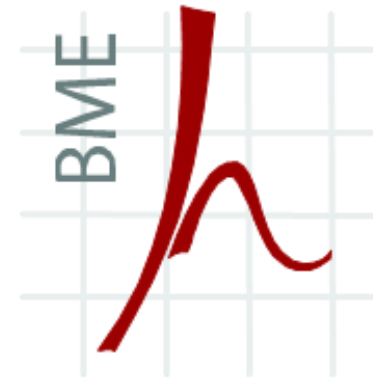




ECONOMIC SURVEILLANCE ON THE WEB: CUTTING-EDGE TECHNOLOGIES IN PROFILING (AND WHAT CAN WE DO AGAINST IT)

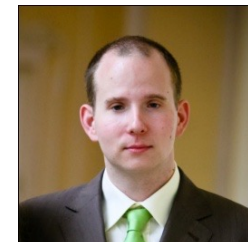
*Living in Surveillance Societies Workshop
October 1-3, 2012*

GULYÁS, Gábor György

Assistant research fellow

Dept. of Telecommunications (BUTE)

gulyasg@hit.bme.hu

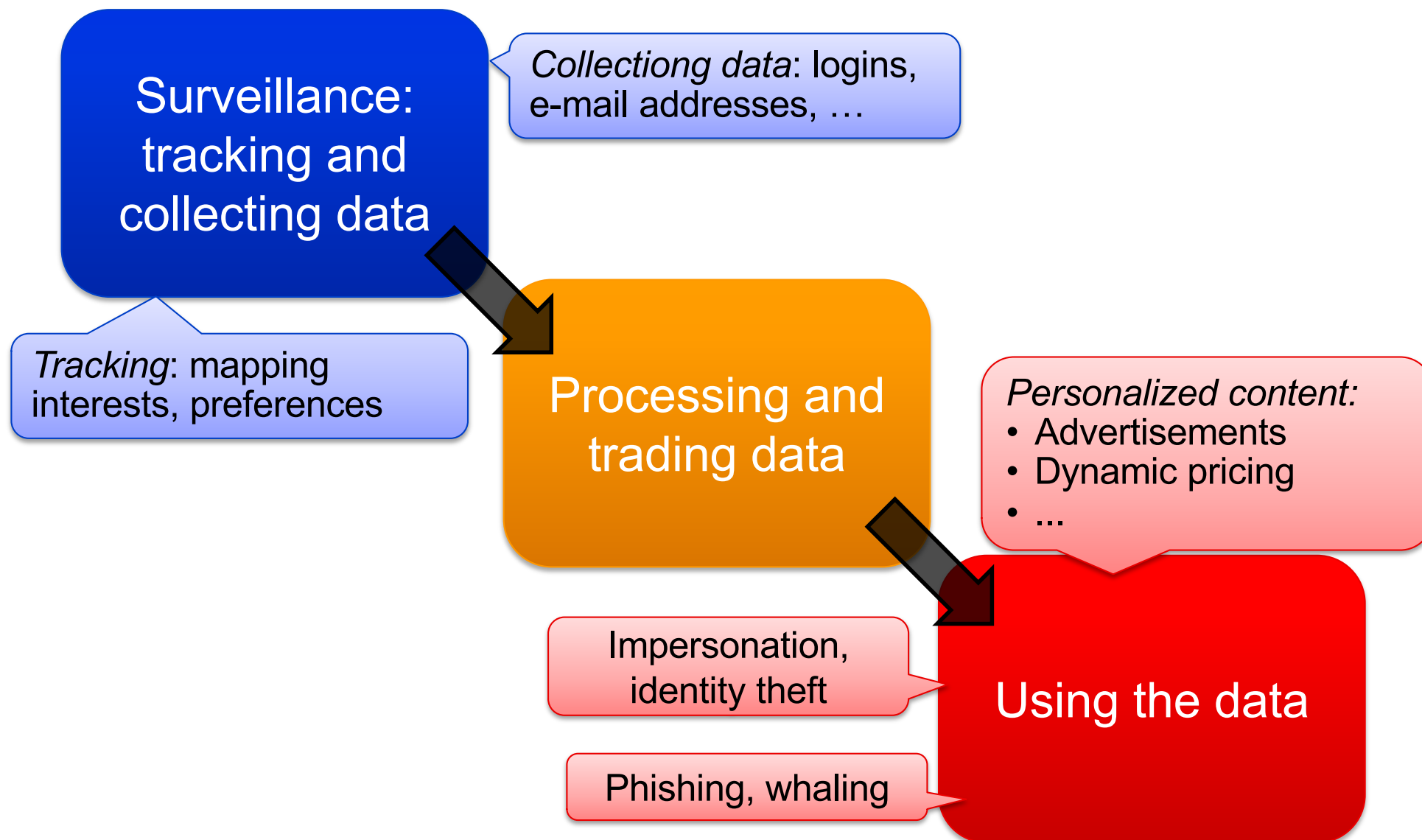


October 2, 2012

Outline (in keywords)



Positioning this talk



Positioning this talk (2)

Who is collecting the data?

	Implicit data	Explicit data
First party	Services of information superpowers	
Third party	Tracking	Public data sources

How serious is it?



The industry is doing well:
\$8.4 billion in Q1 2012!

Continuous increase in

- the *number of covered top sites*,
- the *number of trackers per page* between 2006-2012.

First party trackers:
forced & social networks

Penetration of social
widgets is above 35%

**Broad range of
„infection”:
Alexa top 500 had
7264 instances of
524 trackers**

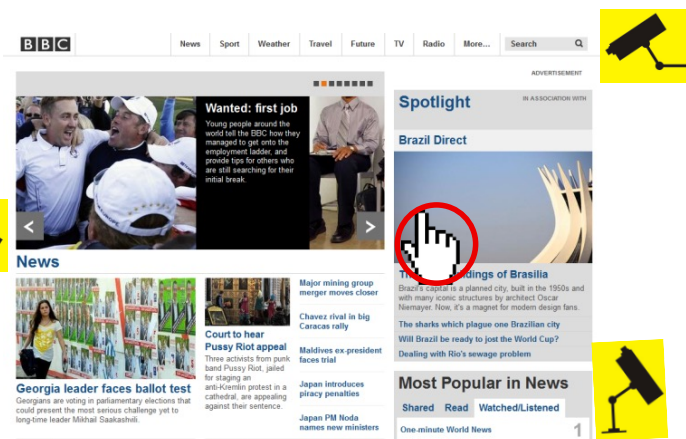
Various other techniques
are used beside cookies.

Cooperation of trackers:
one may imply many!

**Today, it is estimated that 20%+ of
online activities may be monitored!**

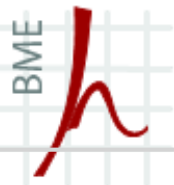
What do they collect?

Visit information



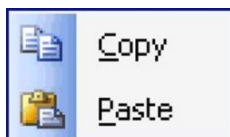
Clickstream





What do they collect? (2)

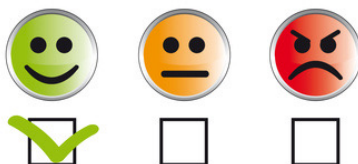
Copy-paste events



Keyboard logging



Content, meaning,

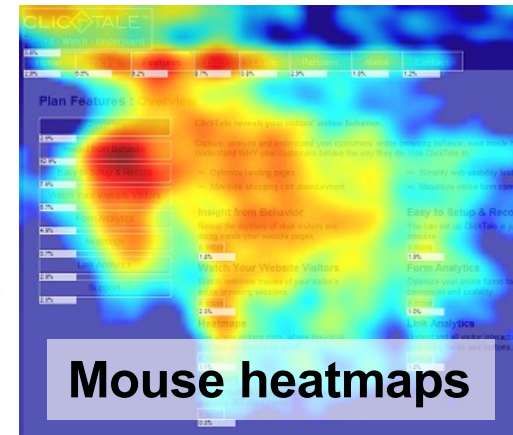


sentiment analysis



What do they collect? (3)

