addition, it will only be allowed to use that content and information in connection with that friend.

Facebook privacy policy evolution – 6

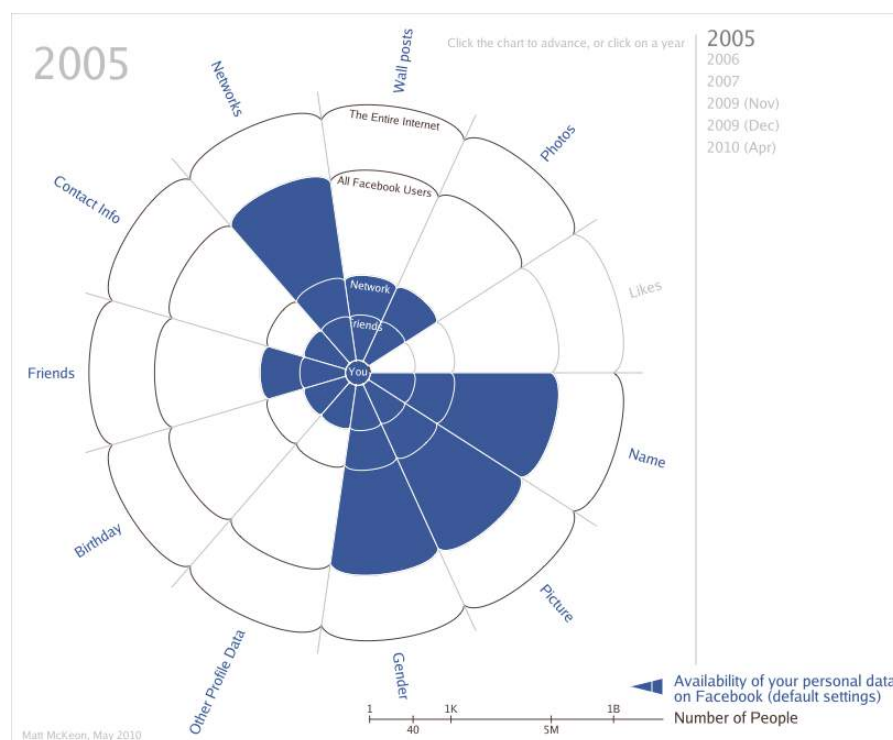
- September 2011

Choosing to make your information public also means that this information can be associated with you (i.e., your name, profile picture, Facebook profile, User ID, etc.) even off Facebook; can show up when someone does a search on Facebook or on a [public search engine](#); will be accessible to the games, applications, and websites you and your friends use; will be accessible to anyone who uses our APIs such as our Graph API. Sometimes you will not be able to select an audience when you post something (like when you write on a Page's wall or comment on a news article that uses our comments plugin). This is because [some types of posts are always public posts](#). When others share information about you, they can also choose to make it public.

