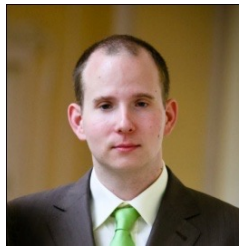
ECONOMIC SURVEILLANCE ON THE WEB: CUTTING-EDGE TECHNOLOGIES IN PROFILING (AND WHAT CAN WE DO AGAINST IT)

Living in Surveillance Societies Workshop

October 1-3, 2012

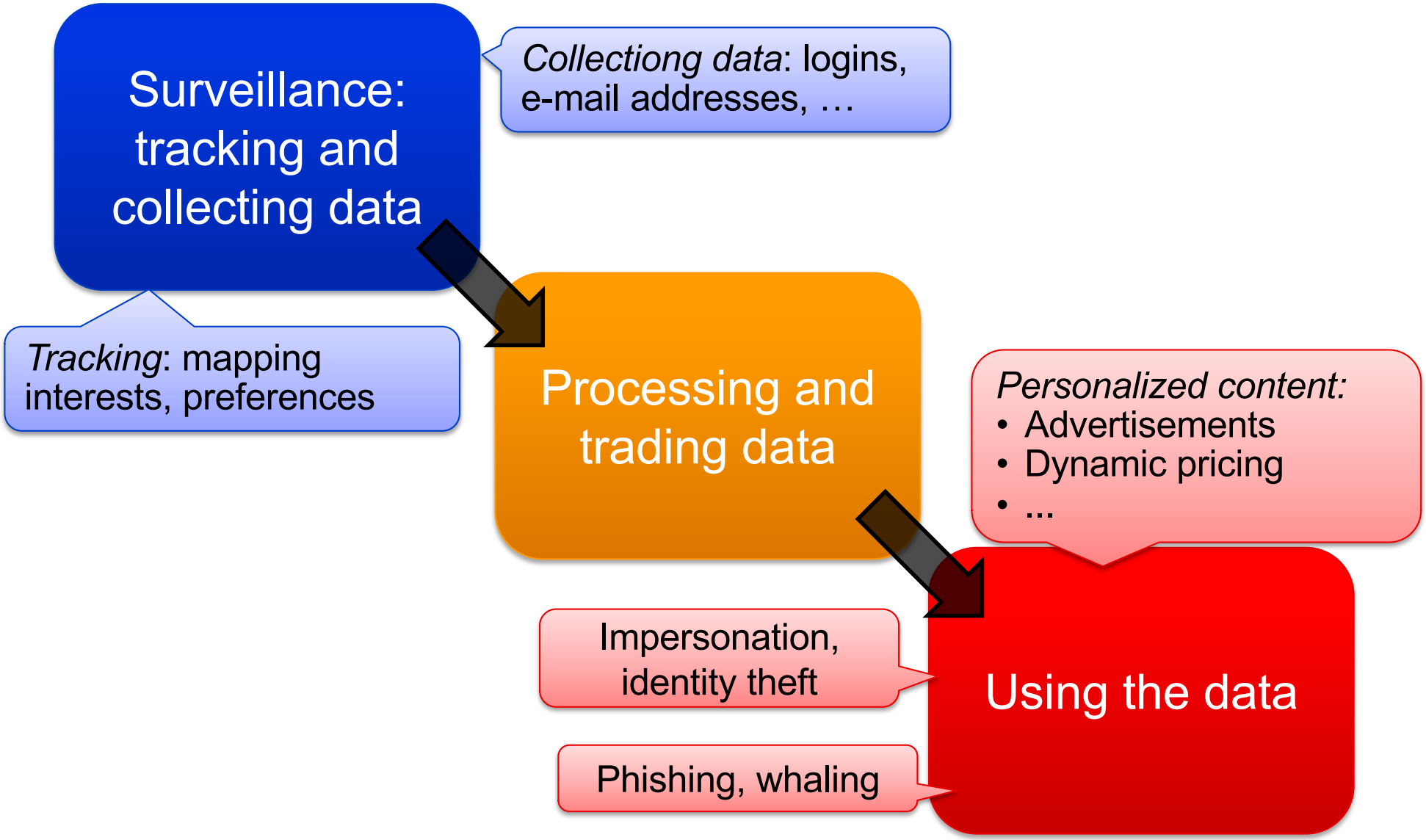
GULYÁS, Gábor György
Assistant research fellow
Dept. of Telecommunications (BUTE)
gulyasg@hit.bme.hu



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

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

Various other techniques are used beside cookies.

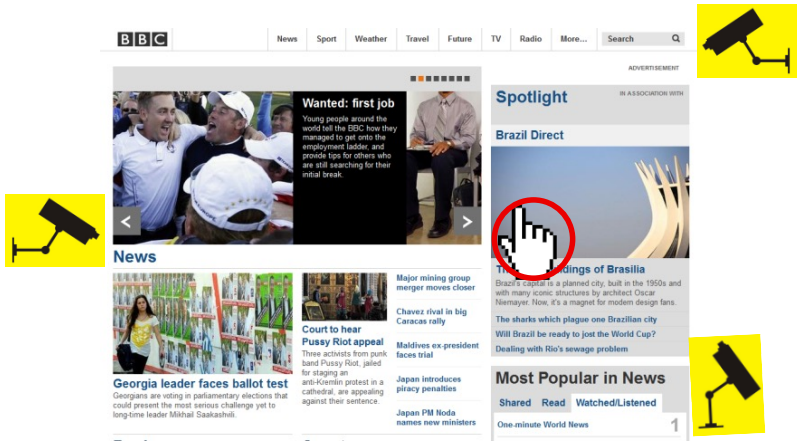
Penetration of social widgets is above 35%

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

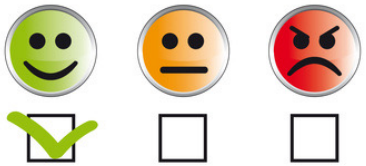
What do they collect? (2)

Copy-paste events

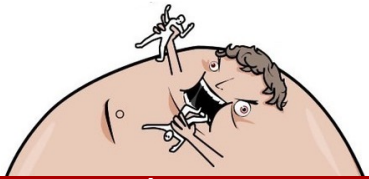
Keyboard logging

RSS tracking

Content, meaning,

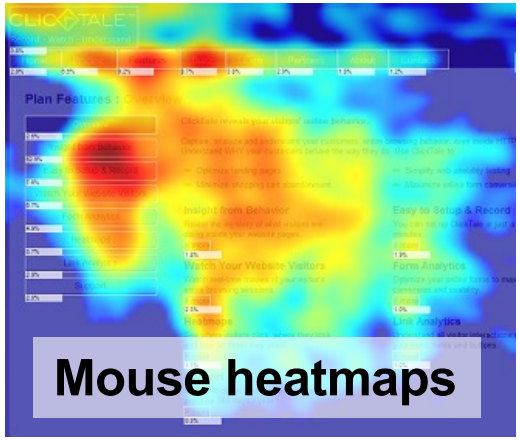


sentiment analysis



What do they collect? (3)

