EECS 388

Introduction to Computer Security

Lecture 20:

Privacy and Anonymity

Nov 21, 2023

Prof. Ensafi

What is Privacy?

Privacy is . . . a complex philosophical and cultural concept.

Seclusion **The right to be let alone.** — *Louis Brandeis (1890)*
Right to *choose seclusion* and limit scrutiny in private settings.

Limits **The ability to limit access.**
Ability to *participate in society* without others collecting data about you.

Control **Control over information.**
Ability to choose when, how, and *what information about you is shared*.

Secrecy **The option of secrecy.** — *Richard Posner (1983)*
Right to *conceal information* that others might use to your disadvantage.

Liberty **A prerequisite for political liberty.**
Essential condition for free speech, association, personal autonomy.

Part 1: Privacy vs. Business

Businesses **track and analyze people's behavior** in order to:

- Deliver targeted advertising
- Recommend products to buy
- Develop better products
- Perform differential pricing
- Obtain competitive intelligence
- Make investment decisions
- Influence public opinion (e.g., political microtargeting)
- Make decisions about you (e.g., employment, insurance, credit)

Companies learn about you through:

- **Data you give them directly** (e.g. search queries, Facebook profile)
- **Data they passively collect** (e.g. geolocation, online behavior)
- **Data they buy from others** (e.g. consumer profiles)

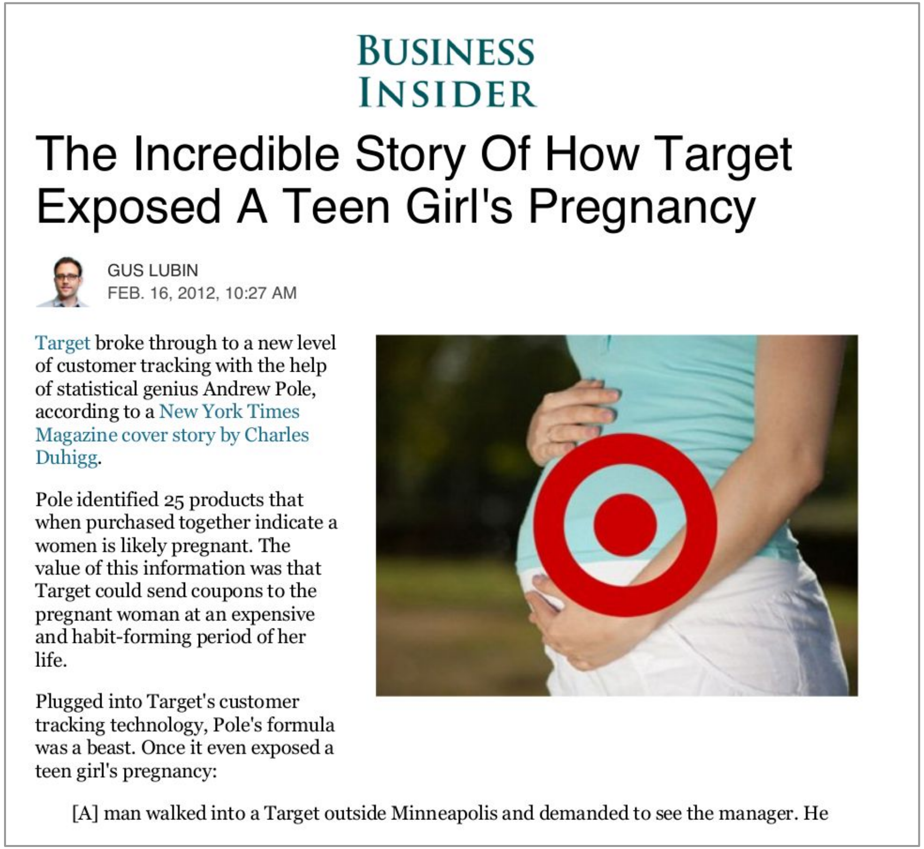
Data collection, resale, aggregation, and use mostly occur without people's knowledge.

Many people find these practices creepy, as they can violate several privacy notions.

Seclusion	**Limits**	**Control**	**Secrecy**	**Liberty**

Tracking Risks: Target

- Using ground truth from customers who signed up for its baby-shower registry, **Target** developed a model to predict when women are pregnant based on their pattern of purchases.
- They mailed a teenage girl baby-related coupons.
- Her dad saw them and fumed at Target that she wasn't pregnant… but she was!



BUSINESS INSIDER

The Incredible Story Of How Target Exposed A Teen Girl's Pregnancy

GUS LUBIN
FEB. 16, 2012, 10:27 AM

Target broke through to a new level of customer tracking with the help of statistical genius Andrew Pole, according to a New York Times Magazine cover story by Charles Duhigg.

Pole identified 25 products that when purchased together indicate a women is likely pregnant. The value of this information was that Target could send coupons to the pregnant woman at an expensive and habit-forming period of her life.

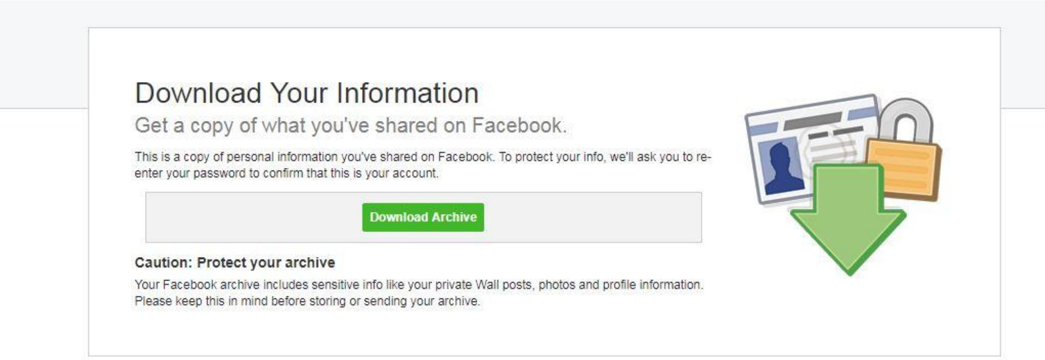
Plugged into Target's customer tracking technology, Pole's formula was a beast. Once it even exposed a teen girl's pregnancy:

> [A] man walked into a Target outside Minneapolis and demanded to see the manager. He

Direct Sharing: Facebook

What kinds of personal data have you given Facebook?

Find out: Use the Download Your Information function.



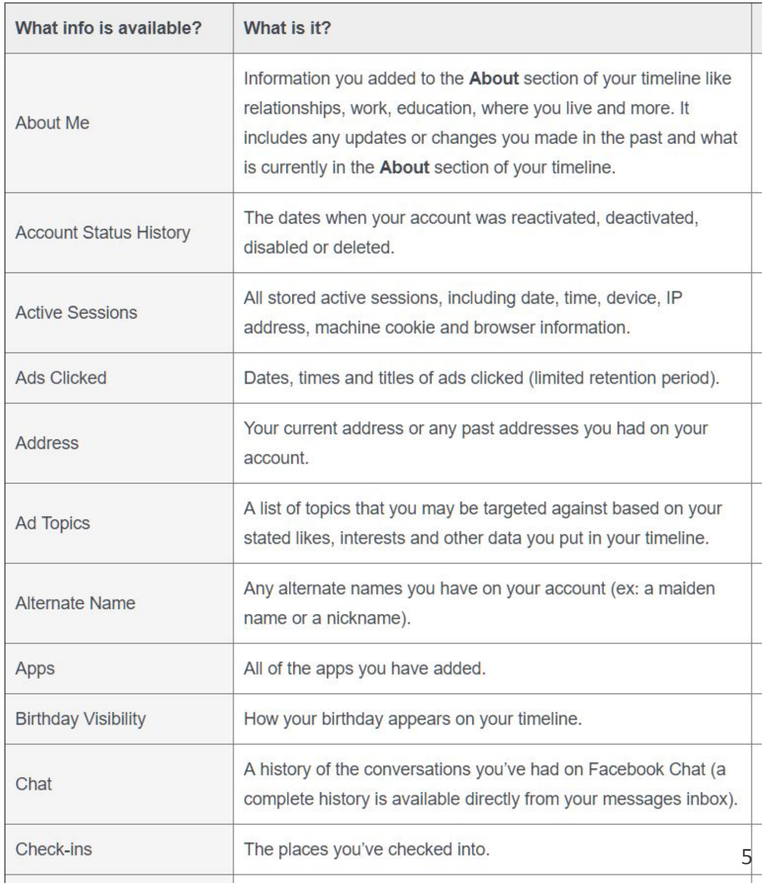
Download Your Information

Get a copy of what you've shared on Facebook.

This is a copy of personal information you've shared on Facebook. To protect your info, we'll ask you to re-enter your password to confirm that this is your account.

