

Who ~~governs~~ controls the Internet?

A case study on using data for the public good

Giovane C. M. Moura

SIDN Labs and TU Delft

Studienstiftung des Deutschen Volkes Summer School

Driebergen, The Netherlands

2023-08-11



Today's Goals

- Understand how the Internet is governed
 - Is there an Internet government?
- Understand how governance differs from Internet control
- Understand how data is being **collected and used** for profit
- Showcase alternative usage of data for the public good
- extra: receive some unsolicited advice



img src: [wallpaperflare](#)

Text in red is clickable, you can get the slides later

Challenge

Diverse crowd:

- 76 students from everywhere:
 - CS, Political Science, Engineering, Intl. Relations, Philosophy, Physics, Math, Social Sciences, Economics, Medicine, Physics
- 26 from College d'Europe:
 - Political Science, International Relations, Law, Economics

Best students from Germany



img src: [Unsplash](#)

- Data Scientist at **SIDN Labs**
 - research team of SIDN, .nl registry
- Assistant Professor at **TU Delft**
- Research focus on **operations**:
 - Internet Security
 - Networking
 - Systems
- PhD (2013, **UTwente**, NL)
- MSc (2008, **UFRGS**, BR)



Presentation @ RIPE86, Rotterdam, May 2023

Governance and Consolidation: end of privacy

Using data for the public good

Counterfeited webshops

Logo Misuse

Wrapping up

Internet Governance

- Processes and policies that shape the use, development, and operations of the Internet.
- Multi-stakeholder approach



Internet Consolidation

- Centralization of power/uses/data in few players
 - Datenkraken
- EU has no big-tech; US and China do



How did we get here?

Internet Governance vs Consolidation

Internet Governance

- Processes and policies that shape the use, development, and operations of the Internet.
- Multi-stakeholder approach



Internet Consolidation

- Centralization of power/uses/data in few players
 - Datenkraken
- EU has no big-tech; US and China do



How did we get here?

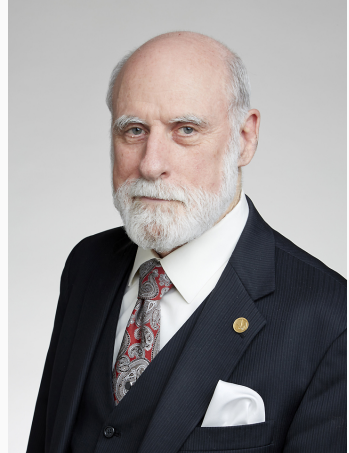
Q: Who *invented* the Internet?

A: Vint Cerf?



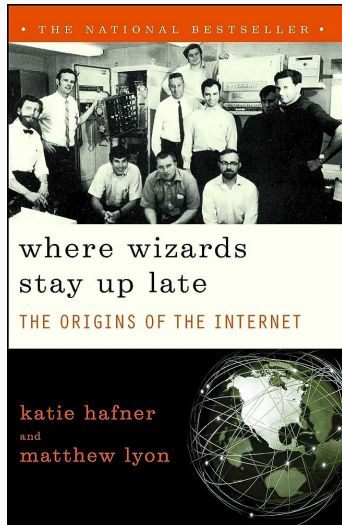
Q: Who *invented* the Internet?

A: Vint Cerf?

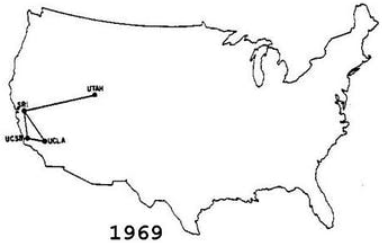


Internet History

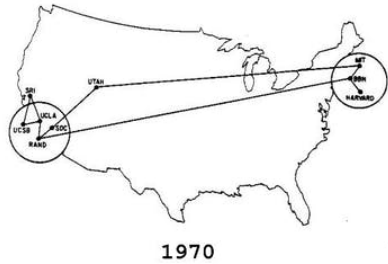
- The Internet has many parents
- Original goal:
 - few uni/research center had their own supercomputers
 - they costed a fortune, were disconnected
 - each uni wanted its own
 - ARPAnet was proposed to connect them
 - **to save money**
 - BBN in Boston built it



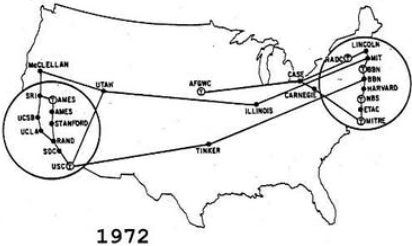
ARPAnet (later became Internet)



