[Ab]using web Intelligence in the Information Age: Data *Secrecy* and *Openness*

Corso di
Sicurezza delle Reti e dei Sistemi Software
aa. 2016/17

Ing. Antonio Pirozzi
antopirozzi[at]gmail{dot}com

whoami

- (Previous) Vulnerability Researcher for Emaze spa
- Research Fellow at University of Sannio
- ISWATlab Co-founder and Researcher
- KOINE Cyber Security Team Leader
- Philosopher

Andrew Grove - Intel

''Only the Paranoid will Survive''

nothing will be as it seems

Outline

Part 1

- The Intelligence Gathering Process
- Mass surveillance programs
- The OSINT Process
- DNS Intelligence
- Personal Information Extraction
- Metadata
- Misc
- Maltego

Part 2

- The web never forgets:
- Browser fingerprinting
- Mobile fingerprinting
- Thwarte Tracking

Intelligence Gathering

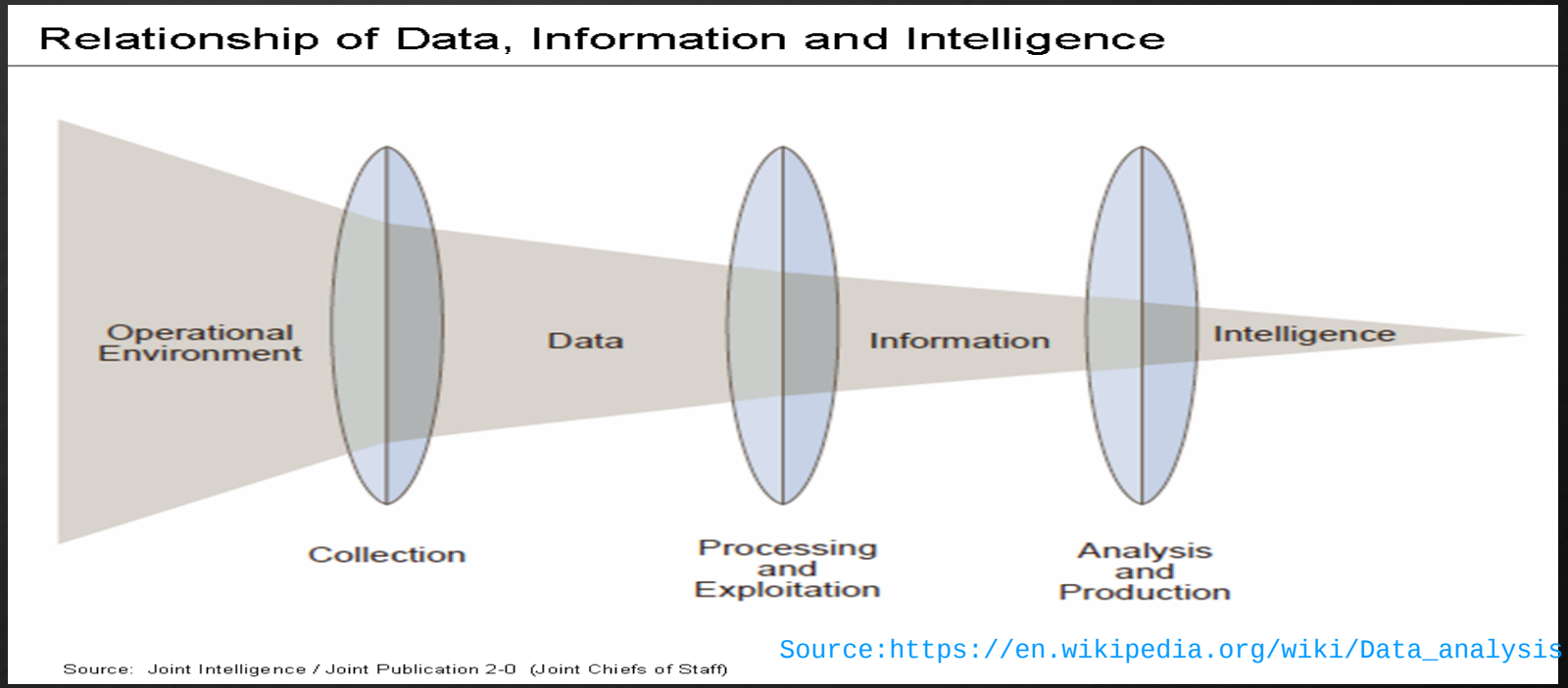
▶ **Open-Source Intelligence (OSINT)** refers to a broad array of information and sources that are **generally available**, including information obtained from the **media** (newspapers, radio, television, etc.), **professional and academic records** (papers, conferences, professional associations, etc.), **and public data** (government reports, demographics, hearings, speeches, etc.).