Caution: Protect your archive

Your Facebook archive includes sensitive info like your private Wall posts, photos and profile information. Please keep this in mind before storing or sending your archive.

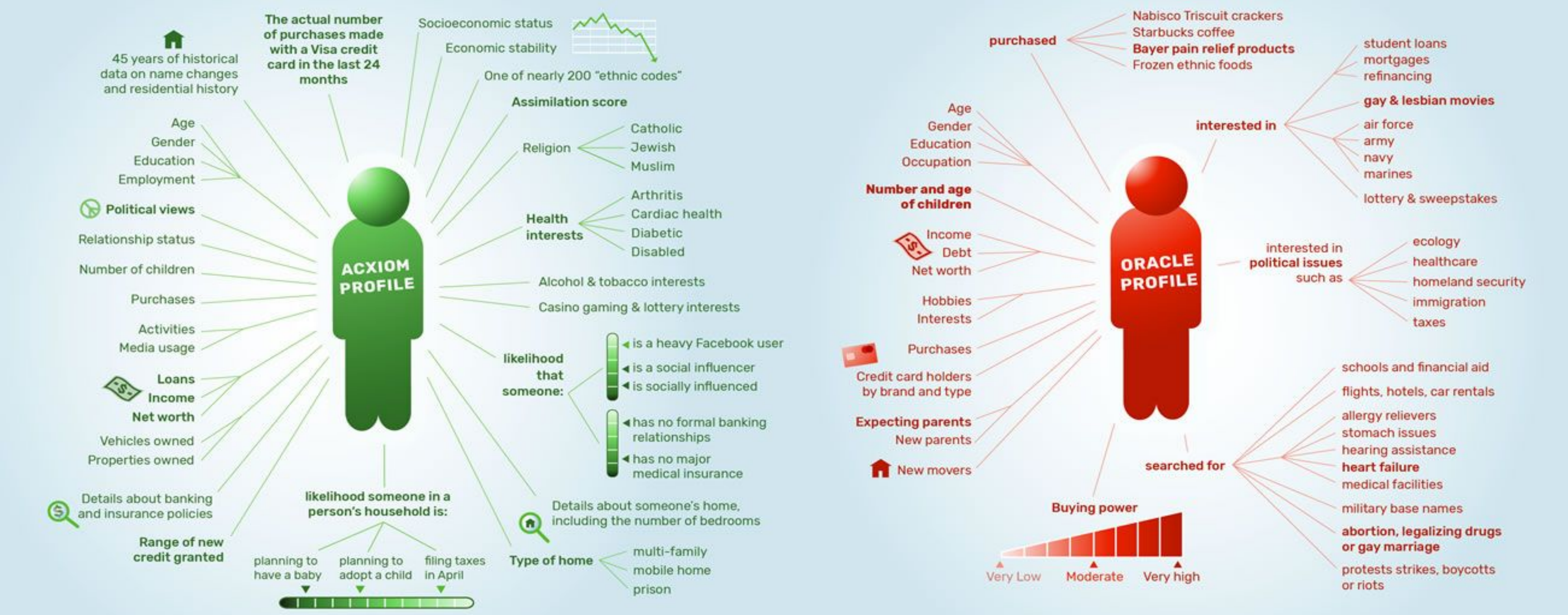


What info is available?	What is it?
About Me	Information you added to the **About** section of your timeline like relationships, work, education, where you live and more. It includes any updates or changes you made in the past and what is currently in the **About** section of your timeline.
Account Status History	The dates when your account was reactivated, deactivated, disabled or deleted.
Active Sessions	All stored active sessions, including date, time, device, IP address, machine cookie and browser information.
Ads Clicked	Dates, times and titles of ads clicked (limited retention period).
Address	Your current address or any past addresses you had on your account.
Ad Topics	A list of topics that you may be targeted against based on your stated likes, interests and other data you put in your timeline.
Alternate Name	Any alternate names you have on your account (ex: a maiden name or a nickname).
Apps	All of the apps you have added.
Birthday Visibility	How your birthday appears on your timeline.
Chat	A history of the conversations you've had on Facebook Chat (a complete history is available directly from your messages inbox).
Check-ins	The places you've checked into.

Data Aggregators and Brokers

Data brokers such as Acxiom and Oracle aggregate data sources to create detailed profiles about billions of people, which they sell to businesses and governments.

> *“our goal is to connect every piece of data with every person on the planet with every available use case that matters … if we accomplish that amazing things happen “*→ Scott Howe CEO Acxiom



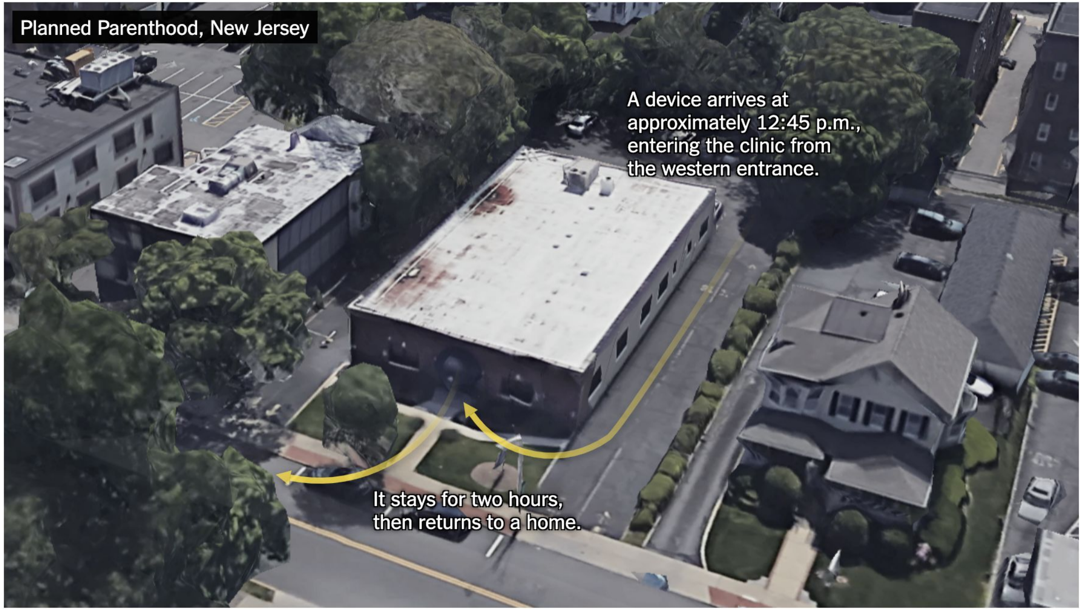
Location Tracking

https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html?module=inline

Weather channel apps bought by IBM! The Wall Street Journal put the deal at over $2 billion.

Your Apps Know Where You Were Last Night, and They're Not Keeping It Secret

Dozens of companies use smartphone locations to help advertisers and even hedge funds. They say it's anonymous, but the data shows how personal it is.



Planned Parenthood, New Jersey

A device arrives at approximately 12:45 p.m., entering the clinic from the western entrance.

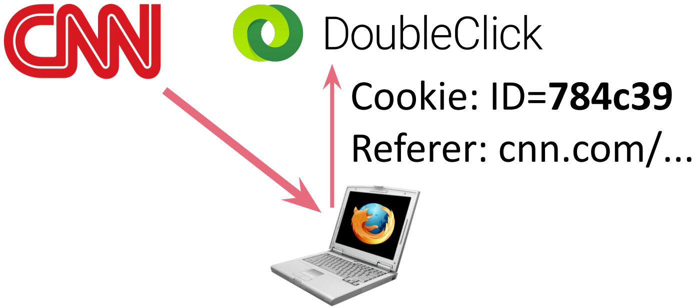
It stays for two hours, then returns to a home.

Tracking Mechanisms: Third-Party Cookies

When a page loads a resource (image, script, iframe) from another site, the browser sends the cookies set by that **third-party origin**.

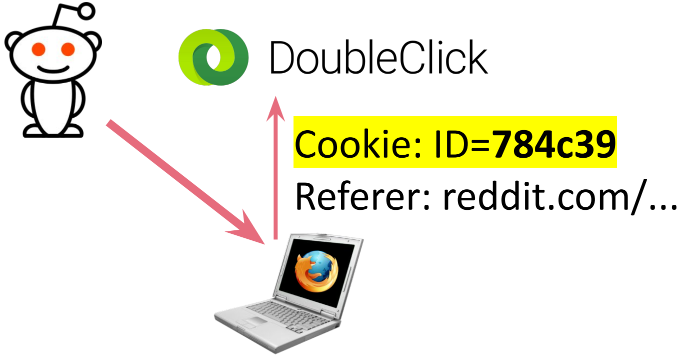
Example: 1. Doubleclick.com sets a **tracking cookie** in your browser when you first load one of its ads.