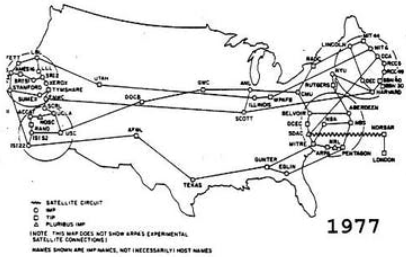
1969



1970



1972



1977

Growing too fast: need some sort of governance

- It was growing too fast
 - Need a way to make sure it kept working
 - Protocols
 - Name space (DNS)
 - IP addresses
 - Open governance models were implemented
- For example, who should manage the .de TLD?
 - Bundesregierung?
 - Siemens?
 - Answer: denic.de
 - BTW, DDR had its .dd TLD
 - you could uni-jena.dd

Growing too fast: need some sort of governance

- It was growing too fast
 - Need a way to make sure it kept working
 - Protocols
 - Name space (DNS)
 - IP addresses
 - Open governance models were implemented
- For example, who should manage the .de TLD?
 - Bundesregierung?
 - Siemens?
 - Answer: denic.de
 - BTW, DDR had its .dd TLD
 - you could uni-jena.dd

Growing too fast: need some sort of governance

- It was growing too fast
 - Need a way to make sure it kept working
 - Protocols
 - Name space (DNS)
 - IP addresses
 - Open governance models were implemented
- For example, who should manage the .de TLD?
 - Bundesregierung?
 - Siemens?
 - Answer: denic.de
 - BTW, DDR had its .dd TLD
 - you could uni-jena.dd

Growing too fast: need some sort of governance

- It was growing too fast
- Need a way to make sure it kept working
 - Protocols
 - Name space (DNS)
 - IP addresses
- Open governance models were implemented
- For example, who should manage the .de TLD?
 - Bundesregierung?
 - Siemens?
 - Answer: denic.de
- BTW, DDR had its .dd TLD
 - you could uni-jena.dd

Growing too fast: need some sort of governance

- It was growing too fast
 - Need a way to make sure it kept working
 - Protocols
 - Name space (DNS)
 - IP addresses
 - Open governance models were implemented
- For example, who should manage the .de TLD?
 - Bundesregierung?
 - Siemens?
 - Answer: denic.de
 - BTW, DDR had its .dd TLD
 - you could uni-jena.dd

Internet Governance

- ICANN and RIRs: preside over IP address space and domain name space
 - in Europe: **RIPE NCC** handles IP addresses (Amsterdam)
 - in Germany: DENIC does .de
 - Netherlands: SIDN does .nl
- IETF: standardize Internet protocols
 - open, compatibility
- Many other bodies

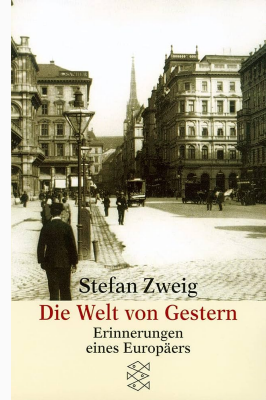
My personal thoughts:

- What would have happened if ArpaNET remained with the DoD?
- Built by other country?
- Current model it is not perfect but it works
- Thanks to open spirit of pioneers
 - Read: **HEISE.de**

Speaking of openness

- Great reading about what happens when group identity is hijacked for evil
 - Specially for College d'Europe folks
- “We live in space age with stone age brains”
- Jared Diamond “The world until yesterday“ book
 - how they hate their neighbors

Masterclass of Humanism



Philosophically speaking

My grandfather (1913)
Offline folks



Generation X
Digital Migrants



Sabrina Parlatore,
MTV Brazil
(I wrote a handwritten
letter to this show)

Generation Z
Digital Natives



Internet consolidation and centralization

- The Internet was very open in the beginning
- Commercial use allowed; emergence of ISPs
- Network effects
 - more users → more value for the websites/platforms
- Economies of scale
 - bigger is cheap – see Amazon cloud AWS
- Buyouts and merges
- Lack of regulatory and laws

