The Basics of

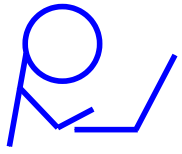
Data Security

Fabian M. Suchanek

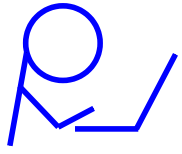
based on "A practical guide to Internet security"

Data

work

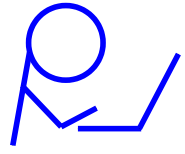


study

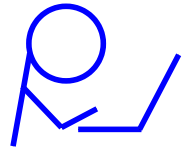


plan vacation

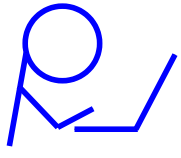
make photo album



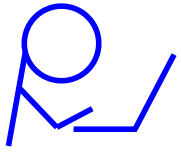
watch movie



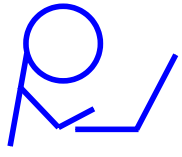
write letter



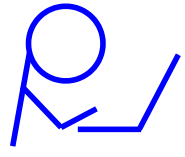
chat with friends



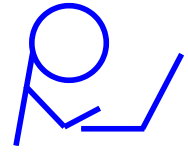
phone with friends



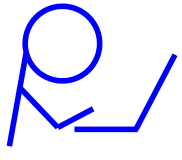
write a book



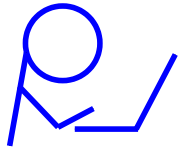
be creative



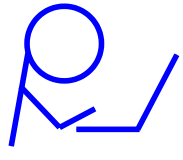
write diary



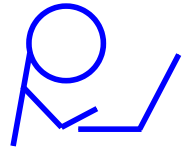
be politically active



play games



learn language



listen to music

Overview

Protecting data against

- yourself
- hackers
- evil interlocutors
- companies
- governments

Who is the biggest enemy to ur data?

hackers

viruses/ransomware

evil governments

big companies

data leaks

Who is the biggest enemy to ur data?

hackers

viruses/ransomware

evil governments

big companies

data leaks

you yourself

What can happen to your data

- theft
- loss
- hazards
- decay
- accidental deletion
- laptop failure

© Tim Gee

```
Hard Disk Error

Please run the Hard Disk Test in System Diagnostics
```

Dear Professors,

Yesterday night my Macbook Pro was stolen. Meaning my projects also. In attachement i put you a proof of the police statement.

2018-02-17

Def: Ransomware

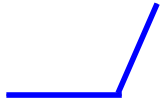
A ransomware is a malicious software program that makes your data unusable by encrypting it, and requests a ransom (=money) to decrypt it.

Cryptolocker: 325m USD paid

WannaCry: 150,000 USD paid, 4b USD in damages

Backup your data!

All important data should live in at least 2 different places.



+

some other place

Solution 1: Cloud Service

A cloud service automatically backs up your data.

(do not copy your data to the cloud service folder, move it there!)

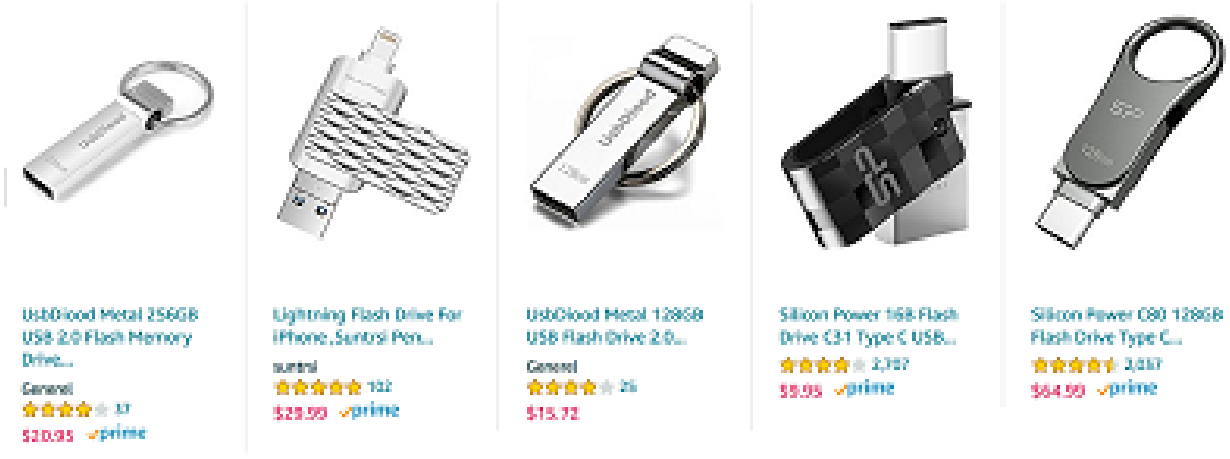
+ hoster

Criteria for selecting a cloud service:

- can you go back in history? ("versioning")
- can you undelete? (= protect against ransomware)
- two factor authentication (see later in this lecture)
- encryption (see later in this lecture)
- Web interface
- Is the client open-source?

Solution 2: USB key

Keep the data on a USB key that is physically stored in a different place.



Amazon

- store the USB key in a different place
- back up your data every few months

Solution 3: Use a source control system such as GIT or SVN,

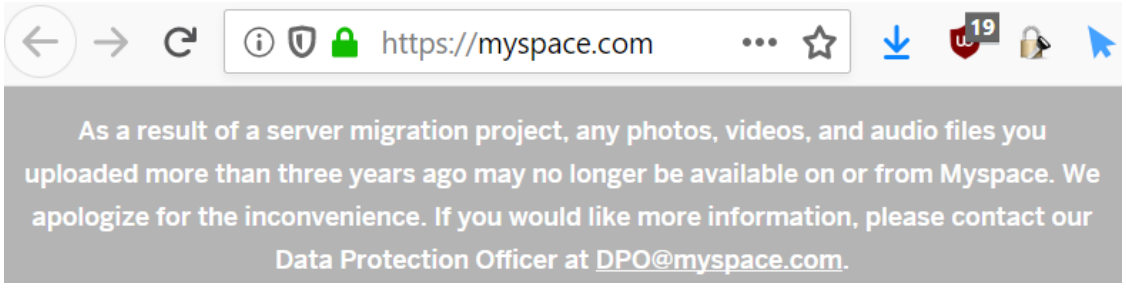
or a remote encrypted backup system such as Duplicity or Borg backup.

>backup

Back up non-file data

- Pictures from your phone
  Either by installing a cloud service, or by copying them to the PC.
  iPhones will automatically backup to iCloud.
- Your SMS (on Android: with an app; on iPhone: iMessages (?))
- Your contacts and calendar on your phone (use CalDav/CardDav)
- Your emails
  Use, e.g., an email client on your computer.
  Otherwise, export your emails from your service provider.
- Your WhatsApp chats (stored only on your phone!)
  Tap the name of the contact, then choose "export".
  Chats may also be backed up when you back up your phone.
- Data that you have stored in online services
  - Facebook
  - Google Drive
  - Trip planners

Back up your Social Media



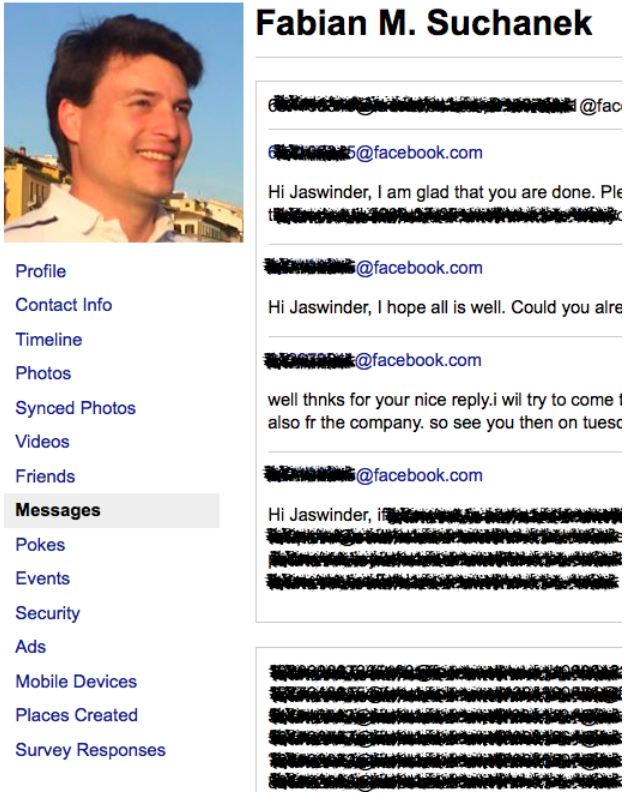
Myspace lost all music that users uploaded 2003 - 2015.

YouTube keeps deleting evidence of Syrian chemical weapon attacks

As Google-owned YouTube continues to delete millions of videos, the Syrian Archive is racing to back-up crucial evidence

Youtube erased Syrian war testimonials – some of which existed nowhere else

Back up your Social Media



You can download
your Facebook data
into neat HTML files.

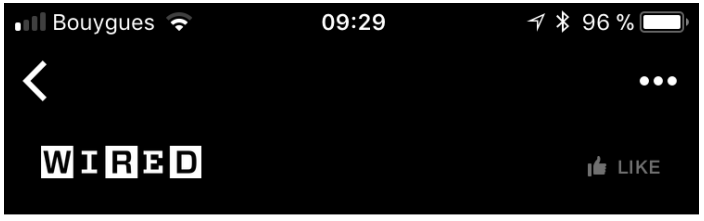
Google can do
the same

>backup

Back up your Passwords

Cryptocurrency exchanges often collapse or are hacked. In February QuadrigaCX, a Canadian exchange, filed for bankruptcy, saying it had lost $165m in deposits when its founder, Gerard Cotton, died, since only he had known the encryption keys protecting Quadrigacx's deposits. But on March 1st Ernst

The Economist, 2019-04-01



LONGFORM

'I Forgot My PIN': An Epic Tale of Losing $30,000 in Bitcoin

A veteran tech journalist tries everything, including hypnosis, to recover a small fortune from a locked bitcoin device.

BY MARK FRAUENFELDER
29 OCTOBER 2017

The Trezor: January 4, 2016: 7.4 BTC = $3,000

In January 2016, I spent $3,000 to buy 7.4 bitcoins. At the time, it seemed an entirely worthwhile thing to do. I had recently started working as a research director at the Institute for the Future's Blockchain Futures Lab, and I wanted

Three solutions:

1) set up an alternate email address
   for resetting the password
   of an online service

2) write it on a piece of paper
   and store it in a secure place

3) put the password in a file,
   encrypt it with a different password,
   give the file to one friend,
   and the password to another friend.

Overview

Protecting data against

- yourself
- hackers
- evil interlocutors
- companies
- governments

Damage by cybercrime

In 2017, 978 million people in 20 countries were affected by cybercrime.

- Having a device infected by a virus or other security threat (53%)
- Experiencing debit or credit card fraud (38%)
- Having an account password compromised (34%)
- Encountering unauthorized access to or hacking of an email or social media account (34%)
- Making a purchase online that turned out to be a scam (33%)
- Clicking on a fraudulent email or providing sensitive (personal/financial information in response to a fraudulent email (32%)

=> Cybercrime victims globally lost $172 billion

[Norton Cyber Security Insights Report 2017 Global Results]

Protecting against hackers

risk = probability of the event × damage

low

high

If a hacker had access to your email, they could

- read all email you have ever written or received (bank, SO, ex,...)
- see all pictures attached to emails that you sent or received
- send emails in your name (e.g., to colleagues or clients)
- post messages in your name on Facebook
- lock you out of your Facebook account (by changing the password)
- lock you out of your email account
- close your email account, close your Facebook account
- mess up your blog
- gain hold of basically all other online accounts.

The main protection is your password.

Popular passwords are bad

Most popular passwords:

123456	abc123
password	admin
12345	121212
12345678	flower
football	passw0rd
qwerty	dragon
1234567890	sunshine
1234567	master
princess	hottie
1234	loveme
login	zaq1zaq1
welcome	password1
solo	

Time.com

A hacker can simply try out all of these passwords

Popular passwords

Common words are bad

- love
- Love
- love you
- …

A hacker can simply try out all words from a dictionary ("dictionary attack")

- l0ve
- 1 l0ve y0u
- …

Replacing letters by numerical counterparts is a known strategy => not safe

In 2006, 55% of MySpace passwords were crackable in 8 hours [Wikipedia]
After the September 11 attacks, the passwords of deceased employees were commercially cracked to allow using their work.

Def: Password strength

The number of possible passwords (combinations) of length n over k characters is k^n.

The password strength is often given "in bits" as the binary logarithm of the number of possible combinations.

Example: 10 letters a-z = 26^{10} combinations (0.1 quadrillions).
Password strength: $log_2(26^{10}) \approx 47$ bits

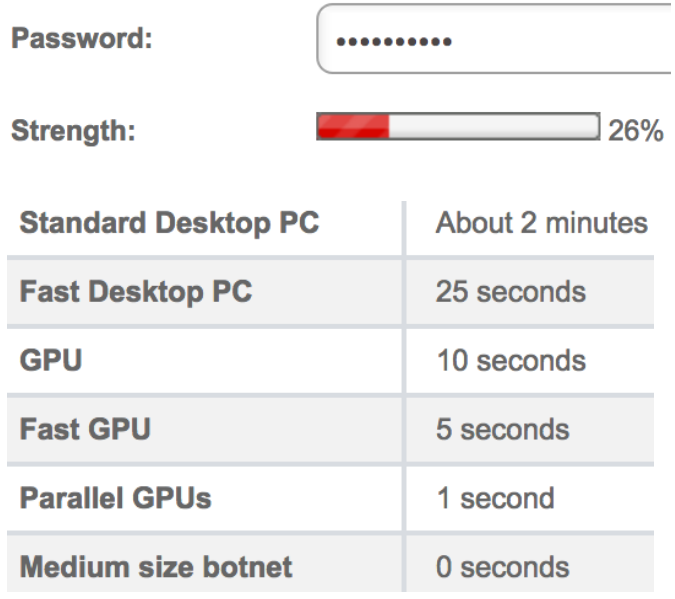
Short passwords are bad

combinations = characters ^ length

A hacker can simply try out all combinations

Example: 10 digits = 10 billion combinations.
A PC can do 100m combinations per second
=> we need only 1.5 minutes to break it

Example: Uber sends out a 4 digit code to verify an account. Generate 1m account requests, verify 100 of them just by chance.



Try it out!
(But not with your real password!)

Personal information is bad

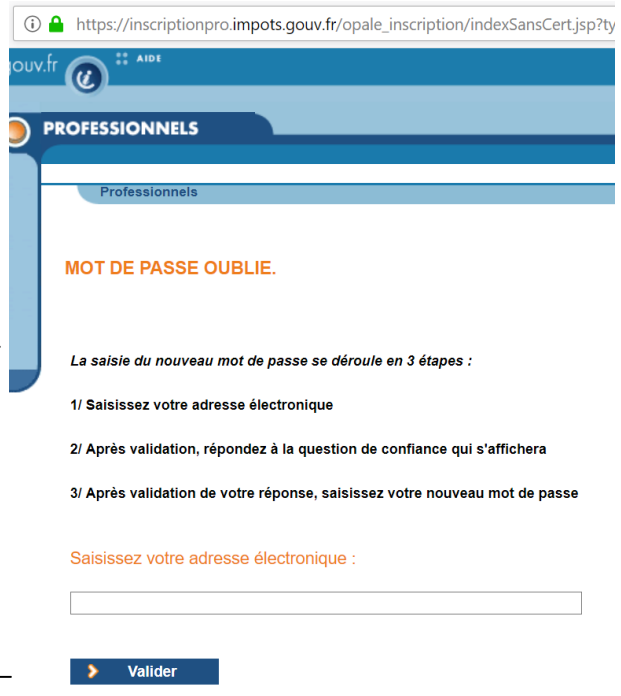
Pet names
A notable date (wedding, b'date)
A family member's birthday
Your child's name
Another family member's name
Your birthplace
A favorite holiday
Your favorite sports team

According to Google

A hacker can find this information in social networks.

Personal information is bad

Pet names
A notable date (wedding, b'date)
A family member's birthday
Your child's name
Another family member's name
Your birthplace
A favorite holiday
Your favorite sports team
Your highschool



A hacker can find this information in social networks.

Julien Romero
PHD Student chez Telecom ParisTech
Paris Area, France | Telecommunications
229 connections

Current	Telecom ParisTech
Previous	Telecom ParisTech, Haufe-umantis AG
Education	Eidgenössische Technische Hochschule Zürich
Websites	Blog

Henri 4
Reçu à Telecom ParisTech, Mathematics, Physics and Computer Science
2011 – 2013

Activities and Societies: Natation, Escalade, Musculation

LinkedIn

Security questions are bad

Security Questions.

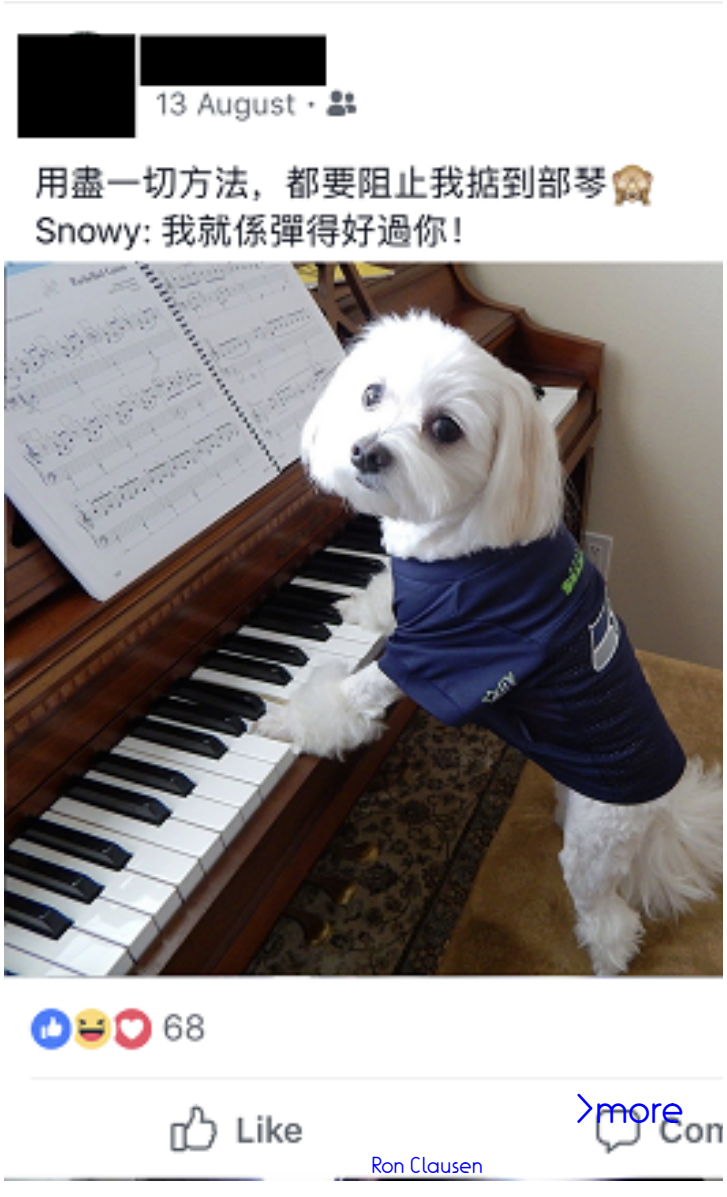
Select three security questions below. These questions will help us verify your identity should you forget your password.