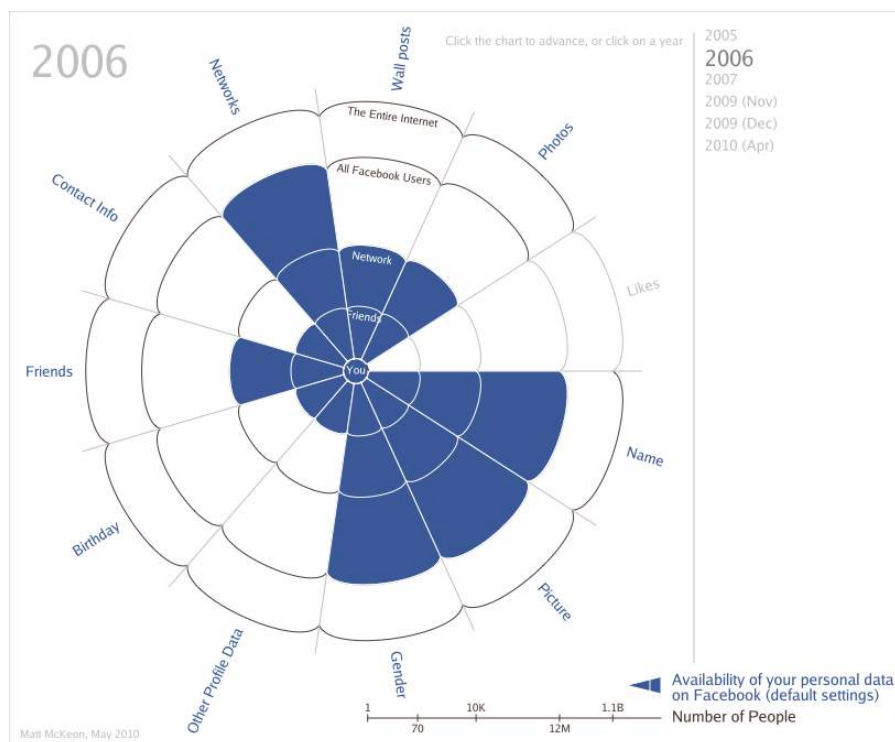
- December 2012

Your trust is important to us, which is why [we don't share information we receive about you with others unless we have: received your permission; given you notice, such as by telling you about it in this policy; or removed your name or any other personally identifying information from it.](#)