2. Your browser sends the **same cookie** each time any site loads any ad from Doubleclick.com.



Cookie: ID=**784c39**
Referer: cnn.com/...

Doubleclick learns that the browser with this tracking ID visited a particular story at CNN.com.



Cookie: ID=**784c39**
Referer: reddit.com/...

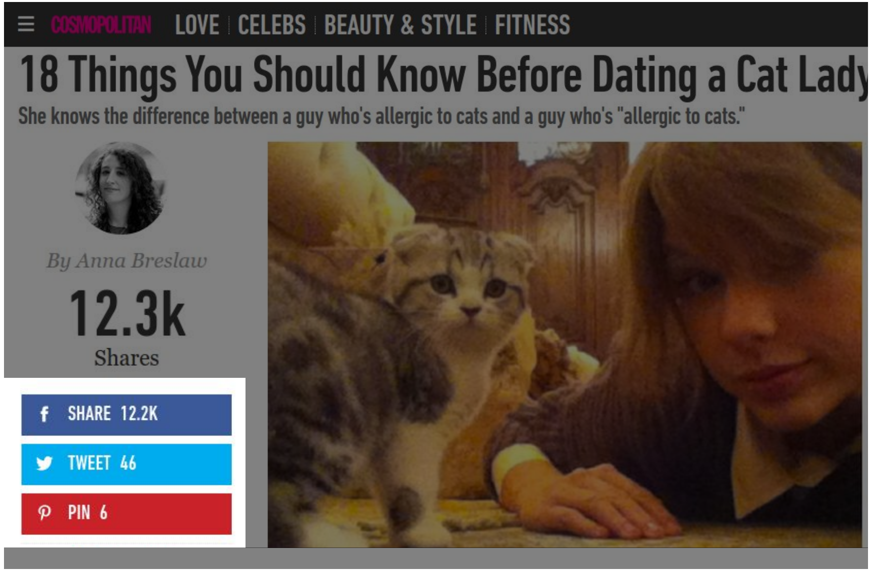
3. When you visit another site with an ad from Doubleclick.com, **same cookie** is sent again. Used to build a profile of the pages you view.

4. If the site knows demographics about you (name, email, age, spending habits, etc.), they can sell to Doubleclick to associate with cookie.

5. Doubleclick uses the entire profile to target ads or could sell it to other data aggregators.

Tracking Mechanisms: Third-Party Cookies

Social media “Like” buttons send third-party cookies when they load. This lets these sites track what pages you visit, **even if you don’t click the buttons**.

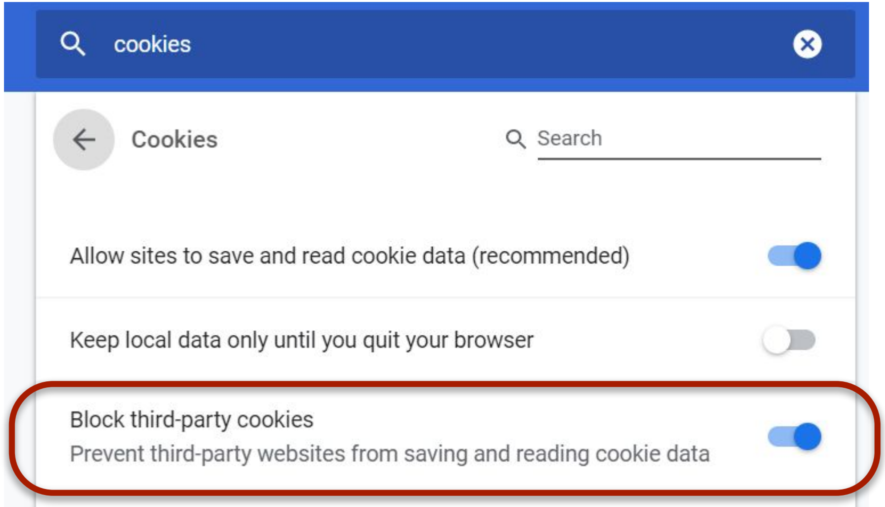


18 Things You Should Know Before Dating a Cat Lady

She knows the difference between a guy who's allergic to cats and a guy who's "allergic to cats."

By Anna Breslaw

12.3k Shares



Allow sites to save and read cookie data (recommended) ●

Keep local data only until you quit your browser ○

Block third-party cookies
Prevent third-party websites from saving and reading cookie data ●

Mitigation: Configure your browser to **disable third-party cookies**.

Deprecating Third-Party Cookies

Apple and Google Are Killing the (Ad) Cookie. Here’s Why

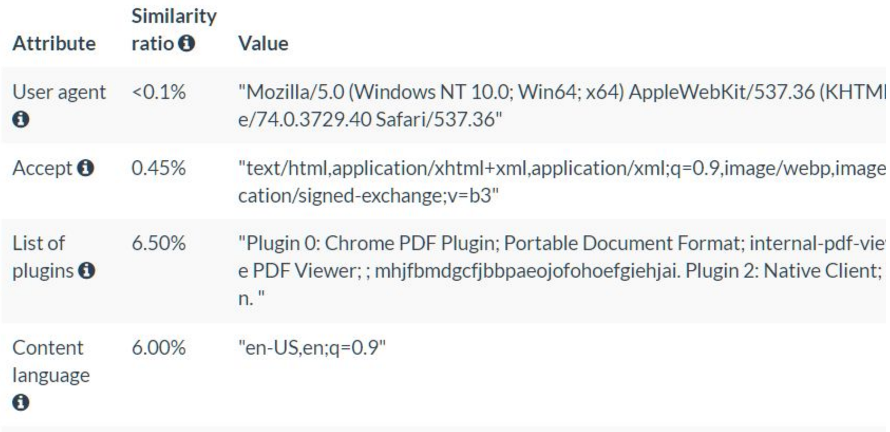
FORBES > SMALL BUSINESS > ENTREPRENEURS

The Slow Death Of Third-Party Cookies

Browser Fingerprinting

Even without cookies, sites can use **browser fingerprinting** to track you.

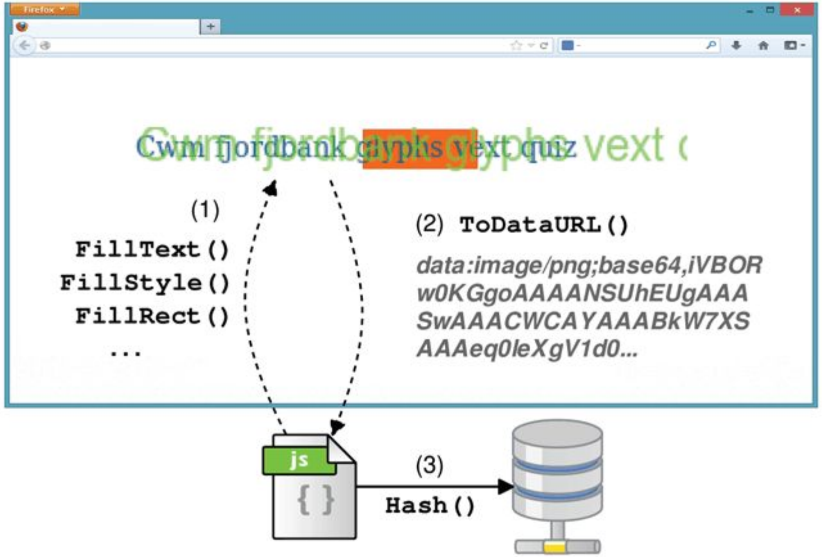
Pages can read hundreds of attributes about your computer configuration. Taken together, they are often globally unique.



Attribute	Similarity ratio ℹ	Value
User agent ℹ	<0.1%	"Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTMI e/74.0.3729.40 Safari/537.36"
Accept ℹ	0.45%	"text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image cation/signed-exchange;v=b3"
List of plugins ℹ	6.50%	"Plugin 0: Chrome PDF Plugin; Portable Document Format; internal-pdf-vie e PDF Viewer; ; mhjfbmdgcfjbbpaeojofohoefgiehjai. Plugin 2: Native Client; n. "
Content language ℹ	6.00%	"en-US,en;q=0.9"

AmIUnique.org tests whether your browser stands out.