Security Question	What was the name of your first pet?
Security Question	What is your dream job?
Security Question	In what city did your parents meet?

A hacker can find this information in social networks.

40% of users fake the answers and then forget them. Most secure questions are also least memorized.

Research by Google



Ron Clausen

Difficult passwords are bad

"Your password must contain at maximum 5 letters,
and must contain a mammal that lives in the sea"

The more annoying passwords restrictions are,
the more likely users are to

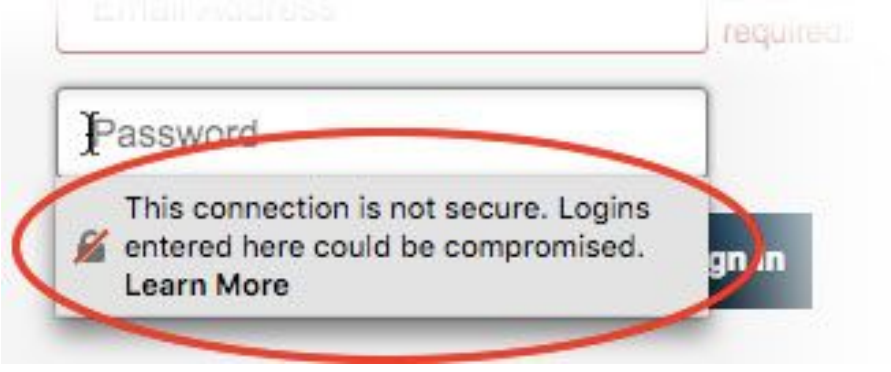
- write the passwords down
- forget the passwords
- use simple variations of the same password

The same applies to passwords that have to be changed regularly.

>more

Using the same password is bad

If the password is compromised on one site



...it allows access to all the others:

Dropbox

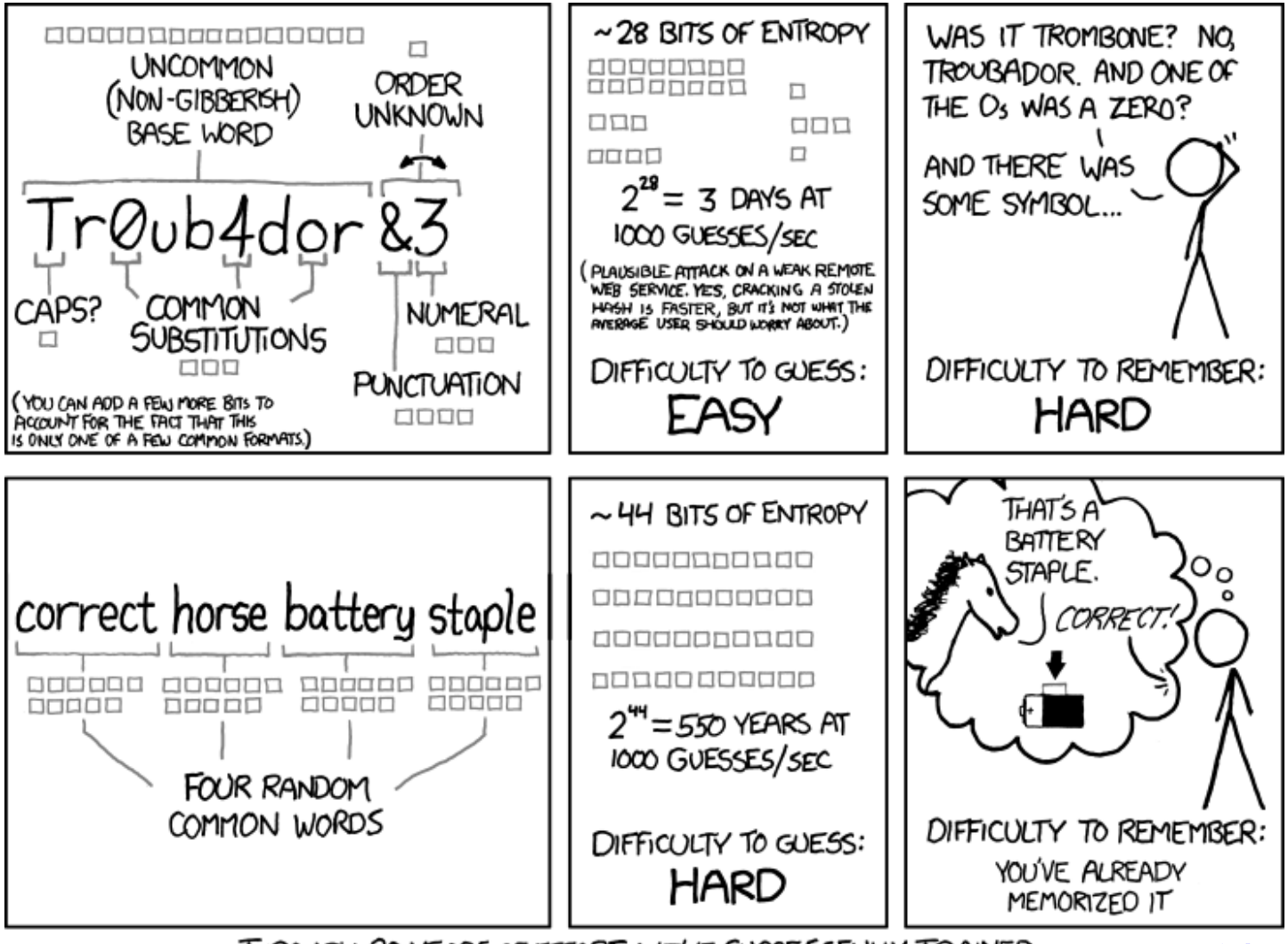
amazon

Solution 1: First letter passwords

"How much money do I have?" -> "Hm$d1h?"

- easier to remember
- high entropy
- recommended strategy

Solution 2: Diceware passwords



THROUGH 20 YEARS OF EFFORT, WE'VE SUCCESSFULLY TRAINED EVERYONE TO USE PASSWORDS THAT ARE HARD FOR HUMANS TO REMEMBER, BUT EASY FOR COMPUTERS TO GUESS.

XKCD

Def: Diceware password

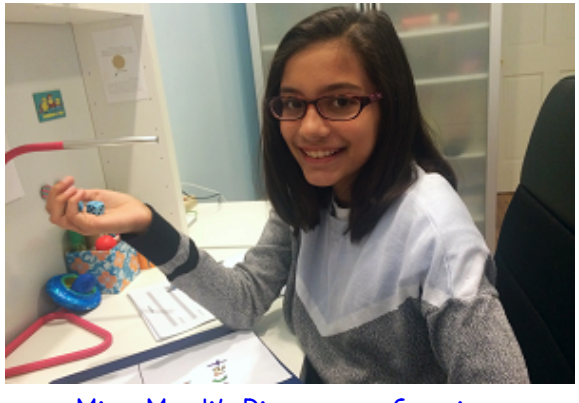
A diceware password of length n is created as follows:

1) Take a list of 6^5 words of some language,
   which are indexed $<0,0,0,0,1>$, $<0,0,0,0,2>$,...
2) Repeat n times
   a) Roll a (physical) dice 5 times,
      obtaining numbers i_1, i_2, i_3, i_4, i_5
   b) Add to your password the word at index
      $<i_1, i_2, i_3, i_4, i_5>$

43136	**mulct**
43141	**mule**
43142	**mull**
43143	**multi**
43144	**mum**
43145	**mummy**
43146	**munch**

EFF Diceware list

Or use a service:



Mira Modi's Diceware Service

combinations = $(6^5)^n$

Solution 3: Password managers (?)

Password managers contain one password per online service, and copy/paste it automatically into the login field.

LastPass •••|  KeePass  1Password dashlane

... or the password manager of your browser

Caveats:

- such services are a target for hackers.
- what if the service gets hacked?
- what if the service leaks data?

Alternatively: use a random password and copy/paste it each time.
You can generate the password online or with pwgen.

When a password is not enough

Mat Honan

amazon

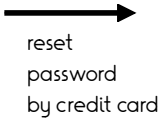
When a password is not enough

Mat Honan

amazon

When a password is not enough

Mat Honan



When a password is not enough

Mat Honan

reset
password
by credit card

reset
password
by backup email

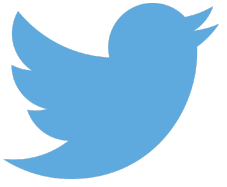
When a password is not enough

Mat Honan

reset
password
by credit card

reset
password
by backup email

reset
password
by backup email



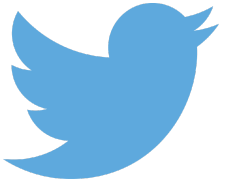
When a password is not enough

Mat Honan

reset password by credit card

reset password by backup email

reset password by backup email



Two-factor authentication

There should be at least 2 independent hurdles to your data!

Def: Two-Factor Authentication

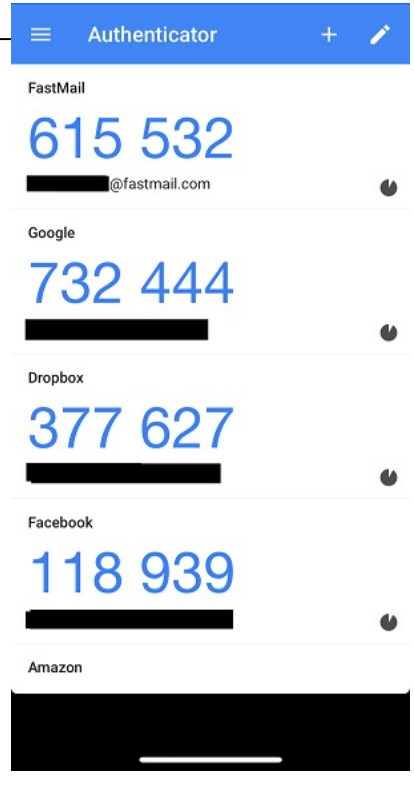
Two Factor Authentication (2FA) method of access control that allows access only if two independent codes ("factors") are entered.

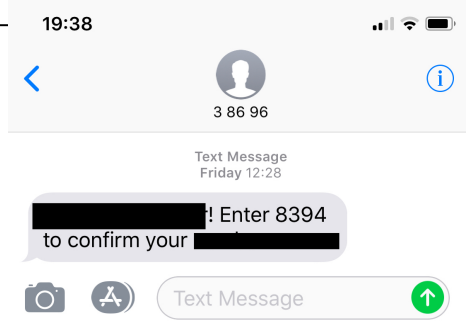
The first factor is usually a password.

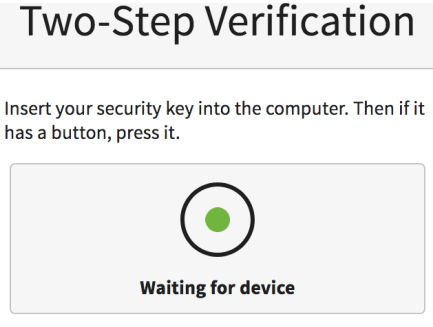
The second factor can come

- from an app on your phone
- from a phone call
- from an SMS
- from a security key

e.g., FreeOTP
or Google Authenticator





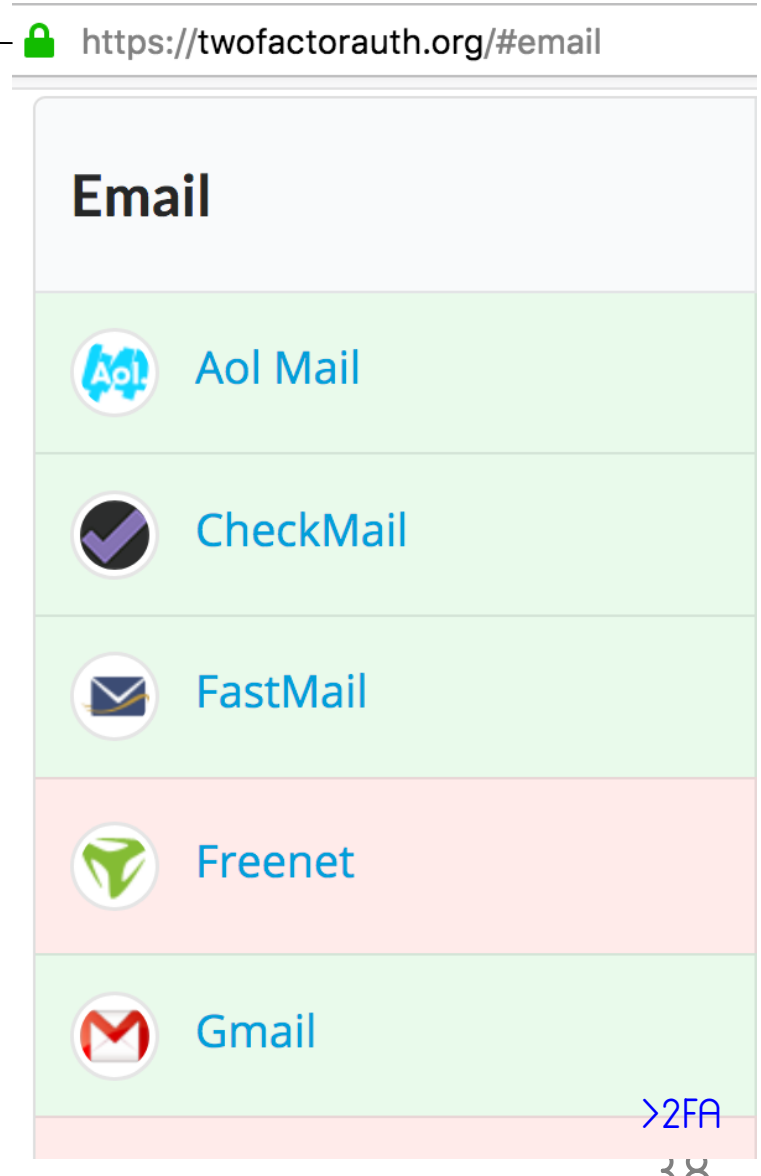


FIDO security key

>2FA

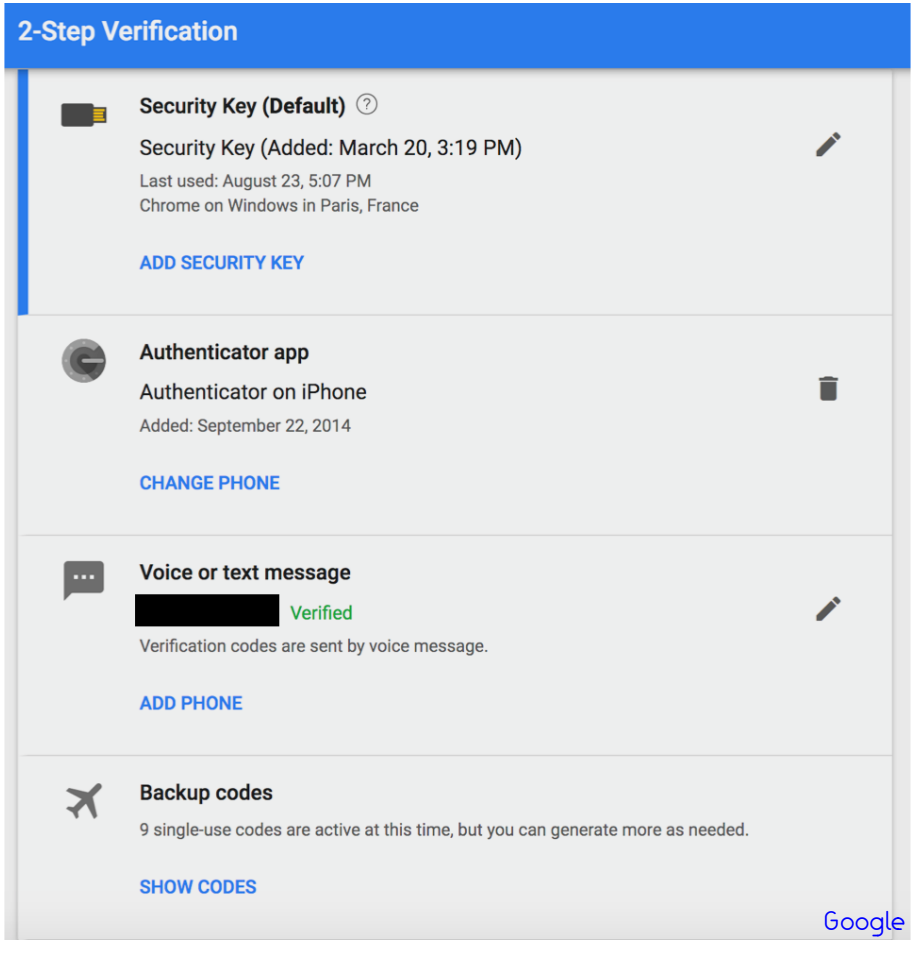
Caveats with two factor authentication

- Not all services support 2FA
- SMS can be intercepted
- With Apple's two factor authentication, any linked device can generate codes => obtaining one device allows messing around with the others



Enable fall-back options!

Never enable 2FA without a fallback.
You risk getting locked out.



Poor man's fall-back: scan the 2FA barcode with the phone of a friend.

Protect your devices

Add the two hurdles also to your devices: Possession + passcode

- Disable notifications on the lock screen.

- Enable passcodes on your laptop
  Consider hard drive encryption, because otherwise the code is useless.
  Macs and Linux can encrypt the drive natively. Windows has Bitlocker.

- Enable passcodes on your phone
  - passcode (cumbersome)
  - lock pattern (easy to copy)
  - fingerprint (great)
  - face id (can be tricked by a picture)

60% of people can reproduce the pattern after seeing it once in 1m distance

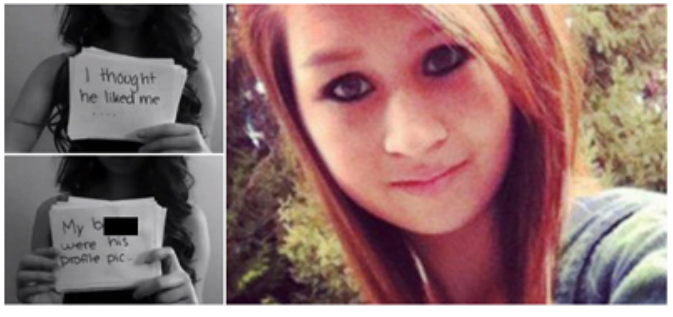
Apple

Protect really sensitive data

Really sensitive data are

- embarrassing pictures of yourself or others
- files that contain passwords
- scans of your passport
- confidential information that you store for others

Such data can be used to embarrass you, to impersonate you, or to blackmail you.



Wikipedia: Suicide of Amanda Todd
Wikipedia: Revenge Porn
Cyberbullying Research Center

Really sensitive data should never live outside protected spaces.

I.e., wherever it is, there should be 3+ hurdles to access it. Encrypt it (1,2).