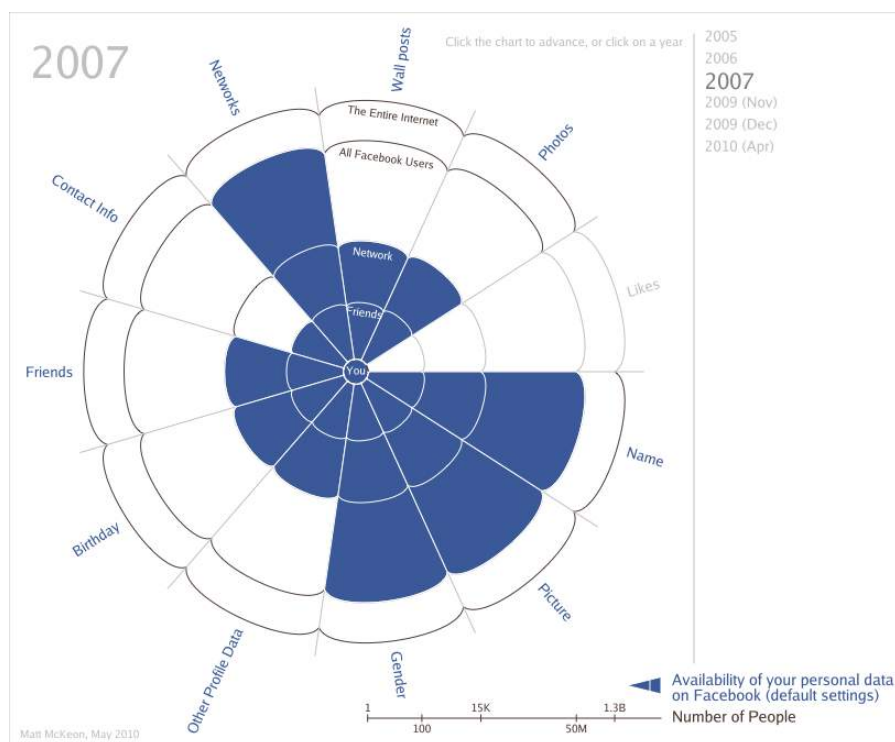
Data availability on Facebook – 1



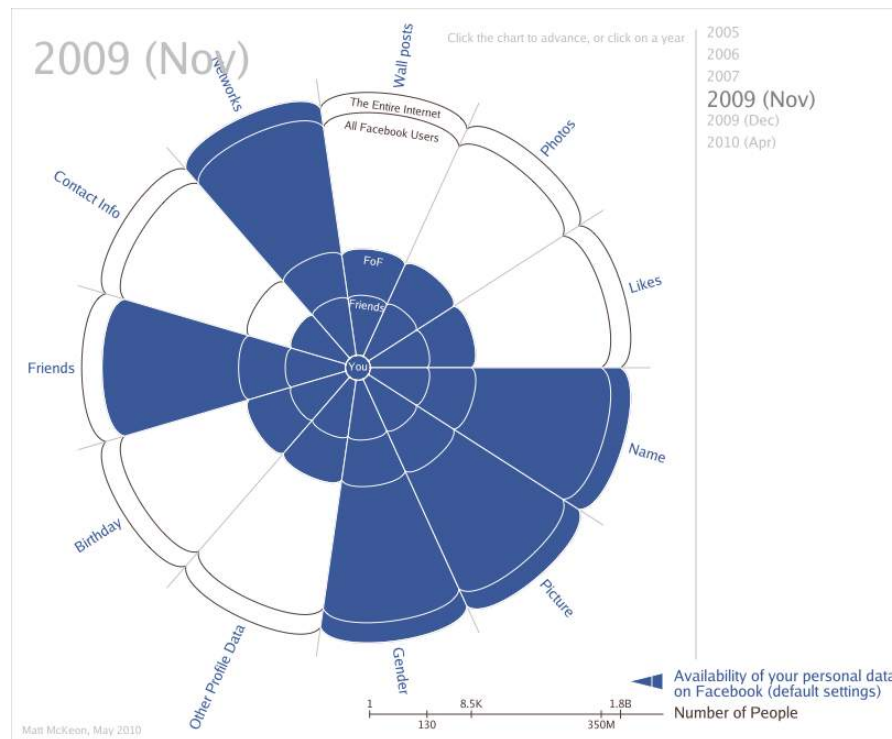
Data availability on Facebook – 2



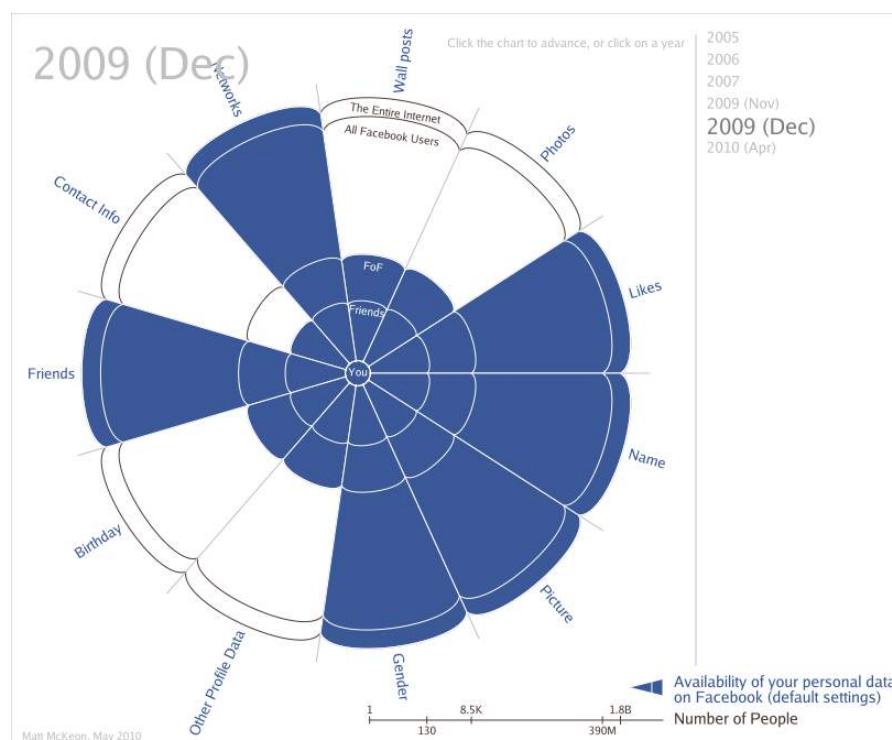
Data availability on Facebook – 3



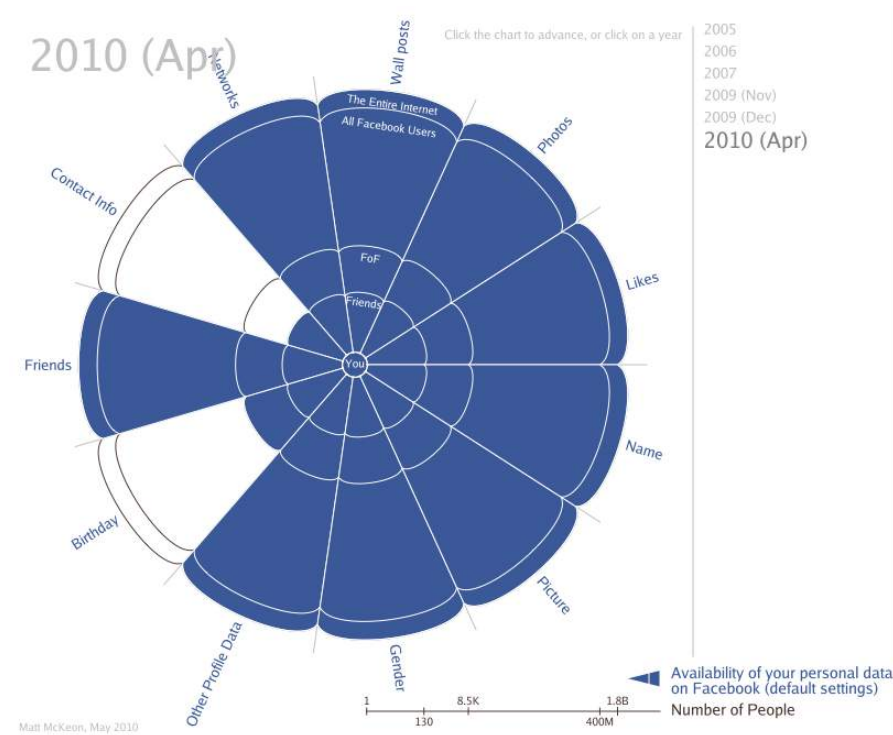
Data availability on Facebook – 4



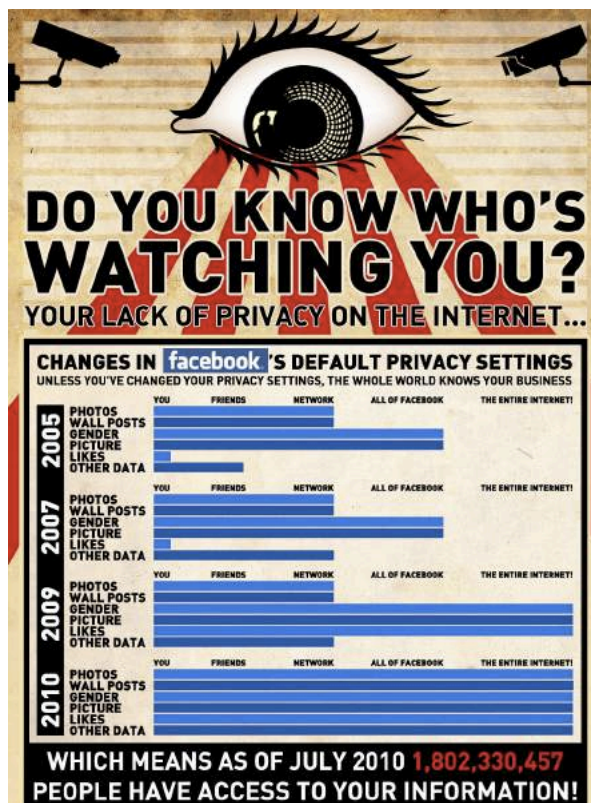
Data availability on Facebook – 5



Data availability on Facebook – 6



My information!



July 2010 \Rightarrow 1,802,330,457 people had the possibility to access **all** information ever published on Facebook!

(<http://www.privacylives.com/2010/07/>)

False sense of security

- Blocking visibility over friends list is **unidirectional**
- Installing a **third-party application** gives to it **access** to users' data
- Restricting visibility over **tagged pictures** only to me \implies they are hidden in my profile, but are **still visible** from other parts of the Social Media (e.g., friends' wall)
- Facebook hides the **page of a picture** rather than the picture itself \implies still possible to access it via its URL
- Deleted resources remain **accessible** via their URLs for **some time**
- ...

Let's delete unnecessary information then?

What happens to content (posts, pictures) that I delete from Facebook?

[Computer Help](#) [Mobile Help](#) ▾

[Share Article](#)

When you choose to delete something you shared on Facebook, we remove it from the site. Some of this information is permanently deleted from our servers; however, some things can only be deleted when you [permanently delete your account](#).

