Forensic analysis of your browser fingerprint

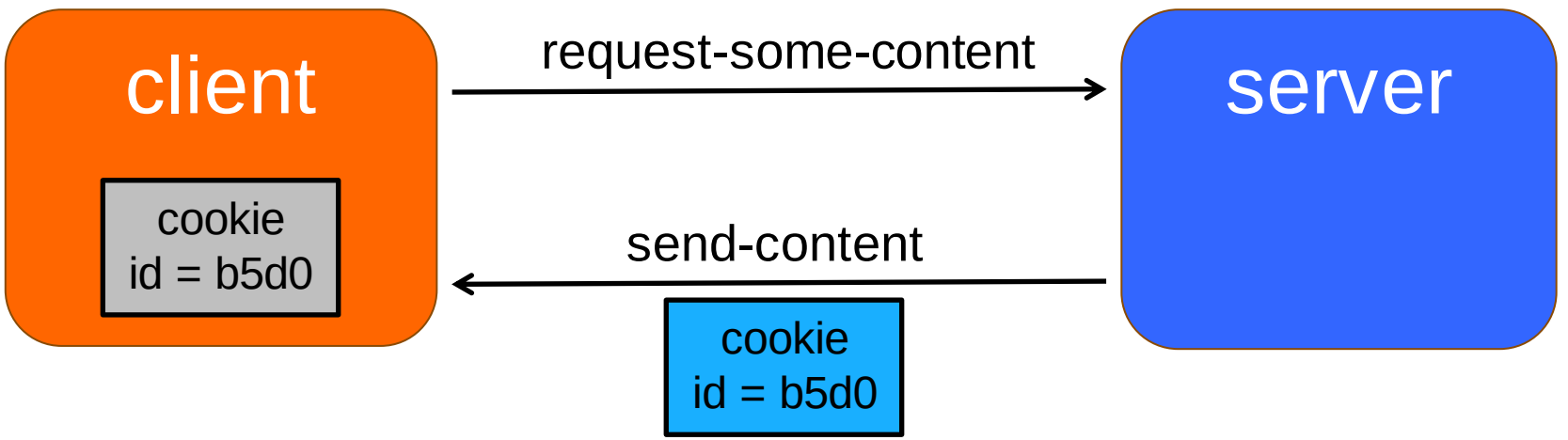
Walter Rudametkin
Inria Spirals & University of Lille
http://rudametw.github.io

Université de Lille

SPIRALS

Inria informatics mathematics

Browser cookies



La tache de Darwin, voyage dans la diversité des formes animales ›

ÉDITION ABONNÉS

Le hasard, une nécessité contre les pirates numériques

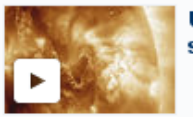
LE MONDE SCIENCE ET TECHNO | 09.06.2014 à 15h41 • Mis à jour le 10.06.2014 à 16h49

Par **David Larousserie**

LECTURE ZEN

Sans que nous le remarquions, l'arsenal technologique des ordinateurs, smartphones et cartes à puce en tout genre ont un point commun avec les

Vidéo



Un "canyon de feu" observé sur le Soleil

La tache de Darwin, voyage dans la diversité des formes animales ›

ÉDITION ABONNÉS

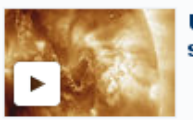
Le hasard, une nécessité contre les pirates numériques

LE MONDE SCIENCE ET TECHNO | 09.06.2014 à 15h41 • Mis à jour le 10.06.2014 à 16h49

Par David Larousserie

Sans que nous le remarquions, l'arsenal technologique des ordinateurs, smartphones et cartes à puce en tout genre ont un point commun avec les