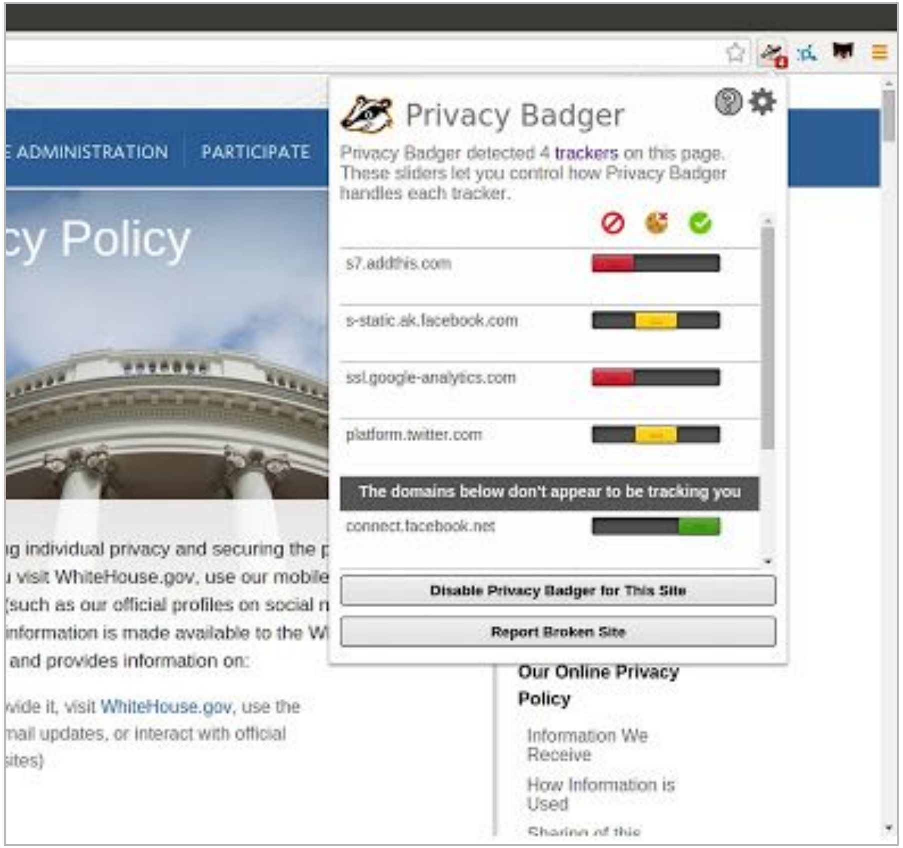
Example: Canvas fingerprinting is a browser fingerprinting method that uses the HTML <canvas> element to detect subtle rendering differences caused by different GPUs and graphics driver versions.



Defending Against Tracking

Privacy Badger is a browser extension created by EFF that uses **behavioral detection** to identify trackers.

Servers that it detects attempting third-party tracking or canvas fingerprinting are added to a blocklist.



Privacy Badger

Privacy Badger detected 4 trackers on this page. These sliders let you control how Privacy Badger handles each tracker.

s7.addthis.com	🟥
s-static.ak.facebook.com	🟨
ssl.google-analytics.com	🟥
platform.twitter.com	🟨

The domains below don't appear to be tracking you

connect.facebook.net	🟩

g individual privacy and securing the p
u visit WhiteHouse.gov, use our mobile
(such as our official profiles on social n
information is made available to the W
and provides information on:

vide it, visit WhiteHouse.gov, use the
nail updates, or interact with official
sites)

Our Online Privacy Policy

Information We Receive

How Information is Used

Sharing of this

Data Anonymization: A Hard Problem

One approach to protecting privacy is **remove identifying information** from datasets you collect or share through a process of **data anonymization**.

Naive approach:

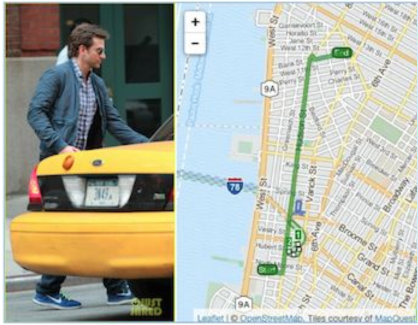
Strip names, addresses, birthdays, etc.

Often not sufficient!

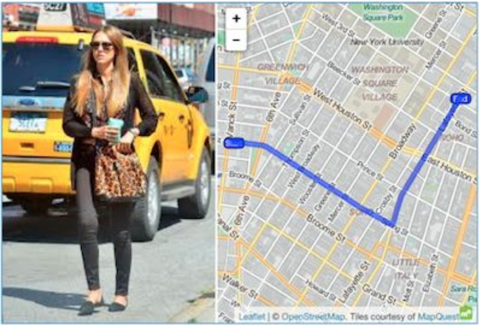
Example: NYC published a database of taxi rides for research purposes [what could go wrong?]

- hash of taxi driver number
- pick up latitude, longitude, and time
- drop off latitude, longitude, and time
- number of passengers

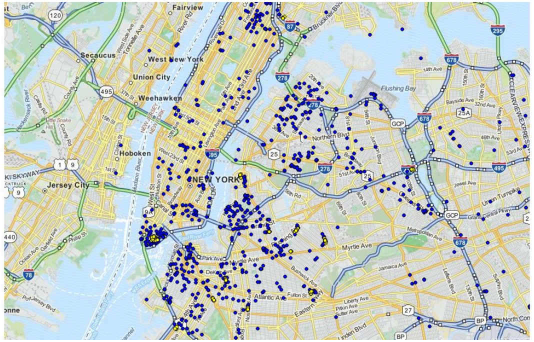
Tracking celebrities who people spotted getting into a cab at a particular place and time.



Bradley Cooper



Jessica Alba



Addresses of pickups for rides that ended at strip clubs.

k-anonymity

Stronger way to anonymize a dataset: ensure each record is **indistinguishable** from at least $k - 1$ other records. Called ***k*-anonymity**.

Must remove **quasi-identifiers** (e.g., zip code, date of birth, gender) or sufficiently reduce their granularity (e.g., use partial zip codes, remove birth month/day).

Many caveats:

- Which attributes form a quasi-identifier?
- What auxiliary information the attacker has?
- Protects **identity disclosure**, not **attribute disclosure** (can learn things about a person other than which record is theirs)
- Tradeoff between privacy and utility.

Example: 3-anonymized table

ZIP	DOD	Disease
902**	1987	Cancer
902**	1987	Cancer
902**	1987	Cancer
902**	1954	Heart Disease
902**	1954	GI Disease
902**	1954	Died of poison apple
904**	1978	Heart Disease
904**	1978	Cancer
904**	1978	Cancer

K-anonymity example details

- We suppose ZIP and DOD are quasi-identifiers, so we ensure that there are at least 3 values with each (ZIP,DOD) pair.
- But with sufficient aux info, disease can be part of a quasi-identifier: suppose attacker knows Alan Turing was the only person in the population to die of a poison apple in 1954. Then his record is identified.
- Attribute disclosure: Suppose you know someone in the population died in 1987. Although you don’t know which record was theirs, you learn they had of cancer.
- Can’t use data to spot cancer cluster within a particular zip code, since we had to reduce their granularity.

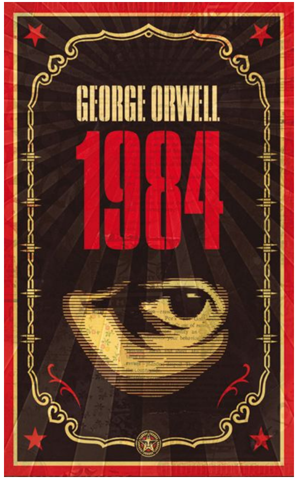
Example: 3-anonymized table

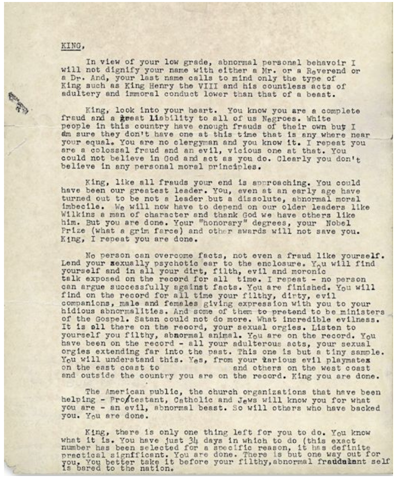
ZIP	DOD	Disease
902**	1987	Cancer
902**	1987	Cancer
902**	1987	Cancer
902**	1954	Heart Disease
902**	1954	GI Disease
902**	1954	Died of poison apple
904**	1978	Heart Disease
904**	1978	Cancer
904**	1978	Cancer

Part 2: Privacy vs. Government

Governments **track and analyze people's behavior** in order to:

- Enforce the law
- Protect the public
- Safeguard national security
- Improve services
- Gain military or economic advantage
- Suppress dissent
- Make decisions about you
  (e.g., security clearance, PreCheck)





KING,

In view of your low grade, abnormal personal behavoir I will not dignify your name with either a Mr. or a Reverend or a Dr. And, your last name calls to mind only the type of King such as King Henry the VIII and his countless acts of adultery and immoral conduct lower than that of a beast.