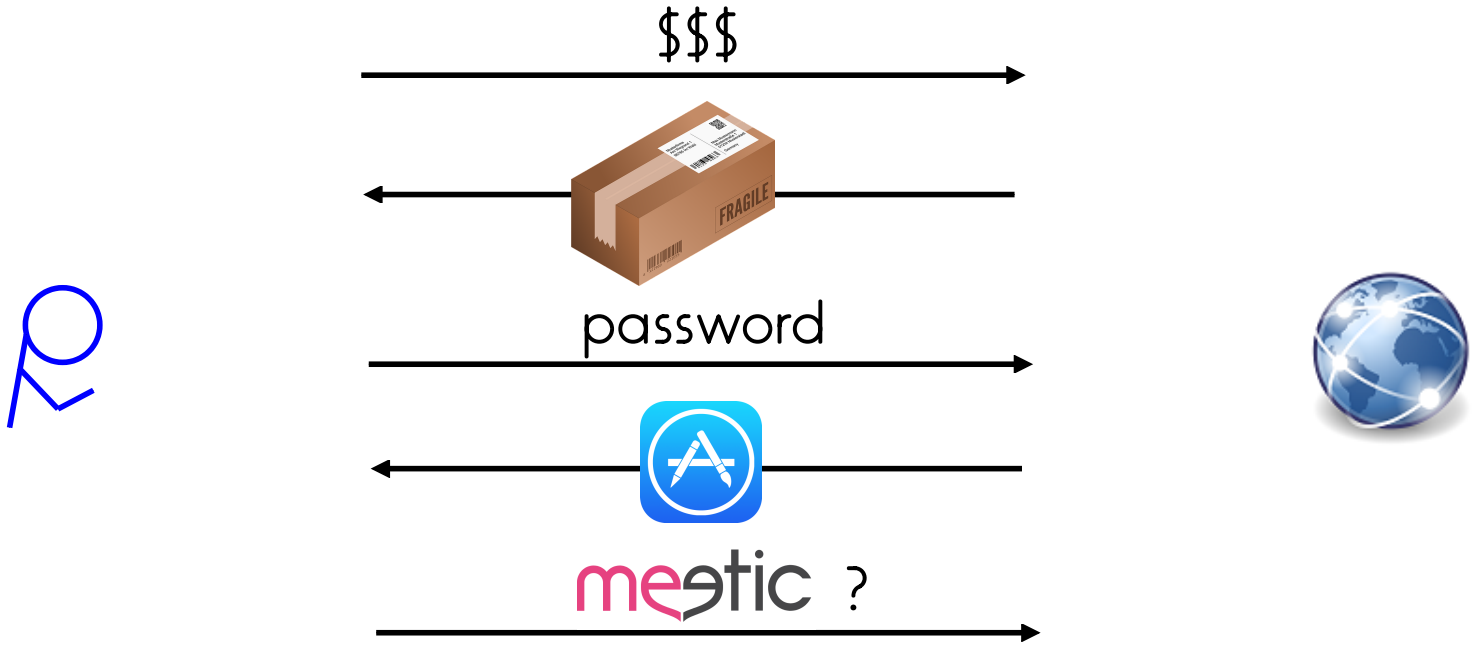
Overview

Protecting data against

- yourself
- hackers
- evil interlocutors
- companies
- governments

Interacting online

We exchange more than just clicks online.



If you're not talking to whom you think you talk, you're in trouble.

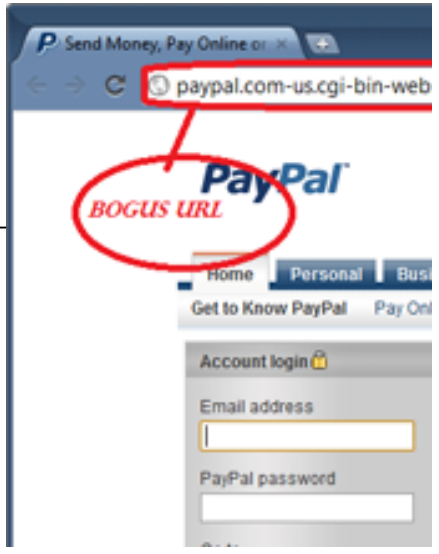
Any serious interaction on the Web should only happen if the identity of your partner has been confirmed by a trusted *third* party.

Bogus URLs

You can be tricked into interacting with a bogus URL by:

- unreadable URLs
- homograph attacks

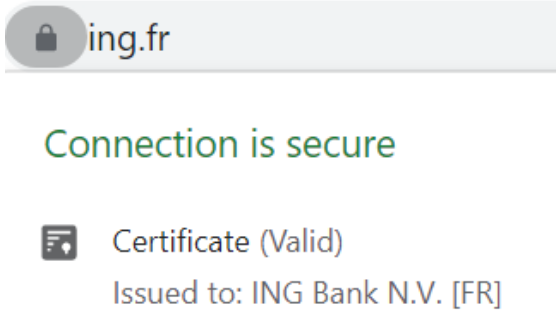
wikipedia.org



The "HTTPS" just means that the traffic between you and the Web site cannot be intercepted by a middle man.
It does not guarantee that the Web site really belongs to the organization you think.

Def: Extended Validation Certificate

An extended validation certificate is a label that means that the identity/owner of the Web page has been confirmed by a third party.



"This Web page is operated by the French company ING Bank".

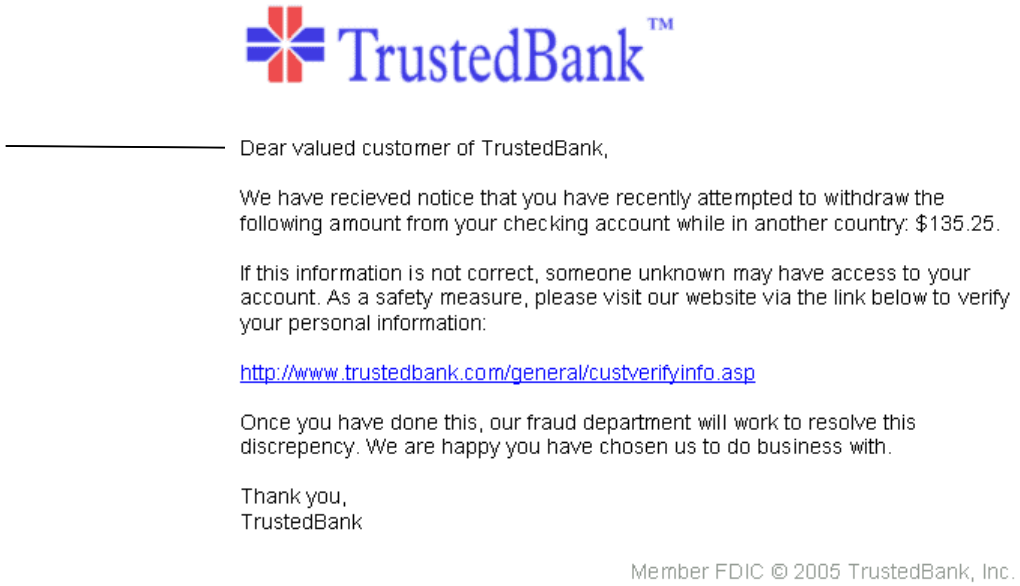
Problem: Users do not verify the EV certificate, and it's easy to register one in another country
=> the trend is against EV

Still: Do not do banking without the EV certificate!

Def: Social Engineering

Social Engineering is the psychological manipulation of people into performing actions or divulging confidential information.

- Phishing (ask for password, pretending to be an authority)
- Vishing (phishing via automated phone message)



TrustedBank™

Dear valued customer of TrustedBank,

We have recieved notice that you have recently attempted to withdraw the following amount from your checking account while in another country: $135.25.

If this information is not correct, someone unknown may have access to your account. As a safety measure, please visit our website via the link below to verify your personal information:

http://www.trustedbank.com/general/custverifyinfo.asp

Once you have done this, our fraud department will work to resolve this discrepency. We are happy you have chosen us to do business with.

Thank you,
TrustedBank

Member FDIC © 2005 TrustedBank, Inc.

Tue, 28 Mar, 14:00

ING Direct: Votre carte est bloquée suite à une suspicion de fraude. Appelez-nous au 0157225400 ou au 0033157225400 depuis l'étranger.

(real...)

- Baiting (leaving hardware around for others to pick it up)

  (In a study, 98% of bait USB keys were picked up, and 45% called home [Wikipedia].)

- Quid pro quo (attacker calling as help desk worker, offering help)

>more

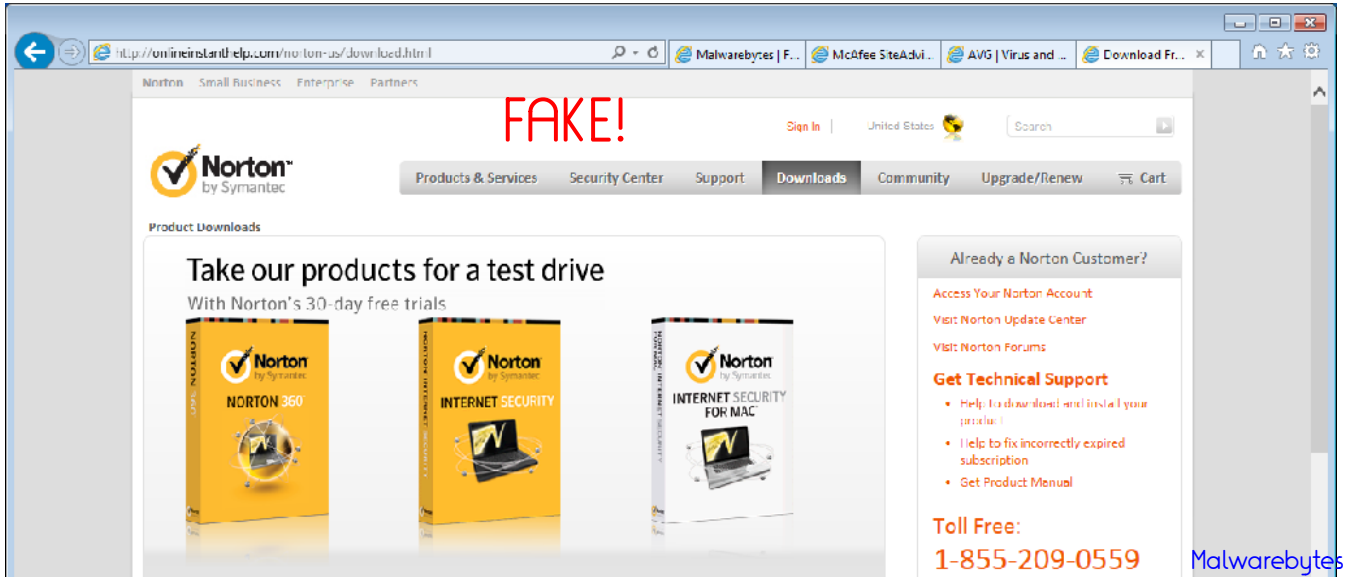
Watch out when downloading

If you download bogus software, you may catch
ad programs, keyloggers, viruses, or ransomware.

Download only

- from the original vendor (with EV certificate)
- if recommended by a reputed third party (computer magazine)
- iPhone apps that have a large number of positive ratings

>more



Malwarebytes

NakedSecurity

Do not trust online acquaintances

Do not send money to, send intimate pictures to, or meet in non-public places with online acquaintances.

Funny examples:

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

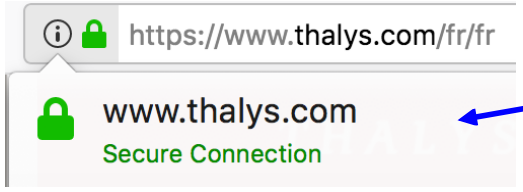
Die PARTEI hacked AFD

Not so funny examples: News24

>more

Oh, and be careful

- Do not click on random advertisements
- Do not open email attachments from unknown senders
- Do not make online payments without HTTPS



Prevents evesdropping, but does not guarantee identity. 50% of phishing Web sites have the padlock!

- Do not buy from online shops, unless
  - they have an extended validation certificate that fits the expectation
  - they are reputed (Amazon, Booking, etc.)
  - they are the official pages of brick-and-mortar shops (ask Google)

On a Mac: Install a virus scanner.
Everywhere: Keep all software updated.

On Windows 8+, the preconfigured virus scanner is generally enough.

Overview

Protecting data against

- yourself
- hackers
- evil interlocutors
- companies
- governments

What the big companies know

Your email provider, your social network and/or your search engine know

- your emails
- your purchases on the Web
- your trips
- your Web searches
- your exact location (if logged in in Maps)
- the people you interact with
- what you like
- when you are online

BUT: The companies deliver a high-quality, free service in return!

What Facebook may know

your personality

substance use

field of study

type (better than your spouse)

impulsivity

political orientation

values

physical health

sensational interests

Try it out with Cambridge tool

depression

gender

likely moving soon

age

receptivity to online insurance offers

"mother type"

education level

type of restaurants

relationship status

pain relief buyers

types of clothing

balance on the credit card

how much money will spend

wants to buy a car

type of vacation

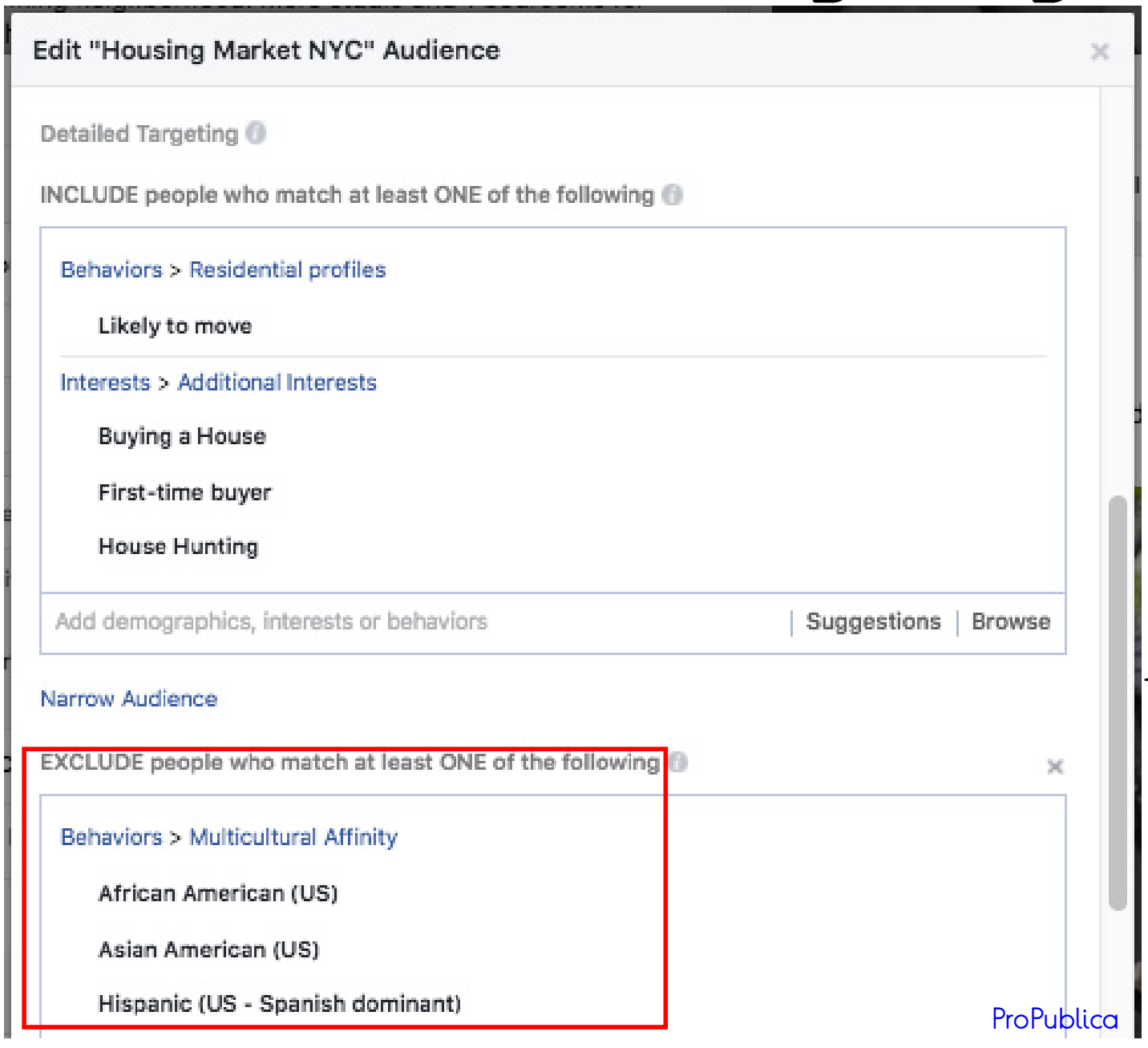
age of car

heavy buying of alcohol

purple = what advertisers can target

>more

Facebook allows targeting ads

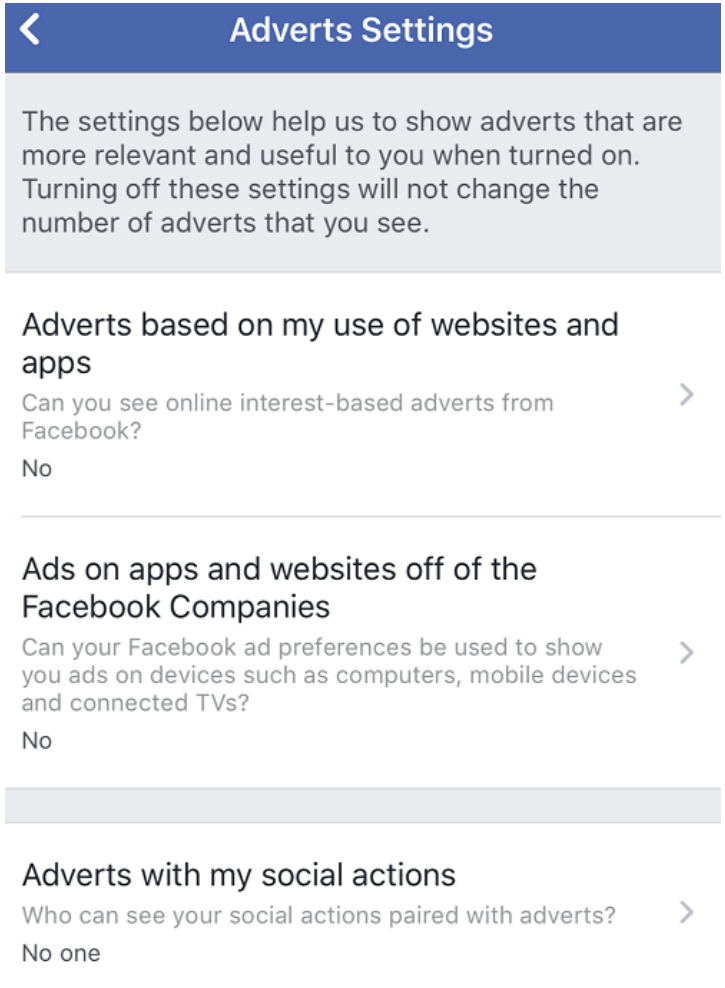


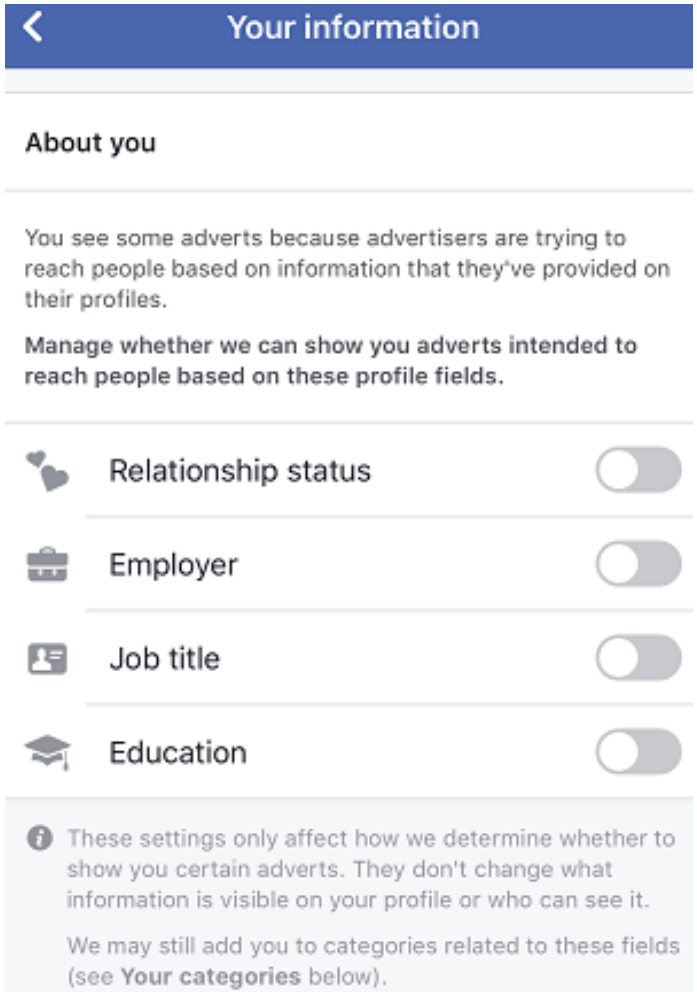
The NGO ProPublica bought ads and asked to exclude African Americans, mothers of high school kids people interested in wheelchair ramps, Jews, expats from Argentina and Spanish speakers.