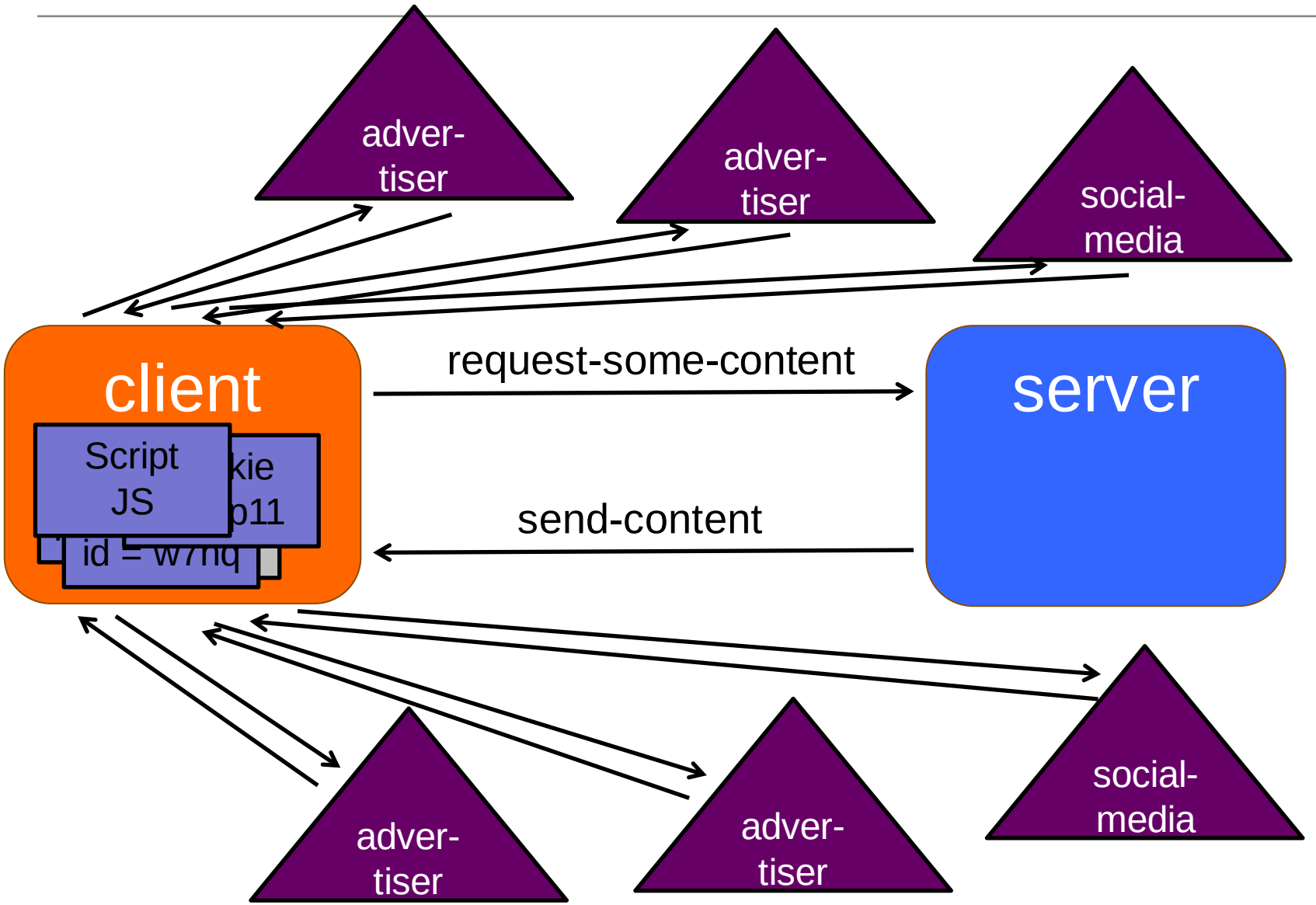
Vidéo



Un "canyon de feu" observé sur le Soleil

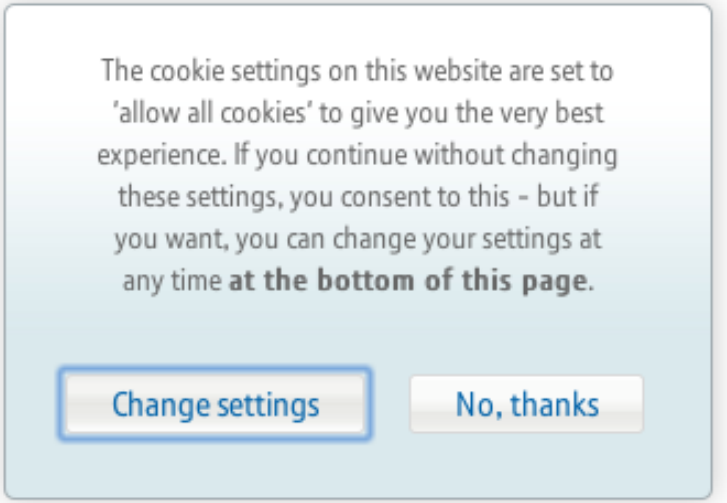
74 trackers (most invisible)

Accuen Media, Acuity Ads, Adap.tv, Adify, Adroit Digital Solutions, AdScale, ADTECH, Advertising.com, Aggregate Knowledge, AppNexus, AT Internet, Atlas, BidSwitch, Casale Media, Cedexis Radar, Chango, ChartBeat, Connexity, Criteo, Datalogix, DataXu, Digilant, Dotomi, DoubleClick, DoubleClick Bid Manager, DoubleClick Spotlight, EQ Advertising, Eulerian, Experian Marketing Services, eyeReturn Marketing, Ezakus, Facebook Connect, Facebook Exchange (FBX), Facebook Social Plugins, Google Adsense, Google Analytics, Improve Digital, Integral Ad Science, Jumptap, Kameleoon, Ligatus, Lijit, Magnetic, Media Innovation Group, Media Optimizer (Adobe), Media6Degrees, MediaMath, Netmining, Neustar AdAdvisor, OpenX, Optimix Media Delivery, Outbrain, OwnerIQ, PubMatic, PulsePoint, Quantcast, Right Media, Rocker Fuel, Rubicon, ScoreCard Research Beacon, SiteScout, Sizmek, SMART AdServer, SpotXchange, TradeDesk, TubeMogul, Turn, Twitter Button, Veruta, Videology, Video Step, Visual Revenu, Yandex.Metrics, Yieldr


adver-
tiser
adver-
tiser
social-
media
client
Script
JS
request-some-content
server
send-content
adver-
tiser
adver-
tiser
social-
media

“Cookie awareness”

- Legal action
  - CNIL, GDPR, ePrivacy
- Extensions
  - 95 add-ons for Firefox for cookies
  - Tor, Ad-Blockers, Anti-trackers
- Browser parameters
  - accept cookies, third-party cookies, etc.
- Private navigation
  - removes cookies after each session
- Browsers
  - Brave



More techniques

- But also ***more agressive*** cookies & fingerprinting
  - Evercookies (multiple cookies), flash cookies (LSO), Permacookie (e.g., Verizon), Google Cookie (used by NSA).
  - Advanced Device Identification using **Browser Fingerprinting**

A browser fingerprint


Attribute	Value
User agent	Mozilla/5.0 (X11; Linux i686; rv:25.0) Gecko/20100101 Firefox/25.0
HTTP accept	text/html, application/xhtml+xml, application/xml;q=0.9,*/*;q=0.8 gzip, deflate en-US,en;q=0.5
Plugins	Plugin 0: QuickTime Plug-in 7.6.6; libtotem-narrowspace-plugin.so; Plugin 1: Shockwave Flash; Shockwave Flash 11.2 r202; libflashplayer.so;
Fonts	Century Schoolbook, Source Sans Pro Light, DejaVu Sans Mono, Bitstream Vera Serif, URW Palladio L, Bitstream Vera Sans Mono, Bitstream Vera Sans, ...
HTTP DoNotTrack	1
Cookies enabled	Yes
Platform	Linux i686
OS	Linux 3.14.3-200.fc20.x86 32-bit
Screen resolution	1920x1080x24
Timezone	-480
DOM Session storage	Yes
DOM Local storage	Yes
I.E. User data	No