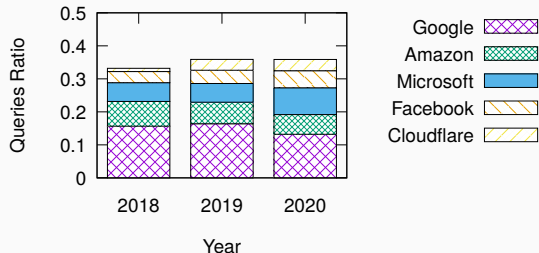


Margrethe Vestager

Internet centralization

- Big tech buys smaller companies
 - prevent innovation
 - Eg: Google/Youtube, Facebook/Instagram
- Network effects
- Traffic/user concentration
 - **Sandvine**: Google, Facebook, Microsoft, Apple, Amazon & Netflix generate almost 57% of Internet traffic

- **Research paper**: 1/3 of .nl traffic from US big tech



Big tech business model

Surveillance Capitalism

Society books

Interview

Shoshana Zuboff: 'Surveillance capitalism is an assault on human autonomy'

Joanna Kavenna

What began as advertising is now a threat to freedom and democracy argues the author and scholar. Time to wake up - and fight for a different digital future



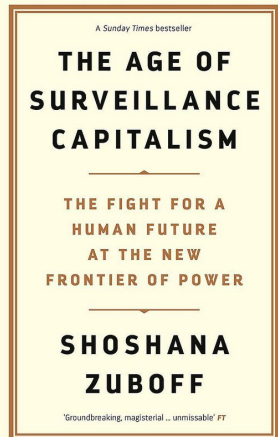
Article: The Guardian

Advertisement: psychological manipulation

The book cover for 'The Attention Merchants' by Tim Wu. The background is bright yellow. At the top, the title 'THE ATTENTION MERCHANTS' is written in large, bold, black capital letters. Below the title, the subtitle 'The Epic Scramble to Get Inside Our Heads' is written in a smaller, bold, black font. The author's name 'TIM WU' is in a white box with black text. Below the name, it says 'Author of The Master Switch'. On the right side, there is a blue speech bubble icon with a white smiley face and a red notification bubble with the number '23'. At the bottom, a quote reads '"DAZZLING." - FINANCIAL TIMES'. A fishing hook is positioned on the right side, with the hook's eye at the top and the handle curving down towards the speech bubble icon.

Surveillance Capitalism

- **New Economic System:** companies extract and monetize vast amounts of personal data
- **Behavior Prediction:** spy on you to **predict** your next steps
- **Selling predictions:** they sell predictions about your behavior
- **Manipulation:** can influence people thoughts and feelings
- **Erosion of democracy:** unchecked power of tech companies influence and shape human behavior could undermine democratic processes, autonomy, and individual agency.



Cloudburst demo tool

- My colleagues at SIDN Labs developed a tool just to raise awareness
- You can connect to their wifi and choose which cloud you want to disable
- You see tons of stuff stop working



Cloudburst

How centralised is your Internet?

Welcome to Cloudburst, where you can experience the effects of the centralisation on the Internet. What happens when a large cloud provider goes offline? Test it here! For more information read our [FAQ](#).

Select the clouds to block and click 'Burst' to activate.

Google

Amazon Web Services

Microsoft

CloudFlare

Akamai

Meta/Instagram

Burst →

What about Die Ärzte?



Not these

What about Die Ärzte?



Not these

What about Die Ärzte?

These



What about Die Ärzte?

- Amazon bought OneMedical for US\$ 3.9B
- 200 clinics; 815k patients (mostly online)
- What will they do with the data?



The image shows a screenshot of a Forbes article. At the top, the Forbes logo is displayed in white on a black background. Below the logo, the breadcrumb navigation reads "FORBES > INNOVATION > HEALTHCARE". Underneath, it says "EDITORS' PICK". The main headline is "Amazon's Latest Healthcare Expansion: Virtual Clinics In All 50 States" in a large, bold, black font. Below the headline, the author's name "Bruce Japsen" is listed as "Senior Contributor" with a small circular icon. A bio line reads "I write about healthcare business and policy". To the right of the bio is a blue "Follow" button. At the bottom right of the article preview, the date and time "Aug 1, 2023, 08:11am EDT" are shown.

Medical data is strictly controlled

- True, but.
- Google Nest microphone:
“ it was a error”

New Technology > Security

Google Calls Hidden Microphone in Its Nest Home Security Devices an 'Error'

The company says its was an oversight, but it does little to stem paranoia.