User profiling – 1

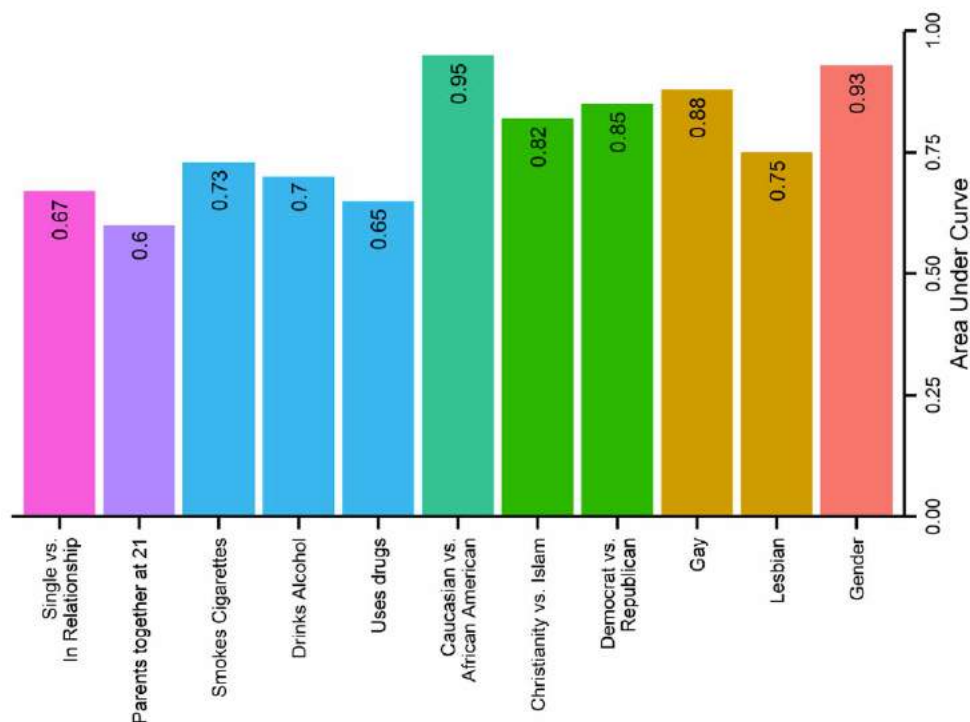
- In usage examples illustrated previously, users can at least discover (even if too late) that their data have been used
- But data can also be used to harm, profile, or manipulate a user without her being even aware



User profiling – 2

- Once a data item is published on Facebook, control over it is lost and it is impossible to know how it will be used
- A 2013 peer-reviewed study showed that Facebook likes can be used to automatically and accurately predict a range of sensitive personal attributes including
 - sexual orientation
 - ethnicity
 - religious and political views
 - personality traits
 - intelligence
 - happiness
 - use of addictive substances
 - parental separation
 - age and gender

User profiling – 3



[M. Kosinski, D. Stillwell, T. Graepel, "Digital records of behavior expose personal traits," PNAS, 110 (15) 5802-5805]

User profiling – 4

- A given message delivered to a certain individual can be more or less effective based on how it is presented
- In marketing and advertising, messages are typically tailored w.r.t. the recipients to maximize their impact
 - informational communication → behavioural communication
- Traditional tailoring has been based on demographics and geographics
- Tailoring based on psychographics has been showed to be more effective
 - personality drives behaviour
 - behaviour drives individual actions, such as voting...

User profiling – 5

- OCEAN model
 - **Openness**: do you enjoy new experiences?
 - **Conscientiousness**: do you prefer plans and order?
 - **Extraversion**: how social you are?
 - **Agreeableness**: do you value others' needs and society?
 - **Neuroticism**: how much do you tend to worry?
- Example: how to nuance a message aimed to defend Amendment II of the US Constitution
 - to a conscientious individual with high neuroticism: a picture of a burglary, with claim "The Second Amendment isn't just a right – It's an insurance policy. Defend the right to bear arms!"
 - to a close and agreeable individual: a picture of a father teaching hunting to his son, with claim "From father to son, since the birth of our Nation. Defend the second amendment."