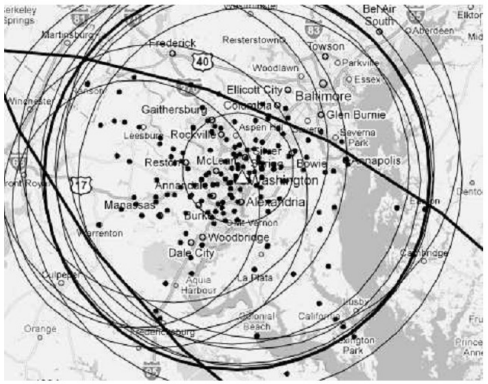
http://goo.gl/4bUVO



Regular techniques

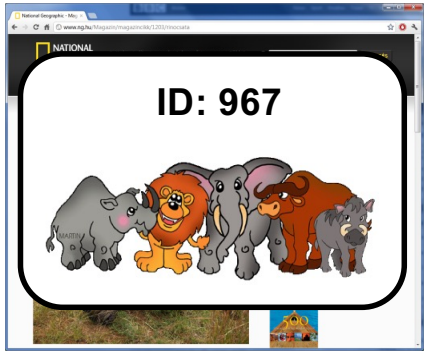
IP based tracking

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



Cookie based tracking

Surfer

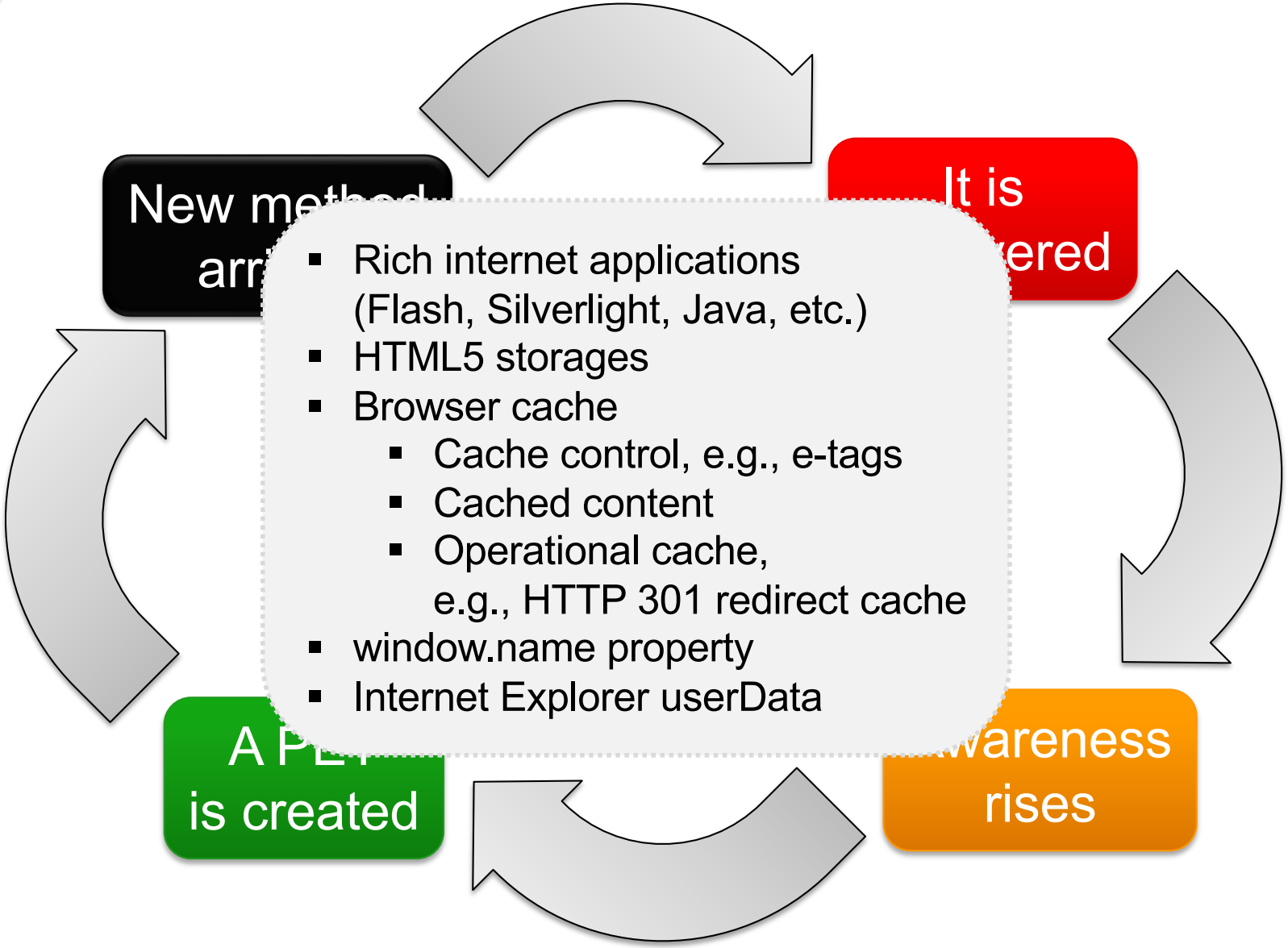


Advertiser

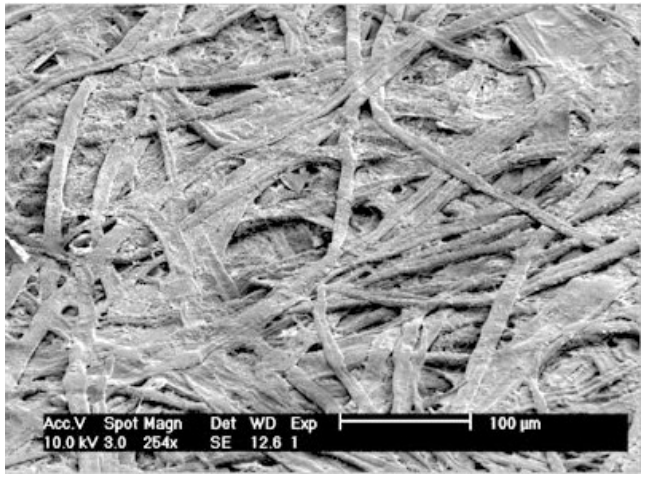
Profile of ID: 967

- Buying apples
- Wild life of the savannah

Regular techniques (2)

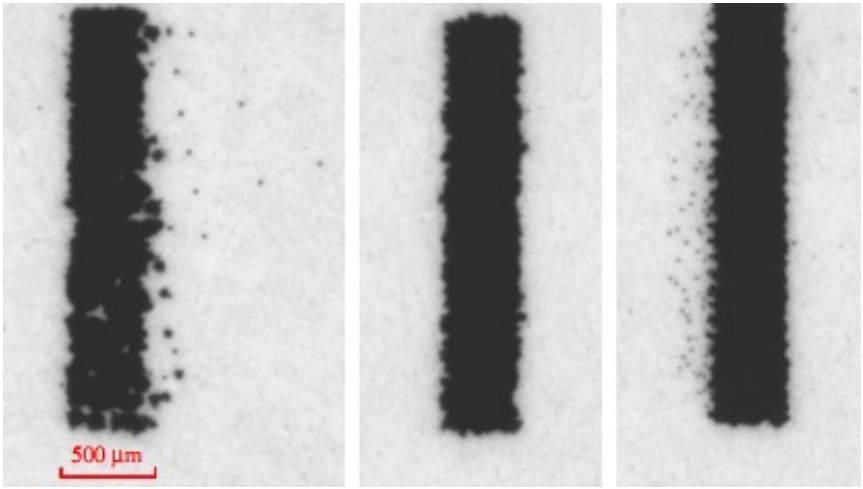


Fingerprinting?