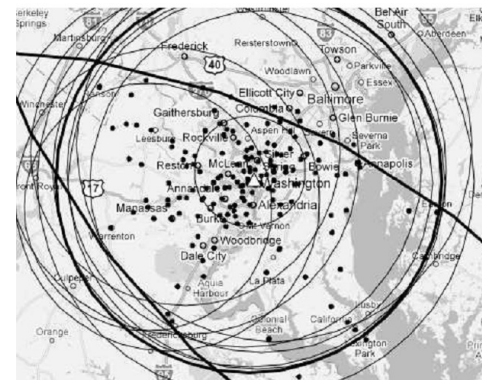
<http://goo.gl/4bUVO>



Regular techniques

IP based tracking

- Not precise → auxiliary information
- Provides geological information (new research indicates <1km accuracy)



Cookie based tracking

Surfer



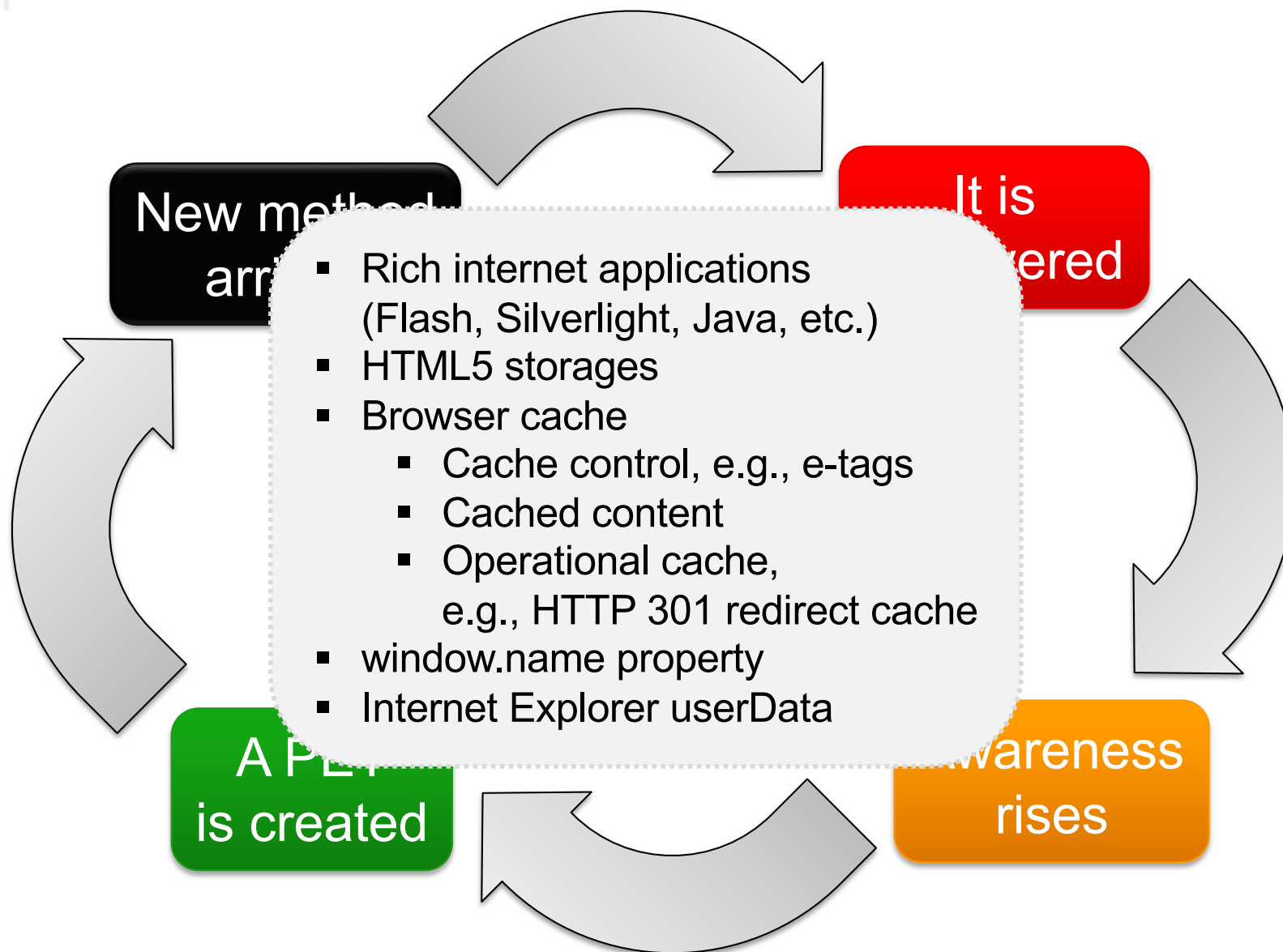
Advertiser



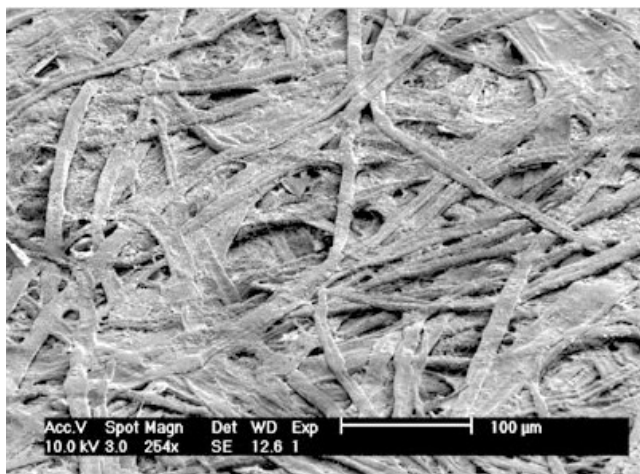
Profile of ID 967

- Buying apples
- Wild life of the savannah

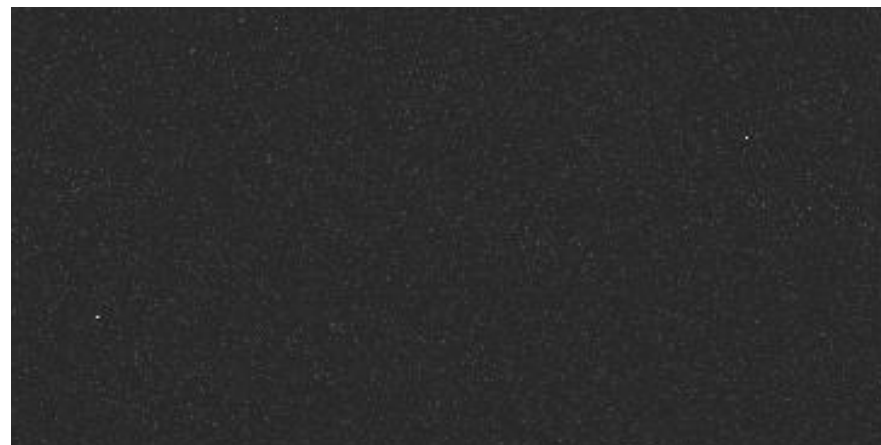
Regular techniques (2)



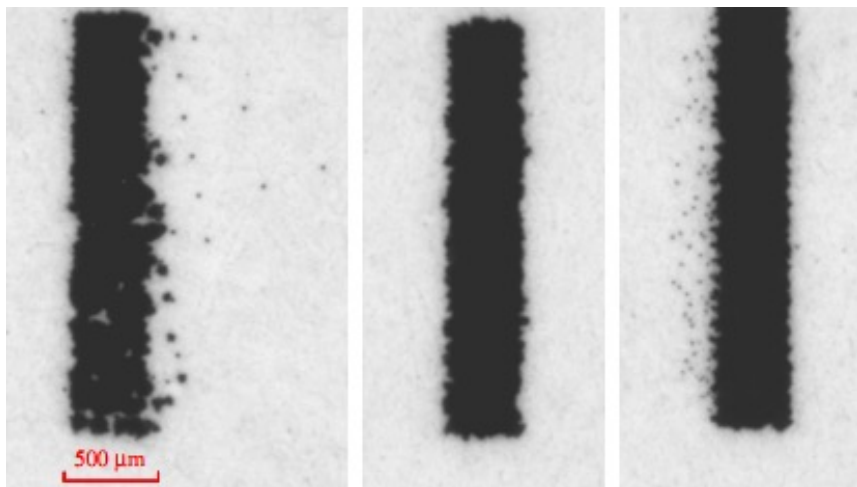
Fingerprinting?