A special attribute: Canvas fingerprint

Browser fingerprinting

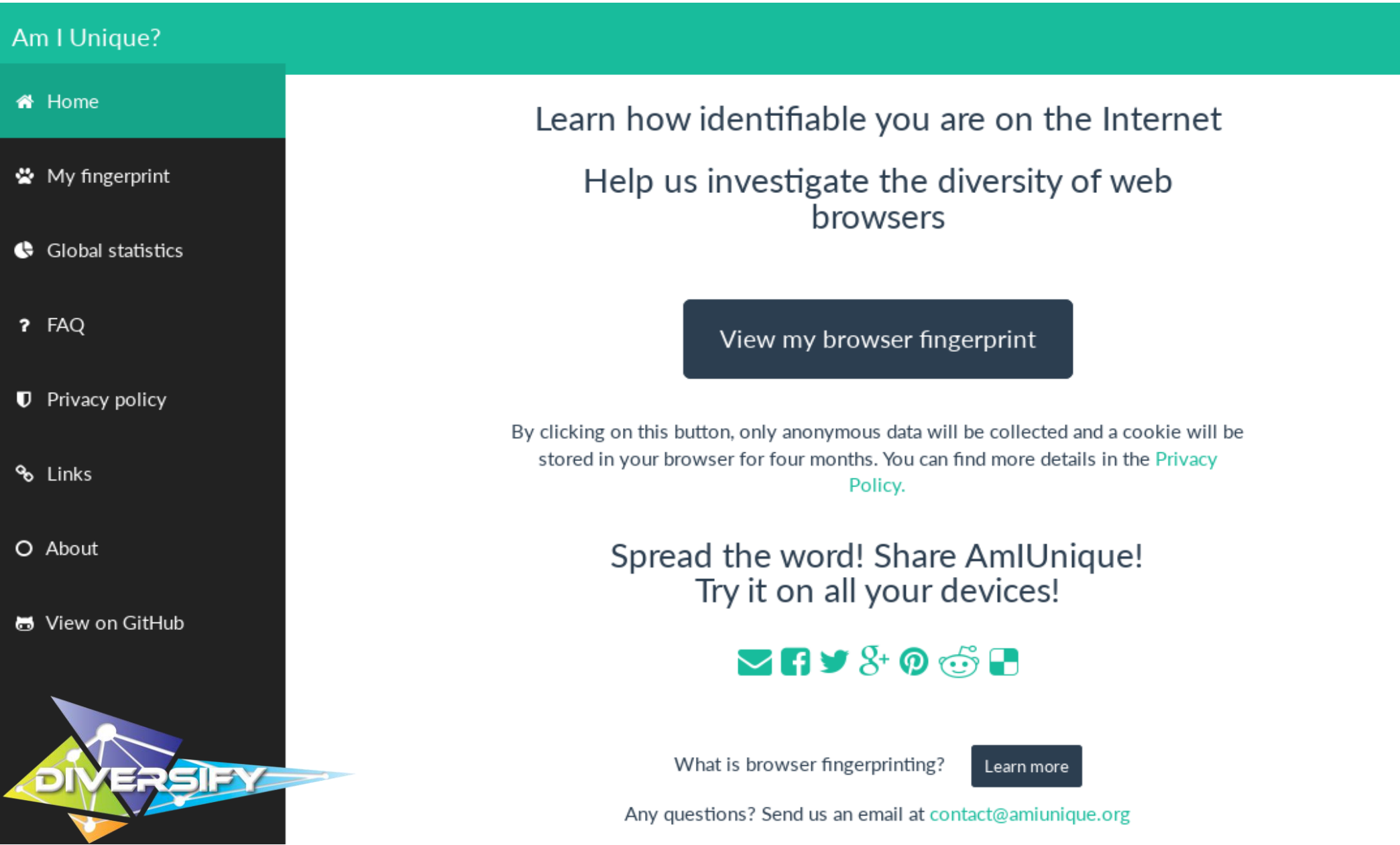
- Began in 2010 and is growing

  [PETS'10[1], CCS'13[2], CCS'14[3], CCS'16[4]]

- Some defenses exist
  - NoScript, randomization/blocking extensions, Brave, …
- Discovering new attributes all the time
  - Searching for unique and stable fingerprints

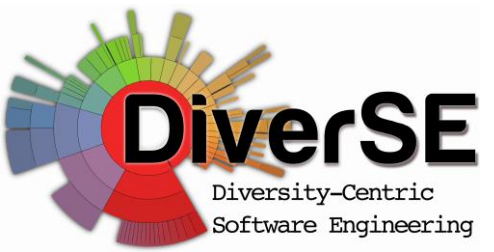
1. Eckersley, P. How unique is your web browser?
2. Acar, G. *et al*. FPDetective.
3. Acar, G., *et al.* The web never forgets: Persistent tracking mechanisms in the wild.
4. Englehardt, S., and Narayanan, A. Online Tracking: A 1-million-site Measurement and Analysis.

https://amiunique.org (2014)



http://amiunique.org (am I unique?)