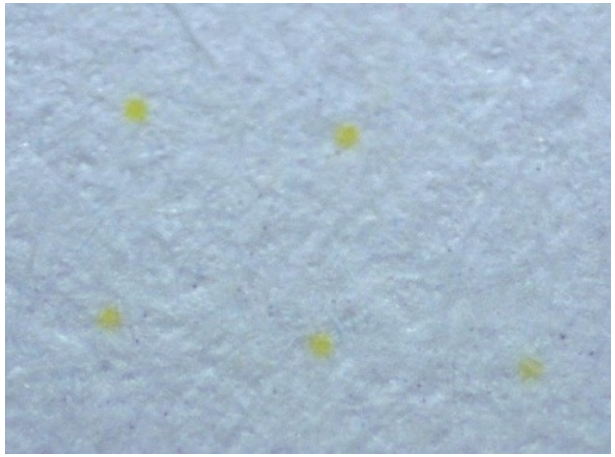


Copy paper

Cameras, scanners



Printers



Unique identifiers added while printing

The Panopticlick project (2010)

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

http://panopticlick.eff.org

Followed by many companies

ThreatMetrix

iovation

41st Parameter

Wolf Software
Part of The Wolf Consortium

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

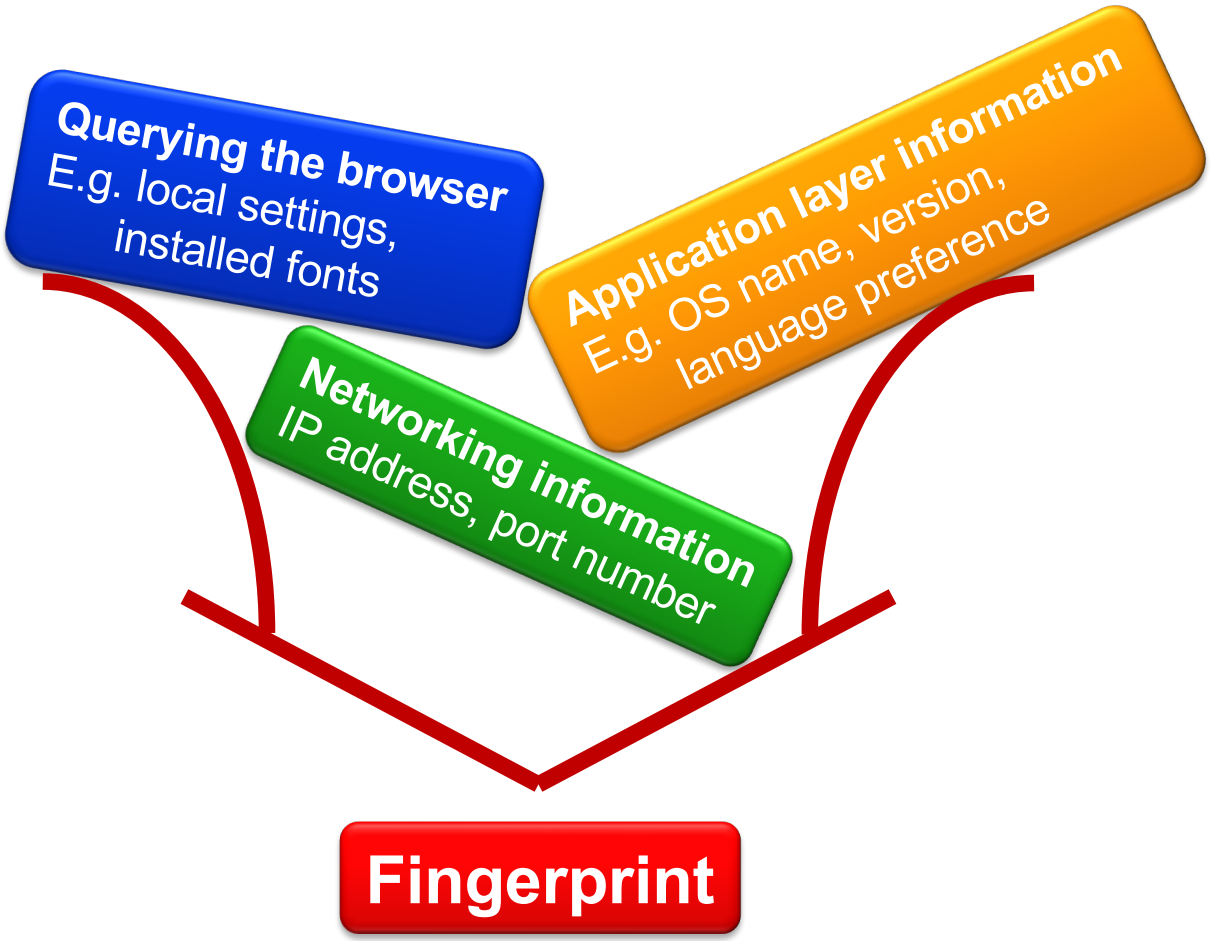
But *how*?

Meanwhile...

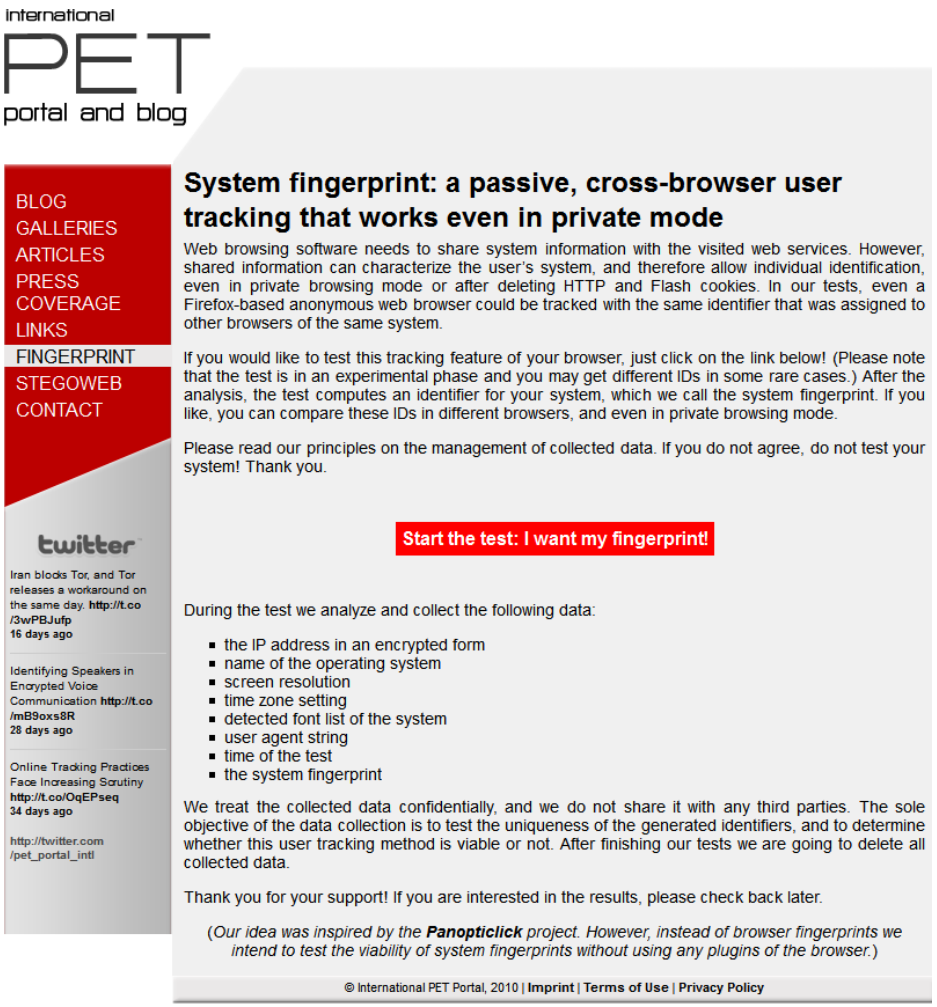
Problems of the Panopticlick project:

- Plugin dependency
- Browser dependency

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



The system fingerprint test



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

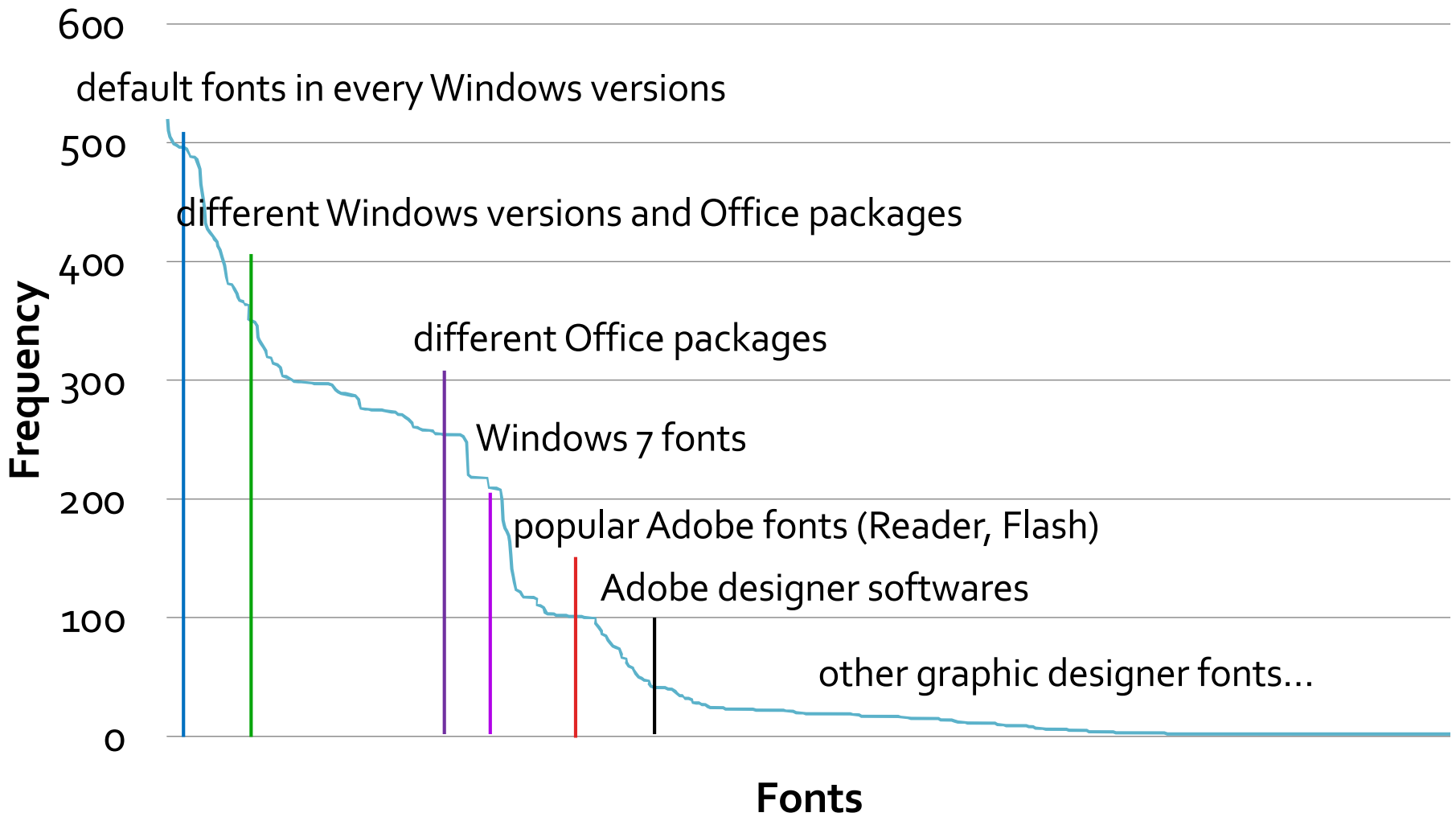
Same fingerprint in all browsers
⇒ pure JavaScript

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

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 aceppted)

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
plugins	15.4
fonts	13.9
video	4.83
supercookies	2.12
http_accept	6.09
timezone	3.04
cookies_enabled	0.353