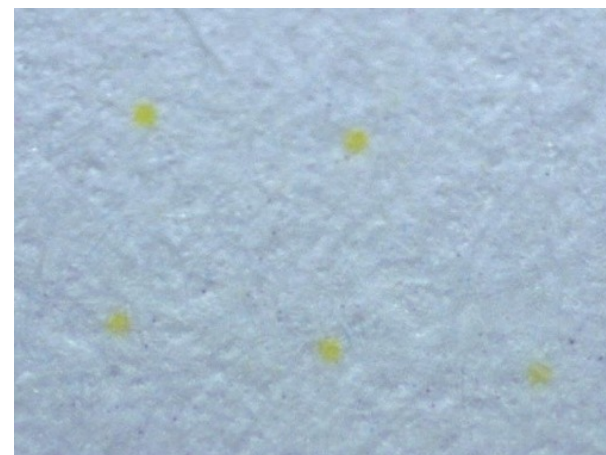
Copy paper



Cameras, scanners



Printers



Unique identifiers added while printing

The Panopticlick project (2010)



A research project of the Electronic Frontier Foundation

Panopticlick

How Unique – and Trackable – Is Your Browser?

Is your browser configuration rare or unique? If so, web sites may be able to track you, *even if you limit or disable cookies.*

Panopticlick tests your browser to see how unique it is based on the **information** it will share with sites it visits. Click below and you will be given a uniqueness score, letting you see how easily identifiable you might be as you surf the web.

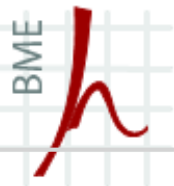
Only **anonymous data** will be collected by this site.



A paper reporting the statistical results of this experiment is now available: **How Unique Is Your Browser?** Proceedings of the Privacy Enhancing Technologies Symposium (PETS 2010), Springer Computer Science.

94.2% of browsers can be fingerprinted!
Changes in fingerprint can be tracked!

<http://panopticlick.eff.org>



Followed by many companies

ThreatMetrix™

iovation®



Wolf Software
Part of The Wolf Consortium

Claiming to identify 99% of online devices (billions!).
Including smart phones, set top boxes, etc.

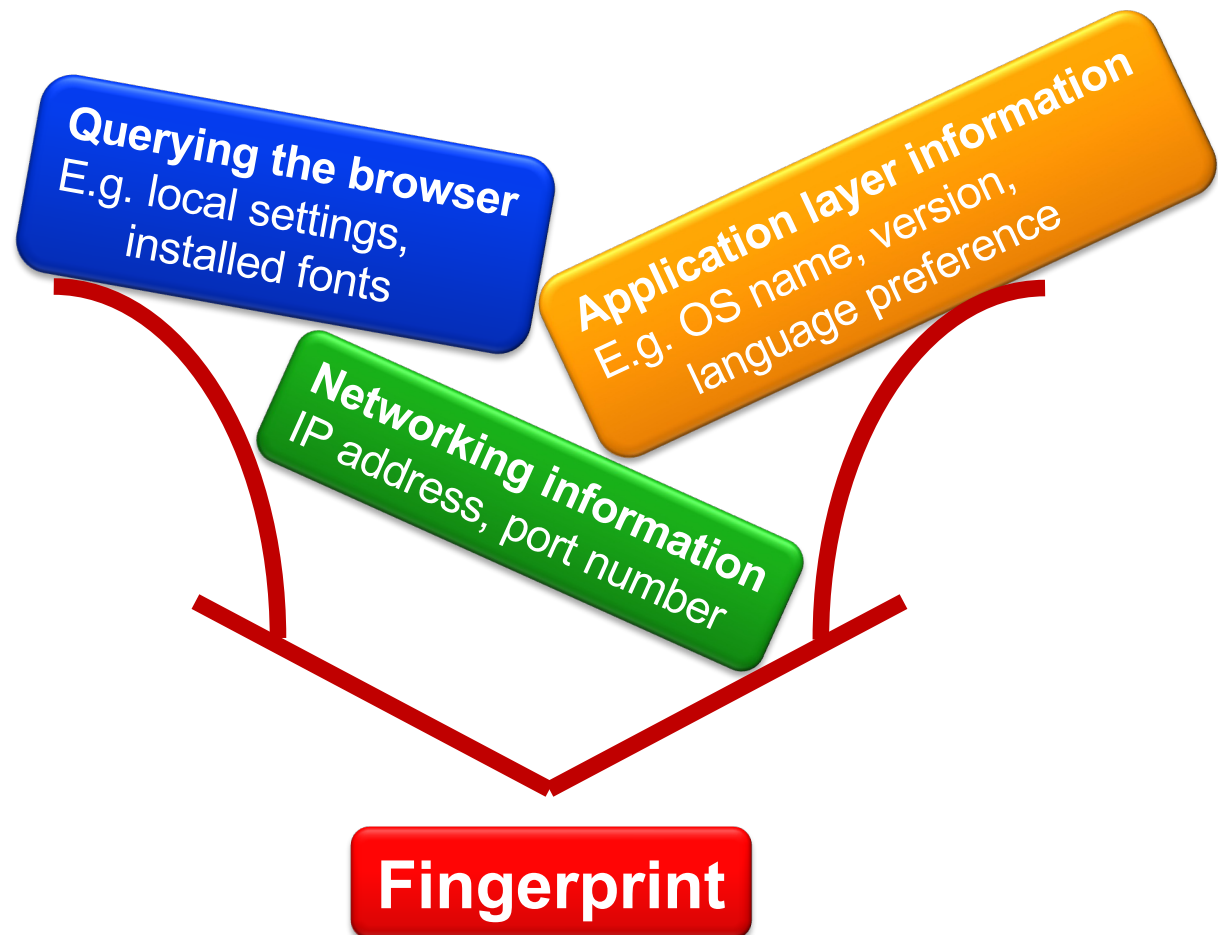
But *how*?

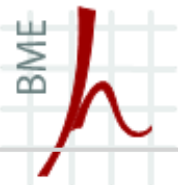
Meanwhile...

Problems of the Panopticlick project:

- Plugin dependency
- Browser dependency

Precise, **if** Flash or Java enabled!





The system fingerprint test

International
PET
portal and blog

BLOG
GALLERIES
ARTICLES
PRESS
COVERAGE
LINKS

FINGERPRINT

STEGOWEB
CONTACT

twitter

Iran blocks Tor, and Tor releases a workaround on the same day: <http://t.co/3wPBJulp> 16 days ago

Identifying Speakers in Encrypted Voice Communication <http://t.co/mB9oxs8R> 28 days ago

Online Tracking Practices Face Increasing Scrutiny <http://t.co/OqEPseq> 34 days ago

http://twitter.com/pet_portal_intl

System fingerprint: a passive, cross-browser user tracking that works even in private mode

Web browsing software needs to share system information with the visited web services. However, shared information can characterize the user's system, and therefore allow individual identification, even in private browsing mode or after deleting HTTP and Flash cookies. In our tests, even a Firefox-based anonymous web browser could be tracked with the same identifier that was assigned to other browsers of the same system.

If you would like to test this tracking feature of your browser, just click on the link below! (Please note that the test is in an experimental phase and you may get different IDs in some rare cases.) After the analysis, the test computes an identifier for your system, which we call the system fingerprint. If you like, you can compare these IDs in different browsers, and even in private browsing mode.

Please read our principles on the management of collected data. If you do not agree, do not test your system! Thank you.

Start the test: I want my fingerprint!

During the test we analyze and collect the following data:

- the IP address in an encrypted form
- name of the operating system
- screen resolution
- time zone setting
- detected font list of the system
- user agent string
- time of the test
- the system fingerprint

We treat the collected data confidentially, and we do not share it with any third parties. The sole objective of the data collection is to test the uniqueness of the generated identifiers, and to determine whether this user tracking method is viable or not. After finishing our tests we are going to delete all collected data.

Thank you for your support! If you are interested in the results, please check back later.