Cit. Fbi.gov

GEOINT
SIGINT
TECHINT
FININT
HUMINT

The Intelligence Gathering Process

The Unstructured becames Structured



Source:https://en.wikipedia.org/wiki/Data_analysis

Who collects and uses Intelligence

- Law Enforcment
- Military
- Criminals
- Spies
- Government
- Journalists
- Business
- Hackers

Before we start.....

I veggenti del terzo millennio

Before we start.

Before we start.

1) Peter Thiel: $ 500.00 Aug. 2004(Bilderberg)

Before we start.

1) Peter Thiel: $ 500.00 Aug. 2004(Bilderberg)
2) $12.7 million came from James Breyer (Accel Partners)

Before we start.

1) Peter Thiel: $ 500.00 Aug. 2004(Bilderberg)
2) $12.7 million came from James Breyer (Accel Partners)

3) Greylock Partners → Howard Cox
$25.5 million

Before we start.

1) Peter Thiel: $ 500.00 Aug. 2004(Bilderberg)
2) $12.7 million came from James Breyer (Accel Partners)

3) Greylock Partners → Howard Cox
$25.5 million

Before we start.

Keyhole Inc

1) Peter Thiel: $ 500.00 Aug. 2004(Bilderberg)
2) $12.7 million came from James Breyer (Accel Partners)

3) Greylock Partners → Howard Cox
$25.5 million

Before we start.

In 2009, Google Ventures and In-Q-Tel invested "under $10 million each

Keyhole Inc

1) Peter Thiel: $ 500.00 Aug. 2004(Bilderberg)
2) $12.7 million came from James Breyer (Accel Partners)

3) Greylock Partners → Howard Cox
$25.5 million

Before we start.

In 2009, Google Ventures and In-Q-Tel invested "under $10 million each

Recorded Future
CREATING AN INSIGHTFUL WORLD

Google

Palantir Technologies

Keyhole Inc

1) Peter Thiel: $ 500.00 Aug. 2004(Bilderberg)
2) $12.7 million came from James Breyer (Accel Partners)

3) Greylock Partners → Howard Cox
$25.5 million

IQT®
IN·Q·TEL

Before we start.

In 2009, Google Ventures and In-Q-Tel invested "under $10 million each

Recorded Future
CREATING AN INSIGHTFUL WORLD

Google

Palantir Technologies

Keyhole Inc

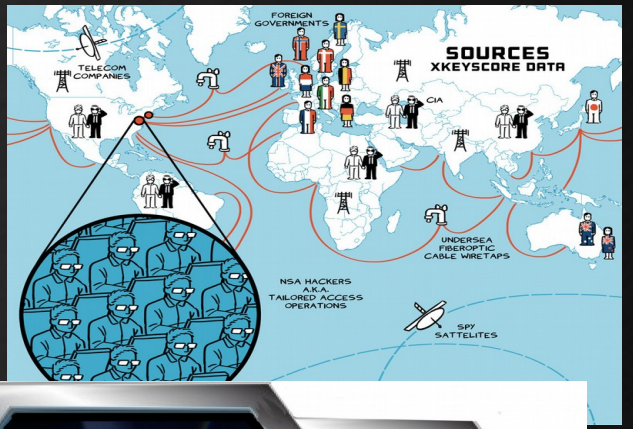
1) Peter Thiel: $ 500.00 Aug. 2004(Bilderberg)
2) $12.7 million came from James Breyer (Accel Partners)