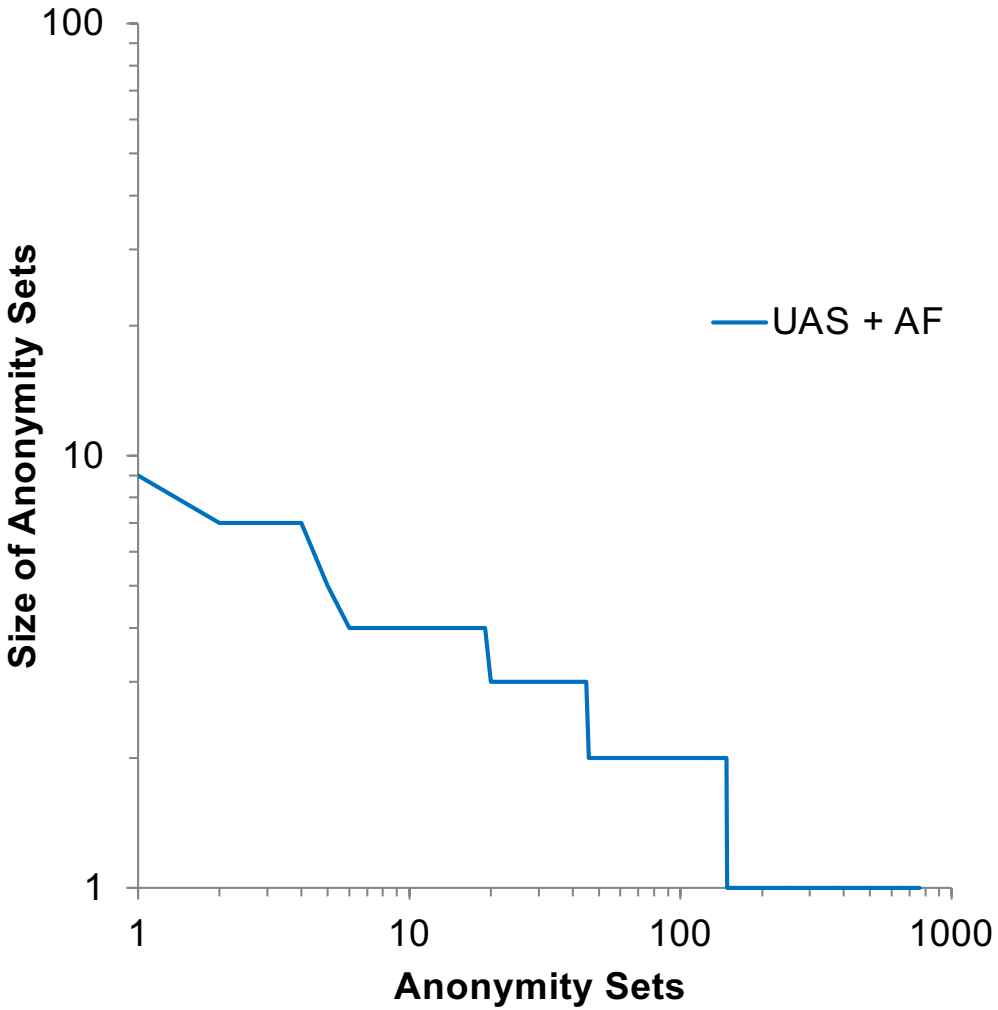
JavaScript

Flash

Java

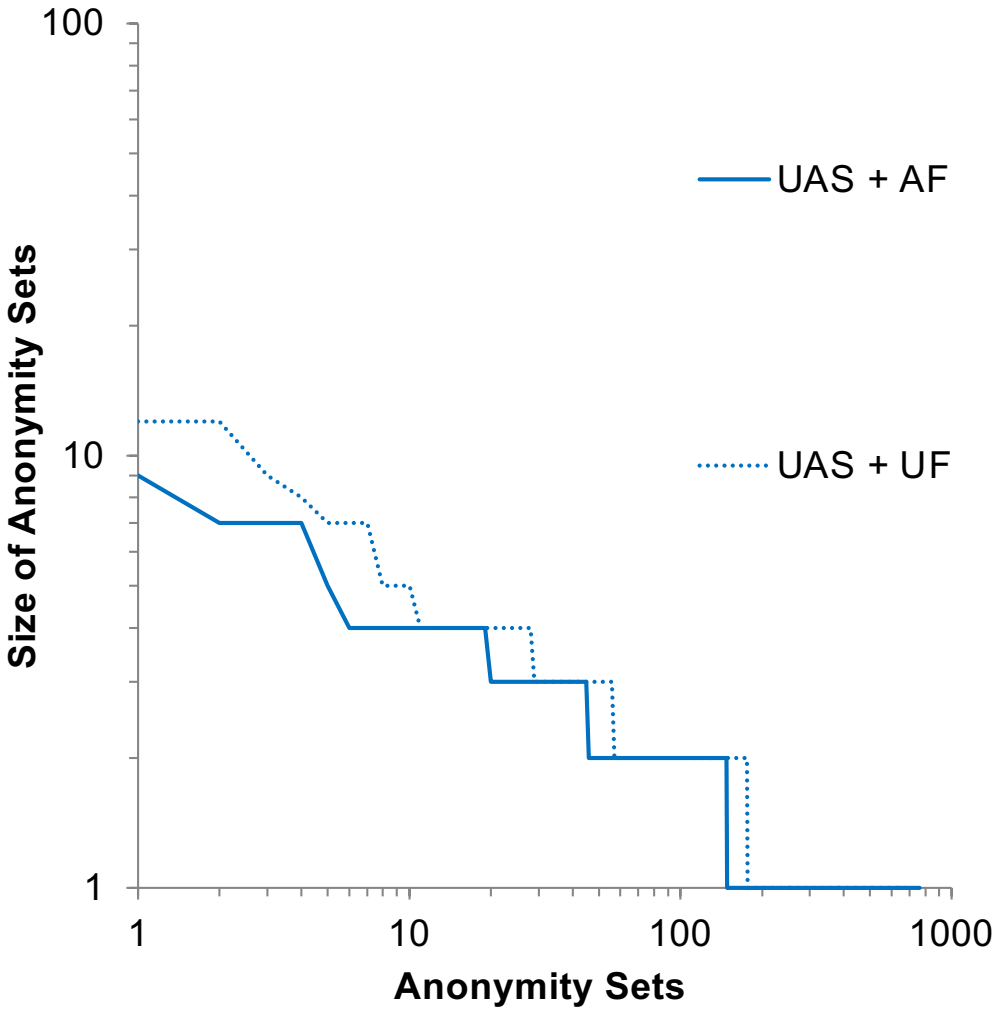
Table 2. Mean surprisal for each variable in isolation

Means of protection (2)



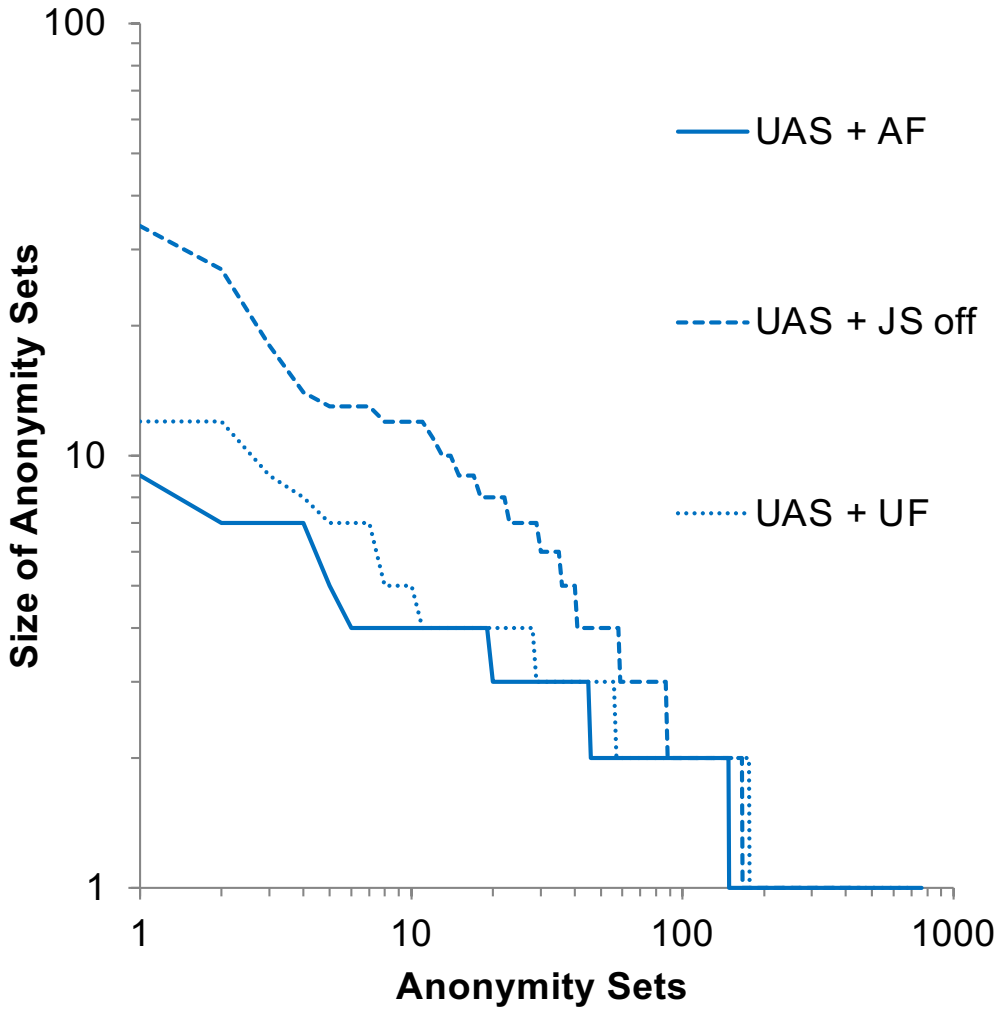
- AF – all fonts
- UF – unified fonts
- JS off – no JavaScript, no fonts at all

Means of protection (2)



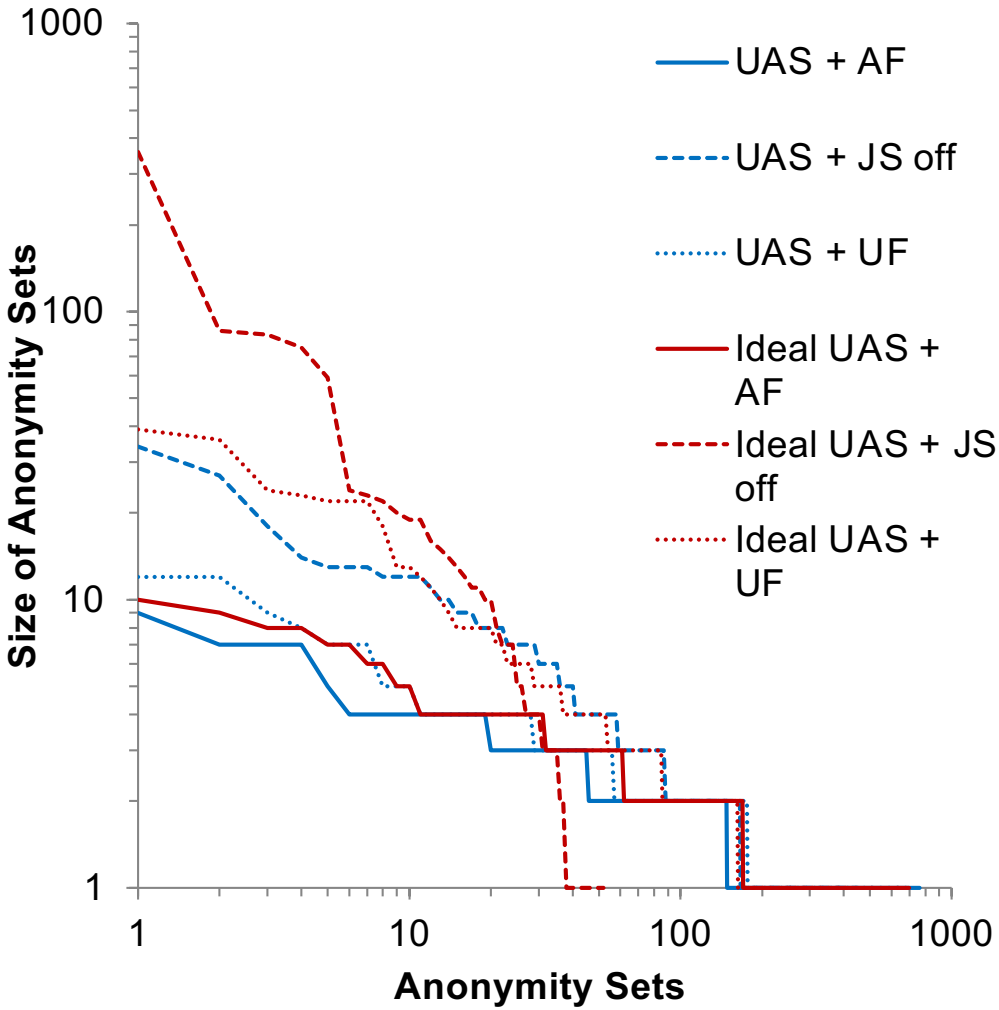
- AF – all fonts
- UF – unified fonts
- JS off – no JavaScript, no fonts at all