*(Our idea was inspired by the **Panoptlick** project. However, instead of browser fingerprints we intend to test the viability of system fingerprints without using any plugins of the browser.)*

© International PET Portal, 2010 | [Imprint](#) | [Terms of Use](#) | [Privacy Policy](#)

Same fingerprint in all browsers
⇒ pure JavaScript

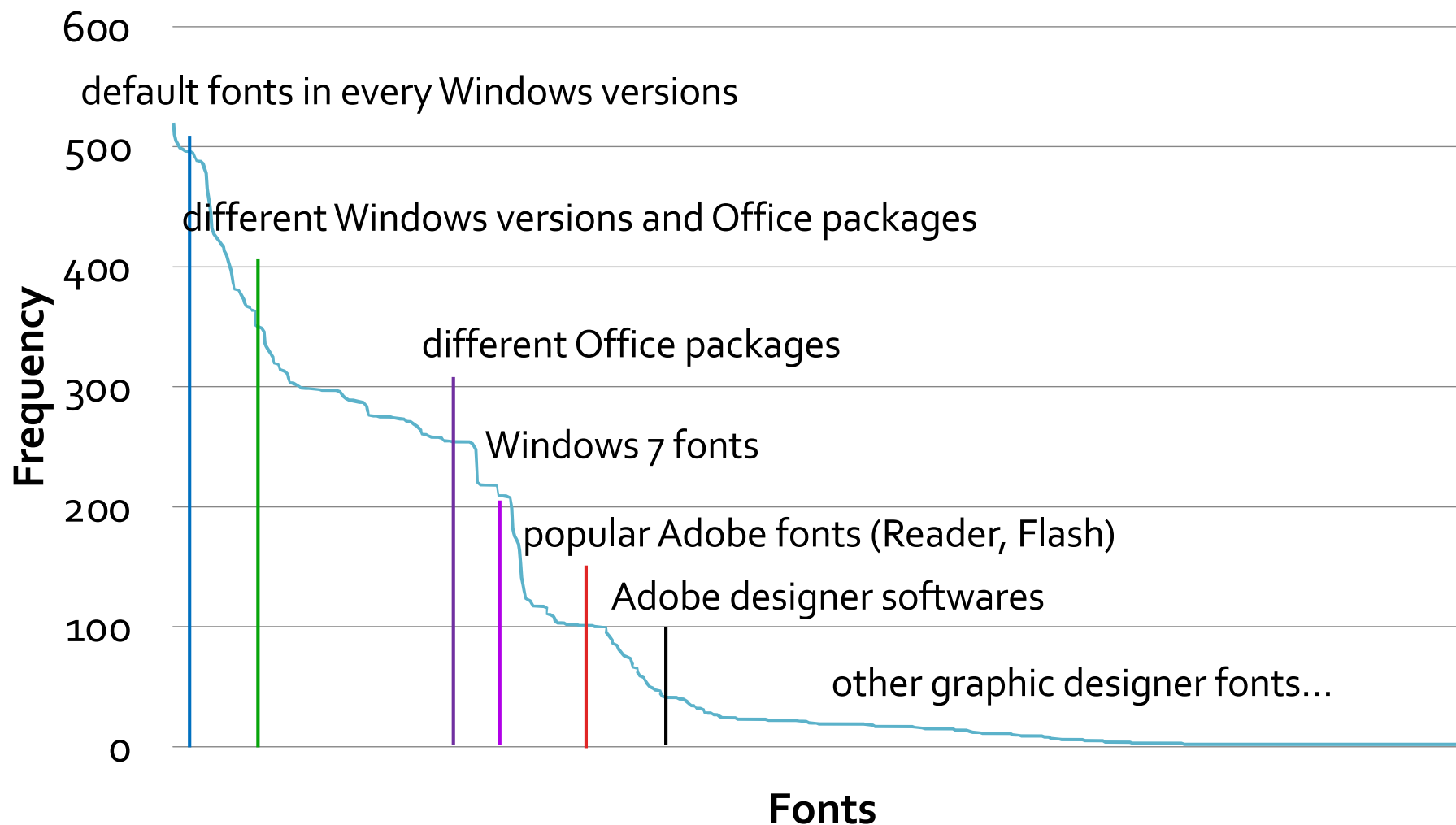
Privacy violation by simple fonts
⇒ a verified entropy source

Works well on Win and OS X:
⇒ the set of installed apps is unique, and rarely changes
⇒ apps install fonts with them

Cross-browser?
⇒ still open question

<http://pet-portal.eu/fingerprint/>

Do fonts indicate installed sw?



The cross-browser fingerprint test 2.0

New font feature set
 ⇒ higher precision expected
 (+ IP not included)

Explicit linking of fingerprints
 ⇒ by an evercookie (if accepted)

Many other small enhancements

Early statistics:

- 14 000+ fingerprints
- Higher user ID entropy
- Smaller UAS entropy

Work is in progress...



<http://fingerprint.pet-portal.eu>

Means of protection

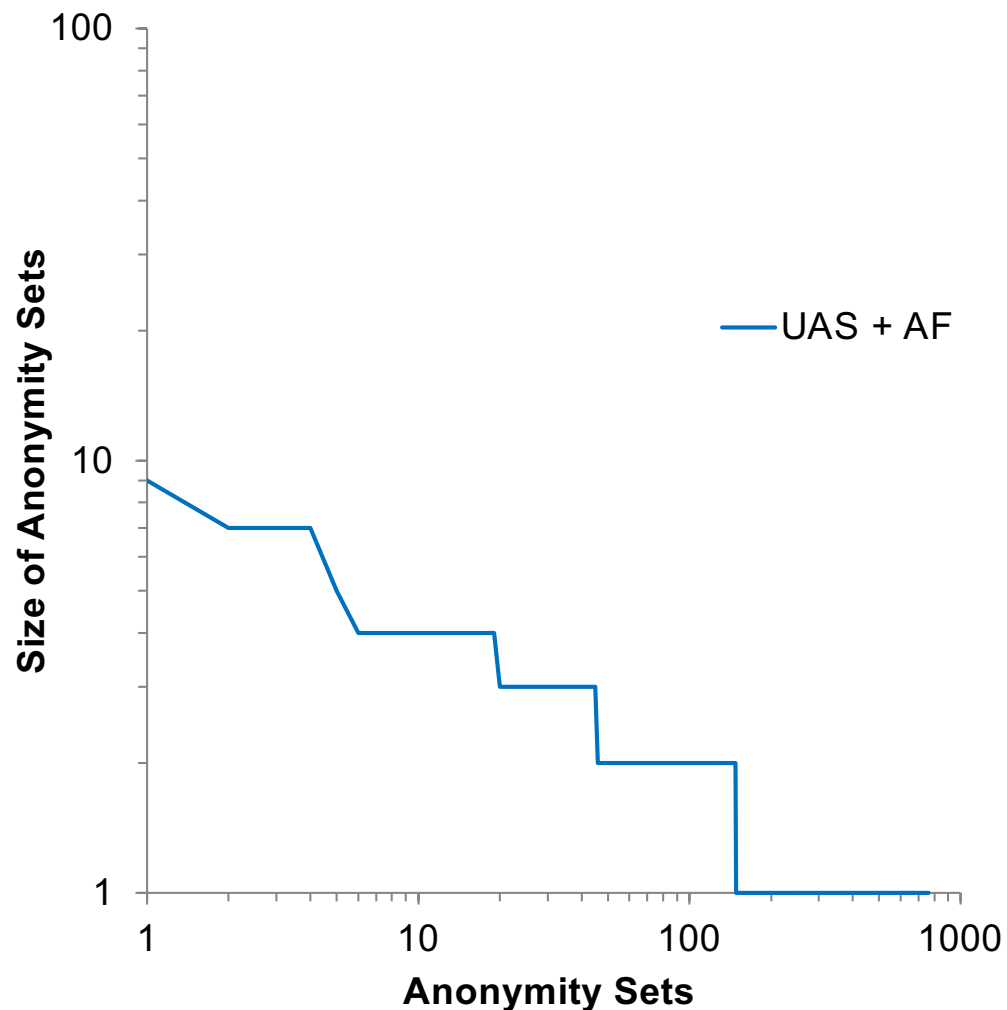
- What are we protecting ourselves from?