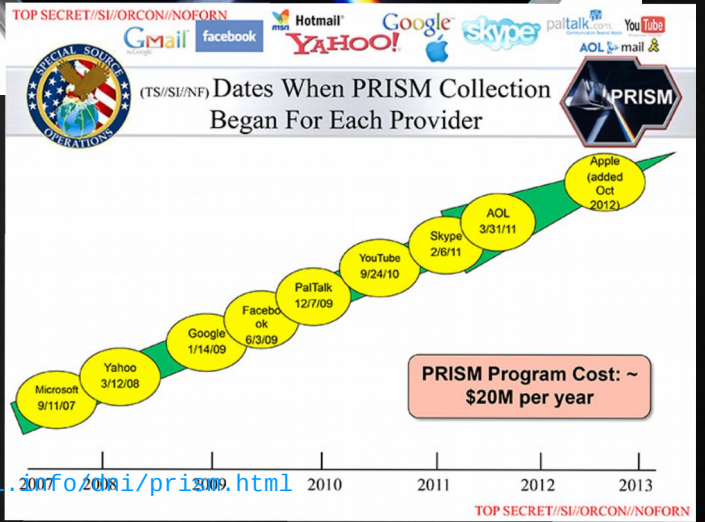
3) Greylock Partners → Howard Cox
$25.5 million

IQT®
IN·Q·TEL

mass surveillance

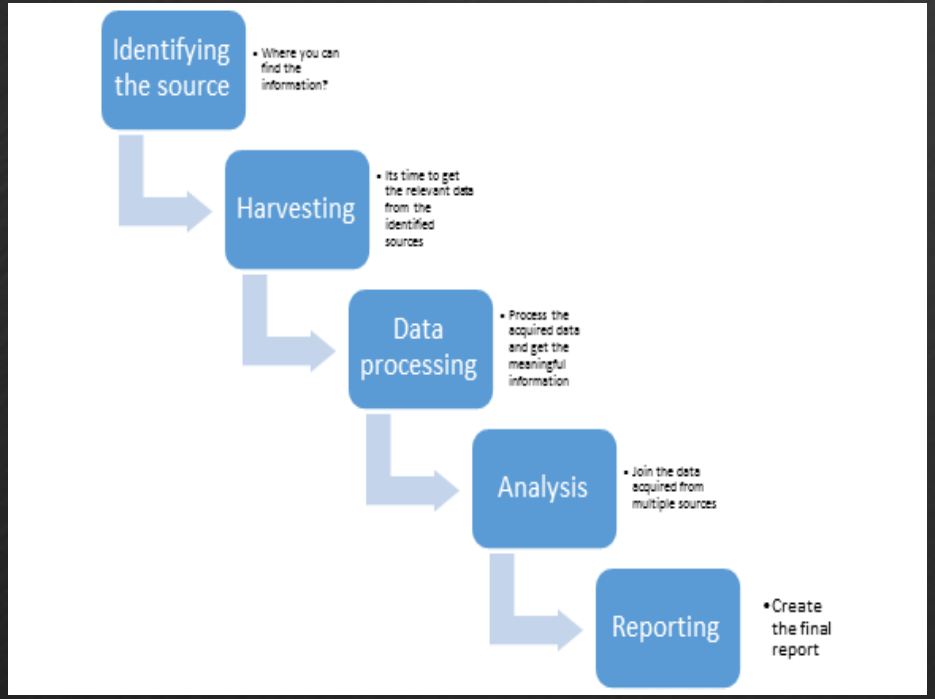
gemalto
security to be free





https://nsa.gov1.info/dni/prism.html

The OSINT Process



- Active harvesting
- Passive harvesting

- Hostname
- Services
- Networks
- SW/HW versions and OS information
- Geo-location
- Network diagram
- Database
- Documents, papers, presentations, spreadsheets and configuration files
- Metadata
- Email and employee search (name and other personal information)
- Technology infrastructure
- IP

DNS Intelligence

- Dig for DNS information
- Cache snooping
- Zone transfer
- Reverse lookup bruteforce
- Wildcard entries
- Fierce, dnsenum, dnsrecon, subroute
- IP to ASN

https://www.robtex.com/

- A (IP address)
- SOA (Start of Authority)
- CNAME (canonical name)
- MX (mail exchange)
- SRV (service)
- PTR (pointer)
- NS (name server)
- axfr

Doxing

Is the Internet-based practice of researching and broadcasting private or identifiable information (especially **personally identifiable information**) about an individual or organization. wikipedia

- Google
- Bing
- Peek you
- Pipl
- Intelius
- Facebook GRAPH
- Shodan

- Name
- Ages
- Email
- Addresses
- Phone numbers
- Photos
- Etc.
- Location
- IP CAMERA & IoT!!!