Means of protection (2)



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

Means of protection (2)

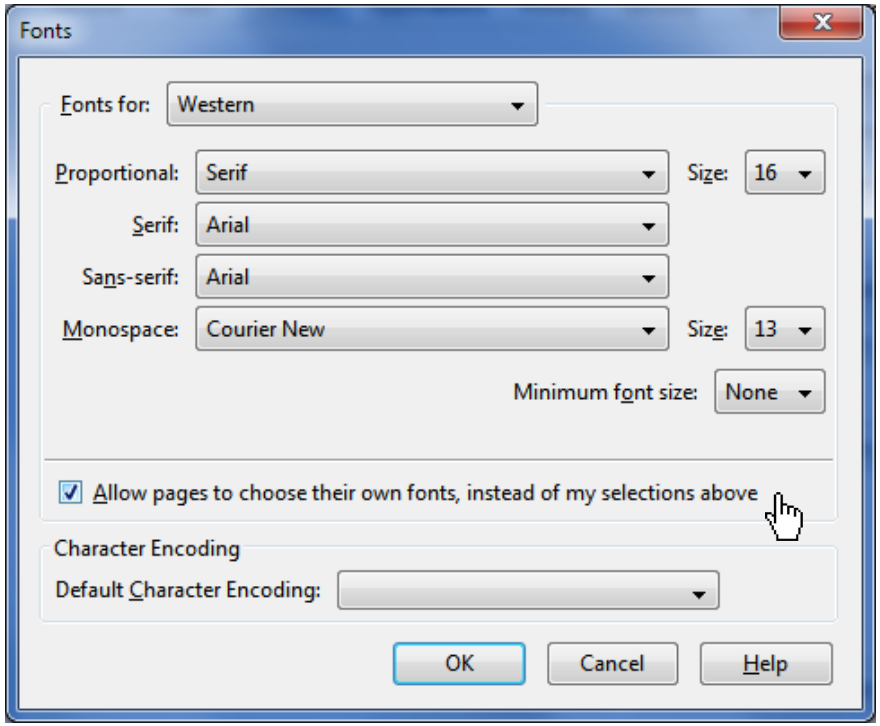


- AF – all fonts
- UF – unified fonts
- JS off – no JavaScript, no fonts at all

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 fo experienced users

Each altered setting may provide information for the fingerprint!

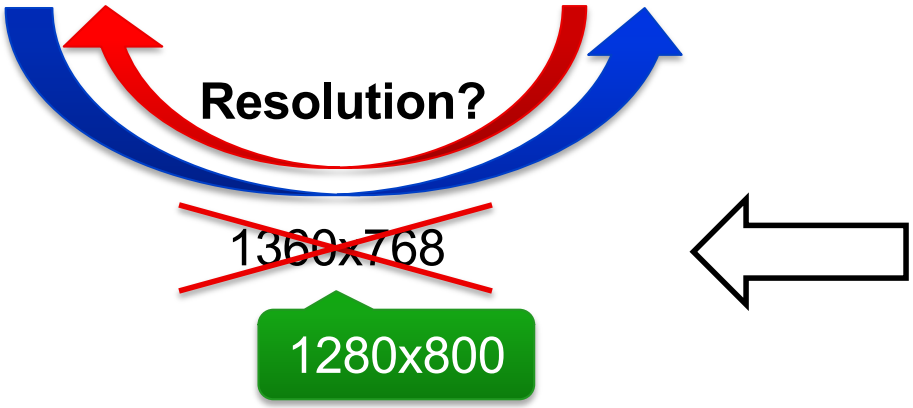


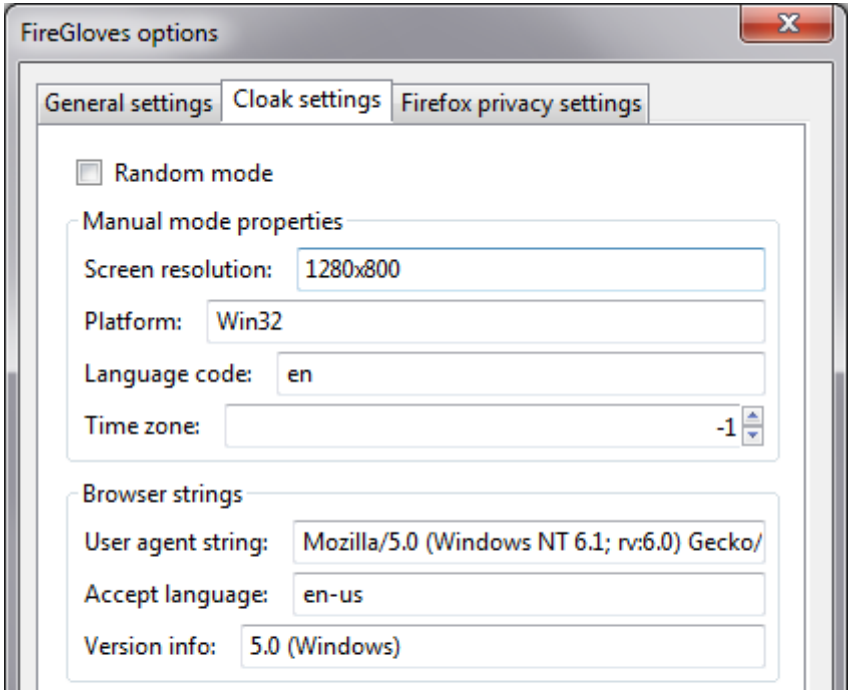
Solution of the Tor Browser

- 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

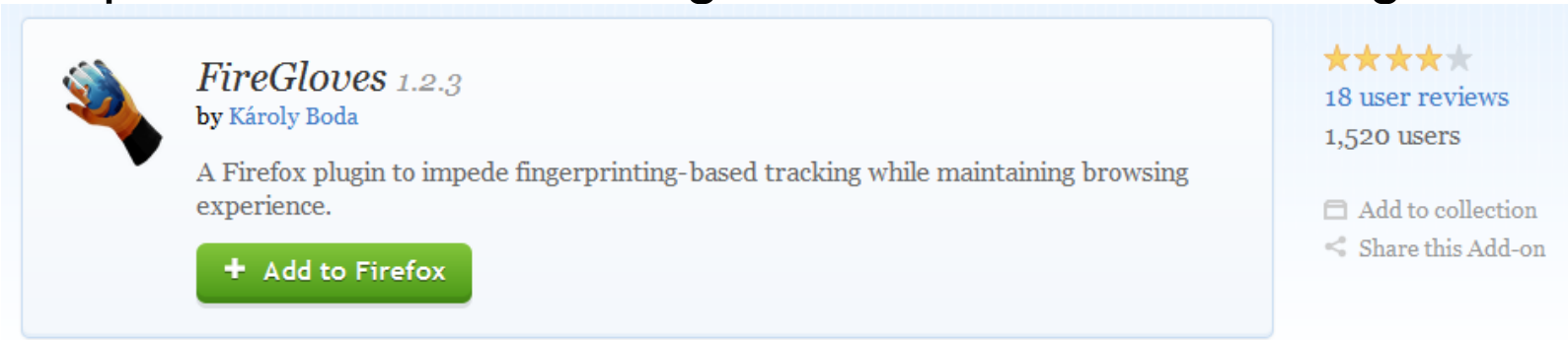

international
PET
portal and blog
Üdvözöljük a Nemzetközi PET Portál és Blogon!
FIATALOKNAK
FELNŐTTEKNEK
PROFIKNAK
twitter
BELÉP
Digitális ujjlenyomat a weben - kíváncsi lennél a sajátodra is?
Szteganográfia a weben
Új tanulmány: az elektronikus személyazonosításról
A Nemzetközi PET Portál és Blog bemutatása