- Inform users
- Study advanced fingerprinting techniques
- Highly visible project
  - Website, browser extensions, research papers, news articles, interviews, and lots of vulgarization
- More than 1 million fingerprints collected



Finding #1: Mobile fingerprints are also unique

- Mobile browser fingerprinting is feasible
  - High uniqueness (depends on make/model)
  - High stability
  - Different attributes (e.g., user agent, emojis)
- As browsers close privacy holes, they also add new APIs
  - Increased attack surface
  - Little thought to privacy

Beauty and the Beast: Diverting modern web browsers to build unique browser fingerprints [IEEE S&P'16]

Emoji fingerprinting [IEEE S&P'16]

(a) Windows 7

(b) Windows 10

(c) Linux

(d) iOS

(e) Firefox OS

(f) Android 4.3 and before

(g) Android 4.4

(h) Android 5.0

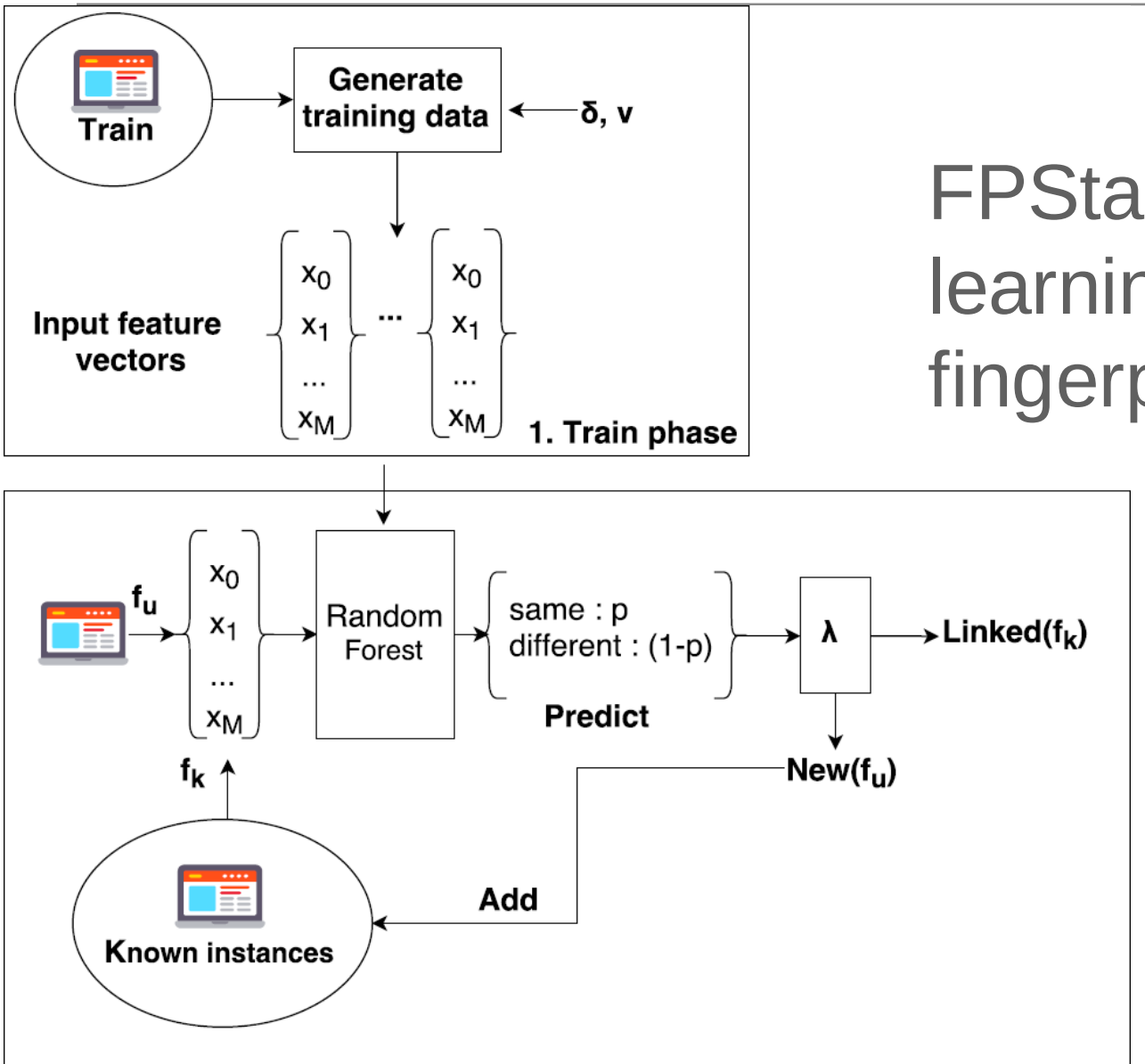
(i) Android on an LG device

(j) Android on a Samsung device

(k) Android on an HTC device

(l) Emoji not supported

Long-term fingerprint tracking

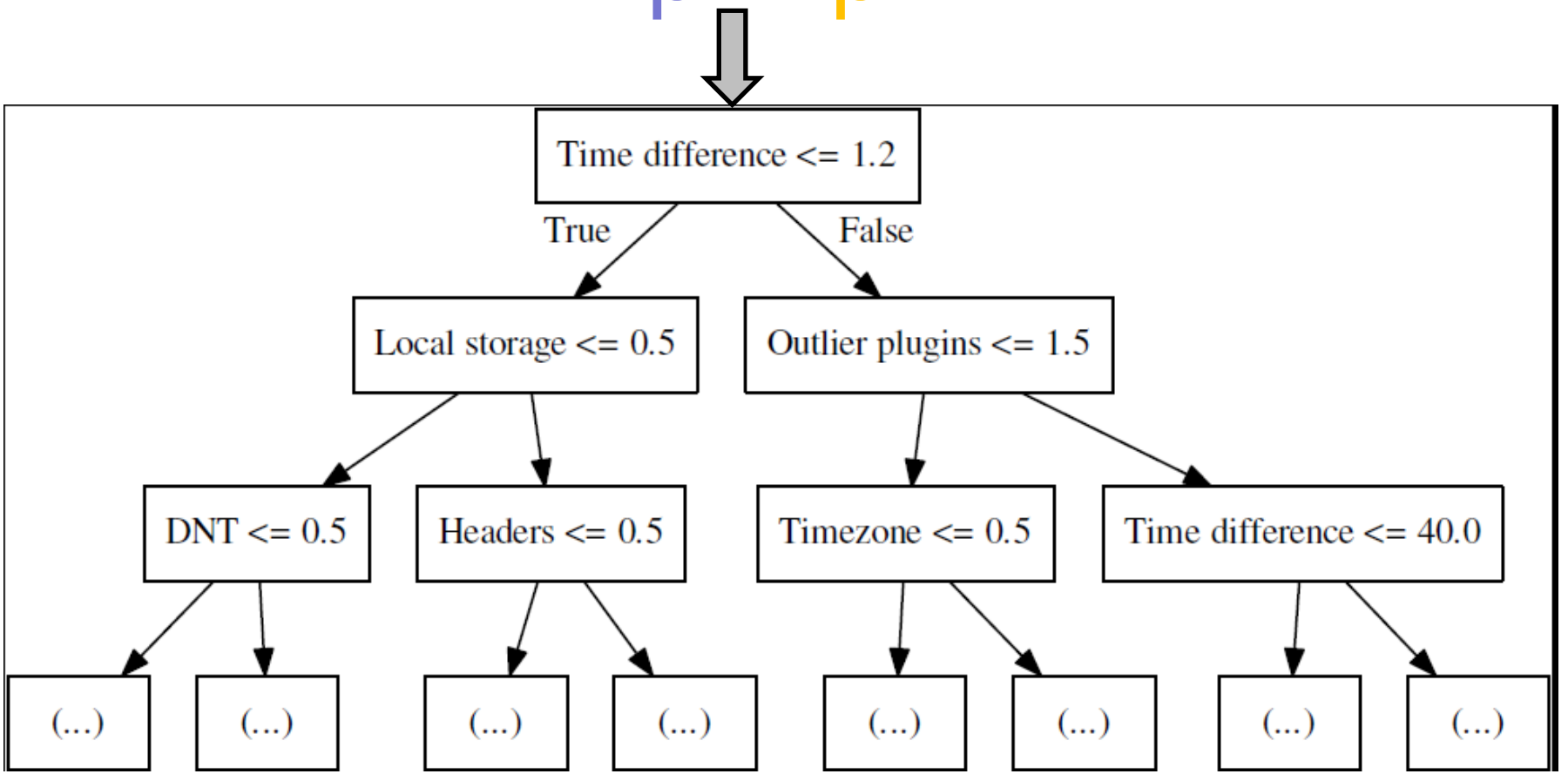


FPStalker: Machine learning for tracking fingerprint evolutions

[IEEE S&P'18]

Example of Decision Tree [IEEE S&P'18]

Fp Fp'



Output: Link Don't Link

Finding #2: Fingerprints can be tracked

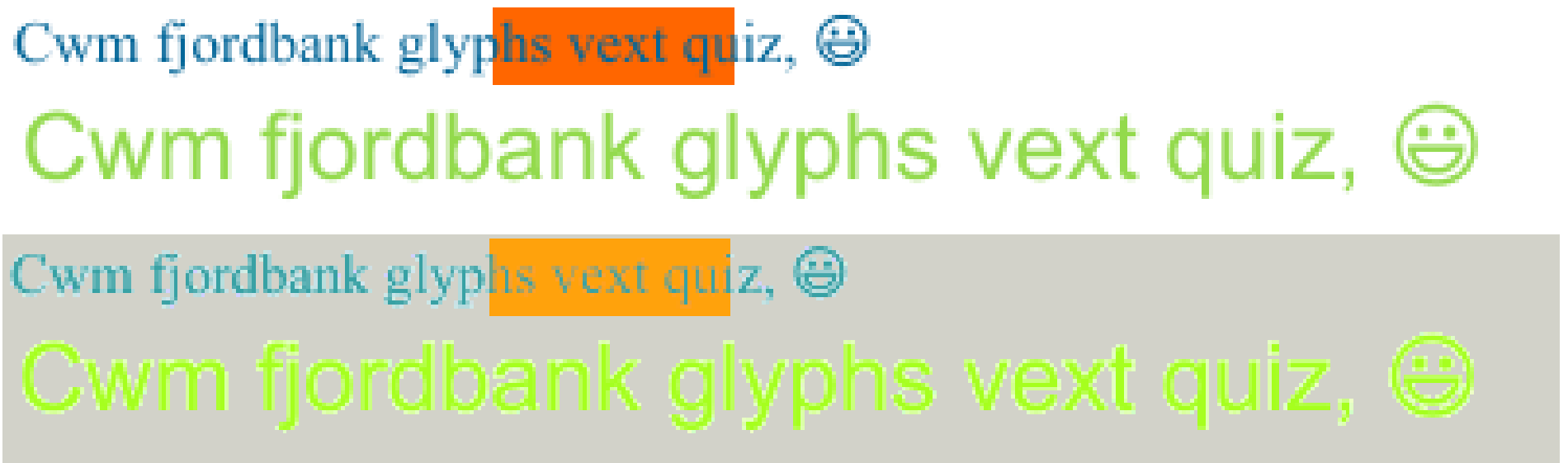
- ~26% of browsers are highly trackable
  - More than 100 days
  - If you try to hide you're worse off
- ~20% of browsers are "*untrackable*"
  - Fingerprints too similar (not enough attributes)
  - Unpredictable evolutions

[IEEE S&P'18]

What about fingerprinting defenses ?