Google advanced operators

- **Web Search:** allinanchor:, **allintext**:, **allintitle**:, **allinurl**:, **cache**:, define:, filetype:, id:, inanchor:, info:, **intext**:, **intitle**:, **inurl**:, link:, **related**:, **site**:
- **Image Search:** allintitle:, allinurl:, filetype:, inurl:, intitle:, site:
- **Groups:** allintext:, allintitle:, author:, group:, insubject:, intext:, intitle:
- **Directory:** allintext:, allintitle:, allinurl:, ext:, filetype:, intext:, intitle:, inurl:
- **News:** allintext:, allintitle:, allinurl:, intext:, intitle:, inurl:, location:, source:
- **Product Search:** allintext:, allintitle:

personally identifiable information

Website investigation:

- https://Pipl.com
- http://www.amazon.com/gp/registry/search/
- http://www.indeed.com/resumes?q=%22PHm%22
- https://www.linkedin.com/vsearch/f?trk=federated_advs&adv=true
- Cached data: http://archive.org/web/, google cache,
- Knowem,
- checkusernames.com/
- https://namechk.com/

US only:

- http://www.classmates.com/,
- https://www.intelius.com/ ,
- http://radaris.com/
- http://www.spokeo.com/
- http://www.peekyou.com/
- http://mugshots.com/

Code:
https://nerdydata.com

Bitcoin :

- https://blockchain.info/address/1Kqzbv4ekpJX3ohYWGEzMqzvf27VjBux35
- https://www.blockseer.com/addresses/1Kqzbv4ekpJX3ohYWGEzMqzvf27VjBux35

Emails:

- https://emailhunter.co
- https://toolbox.googleapps.com

Skype :

- http://www.skresolver.com/

Companies/ business :

- https://findthecompany.com
- http://copyright.gov/
- http://www.keywordspy.com/
- https://www.crunchbase.com
- https://connect.data.com

personally identifiable information

- The Harvester
- Fb
- Social Media
- Recon-ng
- Metadata

Geo-Location

- Gmaps
- Gearth

Private earth observation satellites:

45 cm resolution

GeoEye – 5 satellites: IKONOS, OrbView-2, OrbView-3, GeoEye-1, GeoEye-2
DigitalGlobe – 4 satellites: Early Bird 1, Quickbird, WorldView-1, Worldview-2
Spot Image – 2 satellites: Spot 4, Spot 5

SOCMINT

- FacebookGRAPH
- Fbsleep
- Creepy
- http://www.socialmention.com

Reverse Image Search

- Google Image
- Tineye
- Yandex
- Googles (on G Play Store)
- Findface (on G Play Store)

Ad Analytics

Id <--> domain correlation

- http://spyonweb.com/
- http://sameid.net/
- https://ewhois.com

Google AdSense

ca-pub-9158781978326526

```
data-ad-client="ca-pub-9158781978326526" data-ad-
```

UA-6197637-22

```
,document,'script','//www.google-analytics.com/analytics.js','ga');

e', 'UA-6197637-22', "semrush.com");
re', 'displayfeatures');

ming=function(){function h(d,c,b){var e="XDomainRequest"in window?"
```

Whois intelligence

- Domain related information
- Personal information
- Contact details
- domaintools

- **The registry** – Stemming from ICANN, branching bodies manage top level domains (tld) such as com and org
- **The registrar** – Where we purchase our domain names
- **The registrant** – The purchaser of the domain