ProPublica

>more

How Facebook puts ads



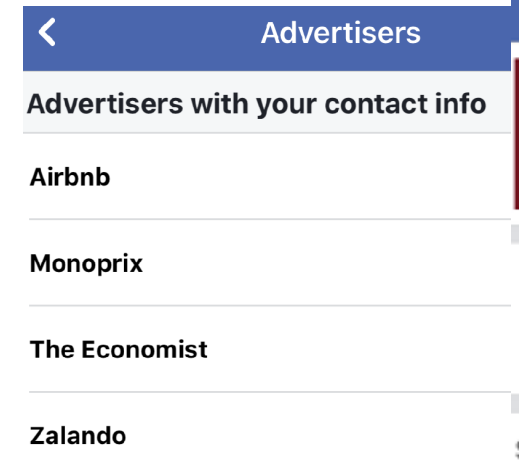


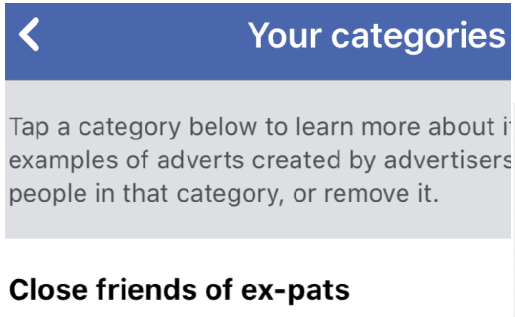
Facebook "Like" buttons trace you even if you don't click them.

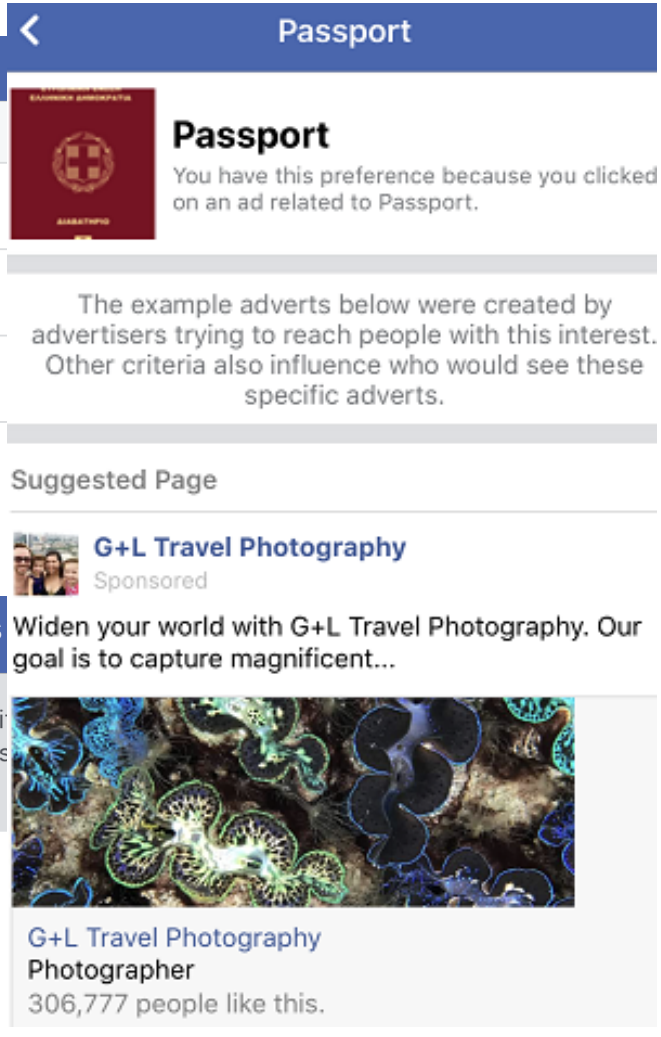
Even if you log out, their cookies remain.

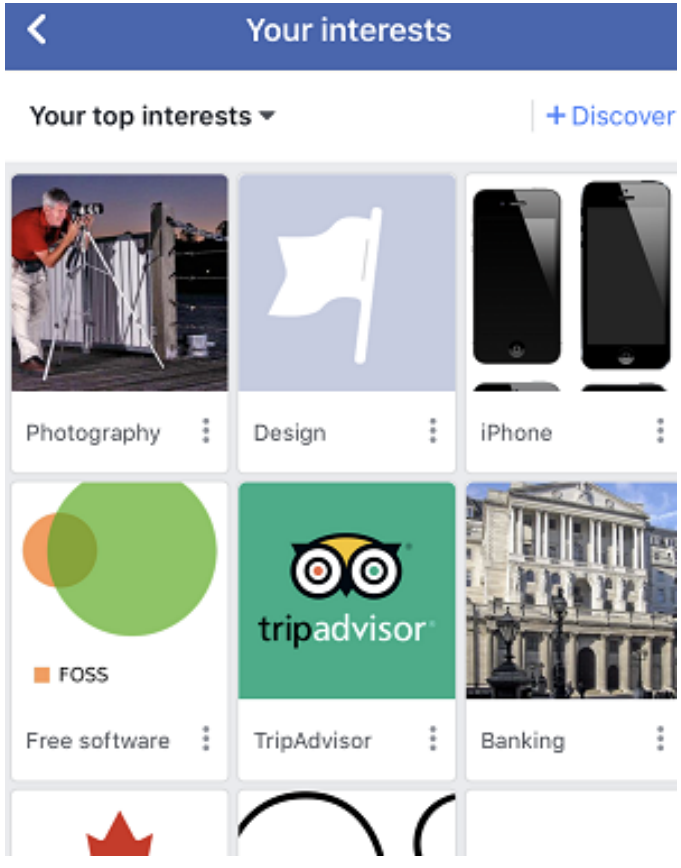
>more

How Facebook puts ads









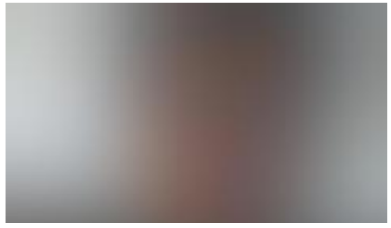
Go check your privacy settings!

>more

What WhatsApp knows

WhatsApp shares your phone number, contact list, and usage data with Facebook. The online time is also publicly available.
The messages themselves are private.

Statistic for WhatsApp user: +316XXXXXXXX



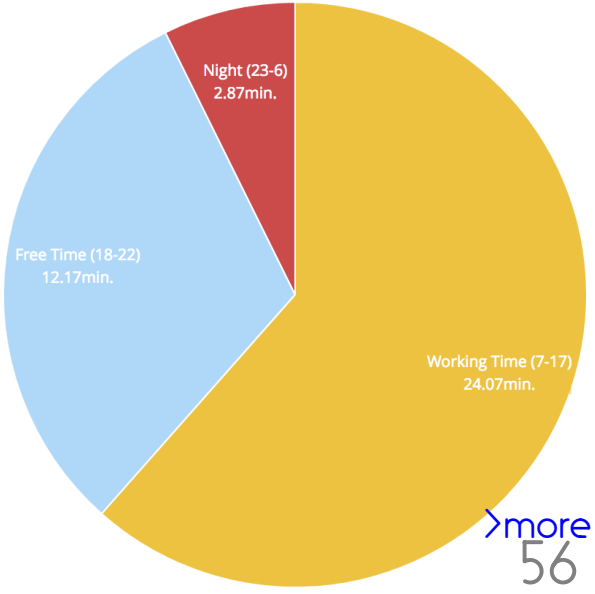
8362 connections 7 days 6:1:21 online
Ø 31.44 connections/day Ø 39:15min. online/day

Statusmessage:

Available

#Weekday	Avg. Online Time
Monday	48.63 min.
Tuesday	44.71 min.
Wednesday	49.92 min.
Thursday	39.15 min.
Friday	39.73 min.
Saturday	31.14 min.
Sunday	20.46 min.

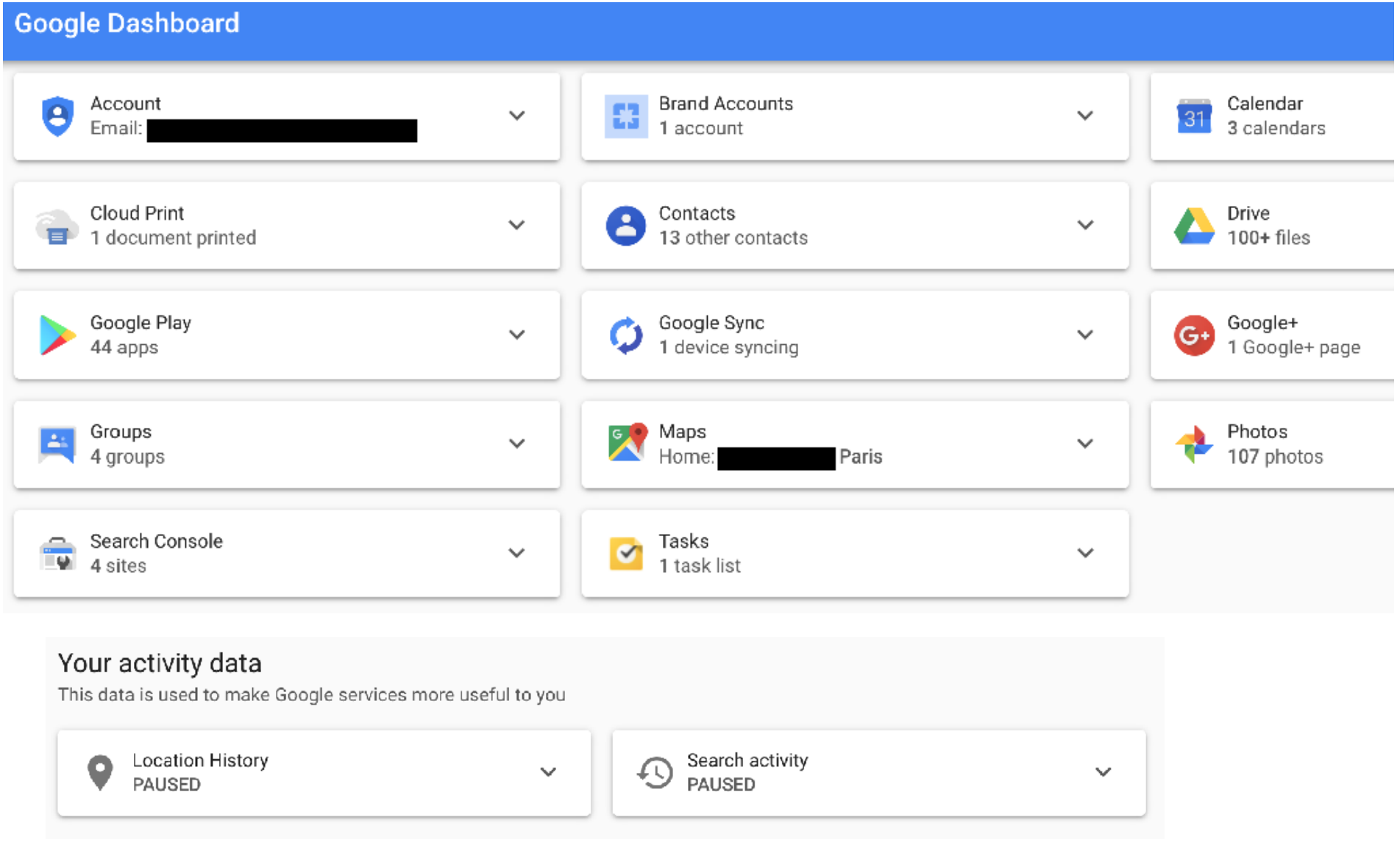
#Hour	Avg. Online Time	Avg. #Connections
0	0.74 min.	0.74
1	0.21 min.	0.25
2	0.05 min.	0.07
3	0.04 min.	0.05
4	0.1 min.	0.09



>more

https://www.onlinestatusmonitor.com

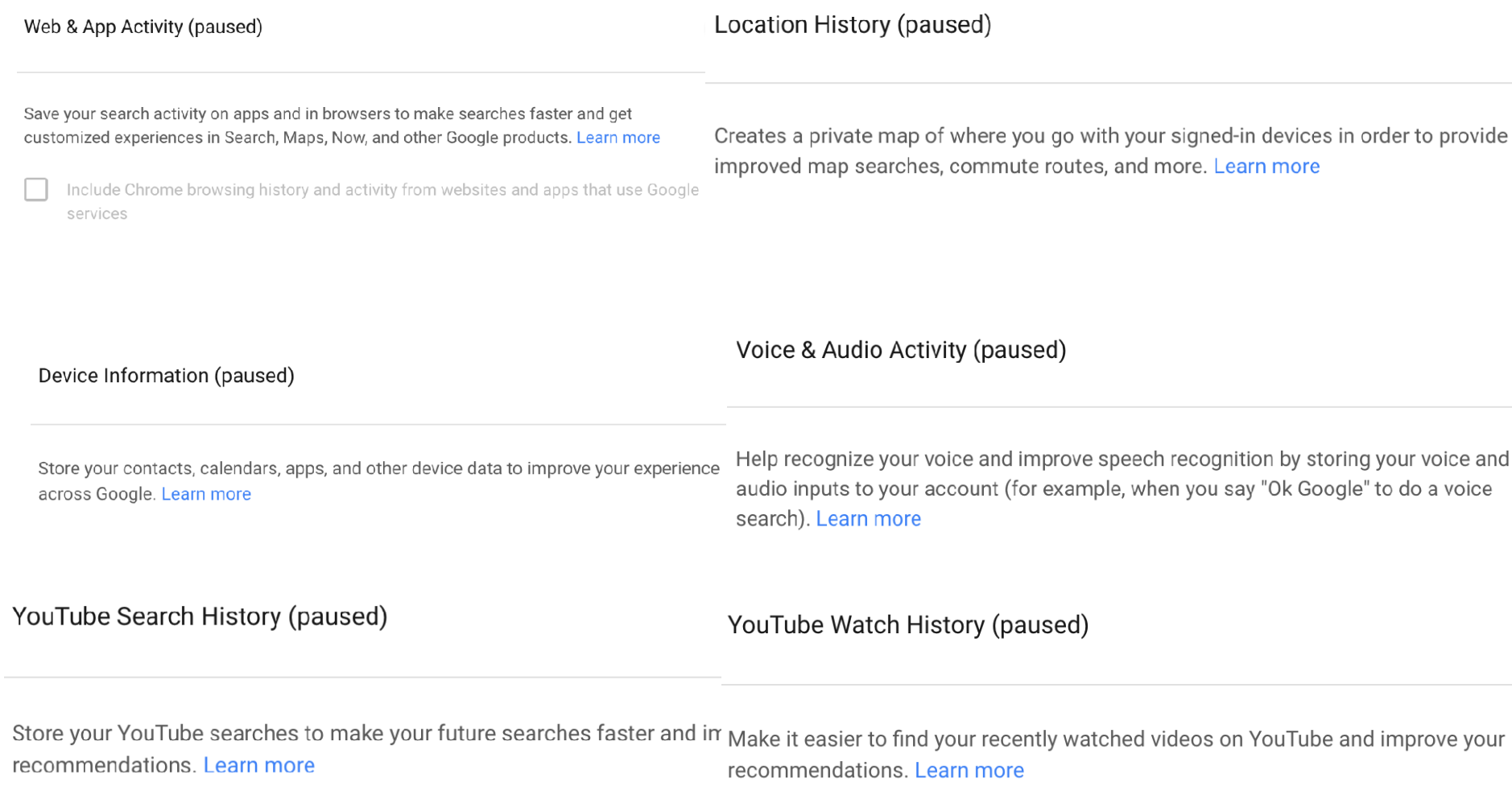
What Google knows



http://google.com/dashboard

>more

What Google knows

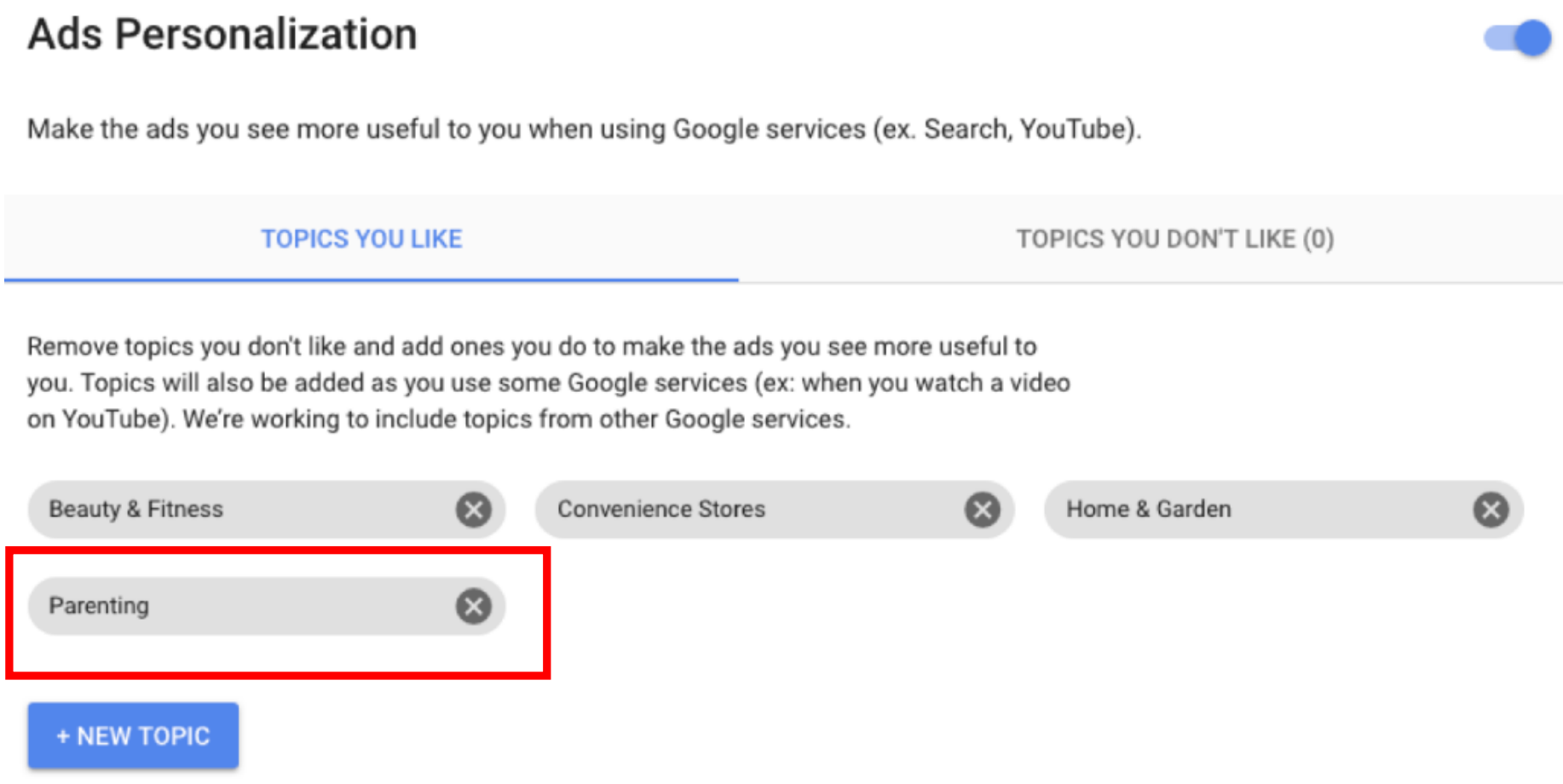


http://google.com/dashboard

Go check your privacy settings!