FP-Scanner : Detecting incoherencies in countermeasures

[USENIX Security'18]



Finding #3: Current defenses are not very effective [USENIX Security'18]

- Current solutions are (very) bad
  - Can be counterproductive
- Privacy is difficult
  - Many side-effects to watch for
  - Spoofing must be coherent throughout multiple channels
  - Browsers are very complex
- The better solutions…
  - Integrated in the browser
  - Block or randomize attributes
  - Reducing APIs/Features

Can we design better tools?

- What do developers need ?
- What kind of feedback to provide?
- Can we automate this ?

Finding #4: We can design better countermeasures but it's hard

- Development and testing phases
  - Uniqueness and stability algorithms
- Automated reports for developers to inform on
  - Uniqueness (entropy)
  - Trackability (stability)
- Automate configuration exploration

FP-Tester : Designing fingerprint-proof browser extensions [IWPE '18]

Summary

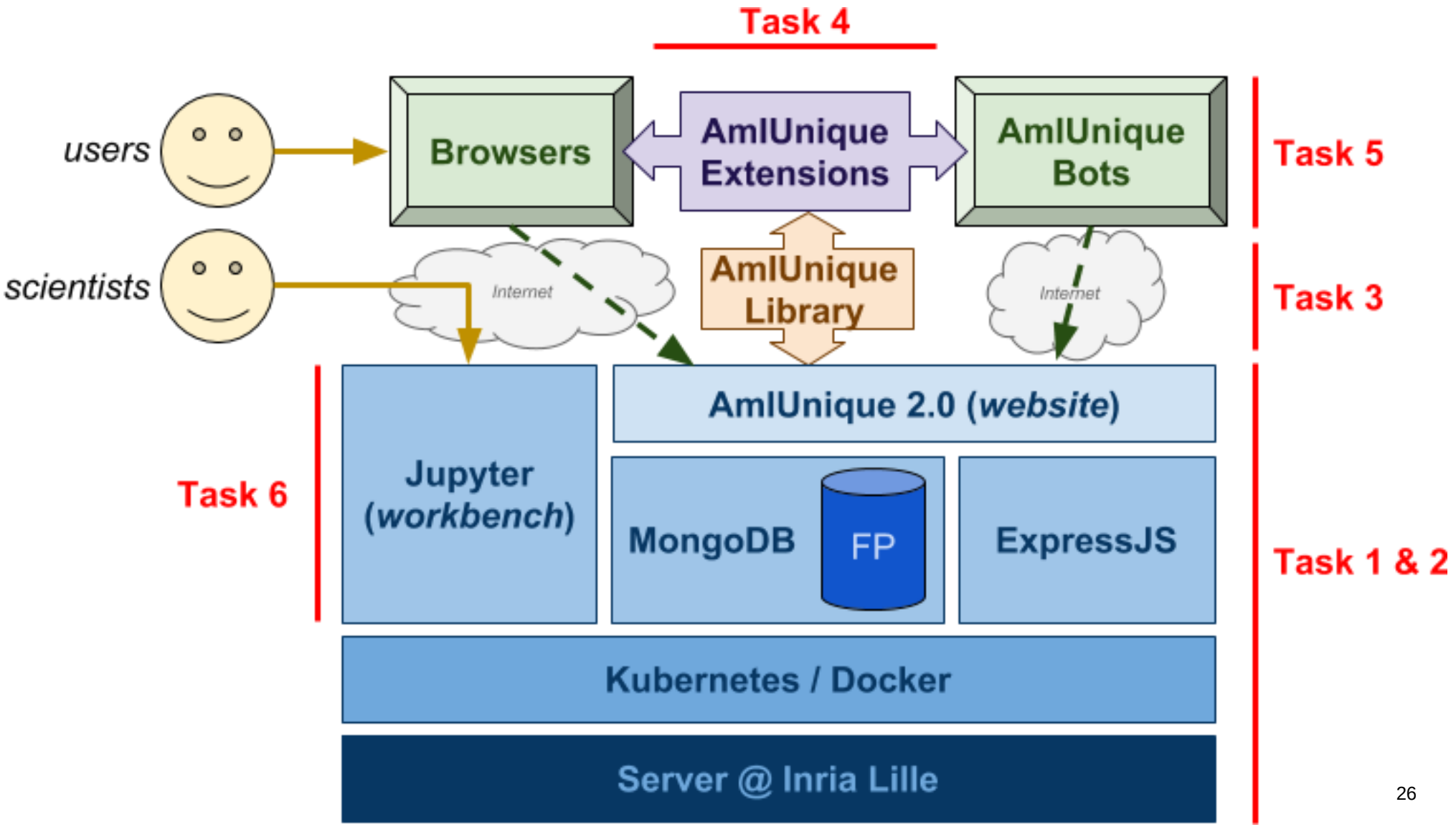
- Fingerprinting is a threat to privacy
  - Difficult to solve (arguably harder than cookies)
  - Use is growing
  - Can be used to track users and to complement cookies
- Currently no good defenses
  - Many are counterproductive
- But… it is possible to build better countermeasures
  - By providing automated tools

Perspectives

- Combine fingerprinting with other techniques
- Explore good uses for fingerprinting
  - Second-tier security, bot detection, fraud detection
    - Verify, block, honey pots, poisoning
- Design and build a fingerprint resilient browser
  - In-browser randomization, attribute blocking, whitelisting
- ADT FingerKit [proposal under review]
  - Spirals (Inria), Indes (Inria), Stony Brook University
  - New AmIUnique project + Extensions.inria.fr
  - Scientific workbench, tools to end-users

QUESTIONS ?