King, look into your heart. You know you are a complete fraud and a great liability to all of us Negroes. White people in this country have enough frauds of their own but I am sure they don't have one at this time that is any where near your equal. You are no clergyman and you know it. I repeat you are a colossal fraud and an evil, vicious one at that. You could not believe in God and act as you do. Clearly you don't believe in any personal moral principles.

King, like all frauds your end is approaching. You could have been our greatest leader. You, even at an early age have turned out to be not a leader but a dissolute, abnormal moral imbecile. We will now have to depend on our older leaders like Wilkins a man of character and thank God we have others like him. But you are done. Your "honorary" degrees, your Nobel Prize (what a grim farce) and other awards will not save you. King, I repeat you are done.

No person can overcome facts, not even a fraud like yourself. Lend your sexually psychotic ear to the enclosure. You will find yourself and in all your dirt, filth, evil and moronic talk exposed on the record for all time. I repeat - no person can argue successfully against facts. You are finished. You will find on the record for all time your filthy, dirty, evil companions, male and females giving expression with you to your hidious abnormalities. And some of them to pretend to be ministers of the Gospel. Satan could not do more. What incredible evilness. It is all there on the record, your sexual orgies. Listen to yourself you filthy, abnormal animal. You are on the record. You have been on the record - all your adulterous acts, your sexual orgies extending far into the past. This one is but a tiny sample. You will understand this. Yes, from your various evil playmates on the east coast to and others on the west coast and outside the country you are on the record. King you are done.

The American public, the church organizations that have been helping - Pro~~s~~testant, Catholic and Jews will know you for what you are - an evil, abnormal beast. So will others who have backed you. You are done.

King, there is only one thing left for you to do. You know what it is. You have just 34 days in which to do (this exact number has been selected for a specific reason, it has definite practical signfficant. You are done. There is but one way out for you. You better take it before your filthy, abnormal fraudulent self is bared to the nation.

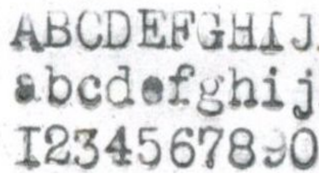
Our relations with companies are often voluntary, with government rarely so.

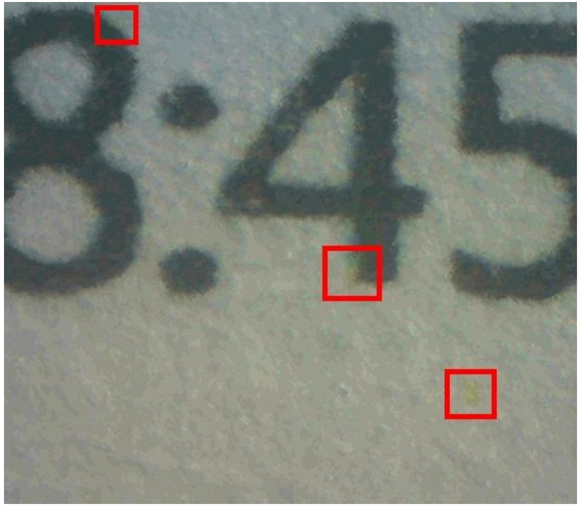
History and literature abound with examples of government spying abuses.

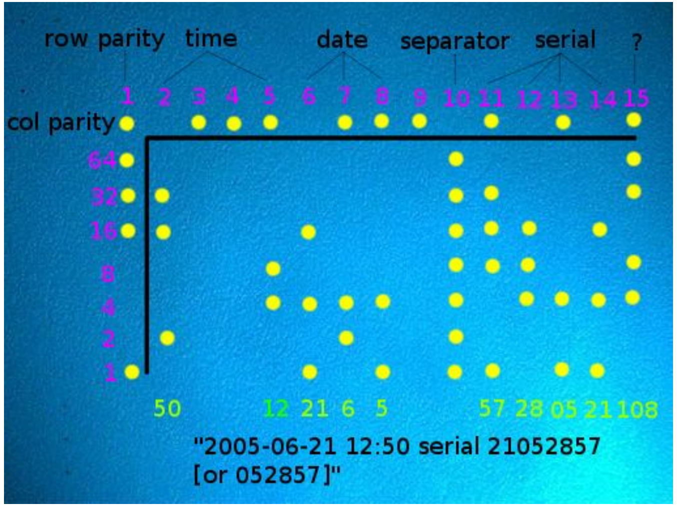
Seclusion **Limits** **Control** **Secrecy** **Liberty**

Tracking Physical Documents

In communist Romania and East Germany, **typewriters had to be registered with the government**, and samples kept on file. This allowed subversive documents to be traced to the owners.



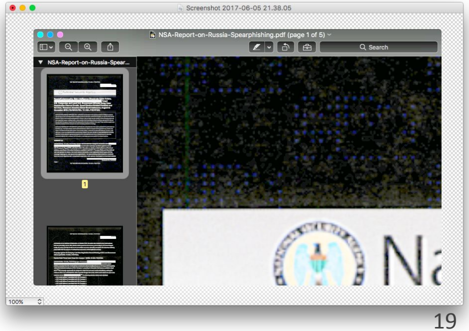




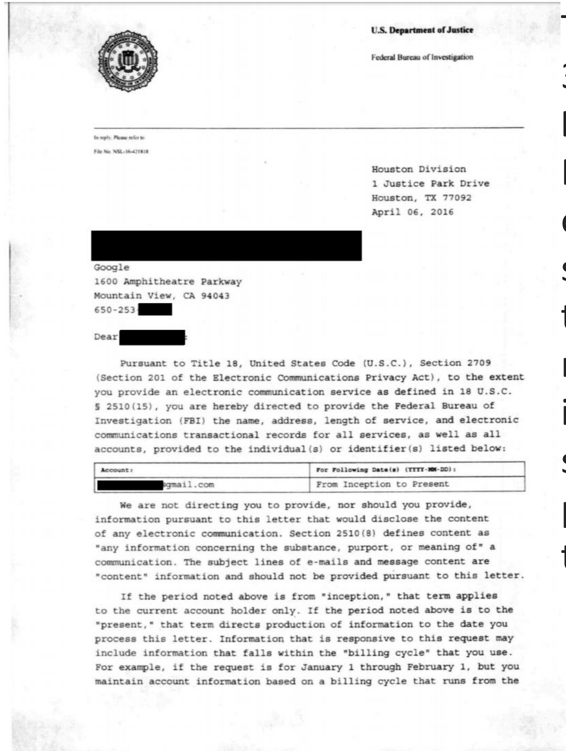
"2005-06-21 12:50 serial 21052857 [or 052857]"

Today, most color printers produce a **Machine Identification Code (MIC)**, a pattern of faint yellow dots on every page. Visible under UV light, this watermark encodes a timestamp and device serial number.

MIC on NSA documents published by journalists in 2017 may have led to the leaker's prosecution.


Screenshot 2017-06-05 21.38.05
NSA-Report-on-Russia-Spearphishing.pdf (page 1 of 5)
Search
NSA-Report-on-Russia-Spear...
1
100%
Na

National Security Orders



U.S. Department of Justice

Federal Bureau of Investigation

In reply, Please refer to

File No. NSL-16-421818

Houston Division
1 Justice Park Drive
Houston, TX 77092
April 06, 2016

[redacted]
Google
1600 Amphitheatre Parkway
Mountain View, CA 94043
650-253-[redacted]

Dear [redacted]:

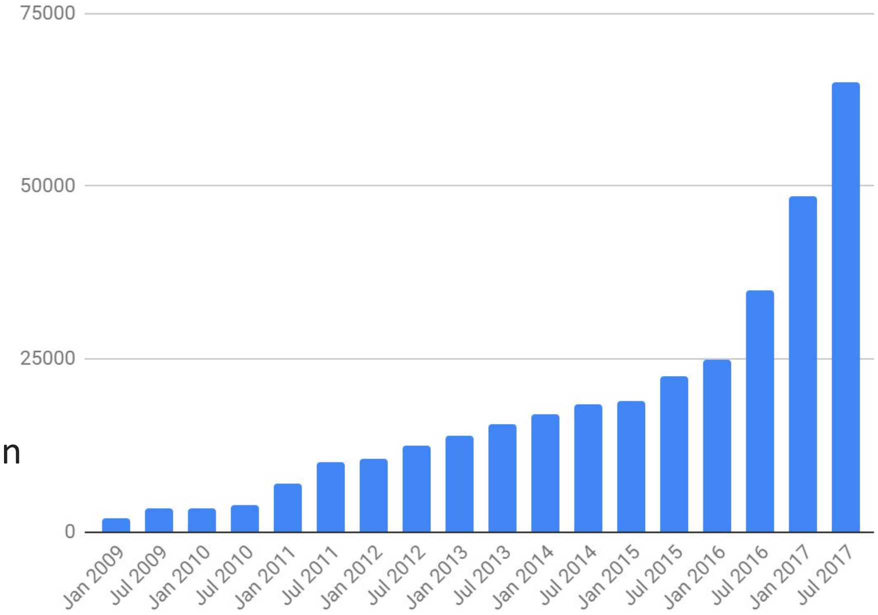
Pursuant to Title 18, United States Code (U.S.C.), Section 2709 (Section 201 of the Electronic Communications Privacy Act), to the extent you provide an electronic communication service as defined in 18 U.S.C. § 2510(15), you are hereby directed to provide the Federal Bureau of Investigation (FBI) the name, address, length of service, and electronic communications transactional records for all services, as well as all accounts, provided to the individual(s) or identifier(s) listed below:

Account:	For Following Date(s) (YYYY-MM-DD):
[redacted]@gmail.com	From Inception to Present

We are not directing you to provide, nor should you provide, information pursuant to this letter that would disclose the content of any electronic communication. Section 2510(8) defines content as "any information concerning the substance, purport, or meaning of" a communication. The subject lines of e-mails and message content are "content" information and should not be provided pursuant to this letter.