>more

What Google knows



http://google.com/dashboard

(Real example in my family, deduced automatically by Google)

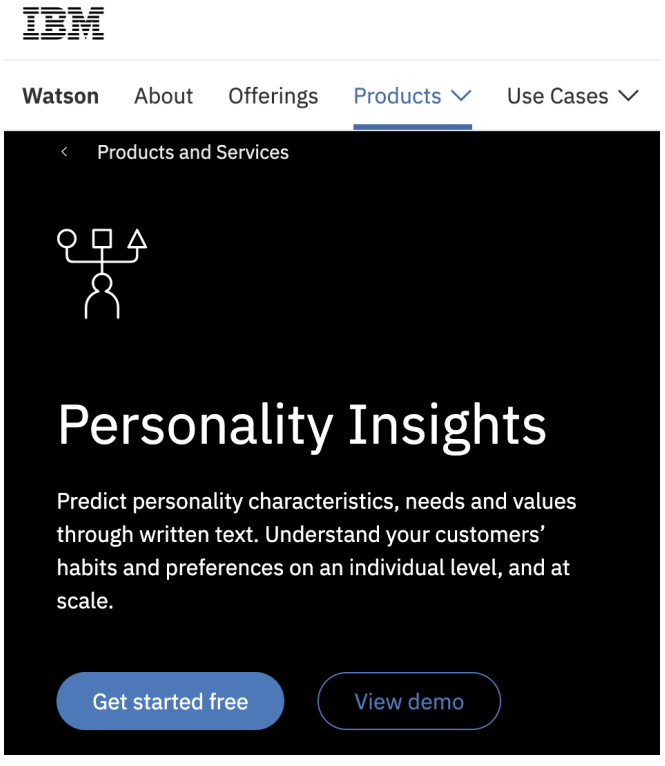
>more

And Apple...

https://appleid.apple.com/

Go and download your data!

And IBM



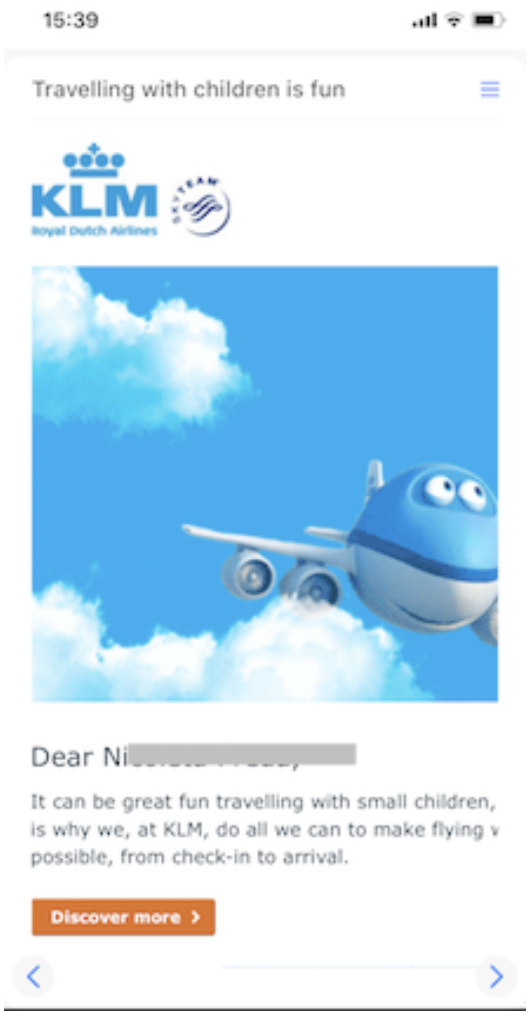
Get detailed personality portraits

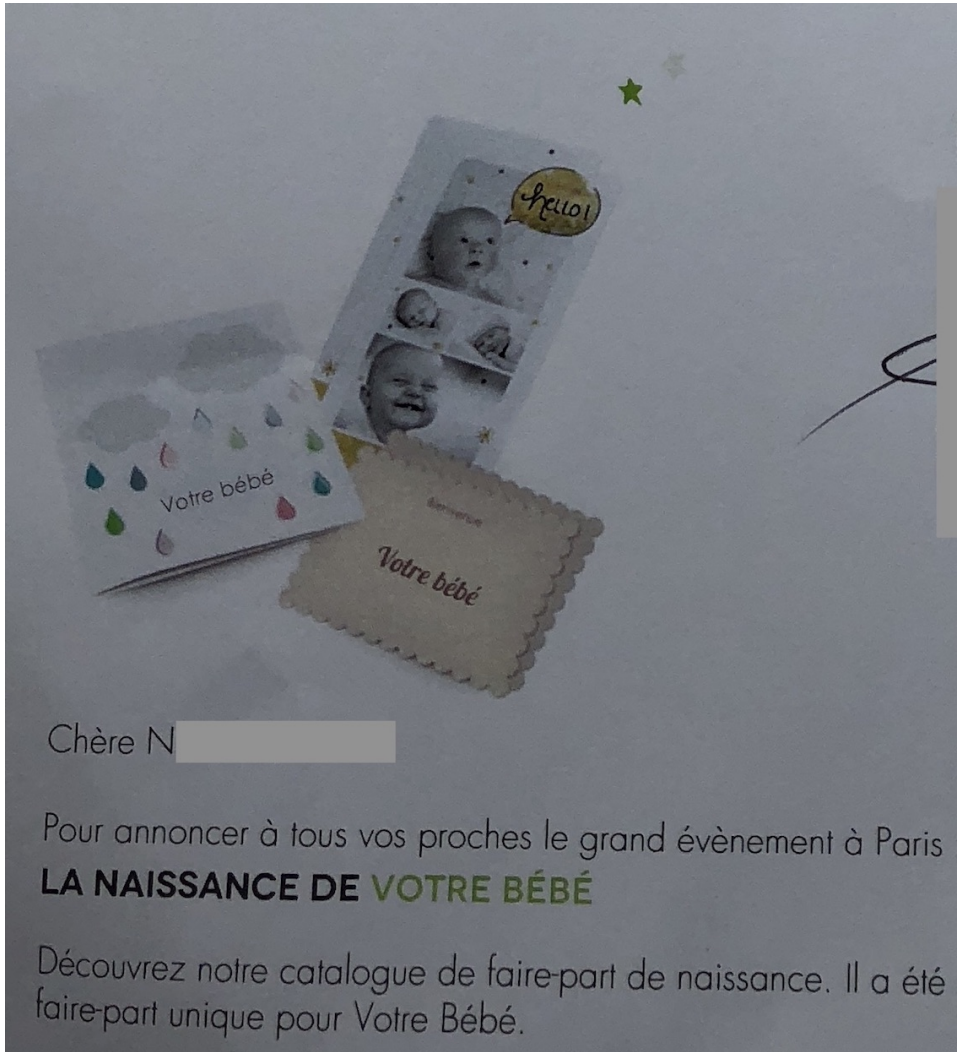
Use linguistic analytics to infer individuals' personality characteristics, including Big Five, Needs, and Values, from digital communications such as email, blogs, tweets, and forum posts.

What it means if they know

You get unsolicited advertisements from companies whom you never told about your life – even before the event happens.

(real examples in my family)





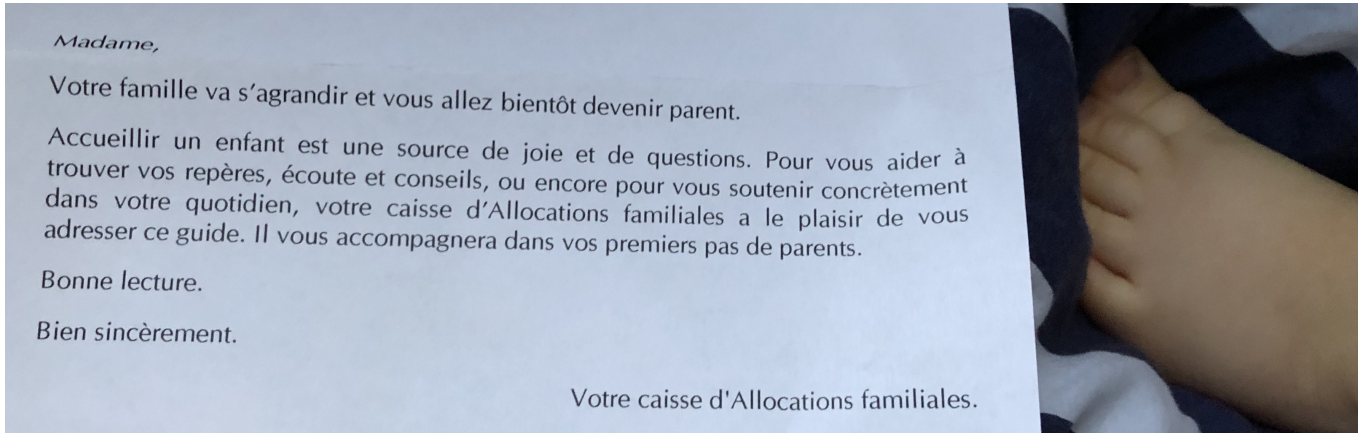
>more

What it means if they know

You get unsolicited advertisements from companies whom you never told about your life – even before the event happens.

(real examples in my family)

...and faster than governments.



Madame,

Votre famille va s'agrandir et vous allez bientôt devenir parent.

Accueillir un enfant est une source de joie et de questions. Pour vous aider à trouver vos repères, écoute et conseils, ou encore pour vous soutenir concrètement dans votre quotidien, votre caisse d'Allocations familiales a le plaisir de vous adresser ce guide. Il vous accompagnera dans vos premiers pas de parents.

Bonne lecture.

Bien sincèrement.

Votre caisse d'Allocations familiales.

What it means if they know

The service providers may also know if you are

- planning a divorce
- having a medical problem
- having an uncommon sexual preference
- under-age and pregnant

What happens if Google takes a political position one day?

This can

- influence the ads you see, even if Google explicitlydisallows it
- make court actions against the service providers difficult (blackmail).

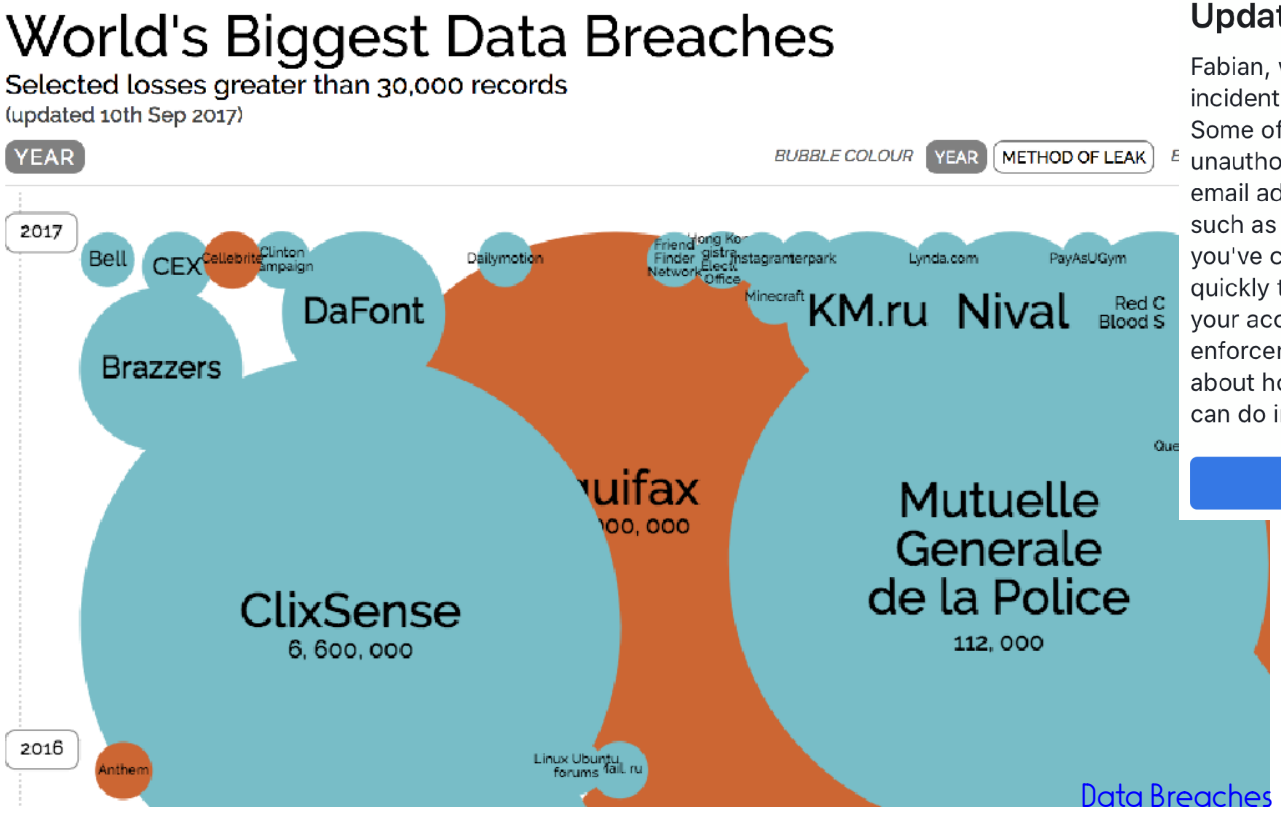
Dubious services providers may sell your information to data brokers, feeding feeds background checks for

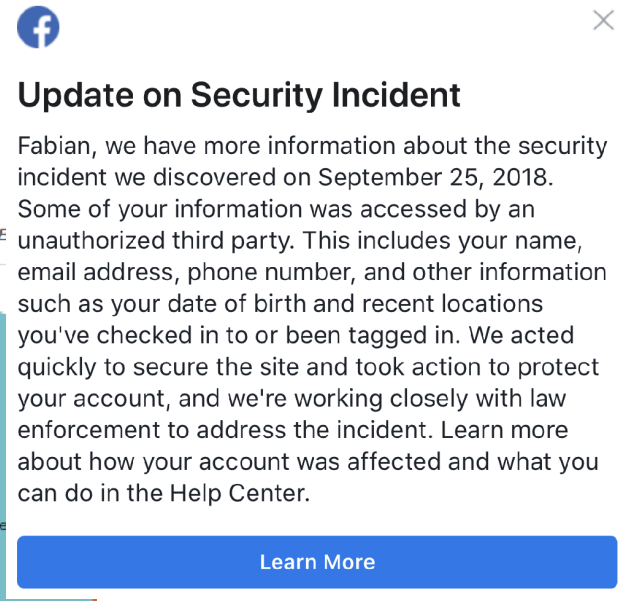
- credit scores
- insurance fees and insurance claims
- advertisements
- hiring decisions

See story of pregnant daughter

>more

Leakage





In 2017, name, birthdates, home addresses, phones, religious affiliations, ethnicities & political biases of 60% of the entire US population leaked to the public.

If such data is out

- somone can impersonate you
- someone can blackmail you
- your credibility suffers

50m Facebook profiles were crawled by Cambridge Analytica to influence elections.

>more

Filter bubbles

Your service provider decides what you get to see.

It wants to keep you happy.

=> It may show you only what you like.
(think about Russia, Islamism, Trump)

=> "intellectual isolation", "echo chamber",
"indoctrination with our own ideas"

=> reduced plurality, reinforced opinions, polarization
"You enter as a vegetarian, you leave as a vegan"

=> "threat to democracy" (Barack Obama)

Wikipedia: Filter bubble

Try it out: search for a controversial topic,
and compare results with a friend.

Societal Degradation

The Humane Technology Center argues that what's best for capturing our attention isn't best for our well-being:

- Snapchat shows how many consecutive days a conversation goes on
  => children measure friendship this way and give their password to friends just to keep up the number
- Facebook creates echo chambers, fragments communities
- YouTube & Netflix autoplay the next video => "Binge watching"
- Providers seed outrage and radicalism to keep users watching

This endangers our mental health, children, society, and democracy.
Plus: the system can be abused by bad advertisers, state agents, or bots to influence what we like and how we think.

Interesting further links:

- One third of people prefer their smartphone over their friends, most used apps
- Uber uses psychological tricks to motivate its drivers.
- Kremlin-linked operation had reached nearly 150 million users with false posts in an effort to sway the 2016 US election

Fake becomes "real"

We can now create fake videos that are nearly indistinguishable from the real thing. Here: futureoffakenews.com

Fake becomes real

Source: Twitter

New York Times, 2019-11-02

Alternative solutions

- Some email providers provide their service for a fee.
  In return they don't use your data for marketing.