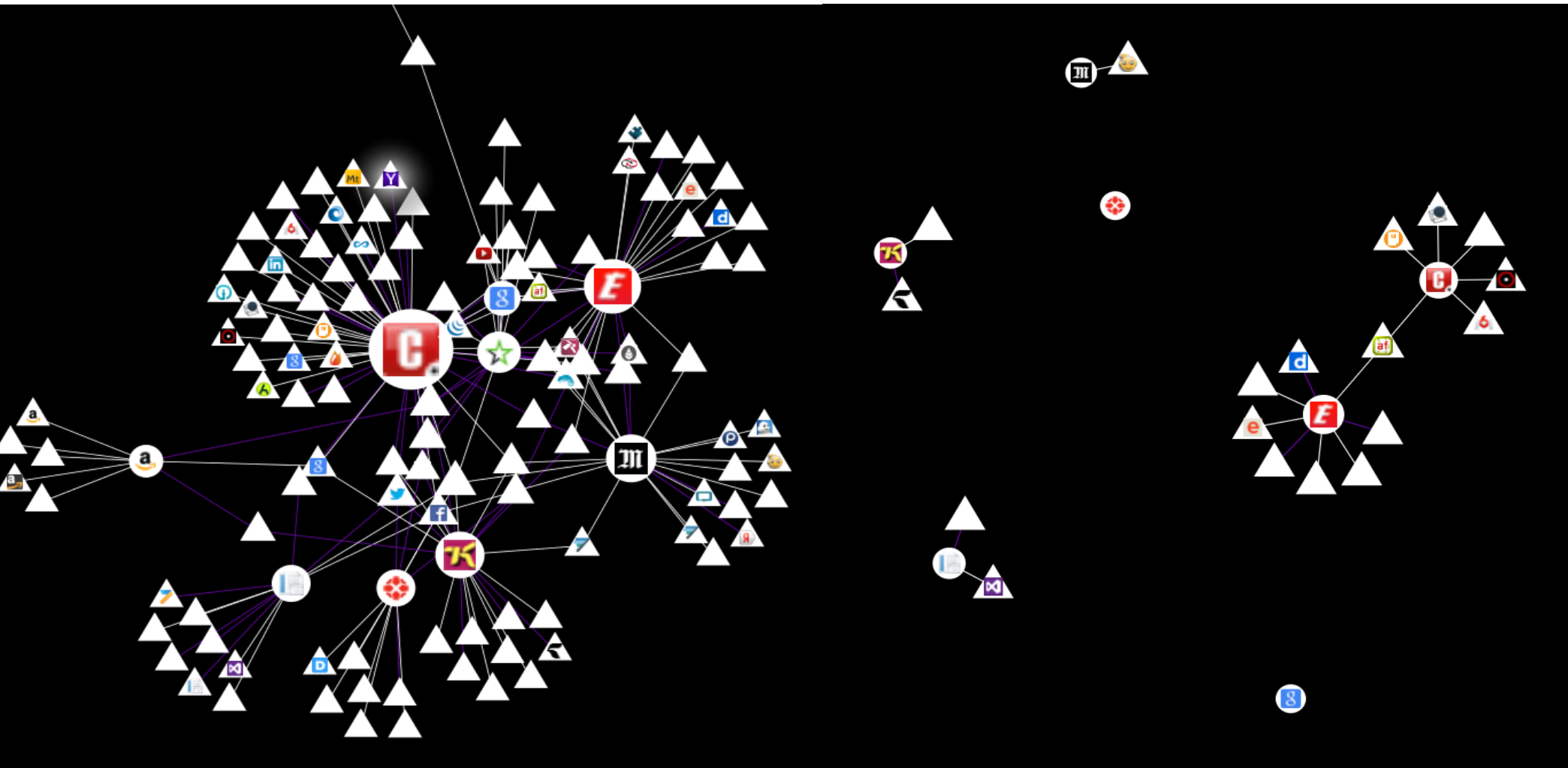
Towards AmIUnique2.0 with the ADT Fingerkit