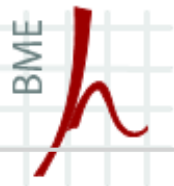
Variable	Entropy (b)	
user_agent	10.0	JavaScript
plugins	15.4	
fonts	13.9	Flash
video	4.83	
supercookies	2.12	Java
http_accept	6.09	
timezone	3.04	
cookies_enabled	0.353	

Table 2. Mean surprisal for each variable in isolation

Means of protection (2)

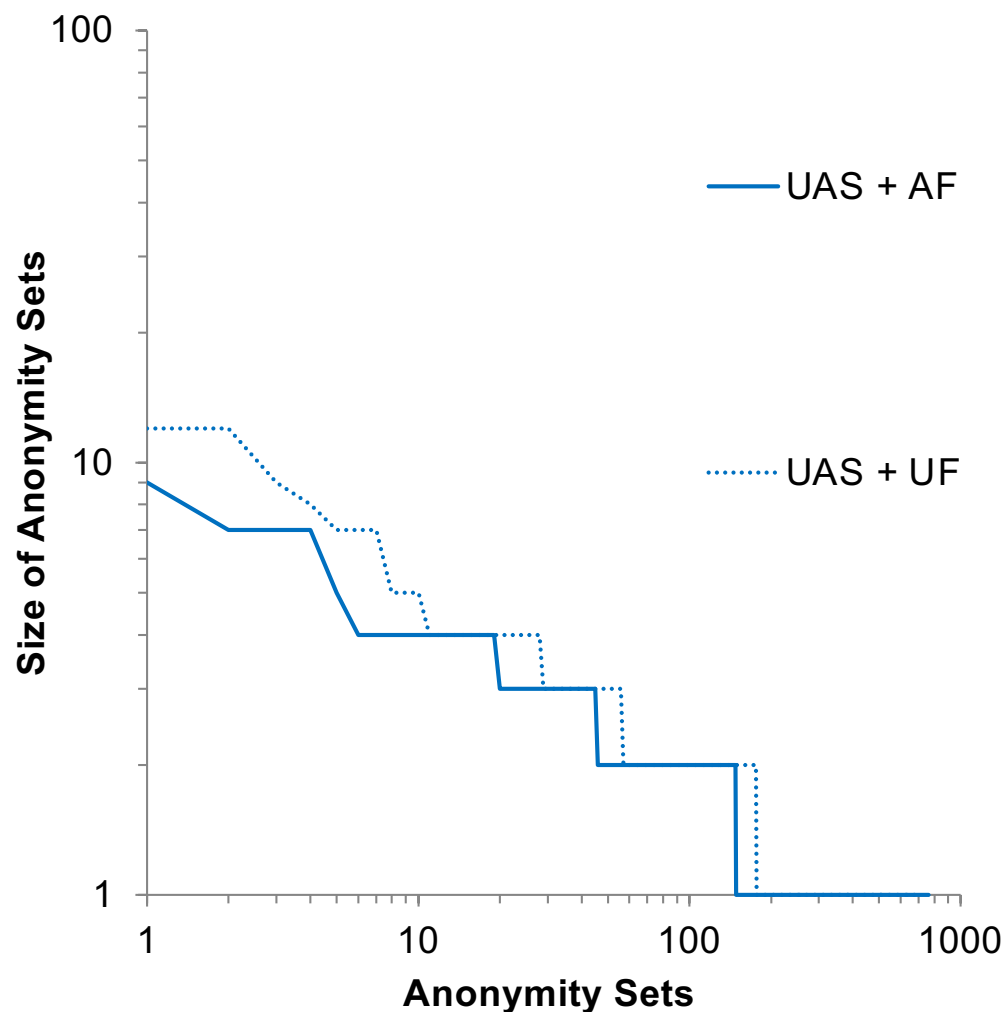
- AF – all fonts
- UF – unified fonts
- JS off – no JavaScript, no fonts at all





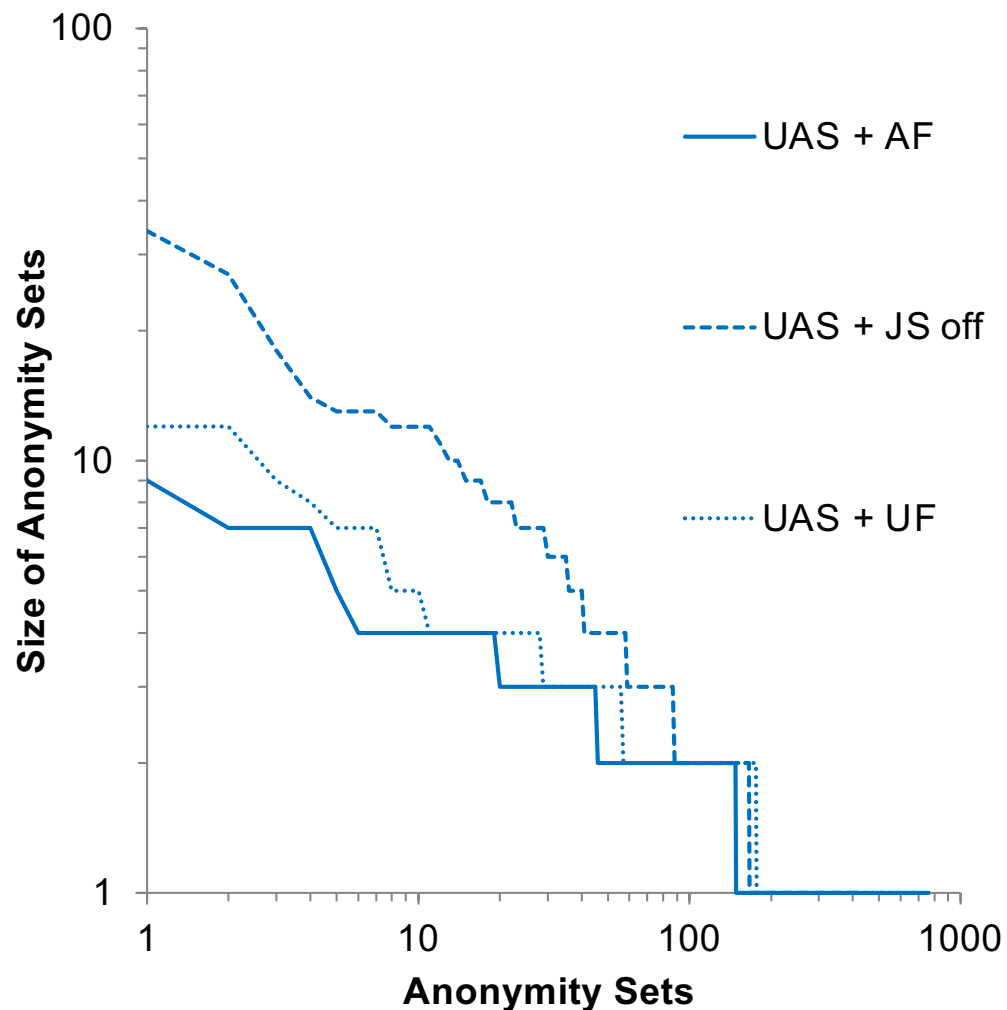
Means of protection (2)

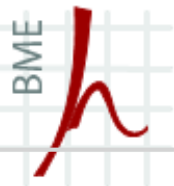
- AF – all fonts
- UF – unified fonts
- JS off – no JavaScript, no fonts at all



Means of protection (2)

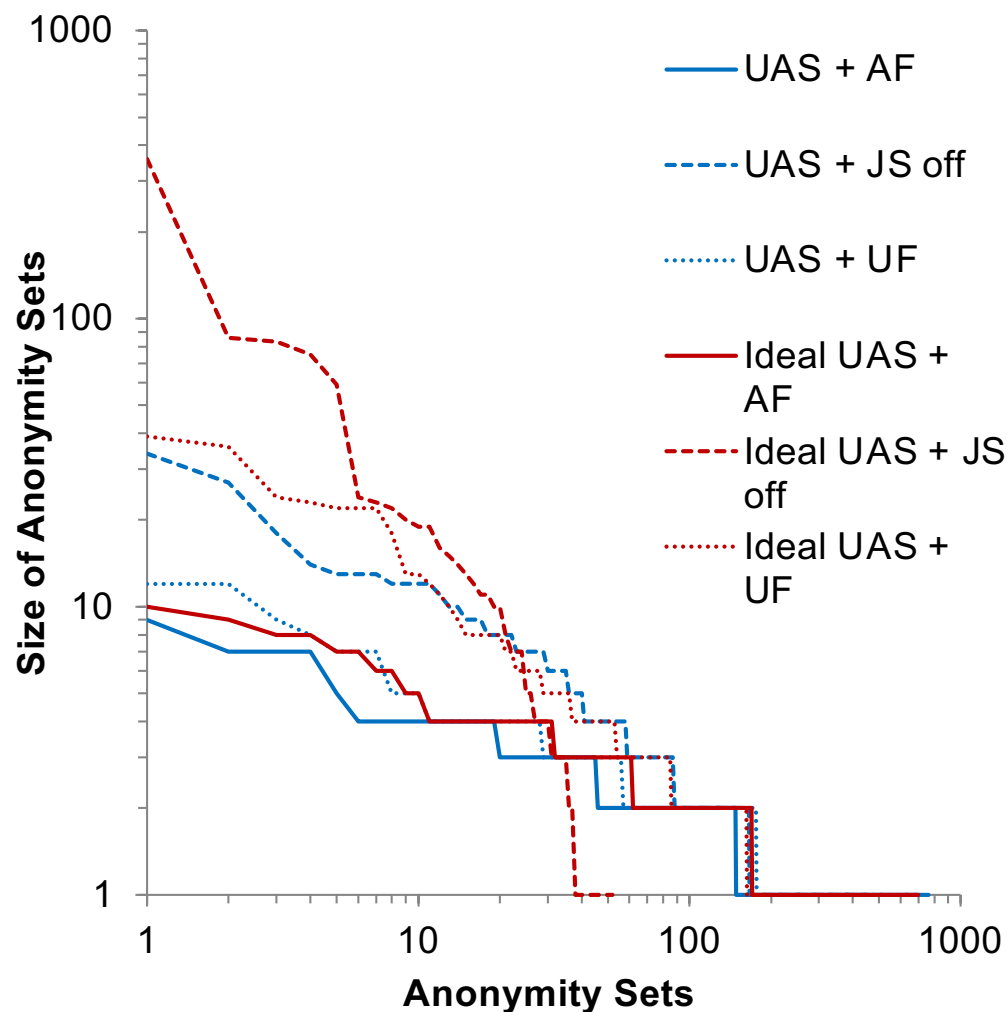
- AF – all fonts
- UF – unified fonts
- JS off – no JavaScript, no fonts at all