Resolution?
1360x768
1280x800


FireGloves options
General settings
Cloak settings
Firefox privacy settings
Random mode
Manual mode properties
Screen resolution: 1280x800
Platform: Win32
Language code: en
Time zone: -1
Browser strings
User agent string: Mozilla/5.0 (Windows NT 6.1; rv:6.0) Gecko/
Accept language: en-us
Version info: 5.0 (Windows)

FireGloves (2)

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



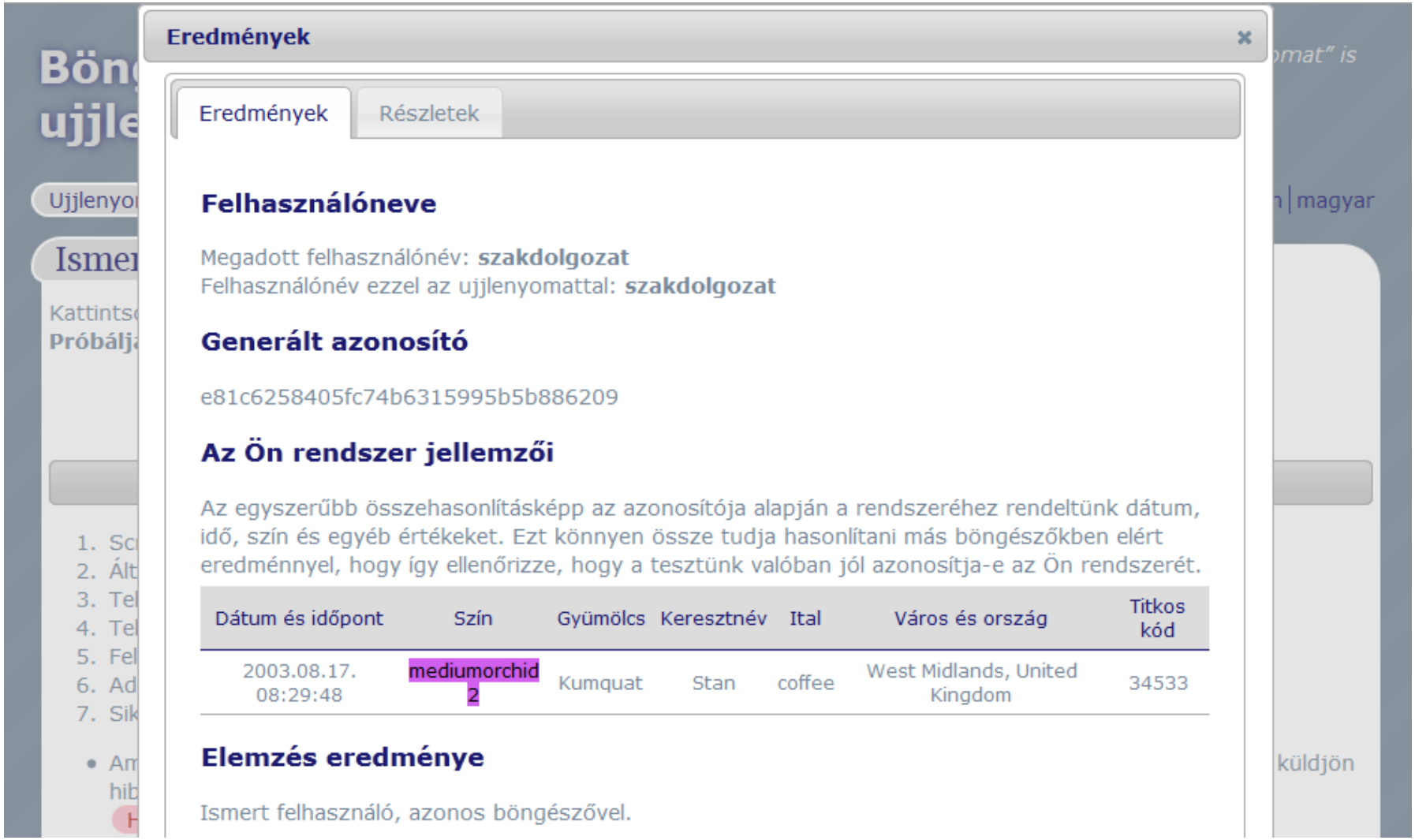
  - 12 000+ downloads, cc. 1500 users currently
  - Feedback on a weekly basis
- BlueCava fails against is it:

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

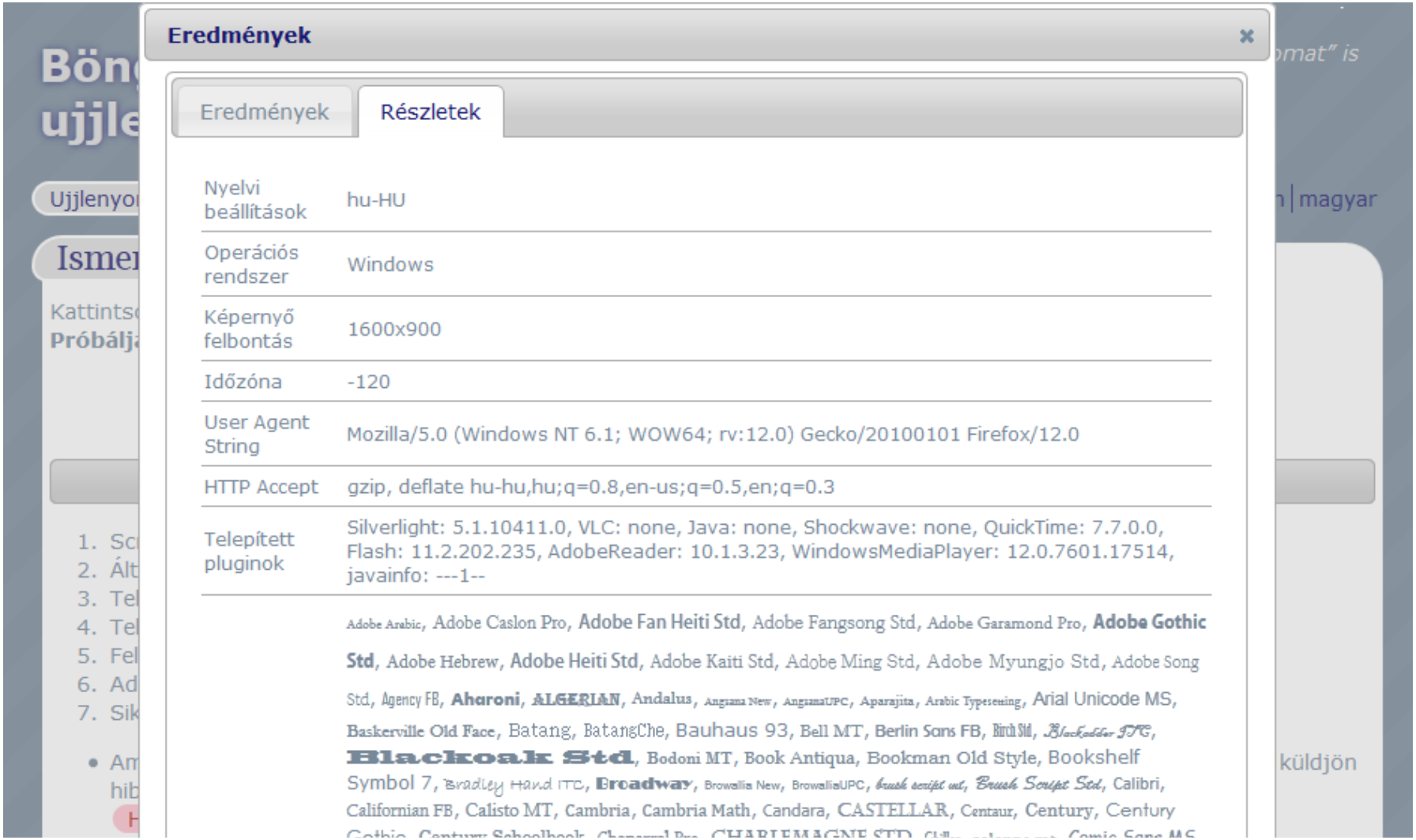
DEMO SECTION

Cross-browser fingerprinting test 2.0

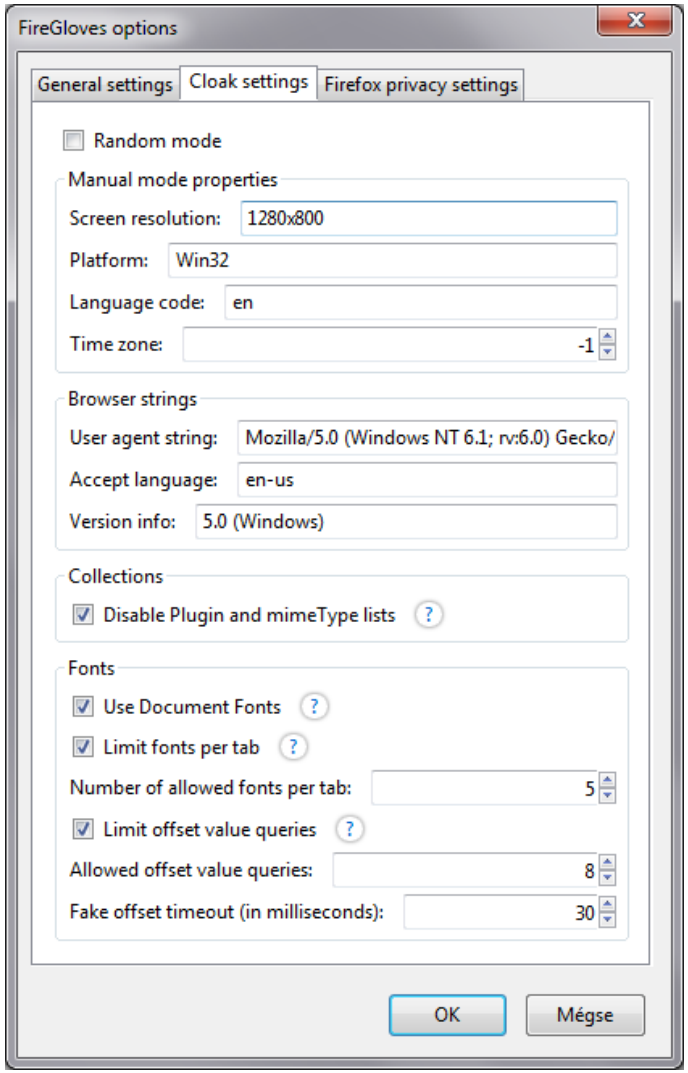
Cross-browser fingerprinting test 2.0 (2)



Cross-browser fingerprinting test 2.0 (3)

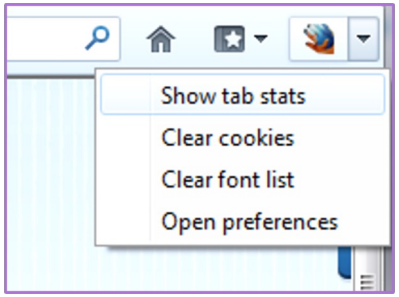


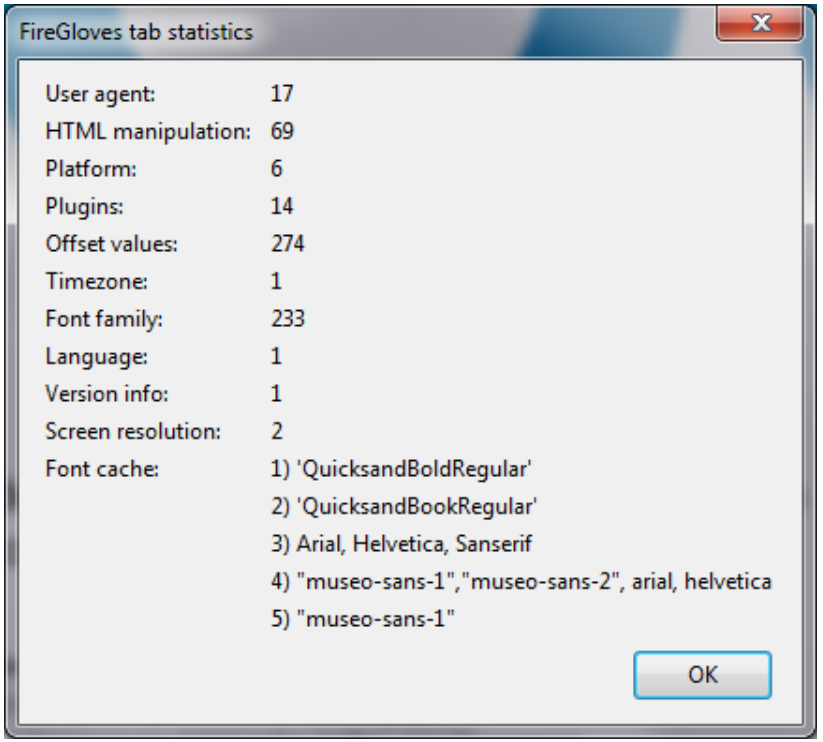
FireGloves 1.2.3



← Privacy and cloaking settings

FireGloves dropdown menu →





← Statistics on what is prevented

Email Path Visualization Project

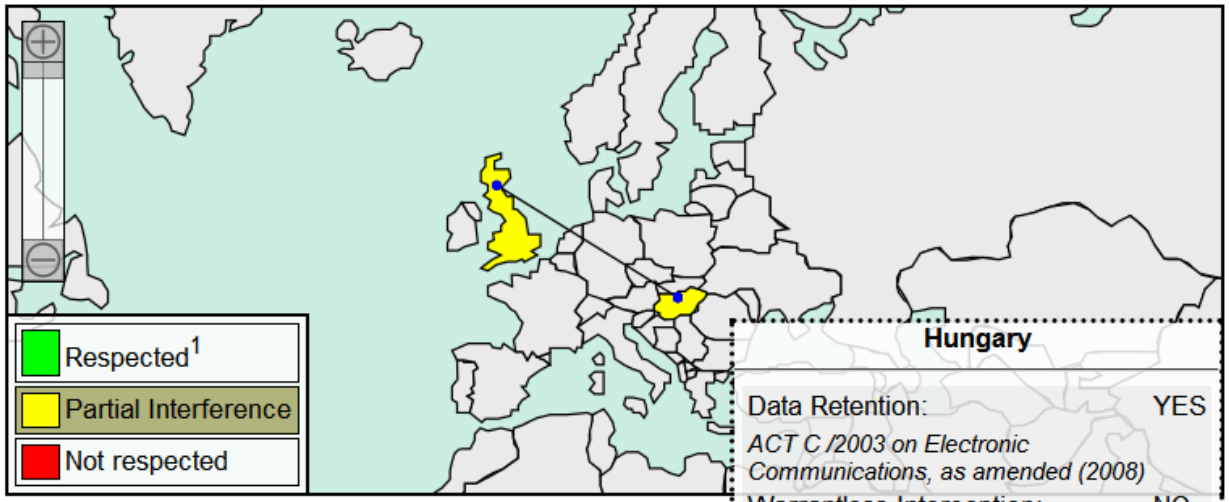
PRIVACY INTERNATIONAL

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	Reserved[2]	Receiver

Questions?

THANK YOU FOR YOUR ATTENTION!

Department of Telecommunications

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