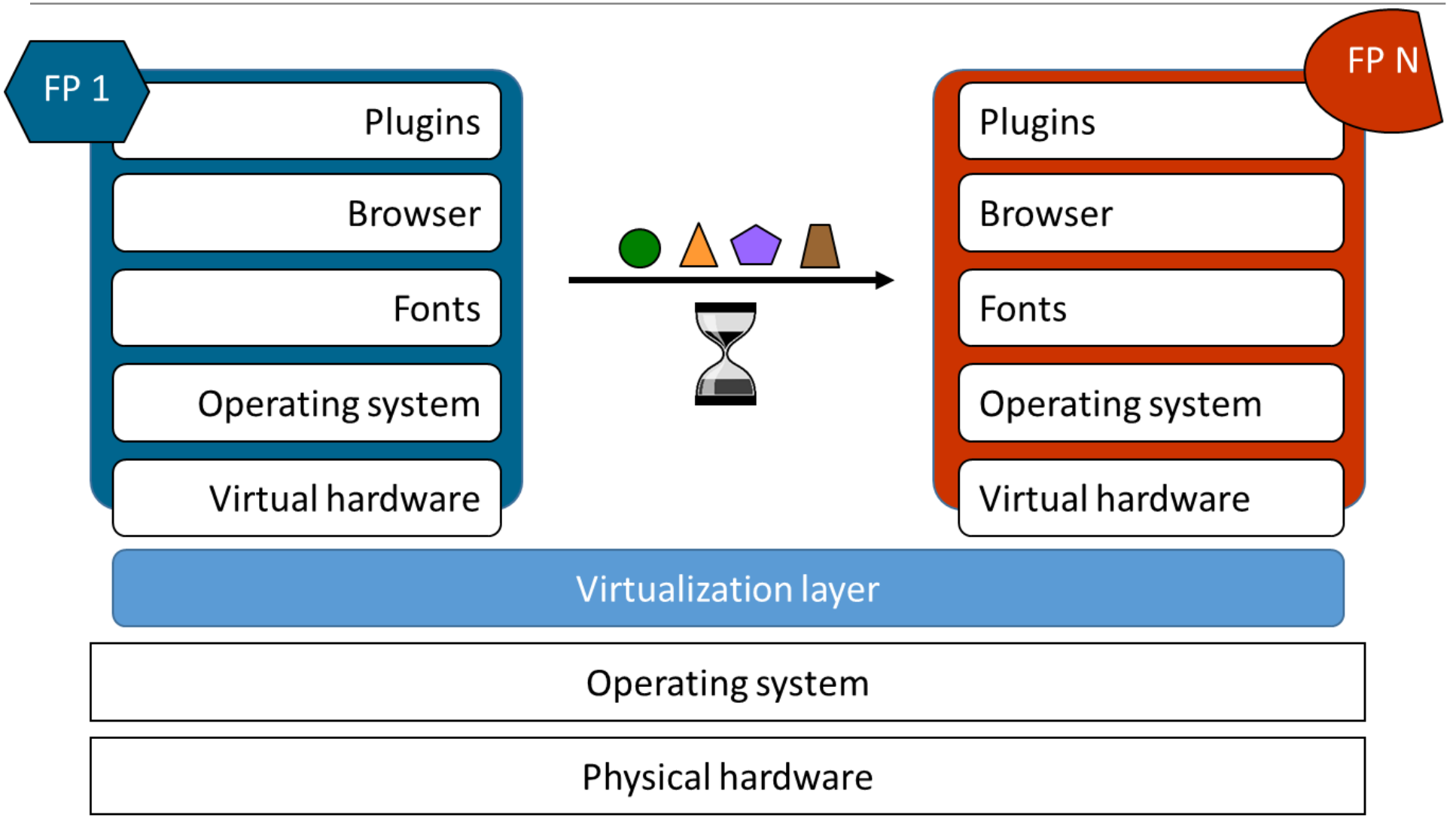
Lightbeam

Without Extensions

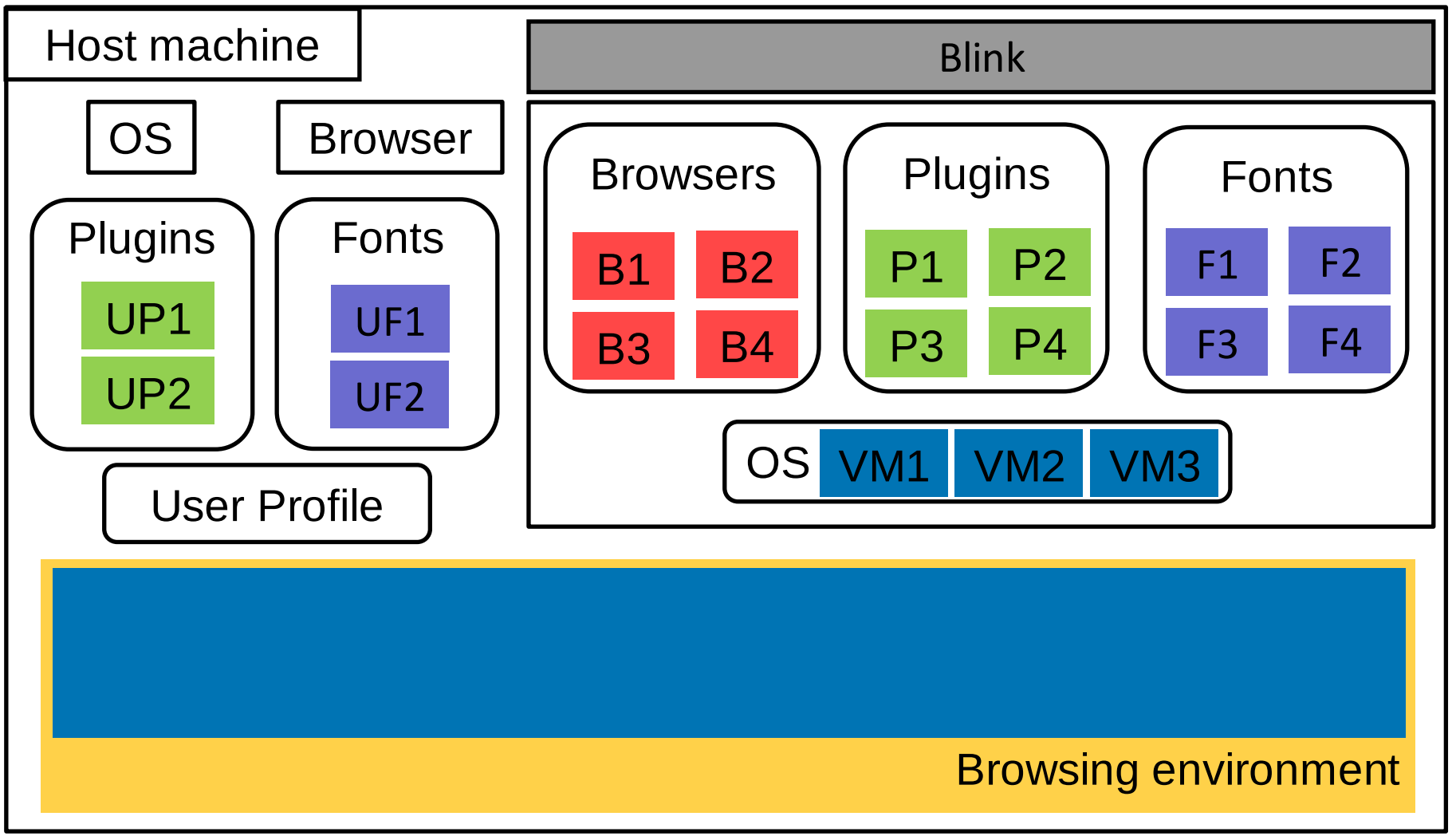
Ghostery + Adblock



Blink: Proactive diversification [SEAMS'14]



Automatic randomization of the platform: Leery mode



Automatic randomization of the platform: Coffee-break mode