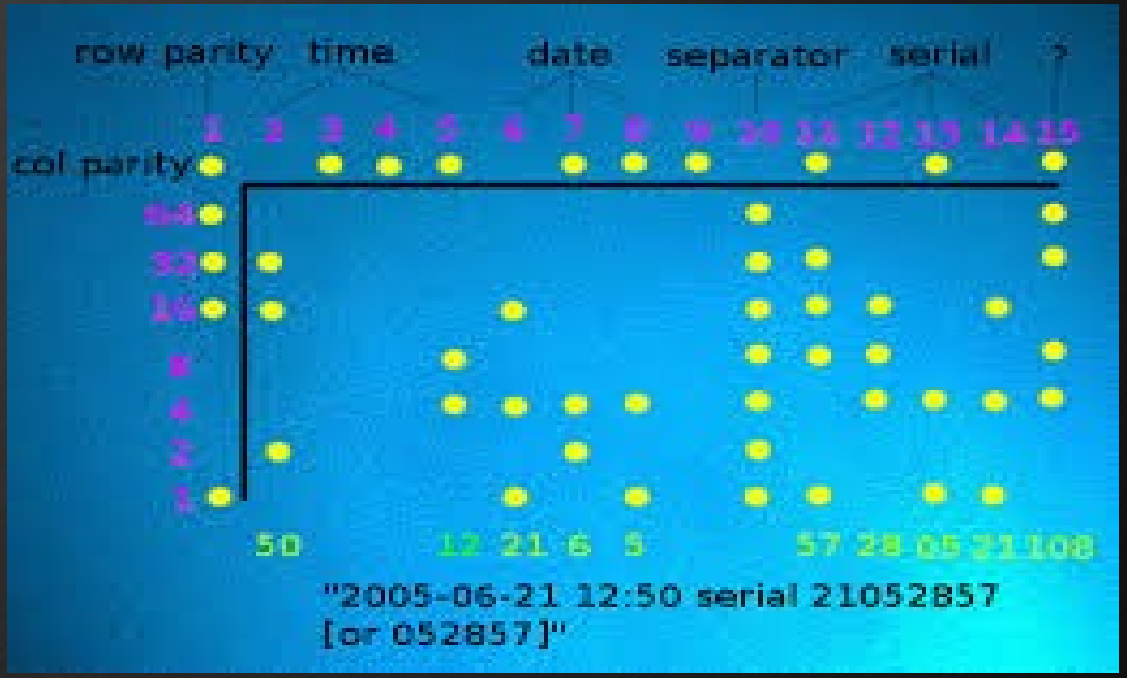
Metadata | 'data about data'

- Foca
- Metagoofil
- Exiftool <exif foto e docs>
http://www.motherjones.com/files/images/war_photo_063010.jpg
- gpg –- with-fingerprint key.asc
- Printers **SRSLY?????**

http://www.repubblica.it/esteri/2015/06/05/news/militante_dell_is_si_fa_un_selfie_e_24_ore_arrivano_3_missili-116096314/