Means of protection (2)

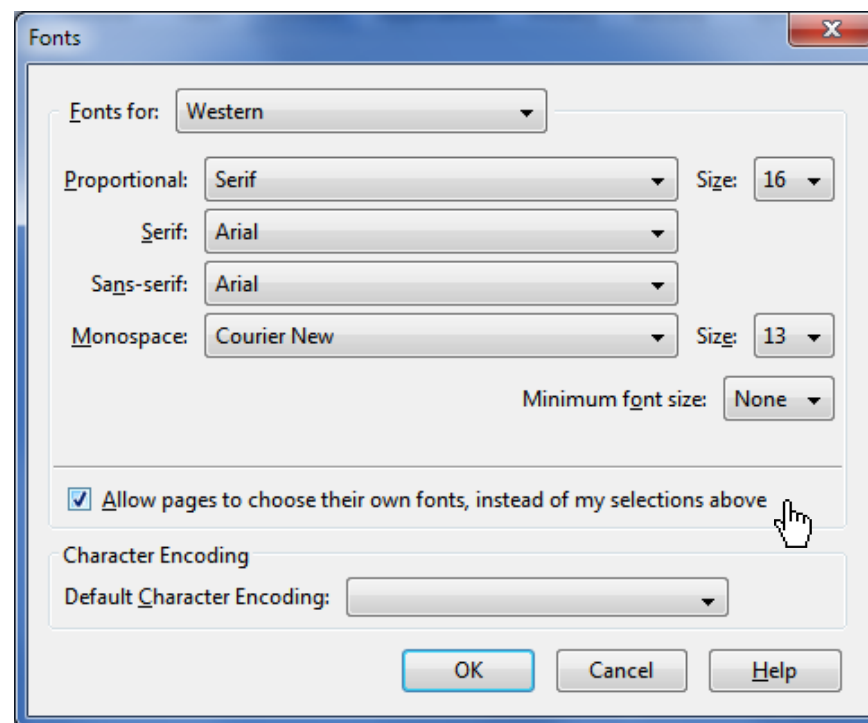
- AF – all fonts
- UF – unified fonts
- JS off – no JavaScript, no fonts at all



Means of protection (3)

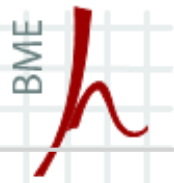
- As a basis: no regular tracking techniques should work.
 - e.g., JondoFox
- Custom builds or using settings
 - Bad user experience
 - Requires serious time investment, and
 - Proposed for experienced users

Each altered setting may provide information for the fingerprint!

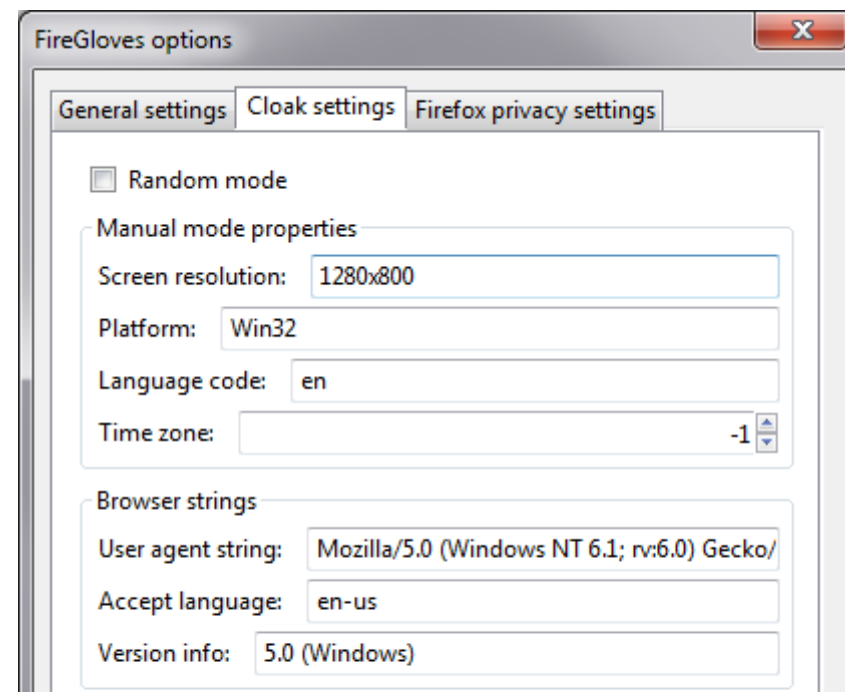
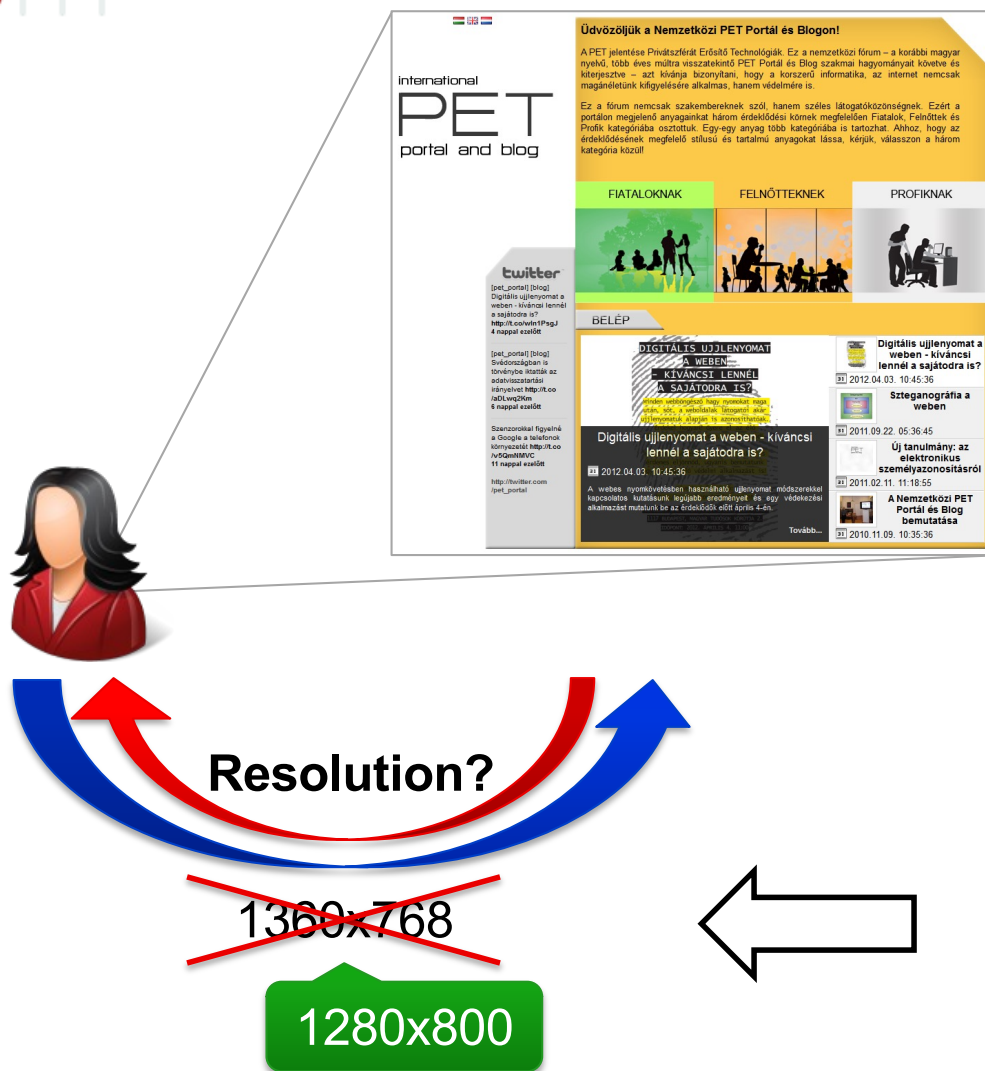