If the period noted above is from "inception," that term applies to the current account holder only. If the period noted above is to the "present," that term directs production of information to the date you process this letter. Information that is responsive to this request may include information that falls within the "billing cycle" that you use. For example, if the request is for January 1 through February 1, but you maintain account information based on a billing cycle that runs from the

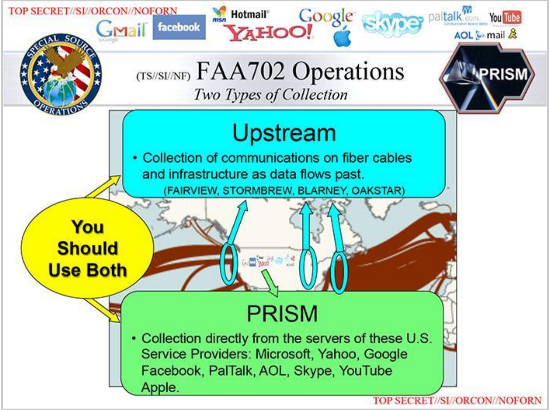
The FBI has issued over 300,000 **National Security Letters (NSLs)** since 2005. NSLs compel a network operator to provide subscriber identity + transaction records relevant to investigations into matters of national security. Providers are often prohibited from disclosing that they received an NSL.



Google users whose data or communications content were provided to the U.S. government under secret orders from the **Foreign Intelligence Surveillance Court**.

Source: Google Transparency Report

Mass Surveillance


Gmail
facebook
msn Hotmail
YAHOO!
Google
skype
paltalk.com
YouTube
AOL mail
PRISM

TOP SECRET//SI//ORCON//NOFORN

(TS//SI//NF) FAA702 Operations

Two Types of Collection

Upstream

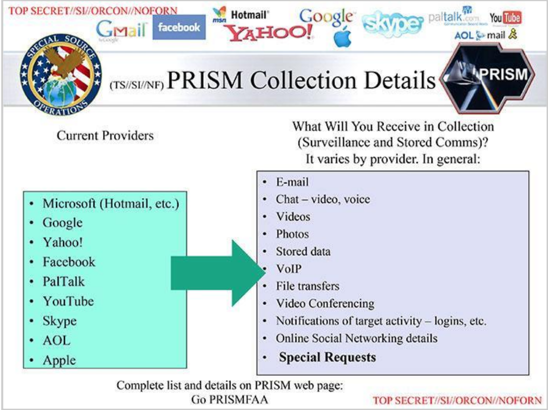
- Collection of communications on fiber cables and infrastructure as data flows past. (FAIRVIEW, STORMBREW, BLARNEY, OAKSTAR)

You Should Use Both

PRISM

- Collection directly from the servers of these U.S. Service Providers: Microsoft, Yahoo, Google Facebook, PalTalk, AOL, Skype, YouTube Apple.

TOP SECRET//SI//ORCON//NOFORN


Gmail
facebook
msn Hotmail
YAHOO!
Google
skype
paltalk.com
YouTube
AOL mail
PRISM

TOP SECRET//SI//ORCON//NOFORN

(TS//SI//NF) PRISM Collection Details

Current Providers

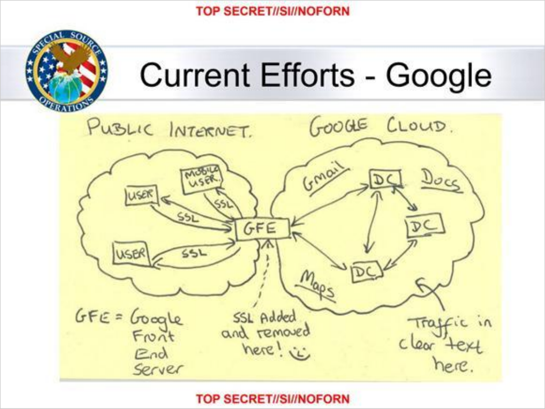
- Microsoft (Hotmail, etc.)
- Google
- Yahoo!
- Facebook
- PalTalk
- YouTube
- Skype
- AOL
- Apple

What Will You Receive in Collection (Surveillance and Stored Comms)? It varies by provider. In general:

- E-mail
- Chat – video, voice
- Videos
- Photos
- Stored data
- VoIP
- File transfers
- Video Conferencing
- Notifications of target activity – logins, etc.
- Online Social Networking details
- **Special Requests**

Complete list and details on PRISM web page:
Go PRISMFAA

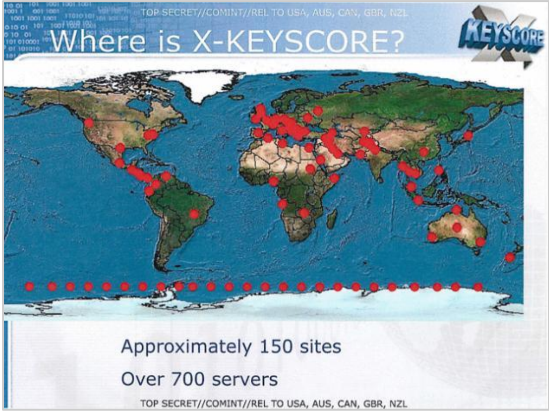
TOP SECRET//SI//ORCON//NOFORN


PUBLIC INTERNET.
GOOGLE CLOUD.
MOBILE USER
USER
USER
SSL
SSL
SSL
GFE
Gmail
DC
Docs
DC
DC
Maps
GFE = Google Front End Server
SSL Added and removed here! ☺
Traffic in clear text here.

TOP SECRET//SI//NOFORN

Current Efforts - Google

TOP SECRET//SI//NOFORN



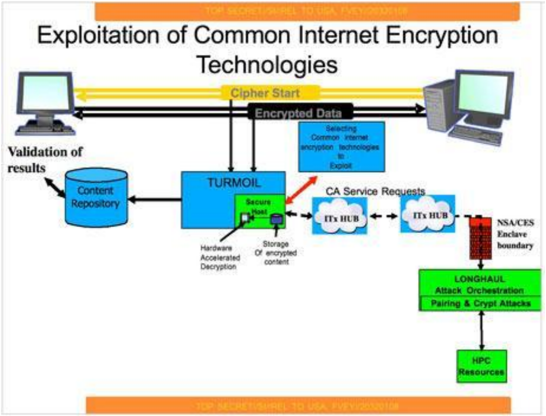
TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL

Where is X-KEYSCORE?