Printers dot CODE

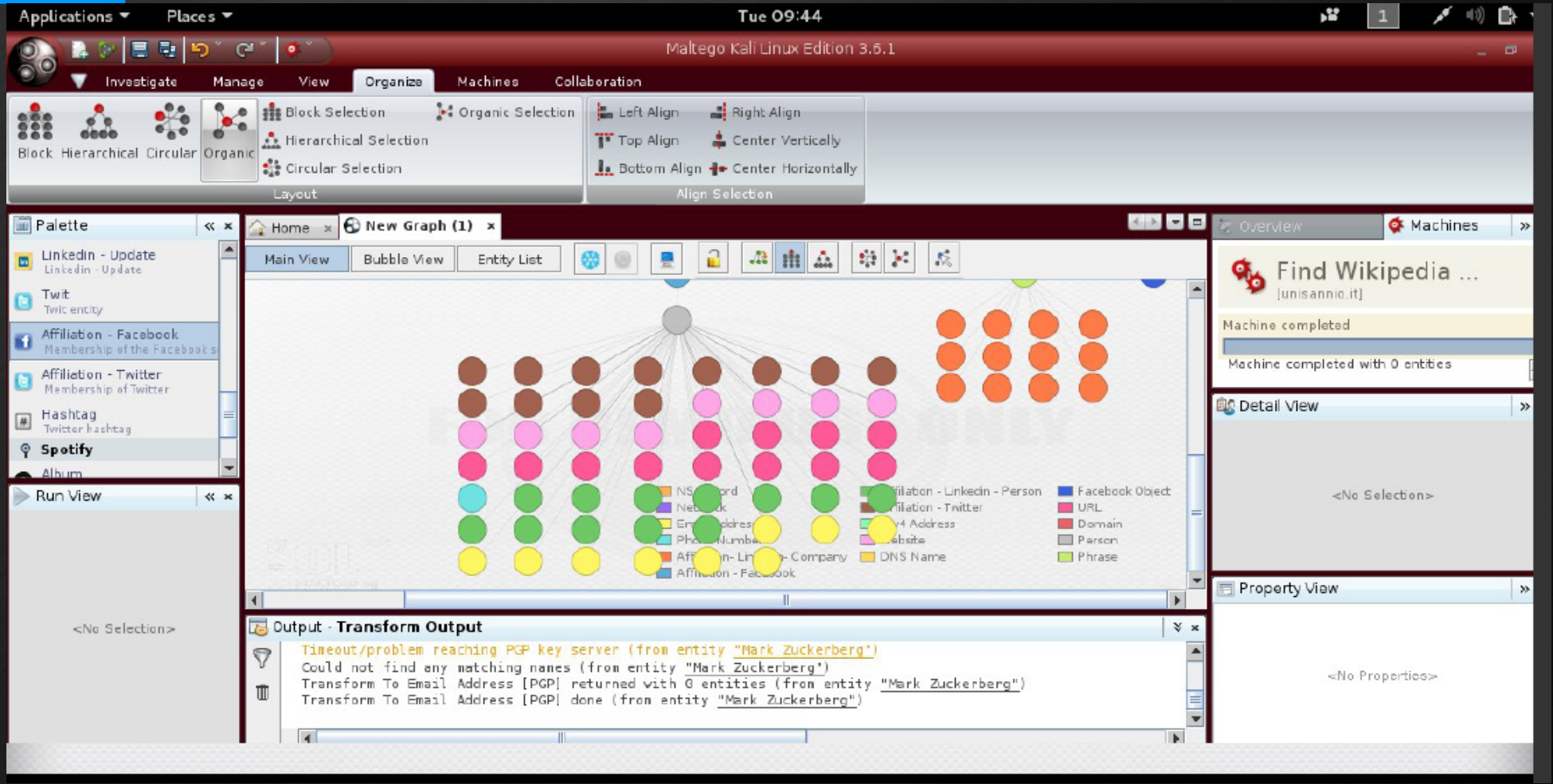
2005 EFF discover it



List of printers:

https://www.eff.org/it/pages/list-printers-which-do-or-do-not-display-tracking-dots

MALTEGO :D emo




PRIVACY
NO PRIVACY ANYWHERE
memegenerator.net

The web NEVER FORGETS:
browser fingerprinting techniques

- HTML5 canvas fingerprinting (obtained in a fraction of a second without user's awareness.) (**Whitehouse.gov,perezhilton.com**)
- cookie syncing
- Evercookies & Respawning (cookies in a web browser that are intentionally difficult to delete / **NSA used it for tracking TOR users**)
- Local shared objects (LSOs) aka **FLASH cookie**

→ More than 94% of Flash - and Java-enabled browsers can be uniquely identified
→ More he changed his plugins and settings relative to the default, the more unique and easily-identifiable his browser becomes

https://securehomes.esat.kuleuven.be/~gacar/persistent/

The web NEVER FORGETS:

browser fingerprinting techniques

- Facebook, Google and Microsoft, will assign a unique identifier to each type of device the user has and link those together to track activity across all of the devices the person uses. These new tracking mechanisms, if they catch on, could be used across each vendor's ecosystem -- and beyond.
- browser and device "fingerprinting" are *cookieless*

https://firstpartysimulator.net/ From EFF
https://amiunique.org

Types of Tracking Solutions on Mobile Platforms

cross-device user identification

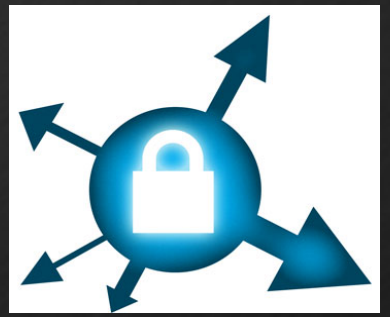
- Client- or Device-specific Ids (Users cannot alter or opt out of tracking) :

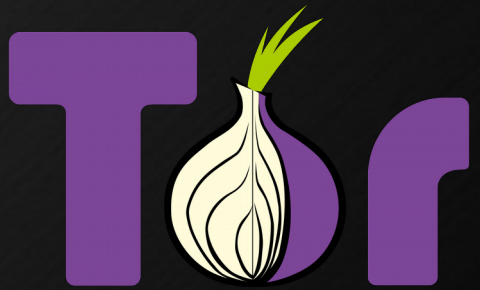
  from Android ID and iOS UDID to **GAID** for Android, **IDFA** for iOS

- SSO
- Cookies on mobile apps

countermeasures

- Do Not Track (DNT)
- Opt-out on DAA
  http://www.aboutads.info/choices/
- **Ghostery** Firefox ext.
- **NoScript** of Course!
- Disable third-part cookies
- AdBlock plus
- HTTPS Everywhere
- VPN, TOR
- Gtranslate





Any Questions?