Solution of the Tor Browser

- Limits the number of available fonts
 - `browser.display.max_font_count = 5`
 - `browser.display.max_font_attempts = 10`
- Available only in the Tor Browser (yet)
- Problem:
 - During a session, multiple fonts can be tested
- Our suggested extension to this solution
 - Font caching, and
 - Per-domain setting

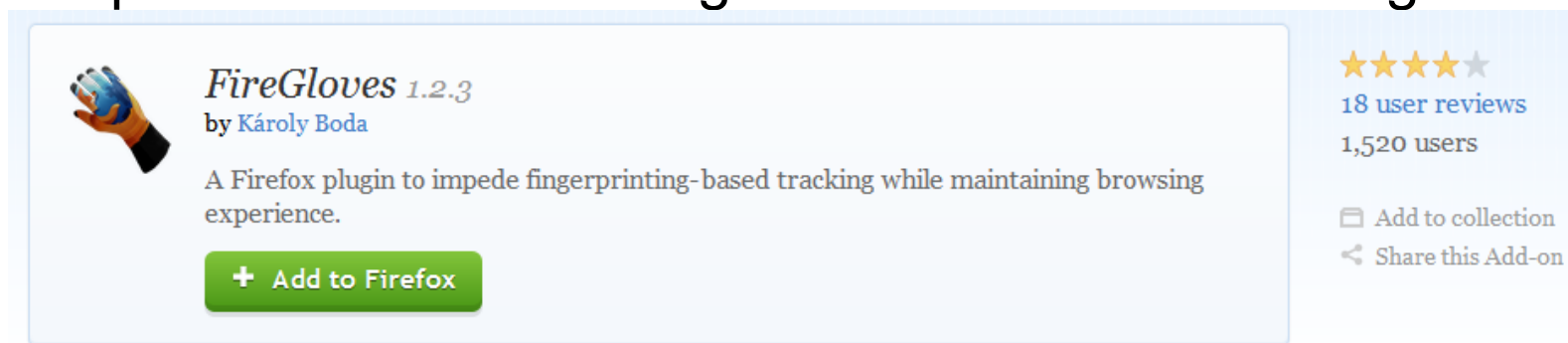


FireGloves



FireGloves (2)

- Proof-of-concept application
 - Modifies the content on-the-fly
 - <https://addons.mozilla.org/en-US/firefox/addon/firegloves/>



- 12 000+ downloads, cc. 1500 users currently
 - Feedback on a weekly basis
- BlueCava fails against it:

first load	2F3D-3810-9863-D286-7989-B4B4-01F2-65BD
reload #1	7E16-E44E-9101-258A-6596-9A2D-B50C-9FB5
reload #2	8D6F-4A0E-2FBF-2B85-5322-8673-D68A-478D



DEMO SECTION

Cross-browser fingerprinting test 2.0

Böngészőfüggetlen ujjlenyomat teszt 2.0



Egy részleges „ujjlenyomat” is
elegendő...

[Ujjlenyomat teszt](#)[Cikkek](#)[GYIK](#)[Partnerek](#)[Kapcsolat](#)[FireGloves](#)[english](#) | [magyar](#)

Ismerje meg ujjlenyomatát!

Kattintson az alábbi gombra, hogy megismerje ujjlenyomatát!

Próbálja ki a tesztet különböző böngészőkben, és nézze meg, hogy ugyanazt az ujjlenyomatot kapja-e!

Ujjlenyomat teszt indítása!

- Amennyiben a teszt indítása után az állapotjelző csík percekig megáll egy helyben (nem fut le a teszt), kérjük, küldjön hibajelentést!
[Hibajelentés küldése](#)
- Hogyha szeretné eltávolítani a teszt közben beállított *evercookie* bejegyzéseket, az alábbi gombra kattintva megteheti.
[Evercookie eltávolítása](#)

Mi az az ujjlenyomat, és miért van erre a kísérletre szükség?

A web böngészése során a legtöbb honlap érdeke, hogy a látogatói viselkedését megfigyelje, elemezze, hogy ezáltal jobb üzleti eredményt érhessen el. [Bővebben »](#)

Cross-browser fingerprinting test 2.0 (2)

Böngésző
ujjle

Ujjlenyom

Ismer

Kattints
Próbálja

1. Sc
2. Ált
3. Tel
4. Tel
5. Fel
6. Ad
7. Sik

• Am
hib

Eredmények

EredményekRészletek

Felhasználóneve

Megadott felhasználónév: **szakdolgozat**
Felhasználónév ezzel az ujjlenyomattal: **szakdolgozat**

Generált azonosító

e81c6258405fc74b6315995b5b886209

Az Ön rendszer jellemzői

Az egyszerűbb összehasonlításképp az azonosítója alapján a rendszeréhez rendeltünk dátum, idő, szín és egyéb értékeket. Ezt könnyen össze tudja hasonlítani más böngészőkben elért eredménnyel, hogy így ellenőrizze, hogy a tesztünk valóban jól azonosítja-e az Ön rendszerét.

Dátum és időpont	Szín	Gyümölcs	Keresztnév	Ital	Város és ország	Titkos kód
2003.08.17. 08:29:48	mediumorchid 2	Kumquat	Stan	coffee	West Midlands, United Kingdom	34533

Elemzés eredménye

Ismert felhasználó, azonos böngészővel.

Cross-browser fingerprinting test 2.0 (3)