Approximately 150 sites

Over 700 servers

TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL

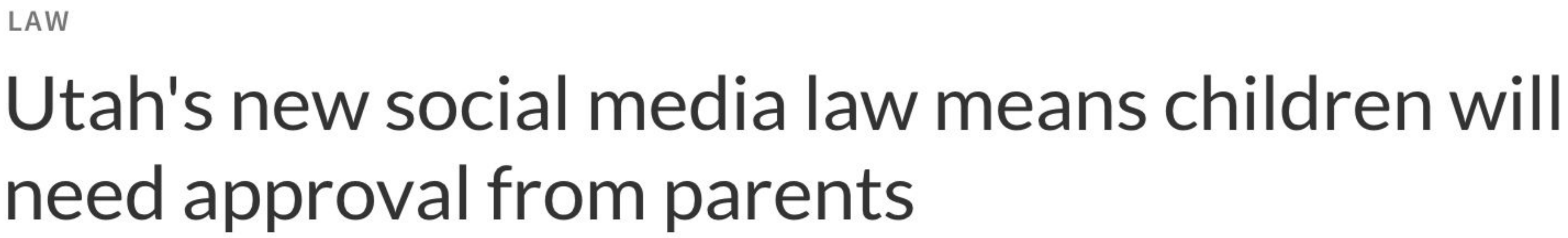


TOP SECRET//SI//REL TO USA, FVEY//20320108

Exploitation of Common Internet Encryption Technologies

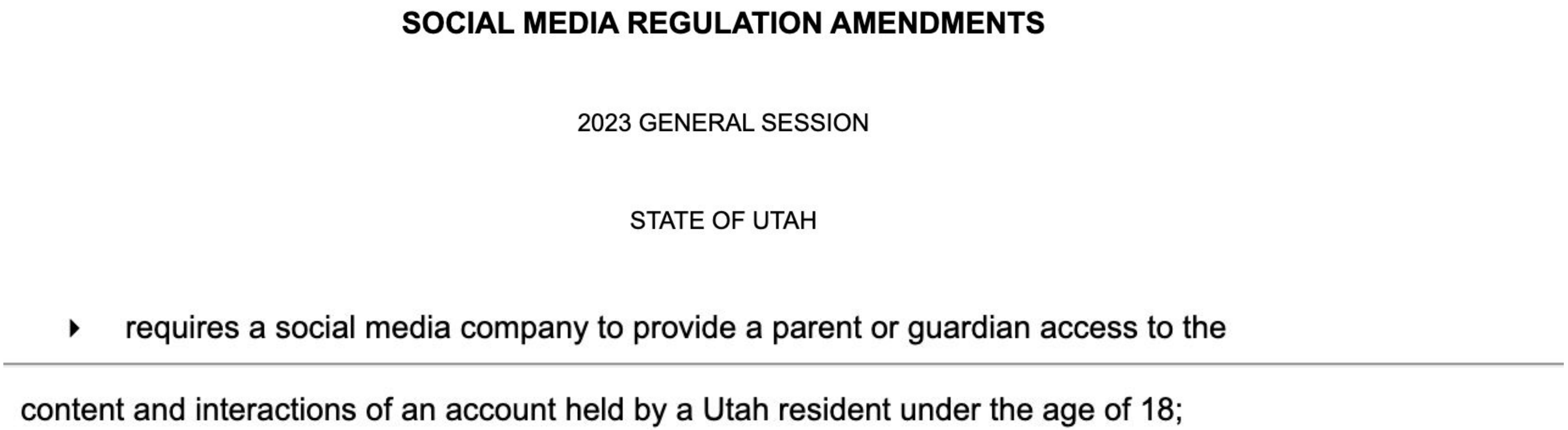
TOP SECRET//SI//REL TO USA, FVEY//20320108

Government vs. Privacy in Social Media



LAW

Utah's new social media law means children will need approval from parents



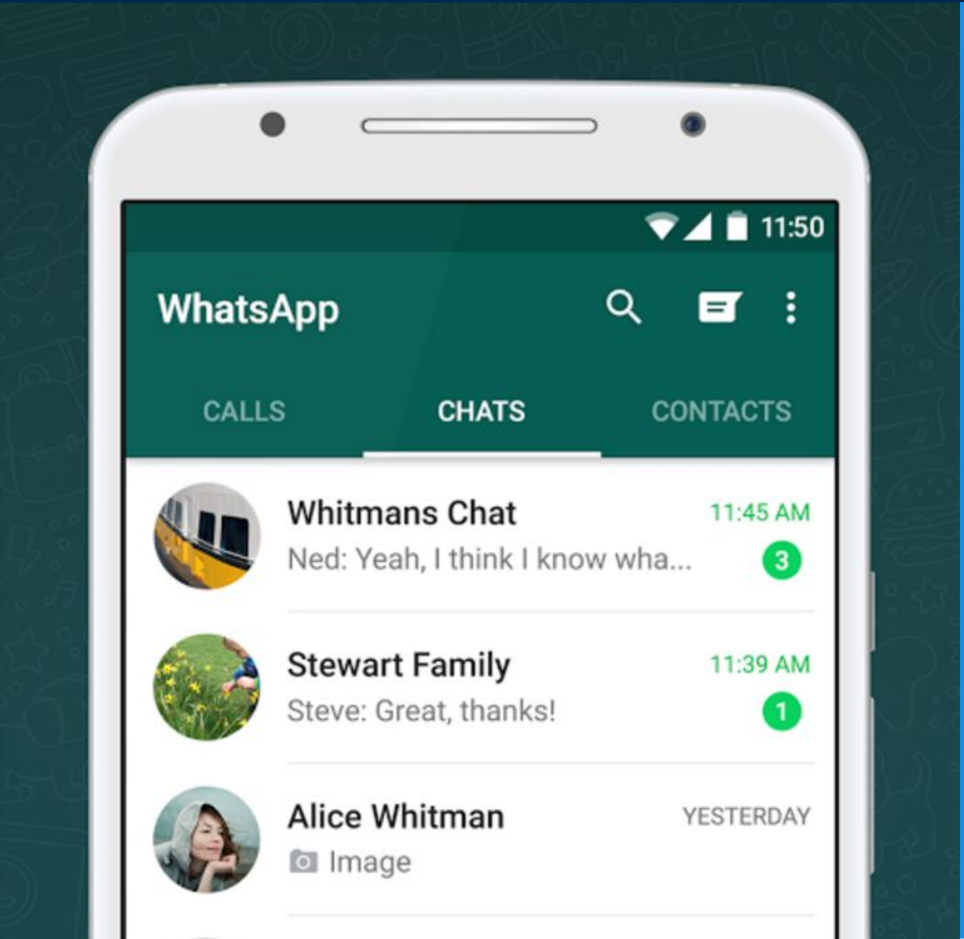
SOCIAL MEDIA REGULATION AMENDMENTS

2023 GENERAL SESSION

STATE OF UTAH

- requires a social media company to provide a parent or guardian access to the content and interactions of an account held by a Utah resident under the age of 18;

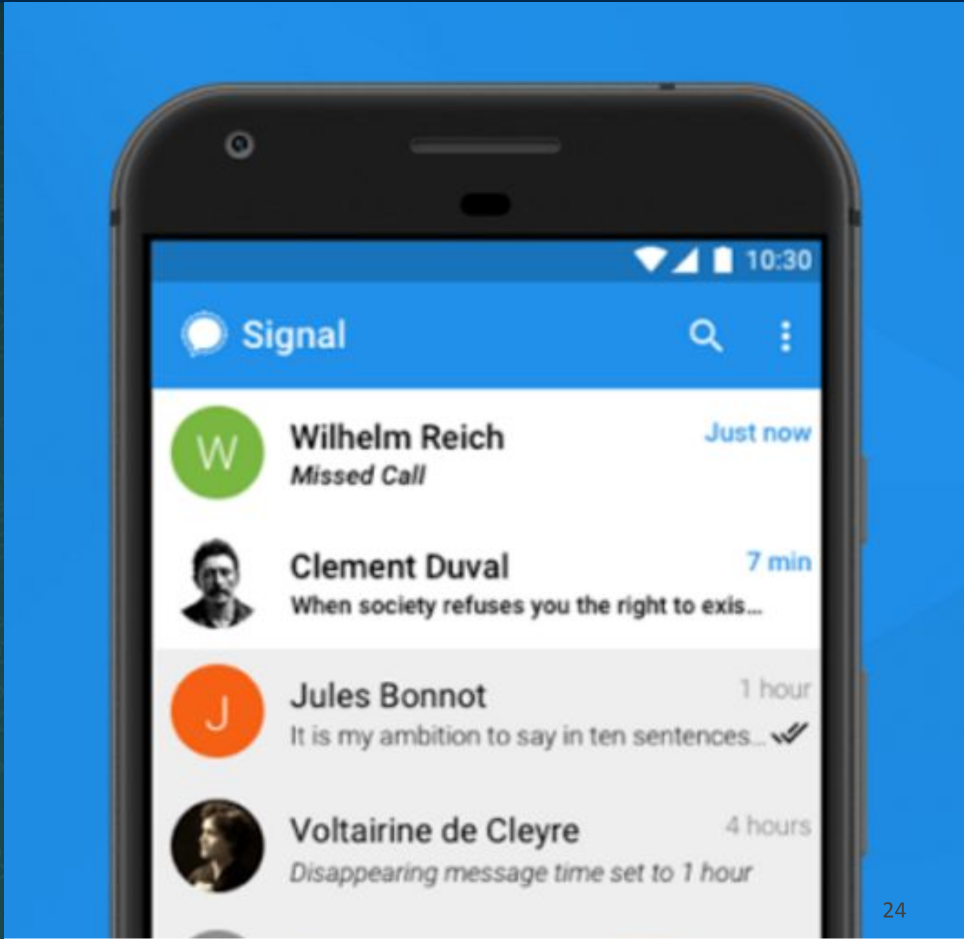
Off-the-Record (OTR) Messaging's Descendants



Whitmans Chat 11:45 AM
Ned: Yeah, I think I know wha... 3

Stewart Family 11:39 AM
Steve: Great, thanks! 1

Alice Whitman YESTERDAY
Image



Wilhelm Reich Just now
Missed Call

Clement Duval 7 min
When society refuses you the right to exis...