User profiling – 6

- The availability of Big Data and sophisticated analytical tools makes easier the prediction of the personality of individuals
- Political consulting companies have used these tools in several electoral races
- e.g., Cambridge Analytica (behavioral microtargeting for election processes) for Ted Cruz Primary Campaign in 2016, with notable results

Facebook–Cambridge Analytica data scandal – 1

- Early 2018: it was revealed the Cambridge Analytica had harvested personal data of millions of people's Facebook profiles without their consent and used it for political purposes
- Major political scandal resulting also in a massive fall in Facebook's stock price

Facebook–Cambridge Analytica data scandal – 2



Facebook–Cambridge Analytica data scandal – 3

- In 2013 a researcher from Cambridge University created a [personality quiz app](#)
- The app was [installed by 300.000 Facebook users](#) who gave permission for accessing their data...
- ...but the app was also [collecting data of those users' friends](#)
- At the end, the app retrieved data from [87M \(!!\)](#) Facebook users

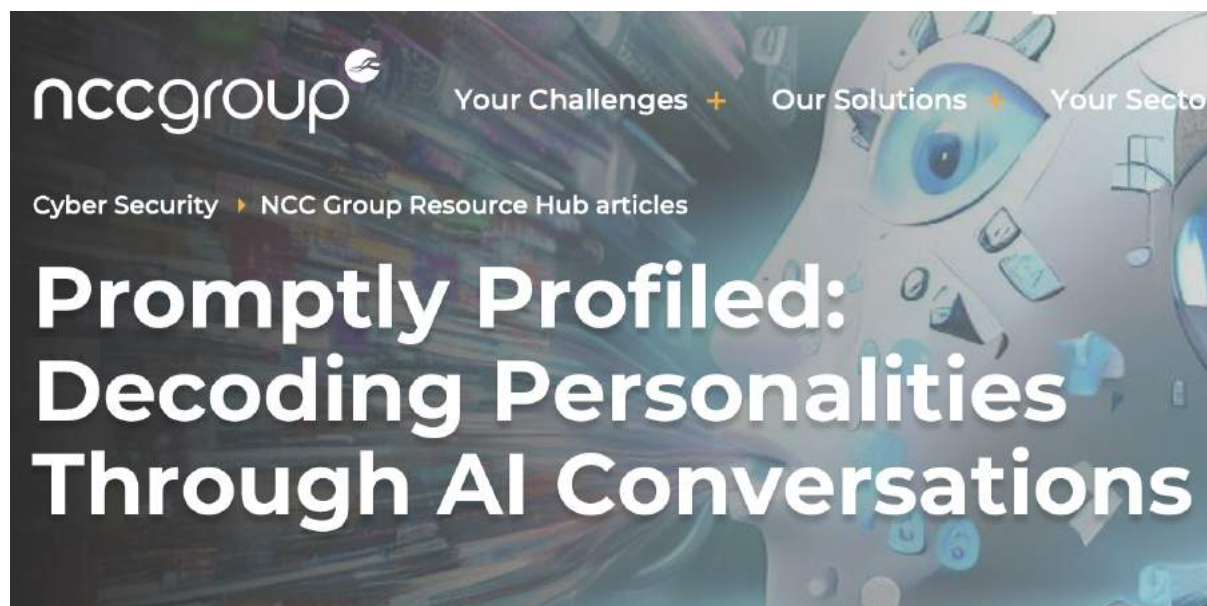
Facebook–Cambridge Analytica data scandal – 3

- The retrieved data were [shared with Cambridge Analytica](#), which used them to create [psychographical profiles](#) of the respondents
 - [segmenting individuals based on their personality](#)
- Profile creation was done by [analysis of](#)
 - [personality test](#) of the 300.000 users
 - [collected Facebook data](#) (e.g., likes) for the remaining ones
- Results of the personality test were [matched to the Facebook activity](#) (e.g., likes) of the same users to [create a map](#)
- The same map has been [applied \(in reverse\)](#) to the remaining users to determine their [personality](#)

Facebook–Cambridge Analytica data scandal – 4

- Profiles allowed for identifying the **most effective** kind of **advertisement** to **persuade** a given individual for a political event
- According to a whistleblower, used in **different political campaigns** including
 - 2016 US presidential elections
 - 2016 Brexit vote
 - ...

... and the future?



the rest is yet to come!



"Before I write my name on the board, I'll need to know how you're planning to use that data."