Böngésző

Ujjeny

Ismer

Kattints

Próbálj

1. Sc

2. Ált

3. Tel

4. Tel

5. Fel

6. Ad

7. Sik

• Am

hib

F

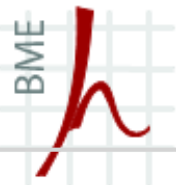
Eredmények

Eredmények

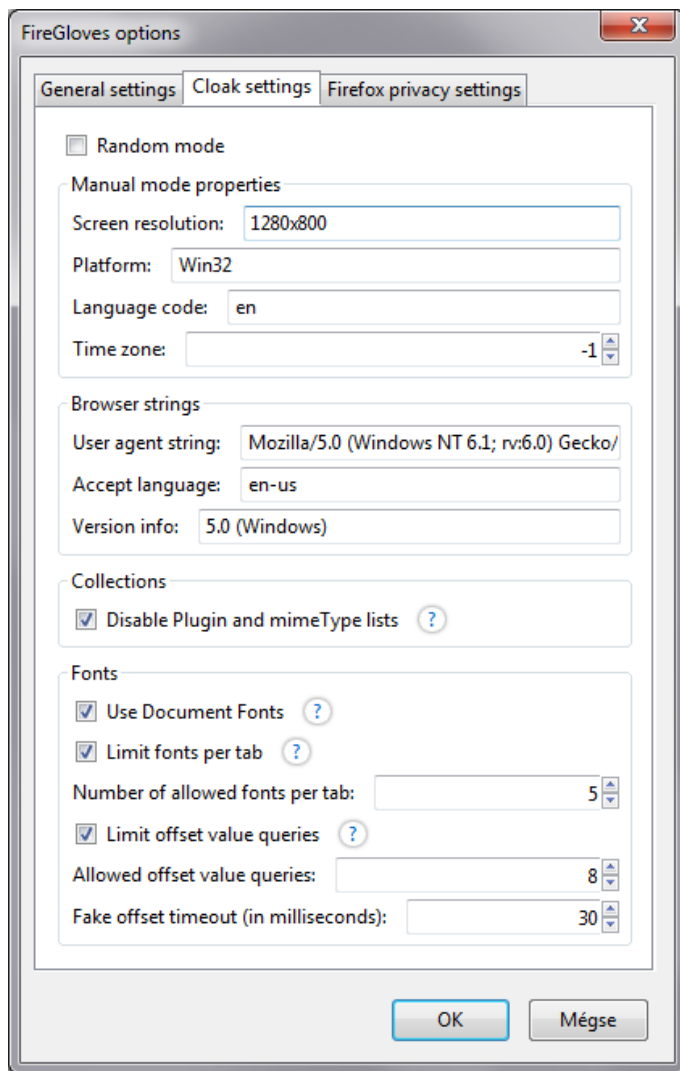
Részletek

Nyelvi beállítások	hu-HU
Operációs rendszer	Windows
Képernyő felbontás	1600x900
Időzóna	-120
User Agent String	Mozilla/5.0 (Windows NT 6.1; WOW64; rv:12.0) Gecko/20100101 Firefox/12.0
HTTP Accept	gzip, deflate hu-hu,hu;q=0.8,en-us;q=0.5,en;q=0.3
Telepített pluginok	Silverlight: 5.1.10411.0, VLC: none, Java: none, Shockwave: none, QuickTime: 7.7.0.0, Flash: 11.2.202.235, AdobeReader: 10.1.3.23, WindowsMediaPlayer: 12.0.7601.17514, javainfo: ---1--

Adobe Arabic, Adobe Caslon Pro, Adobe Fan Heiti Std, Adobe Fangsong Std, Adobe Garamond Pro, **Adobe Gothic Std**, Adobe Hebrew, Adobe Heiti Std, Adobe Kaiti Std, Adobe Ming Std, Adobe Myungjo Std, Adobe Song Std, Agency FB, **Aharoni**, **ALGERIAN**, Andalus, Angama New, AngamaUPC, Aparajita, Arabic Typesetting, Arial Unicode MS, Baskerville Old Face, Batang, BatangChe, Bauhaus 93, Bell MT, Berlin Sans FB, Birch Std, *Blackadder JTC*, **Blackoak Std**, Bodoni MT, Book Antiqua, Bookman Old Style, Bookshelf Symbol 7, *Bradley Hand ITC*, **Broadway**, Brownlie New, BrownlieUPC, *brush script int*, *Brush Script Std*, Calibri, Californian FB, Calisto MT, Cambria, Cambria Math, Candara, CASTELLAR, Centaur, Century, Century Gothic, Century Schoolbook, Chancery L, CHARLEMAGNE STD, *Chicle*, *Comic Sans MS*

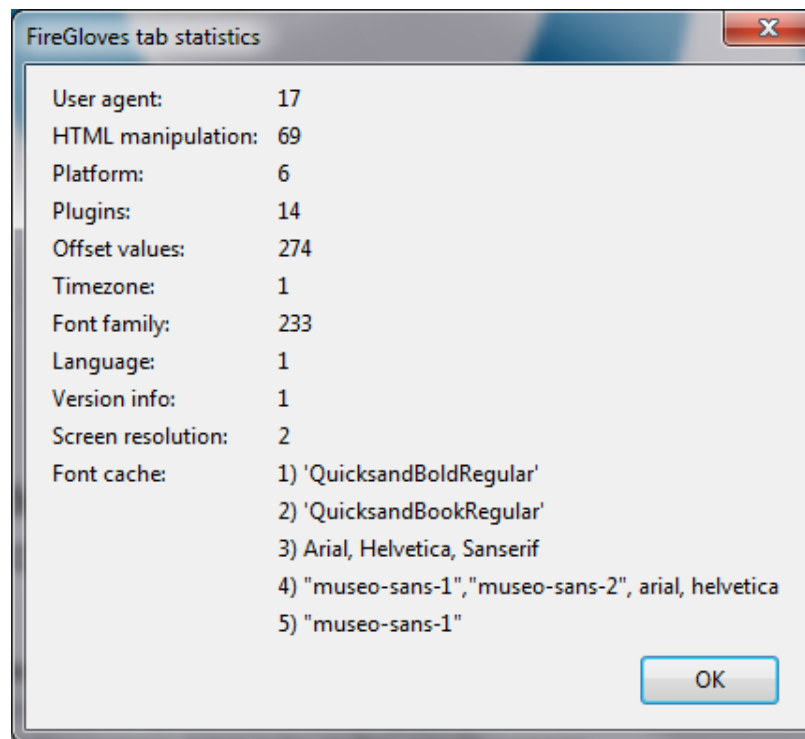
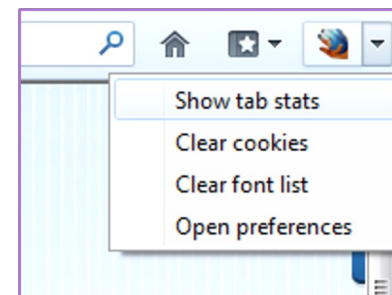


FireGloves 1.2.3



← Privacy and
cloaking settings

FireGloves
dropdown menu →



←
Statistics
on what is
prevented



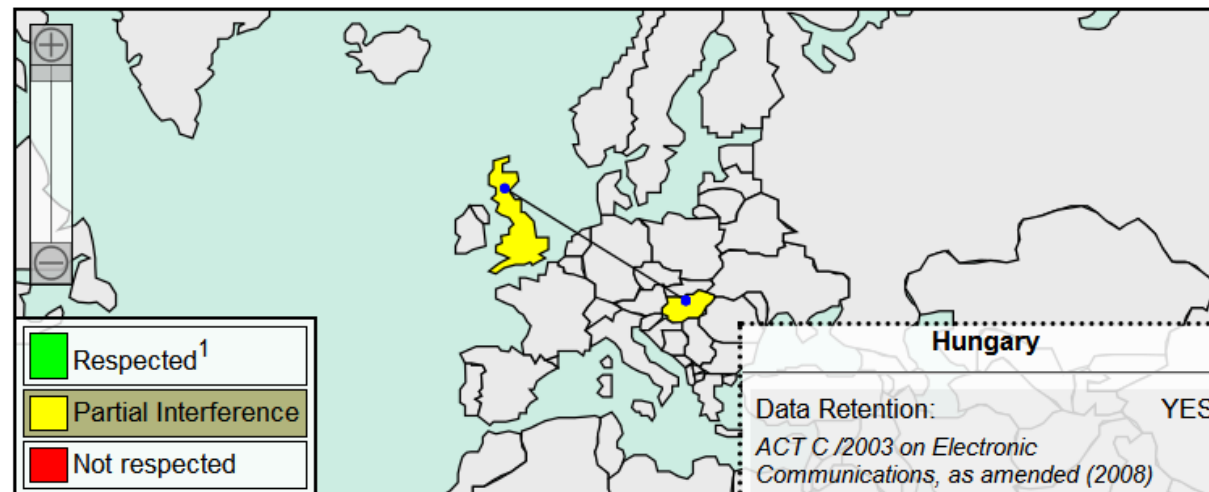
Save as HTML

☐ Hide IPs

Email Privacy Status:

communication surveillance in countries through which email was routed
(based on routing path simulation)

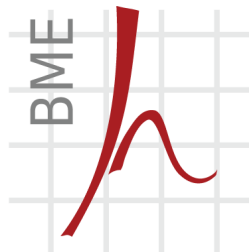
From: e.a.carroll@stir.ac.uk To: Christel.Backman@sociology.gu.se



IP Address	Country
Not available	
139.153.13.210	United Kingdom
139.153.13.111	United Kingdom
139.153.13.117	United Kingdom
88.151.101.251	Hungary
88.151.96.218	Hungary
81.183.2.165	Hungary
81.183.0.147	Hungary
81.183.2.4	Hungary
84.1.82.250	Hungary
84.1.82.250	Hungary
192.168.1.1	Reserved²
	Receiver

Questions?

THANK YOU FOR YOUR ATTENTION!



Department of
Telecommunications

GULYÁS, Gábor György
Assistant research fellow
Dept. of Telecommunications (BUTE)
gulyasg@hit.bme.hu