Jules Bonnot 1 hour
It is my ambition to say in ten sentences...

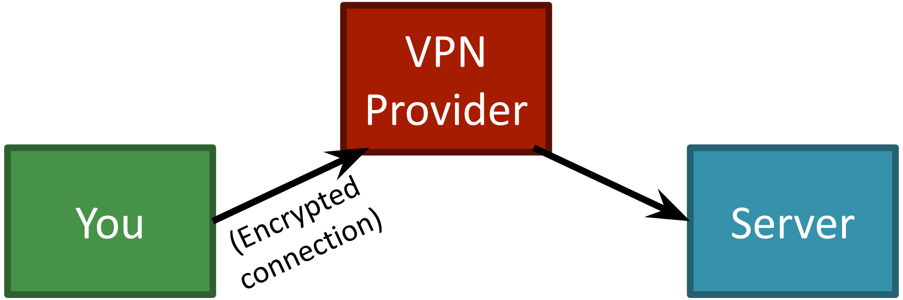
Voltairine de Cleyre 4 hours
Disappearing message time set to 1 hour

Anonymous Networking

Anonymous networking techniques provide the ability to communicate online where the identity of the source and/or destination host are concealed.

> Note: *Confidentiality* is about content, whereas anonymity is about identity.

Internet anonymity is challenging, because every TCP connection must have a valid source and destination IP address.



Naive approach: Tunnel connection through a **Virtual Private Network (VPN)** or **proxy server**.

- Your traffic appears to originate at the VPN, so **Server** doesn't learn your real IP address.
- Your connection to the VPN is encrypted, so **your ISP** doesn't know the Server's address.
- But the **VPN provider** knows who you are and where you're going! Can be dishonest, hacked, or compelled to compromise anonymity.

Anonymous Networking: Tor

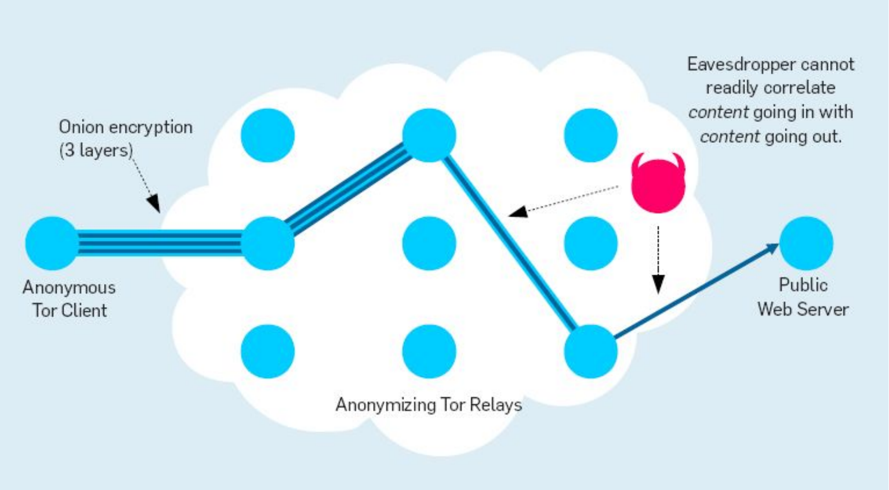
Better approach: Tor

A distributed anonymity service based on a network of volunteer-run relay servers located all over the world.

Connection travels over a randomly selected path of three relays: **entry node**, **middle node**, and **exit node**.

Onion-routing technique uses nested encrypted tunnels. Each node along path removes the outer layer of encryption and forwards the contents.

No Tor relay knows both the client's address and the destination address.



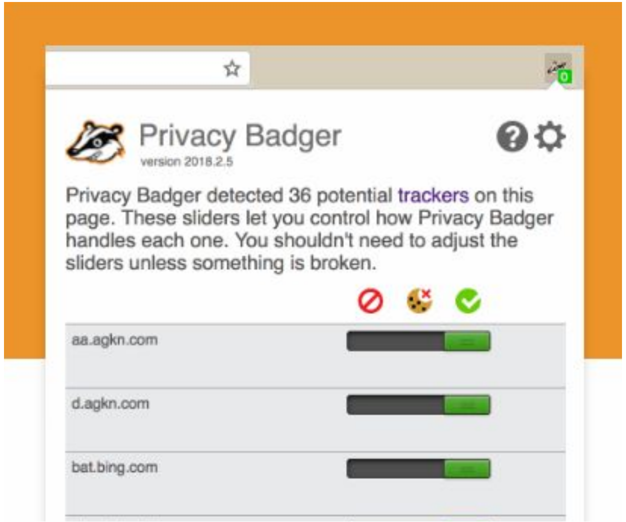
Entry node: knows client's address and identity of middle node, but not destination.

Exit node: knows a Tor client is connecting to destination, doesn't know client's address.

Destination: knows a Tor user is connecting.

Try It Yourself!

Experiment with these privacy tools and defenses:



Privacy Badger

version 2018.2.5

Privacy Badger detected 36 potential trackers on this page. These sliders let you control how Privacy Badger handles each one. You shouldn't need to adjust the sliders unless something is broken.

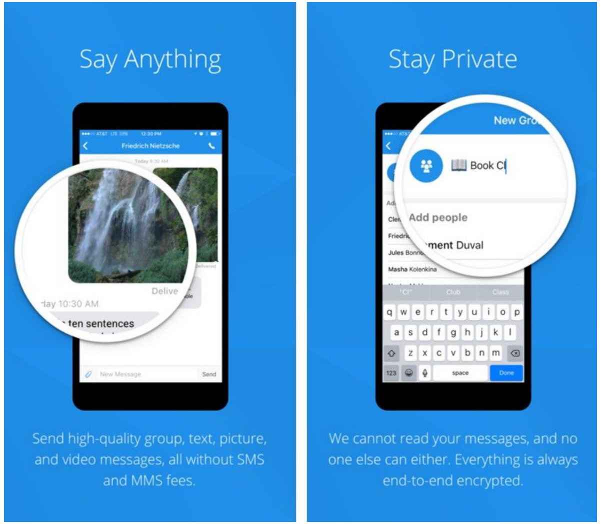
aa.agkn.com

d.agkn.com

bat.bing.com

Privacy Badger

https://eff.org/privacybadger



Say Anything

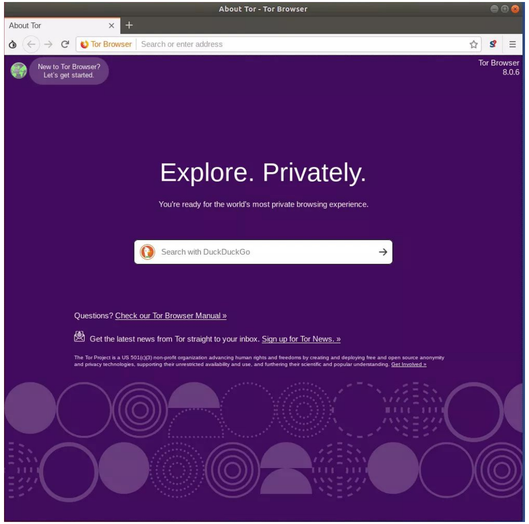
Send high-quality group, text, picture, and video messages, all without SMS and MMS fees.

Stay Private

We cannot read your messages, and no one else can either. Everything is always end-to-end encrypted.

Signal

https://signal.org



About Tor - Tor Browser

About Tor

Tor Browser | Search or enter address

New to Tor Browser? Let's get started.

Tor Browser 8.0.6

Explore. Privately.

You're ready for the world's most private browsing experience.

Search with DuckDuckGo

Questions? Check our Tor Browser Manual »

Get the latest news from Tor straight to your inbox. Sign up for Tor News. »

The Tor Project is a US 501(c)(3) non-profit organization advancing human rights and freedoms by creating and deploying free and open source anonymity and privacy technologies, supporting their unrestricted availability and use, and furthering their scientific and popular understanding. Get Involved »

Tor Browser

https://torproject.org

Coming Up

Reminders:

Lab 5 due 6 p.m at Nov 30

Forensics Project due 6 p.m at Dec. 7

Thursday

Thursday, Nov. 23

No lecture

Thanksgiving break

Tuesday, Nov. 28

Only via Zoom

Side Channels

Guest Lecture by Andrew Kwong