Criteria for choosing:

- two-factor authentication
- reputation
- Qualis rating

Alternative solutions

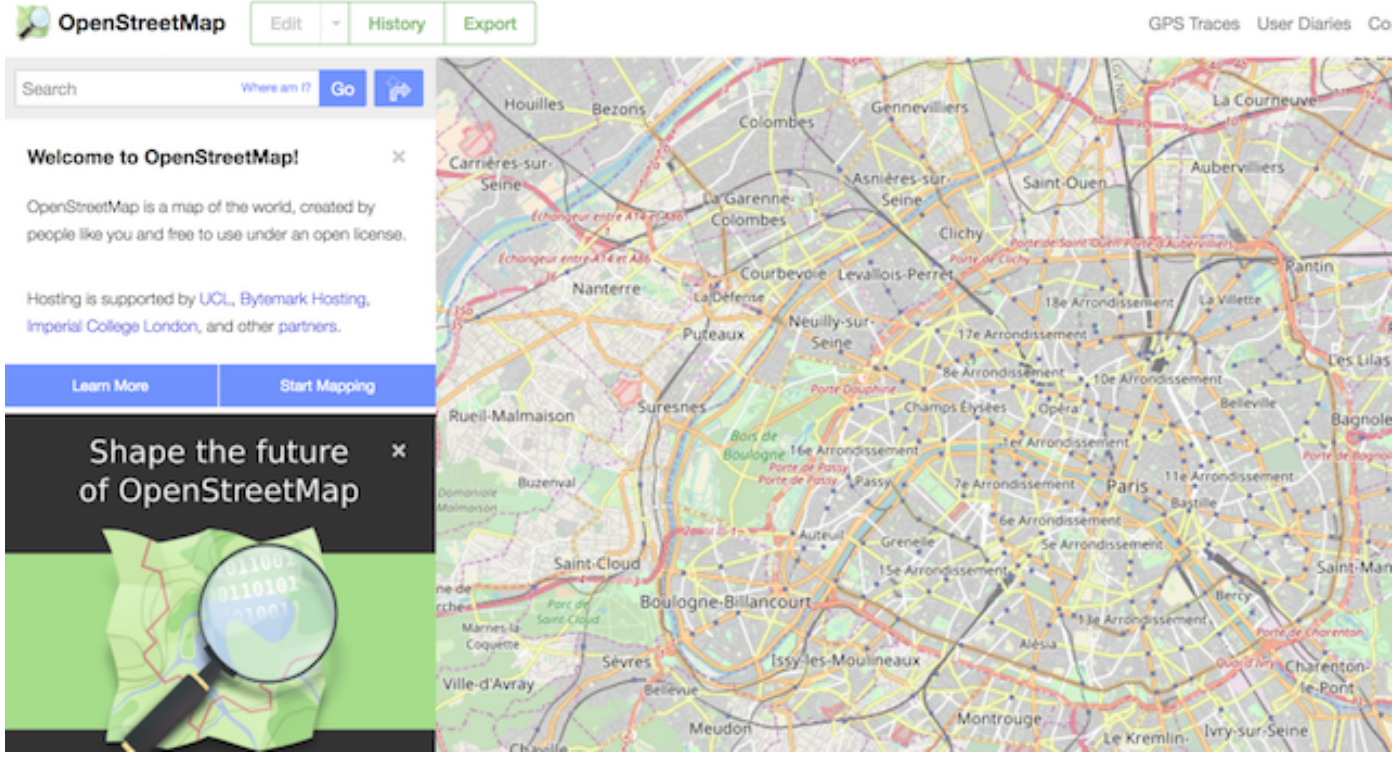
- Some email providers provide their service for a fee.
  In return they don't use your data for marketing.
- WebRTC is a free protocol for voice and video calls over the Web,
  which works without creating an account or installing software.
  => a priori no sharing of data with NSA, no data collection

See here for a list

>more

Alternative solutions

- Some email providers provide their service for a fee.
  In return they don't use your data for marketing.
- WebRTC is a free protocol for voice and video calls over the Web
- OpenStreetMaps is a collaborative open map project, and there are mobile phone apps for it.



Give a chance to the small ones

- Some email providers provide their service for a fee.
  In return they don't use your data for marketing.
- WebRTC is a free protocol for voice and video calls over the Web
- OpenStreetMaps is a collaborative open map project
- A number of search engines aim to protect your privacy.

Give a chance to the small ones

- Some email providers provide their service for a fee.
  In return they don't use your data for marketing.
- WebRTC is a free protocol for voice and video calls over the Web
- OpenStreetMaps is a collaborative open map project
- A number of search engines aim to protect your privacy.
- There are a number of smaller social networks,
  and some initiatives to make social network activity
  provider-independent

identi.ca

ActivityPub
(W3C recommendation)

(Mastodon)

See Degooglisons for a list of alternatives

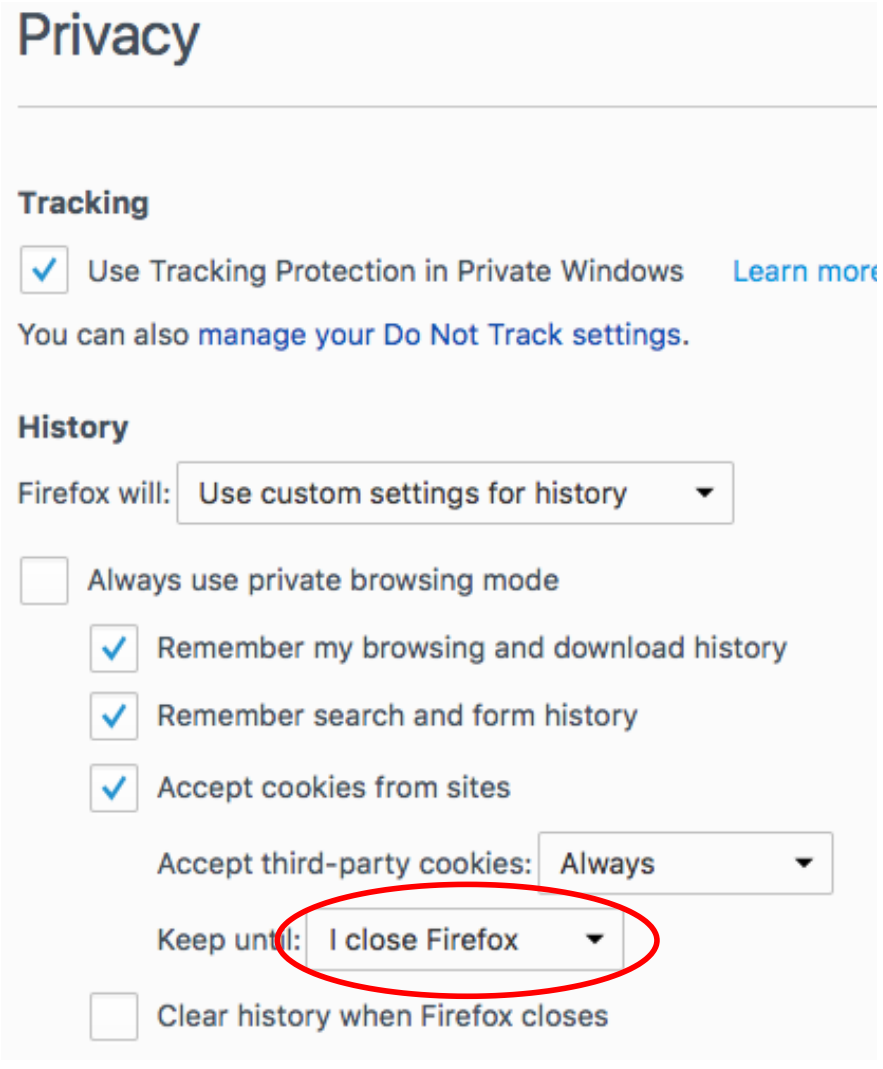
>more

Check your browser settings

You can

- verify your browser privacy settings (in particular Web auto-completion)
- clear cookies upon closing

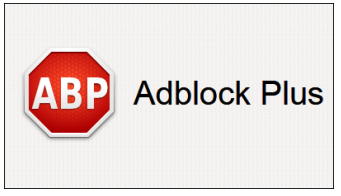
Also check your phone settings.
Third-party tracking is ubiquitous in phone apps.



Check your plugins

You can

- remove toolbar plugins
  (these often just collect information and send it to the provider)
- use privacy plugins



(has deal with advertisers)

See how companies can track you:

- https://fingerprint.pet-portal.eu/
- https://browserleaks.com/canvas
- https://panopticlick.eff.org/

Watch out with Browser Referrals

Via the referrer header, the browser tells a Web site which Web site you came from. This may include form data.

https://www.healthcare.gov/see-plans/85601/results/?
county=04019&age=40&smoker=1&parent=1
&pregnant=1&mec=&zip=85601&state=AZ
&income=35000& &step=4?

Overview

Protecting data against

- yourself
- hackers
- evil interlocutors
- companies
- governments

Protecting against the government?

- if you are a political activist in a less democratic country
- if you want to guard against social scoring (in China, your social media interactions will determine your rights.)
- if you generally don't like the government spying on you

In the United Arab Emirates, showing sympathy to Qatar (also on social media) is punishable by up to 15 years in prison.

In some cities in China, jaywalkers are identified by face recognition and then publicly shamed.

->Email

Protecting against the government?

- if you are a political activist in a less democratic country
- if you want to guard against social scoring (in China, your social media interactions will determine your rights.)
- if you generally don't like the government spying on you

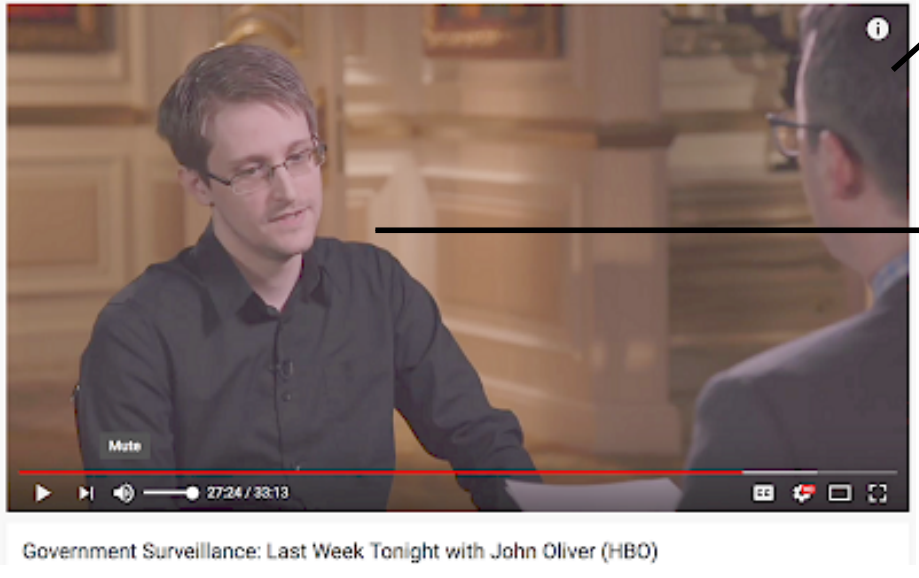
Government Surveillance: Last Week Tonight with John Oliver (HBO)

Last Week Tonight with John Oliver: Government Surveillance

If I had knowledge that the US government had a picture of my d*ck, I would be very p*ssed.

Protecting against the government?

- if you are a political activist in a less democratic country
- if you want to guard against social scoring (in China, your social media interactions will determine your rights.)
- if you generally don't like the government spying on you



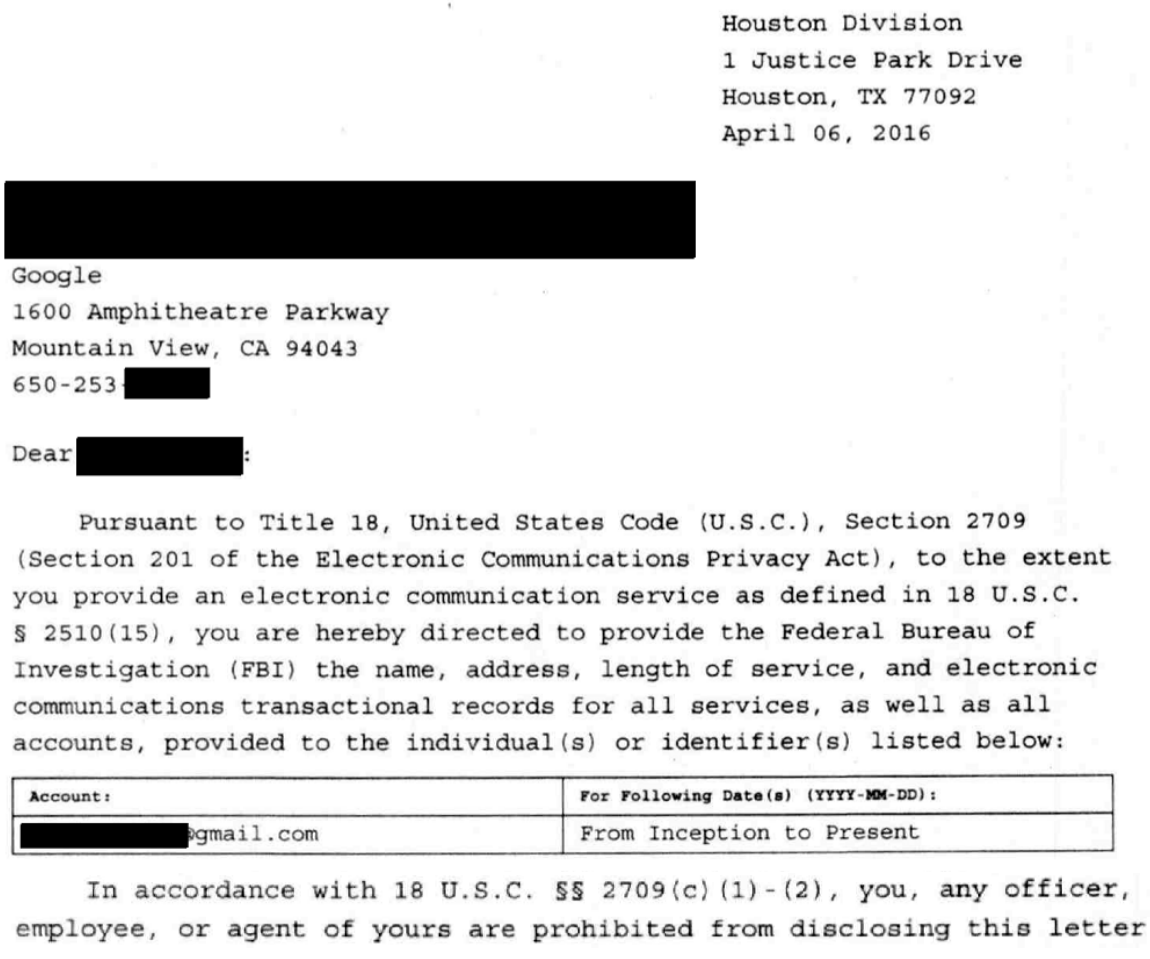
Last Week Tonight with John Oliver: Government Surveillance

Can they see my d*ck?

Yes. [...]
Anytime you have your picture on gmail [...], your junk ends up in the[ir] database.

National Security Letters

The NSA can request that a service provider turns over client data without telling the client about it.



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

In accordance with 18 U.S.C. §§ 2709(c)(1)-(2), you, any officer, employee, or agent of yours are prohibited from disclosing this letter

NSL to Google

Content requests

The FISA Amendments Act, passed in 2008, authorizes the government to require US companies to provide information and the content of communications associated with the accounts of non-U.S. citizens or non-lawful permanent residents who are located outside the US.

Content requests

A content request implicates content held in a user's account, such as Gmail messages, documents, photos, and videos on YouTube.

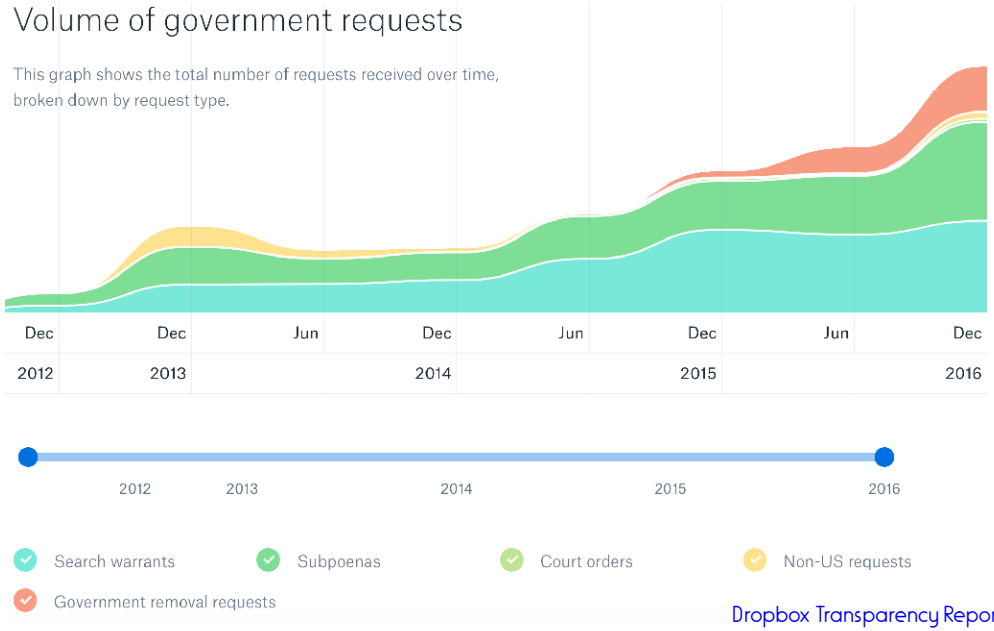
Reporting period	Number of requests	Users/accounts
Jul 2016 – Dec 2016	500 – 999	35000 – 35499
Jan 2017 – Jun 2017	Data subject to six month reporting delay	Data subject to six month reporting delay

Google Transparency Report

Content requests

Dropbox guiding principles

- "Online services should be allowed to report the exact number of government data requests received"
- "Government data requests should be limited to specific people and investigations"
- "Governments should never install backdoors into online services or compromise infrastructure to obtain user data"



Dropbox Transparency Report

≈ 1000

Using government-style protection

You may also want to use this type of protection for

- personal problems you want to keep secret
- business secrets that the competition may not know
- activity that is (moral but) illegal in your country
- communication with whistle-blowers
- traits that are despised by society (e.g, being atheist in Bangladesh)
- activism against powerful or violent people
- sharing data that can be abused: credit card numbers, scans of your personal id card, etc.

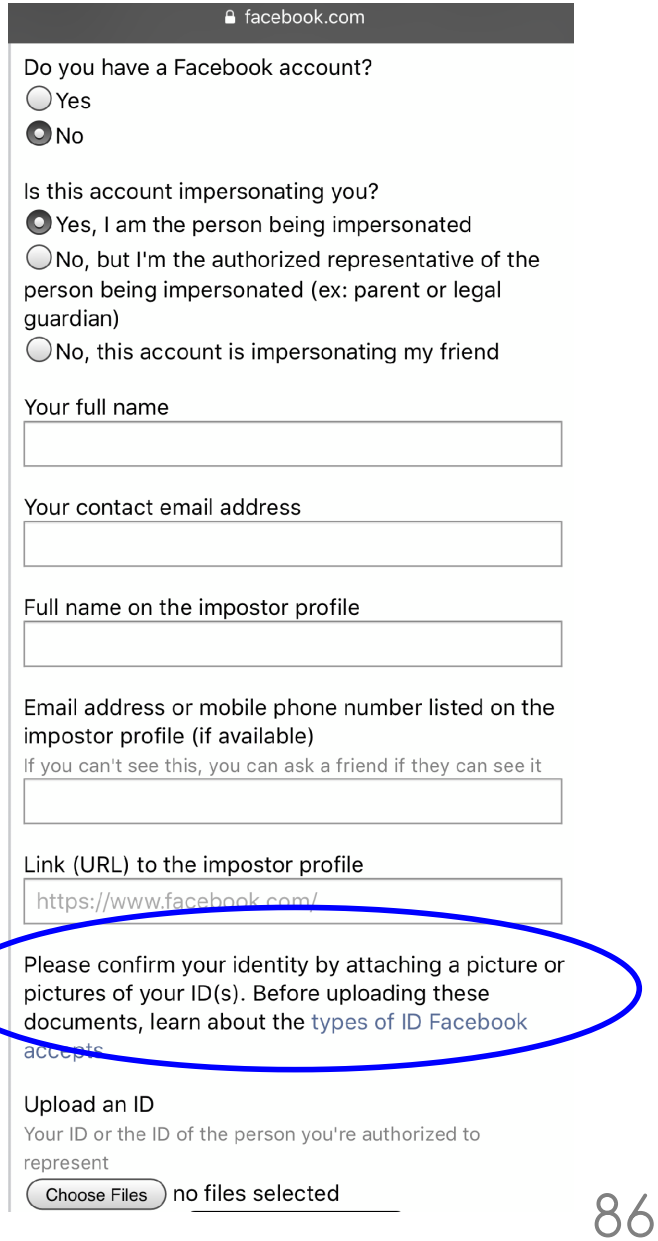
...or if the law requires you to encrypt your data:

- lawyers, notaries, tax consultants
- doctors, pharmacists
- owners of critical infrastructure

Using government-style protection

Scans of your personal id
are often used to verify
your identity...

...for example to identify
imposter accounts on Facebook.



So don't give away scans of your id
easily!

Oh, and don't send them by email!

Def: End-to-end encrypted storage

An end-to-end encrypted cloud service encrypts your data on your device before uploading it into the cloud.

- The service provider can't read it (and cannot hand it to the gov.)
- Nobody in the middle can read it
- The service provider cannot hand it to a government
- If the service provider is hacked, the data is of no use

Disadvantages:

- usually less mainstream, more cumbersome providers
- usually no means to reset your password
- no added data services

Watch out for:

- two-factor authentication
- ability to undelete

Phone, SMS, Chats

- the government can have access to the phone meta data
- SMS can be intercepted
- chats can be read by the service provider (and handed over to the government), unless they're end-to-end encrypted.

End-to-end encrypted chat systems include the following:

WhatsApp

BUT: see privacy

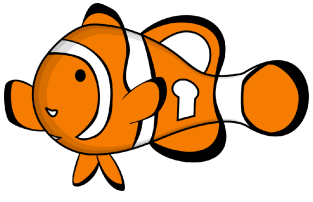
iMessage

Telegram

Signal Private Messenger

(recommended by Edward Snowden)

Single provider



OMEMO + XMPP

ChatSecure

Open to several providers

Email encryption

Solution 0: send encrypted cloud storage link ("share link")

- Advantage: works out of the box if you have such a cloud storage
- Disadvantage: cumbersome for sender and recipient

Solution 1: encrypted email services.

- Advantage: easy to use
- Disadvantage: provider dependency (works only inside 1 provider)

Solution 2: SMIME

- Advantage: well implemented
- Disadvantage: relies on central authority

Solution 3: Public/private keys

- Advantage: provider-independent, decentralized
- Disadvantage: cumbersome to use

Def: Symmetric key algorithm

A symmetric key algorithm is the following metjod for exchanging an emcrypted message between Alice and Bob:

1) Alice and Bob agree secretly on a secret password p
2) When Alice wants to send a message, she encrypts it with p, and Bob decrypts it with p

Caesar cipher algorithm: p is an integer number, and each letter of the original message is shifted by p before transmission. E.g., for $p=4$, "A" becomes "E", "Y" becomes "C", etc. => Can be cracked easily

Used until 1915 by the Russian Army, and also in 2011 by an Islamic terrorist.

>DH

Problem with all of these: Alice and Bob have to agree secretly on p!

Def: Symmetric key algorithm

A symmetric key algorithm is the following metjod for exchanging an emcrypted message between Alice and Bob:

1) Alice and Bob agree secretly on a secret password p
2) When Alice wants to send a message, she encrypts it with p, and Bob decrypts it with p

Caesar cipher algorithm: p is an integer number, and each letter of the original message is shifted by p before transmission. E.g., for $p=4$, "A" becomes "E", "Y" becomes "C", etc. => Can be cracked easily

Used until 1915 by the Russian Army, and also in 2011 by an Islamic terrorist.

Vigenère cipher: p is a word, and the letter at position i of the original message is shifted by $p_{i \bmod |p|}$. E.g., for $p=$"DIG", the letter at position 1 is shifted by 4, at position 2 by 9, at position 3 by 7, then again by 4, 9, 7 etc.

Today: Advanced Encryption Standard (AES)

>DH

Problem with all of these: Alice and Bob have to agree secretly on p!

Binary primitive root

An odd number p has a binary primitive root, if

$\forall\, a \in \{1,\ldots,p-1\}: \exists\, k : (2^k \bmod p) = a.$

(The full definition extends to roots that are not binary.)

Example: $p=5$ has a binary primitive root, because

for any $a=1\ldots p-1$	I can choose k	so that $(2^k \bmod p)=a$.
1	0	$2^0=1,\ (1 \bmod 5)=1$
2	1	$2^1=2,\ (2 \bmod 5)=2$
3	3	$2^3=8,\ (8 \bmod 5)=3$
4	2	$2^2=4,\ (4 \bmod 5)=4$
what I want to encode	what I send instead	How to get back to the original number

Other examples are $p=3, 9, 11, 25, 27, \ldots$

Def: Diffie-Hellman key exchange

The Diffie-Hellman[-Merkle] key exchange is the following method for sending encrypted messages between Alice and Bob

1) Alice and Bob publicly agree on a prime number p
   with a binary primitive root (or any other public primitive root)
2) Alice chooses a secret number $a<p$ and sends $A=2^a \bmod p$
3) Bob chooses a secret number $b<p$ and sends $B=2^b \bmod p$
4) Alice computes $s_a=B^a \bmod p$
5) Bob computes $s_b=A^b \bmod p$

Now, $s_b = A^b \bmod p = (2^b \bmod p)^a \bmod p = 2^{a\times b} \bmod p$
$= (2^a \bmod p)^b \bmod p = B^a \bmod p = s_a.$

Hence, s_a can serve for symmetric encryption.

Def: Diffie-Hellman key exchange

The Diffie-Hellman[-Merkle] key exchange is the following method for sending encrypted messages between Alice and Bob

1) Alice and Bob publicly agree on a prime number p $p=11$
   with a binary primitive root (or any other public primitive root)
2) Alice chooses a secret number $a<p$ and sends $A=2^a \bmod p$ $a=9,\ A=6$
3) Bob chooses a secret number $b<p$ and sends $B=2^b \bmod p$ $b=5,\ B=10$
4) Alice computes $s_a=B^a \bmod p$ $s_a=10$
5) Bob computes $s_b=A^b \bmod p$ $s_b=10$

Now, $s_b = A^b \bmod p = (2^b \bmod p)^a \bmod p = 2^{a\times b} \bmod p$

$= (2^a \bmod p)^b \bmod p = B^a \bmod p = s_a.$

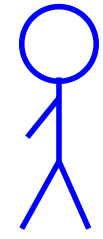
Hence, s_a can serve for symmetric encryption.

RSA

RSA (Rivest-Shamir-Adleman) is a method for sending an encrypted message that works as follows (simplified):

1. Choose your public key pair n, e

Public key:
e=7, n=33



Barack

RSA

RSA (Rivest-Shamir-Adleman) is a method for sending an encrypted message that works as follows (simplified):

1. Choose your public key pair n, e
2. Compute your private key d such that

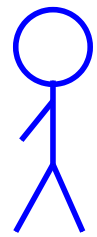
$\forall\, m\colon (m^{e\times d}\ mod\ n) = m\ mod\ n$

Public key:

$e=7$, $n=33$

Private key:

$d=3$



Barack

RSA

RSA (Rivest-Shamir-Adleman) is a method for sending an encrypted message that works as follows (simplified):

1. Choose your public key pair n, e
2. Compute your private key d such that
   $\forall\, m$: $(m^{e\times d}\ mod\ n)=m\ mod\ n$
3. The sender encrypts a message m as $c=m^{e}\ mod\ n$

Public key:

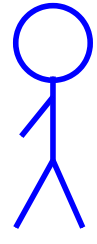
$e=7$, $n=33$

Private key:

$d=3$

Angela

My message: $m=2$

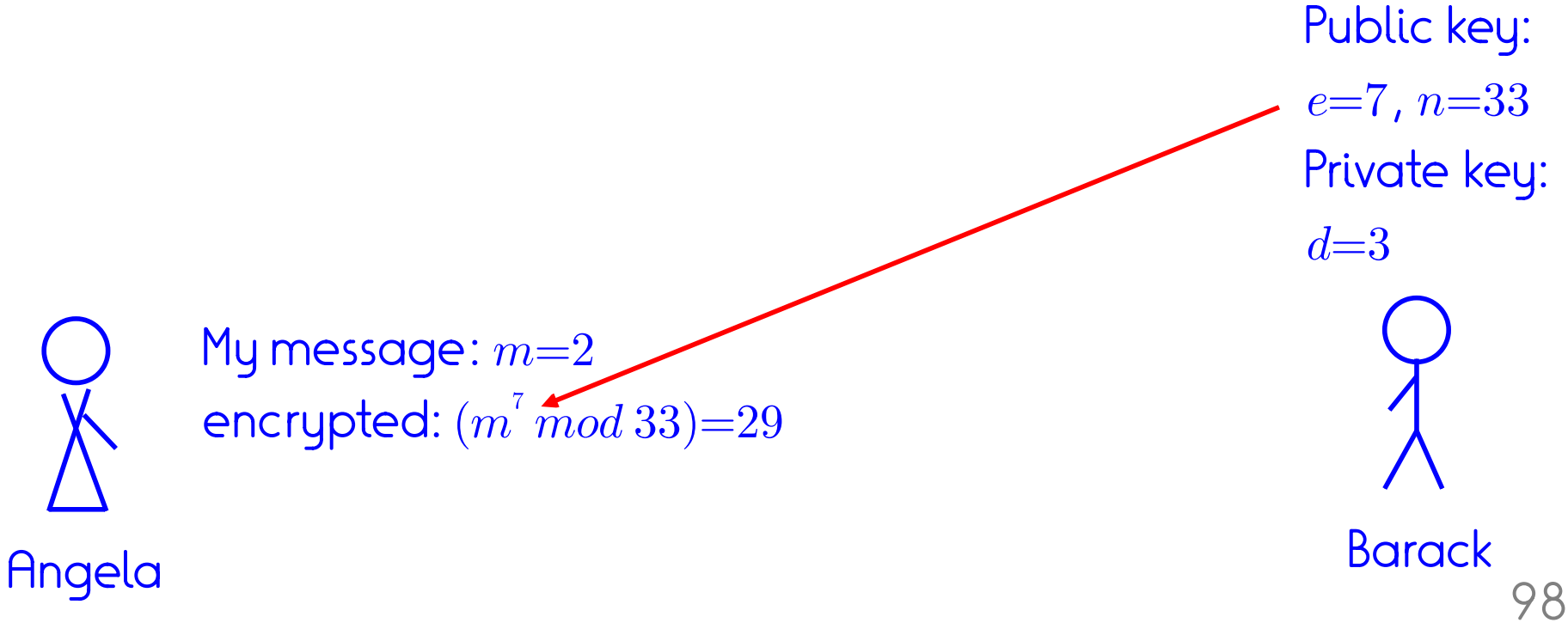


Barack

RSA

RSA (Rivest-Shamir-Adleman) is a method for sending an encrypted message that works as follows (simplified):

1. Choose your public key pair n, e
2. Compute your private key d such that
   $\forall\, m$: $(m^{e\times d}\ mod\ n)=m\ mod\ n$
3. The sender encrypts a message m as $c=m^{e}\ mod\ n$



RSA

RSA (Rivest-Shamir-Adleman) is a method for sending an encrypted message that works as follows (simplified):

1. Choose your public key pair n, e
2. Compute your private key d such that
   $\forall\, m$: $(m^{e\times d}\ mod\ n)=m\ mod\ n$
3. The sender encrypts a message m as $c=m^{e}\ mod\ n$
4. The encrypted message c is sent

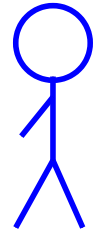
Public key:
$e=7$, $n=33$
Private key:
$d=3$

Angela

My message: $m=2$
encrypted: $(m^{7}\ mod\ 33)=29$

29 received

Donald cannot understand 29



Barack

Def: RSA

RSA (Rivest-Shamir-Adleman) is a method for sending an encrypted message that works as follows (simplified):

1. Choose your public key pair n, e
2. Compute your private key d such that
   $\forall\, m\colon (m^{e\times d} \bmod n) = m \bmod n$
3. The sender encrypts a message m as $c = m^{e} \bmod n$
4. The encrypted message c is sent
5. You decrypt the message c as
   $c^{d} \bmod n = (m^{e} \bmod n)^{d} \bmod n = m^{e\times d} \bmod n = m$

Public key:
$e=7$, $n=33$
Private key:
$d=3$

Angela

29 received,
compute
$(29^{3} \bmod 33) = 2$

Barack

>details

Def: RSA

RSA (Rivest-Shamir-Adleman) is a method for sending an encrypted message that works as follows (simplified):

1. Choose your public key pair n, e
2. Compute your private key d such that

$\forall\, m\colon (m^{e\times d}\ mod\ n)=m\ mod\ n$

How do you find such d?
And why can the attacker not also find that d?

Public key:
$e=7$, $n=33$
Private key:
$d=3$

Angela

29 received,
compute
$(29^3\ mod\ 33)=2$

Barack

RSA in detail

RSA (Rivest-Shamir-Adleman) is the following encryption method:

1. Choose secret primes p,q, compute $n=p\times q$, compute $\phi(n)=(p-1)\times(q-1)$ — $p=3$, $q=11$, $n=33$; $\phi(n)=2\times 10=20$
2. Choose a number $e<\phi(n)$ co-prime with $\phi(n)$, publish e and n. — usually e is prime; $e=7$
3. Compute your private key d such that $((e\times d)\ mod\ \phi(n))=1$. — $d=3$, because $((7\times 3)\ mod\ 20)=1$

   By Euler's theorem, this implies

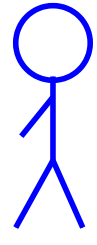
   $\forall\, m$: $(m^{e\times d}\ mod\ n)=(m\ mod\ n)$ — $\forall\, m$: $(m^{7\times 3}\ mod\ 33)=(m\ mod\ 33)$
4. Encrypt m as $c=m^{e}\ mod\ n$, decrypt c as $c^{d}\ mod\ n = m^{e\times d}\ mod\ n = m$

Angela

My message: $m=2$

encrypted: $(m^{7}\ mod\ 33)=29$

29 received, compute $(29^{3}\ mod\ 33)=2$



Barack

RSA in detail

RSA (Rivest-Shamir-Adleman) is the following encryption method:

1. Choose secret primes p,q, compute $n=p\times q$, compute $\phi(n)=(p-1)\times(q-1)$

   $p=3$, $q=11$, $n=33$

   $\phi(n)=2\times 10=20$

2. Choose a number $e<\phi(n)$ co-prime with $\phi(n)$, publish e and n.

   usually e is prime

   $e=7$

3. Compute your private key d such that

   $((e\times d)\ mod\ \phi(n))=1$.

   By Euler's theorem, this implies

   $\forall\, m$: $(m^{e\times d}\ mod\ n)=(m\ mod\ n)$

   d can be computed only by knowing $\phi(n)$, i.e. by knowing p and q.

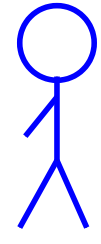
4. Encrypt m as $c=m^e\ mod\ n$, decrypt c as $c^d\ mod\ n = m^{e\times d}\ mod\ n = m$

Angela

My message: $m=2$

encrypted: $(m^7\ mod\ 33)=29$

29 received, compute $(29^3\ mod\ 33)=2$



Barack

RSA in detail

RSA (Rivest-Shamir-Adleman) is the following encryption method:

1. Choose secret primes p,q, compute $n=p\times q$, compute $\phi(n)=(p-1)\times(q-1)$

   $p=3$, $q=11$, $n=33$

   $\phi(n)=2\times 10=20$

2. Choose a number $e<\phi(n)$ co-prime with $\phi(n)$, publish e and n.

   usually e is prime

   $e=7$

3. Compute your private key d such that

   $((e\times d)\ mod\ \phi(n))=1$.

   By Euler's theorem, this implies

   $\forall\, m$: $(m^{e\times d}\ mod\ n)=(m\ mod\ n)$

4. Encrypt m as $c=m^e\ mod\ n$, decrypt c as $c^d\ mod\ n = m^{e\times d}\ mod\ n = m$

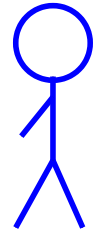
In practice, m is a random symmetric password that is used to encrypt the payload.

Angela

My message: $m=2$

encrypted: $(m^7\ mod\ 33)=29$

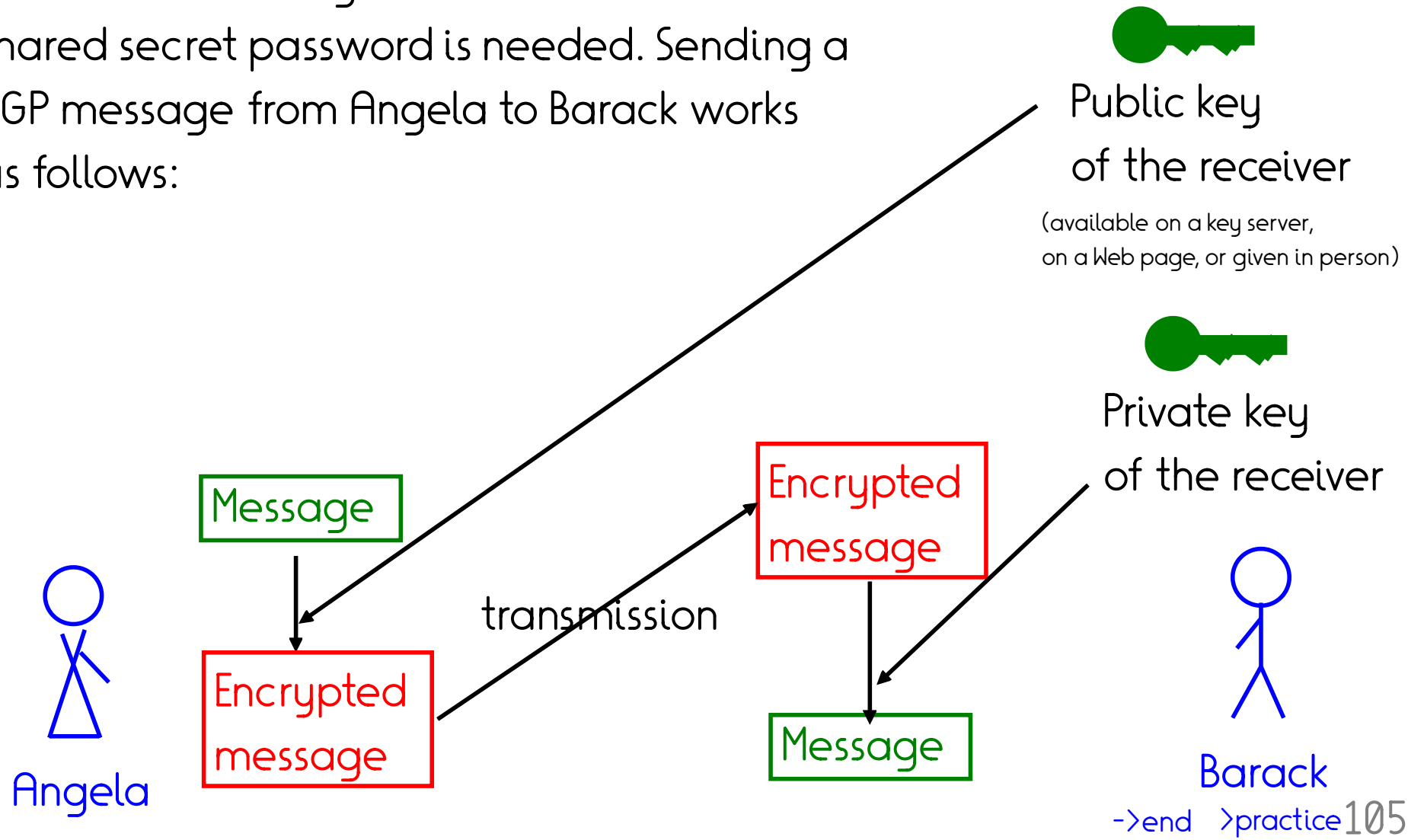
29 received, compute $(29^3\ mod\ 33)=2$



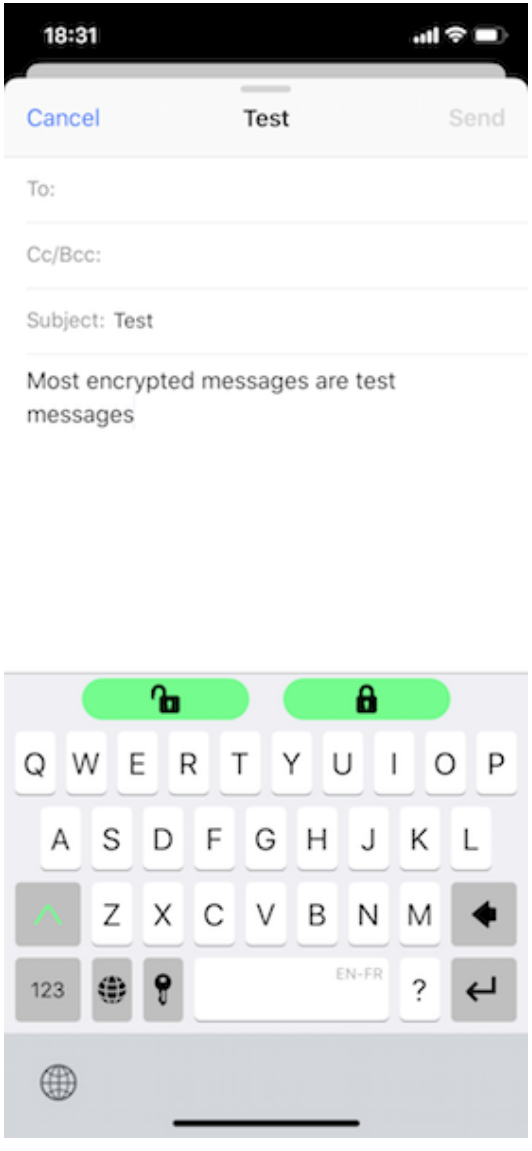
Barack

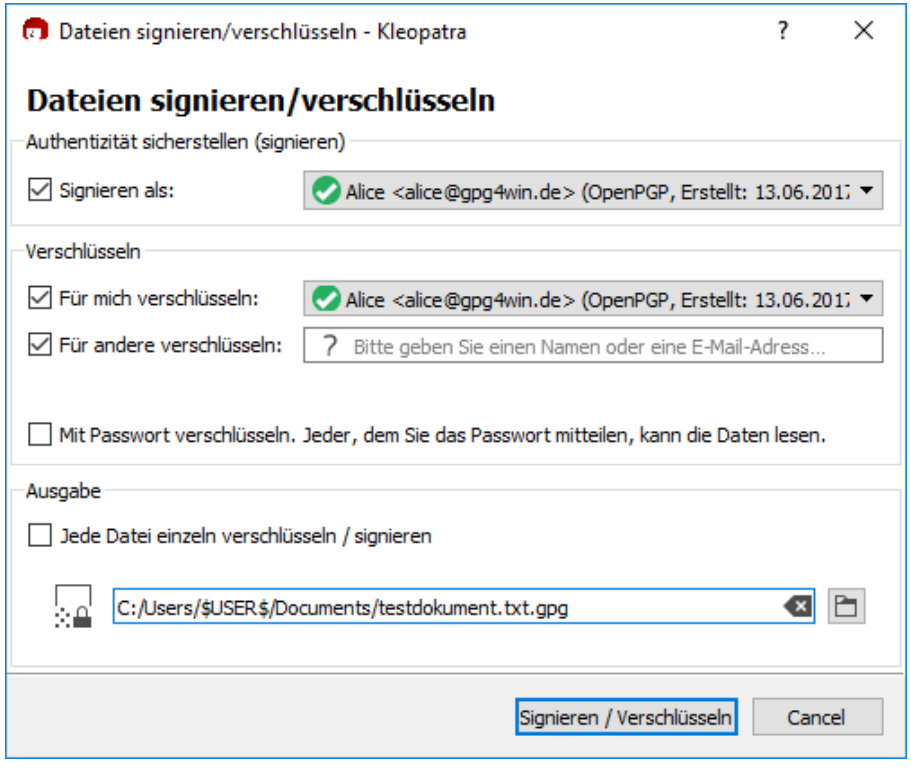
Def: PGP

PGP (Pretty Good Privacy) is an encryption program that implements RSA. The main advantage of PGP (and RSA) is that no shared secret password is needed. Sending a PGP message from Angela to Barack works as follows:



PGP encryption in practice

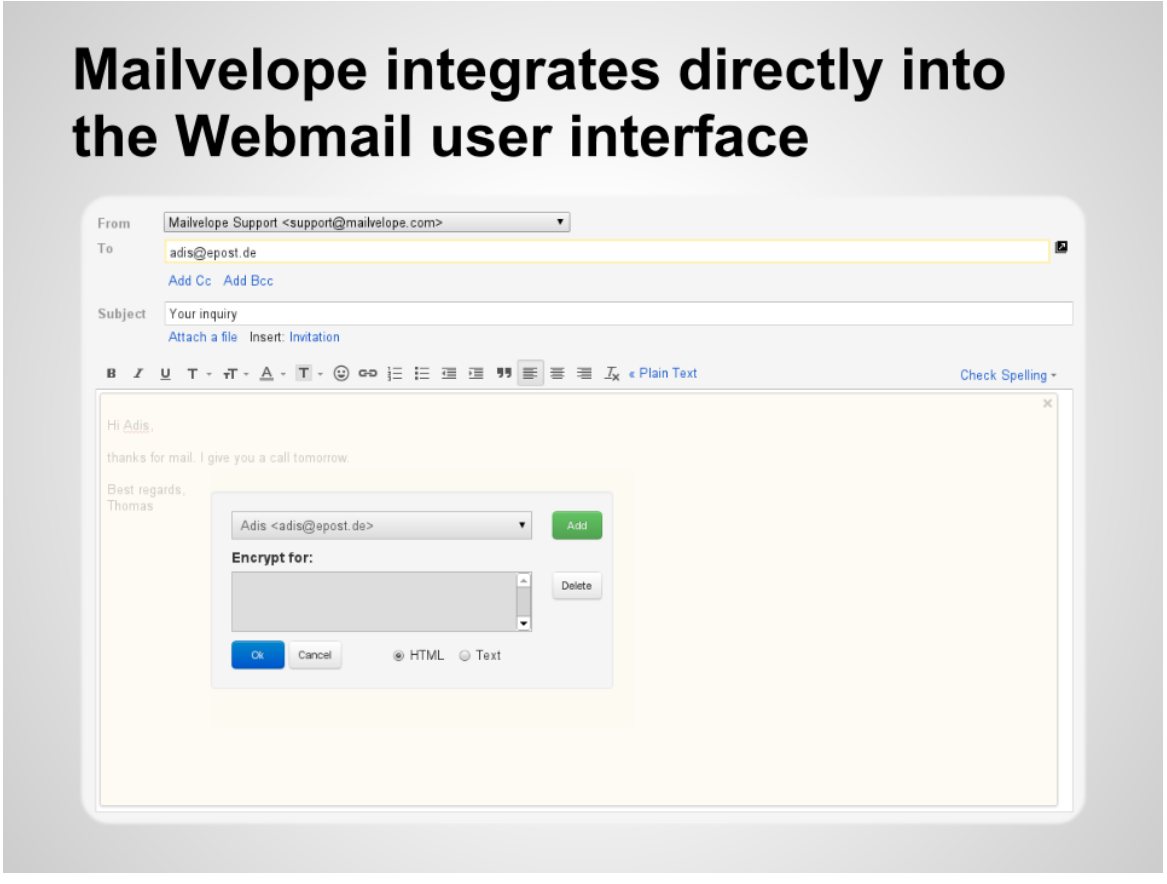




On a Windows machine,
e.g., with PGP4win

On a phone, e.g.,
with PGPeverywhere or CanaryMail

PGP encryption in practice

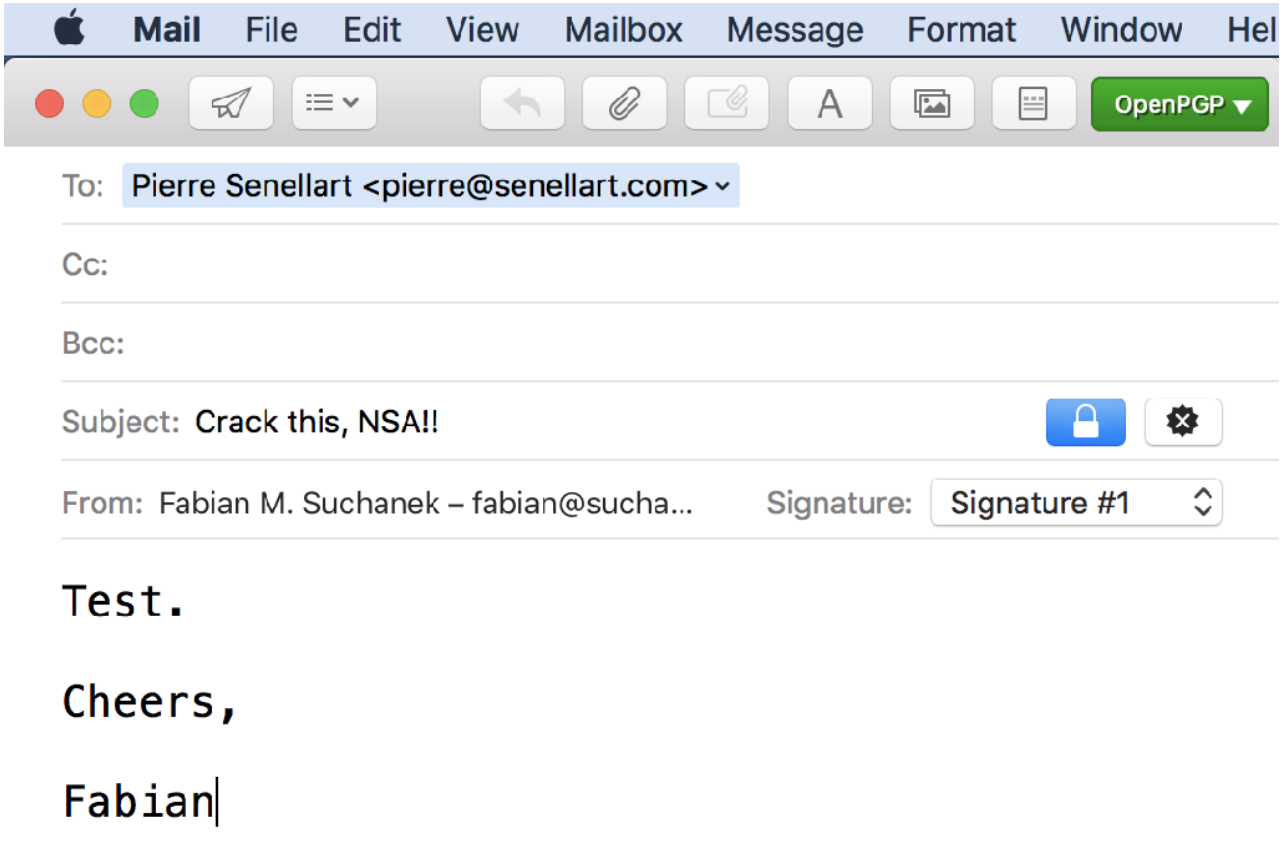


In a browser, e.g.
with Mailvelope

PGP encryption in practice

In Thunderbird, e.g., with the Enigmail plugin

PGP encryption in practice

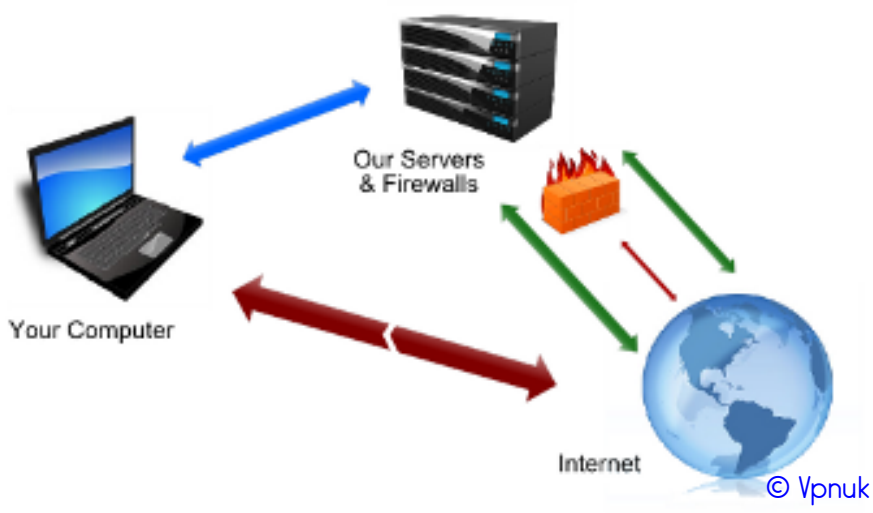


On a Mac, e.g., with GPGSuite

->end

Def: Virtual Private Networks

A Virtual Private Network (VPN) is a software that encapsulates your data from your computer to the Internet in a secure tunnel.



=> Nobody can see it's you who accesses the page

Useful if

- you want to access a page that is blocked in your country
- you want to hide that you're accessing the page
- you are using an unprotected Wifi

Virtual Private Networks

A Virtual Private Network (VPN) is a software that encapsulates your data from your computer to the Internet in a secure tunnel.

=> Nobody can see it's you who accesses the page

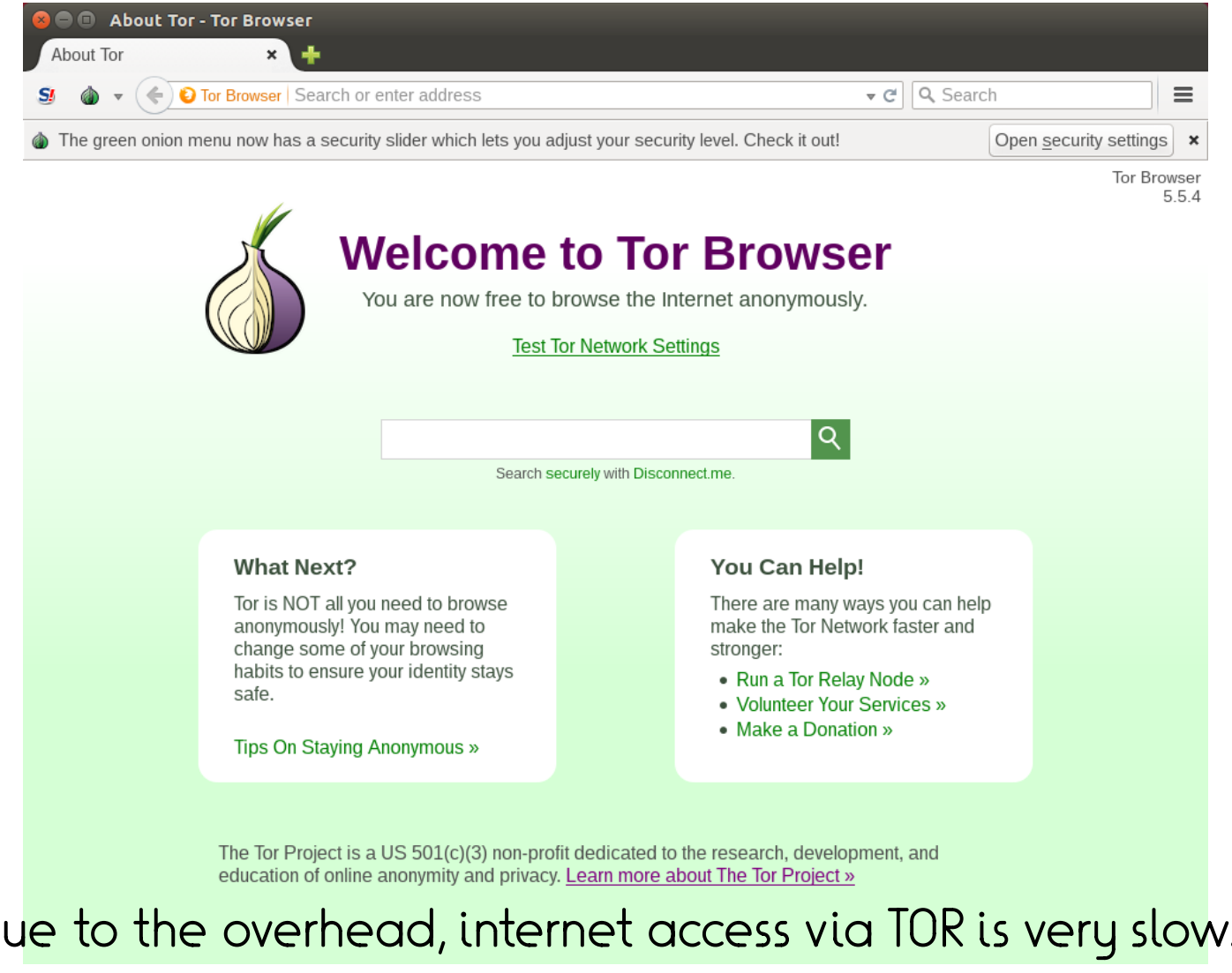
But:

- the Internet access is usually slower
- VPNs are usually not for free
- some free VPNs sell your data for marketing purposes
- if you want to protect the data, not the visit, HTTPS is sufficient

->end

TOR

The TOR browser routes your queries through a distributed network, thus thwarting any tracking.



But: Due to the overhead, internet access via TOR is very slow.

->end

Happy encrypting :-)

Protecting data against

- yourself
- hackers
- evil interlocutors
- companies
- governments