 BY SAM BLUM PUBLISHED: FEB 21, 2019

 SAVE ARTICLE

Previous Examples

- Google Maps collected Wi-Fi network names AND data
- Discovered by by Germany's data protection authority
- Google: "It was a mistake"

Google admits collecting Wi-Fi data through Street View cars

German request for data audit reveals the web giant 'accidentally' stored payload information from open networks



Link: [The Guardian](#)

GDPR and Spotify

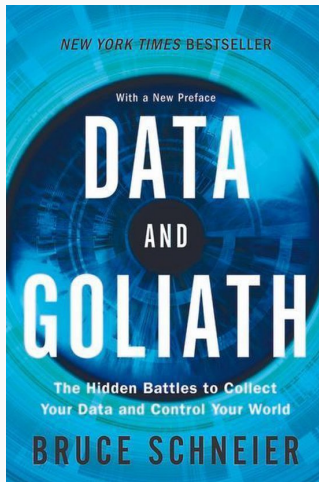
1. Who here uses Spotify?
2. Who knows what GDPR is?
3. Who knew you could ask them for your data?



link: [here](#)

We live in the privacy dark ages

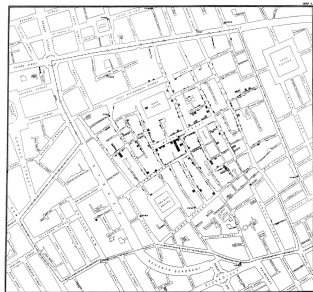
For those interested in the topic



Data is the new oil

- Surveillance capitalism erodes privacy in spying people for profit
- They (can) manipulate peoples feelings and emotions
- Are there alternative models to that, of using data for the public good?

Data use for public good: **John Snow's maps of cholera outbreak** in London: showed germ-contaminated water was the source



Governance and Consolidation: end of privacy

Using data for the public good

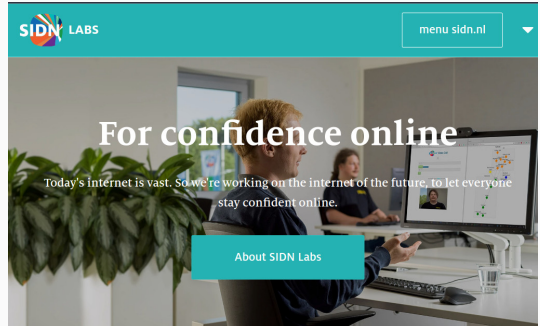
Counterfeited webshops

Logo Misuse

Wrapping up

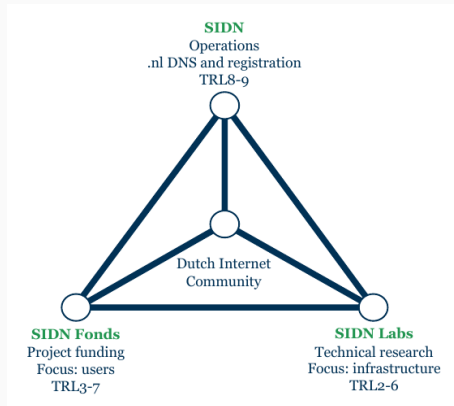
Using data for the public good

- Data can also be used for the public good
- I will show one example of how we do it at SIDN Labs

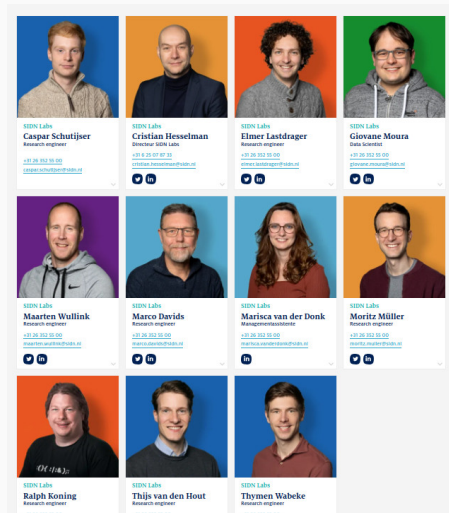


Background on SIDN

- SIDN operates the .nl TLD
- Not-for-profit private organization for the benefit of Dutch society (public role)
- Services we provide:
 - domain name registration
 - domain name lookups (DNS servers)



- Research arm of SIDN
- Goal: increase Internet infrastructure security through technical research
 - Focus on the Netherlands but also globally
- Three main areas:
 - Domain name security
 - Infrastructure security
 - New technologies
- Deliverables: academic papers, systems, software, standardization
- Bridge between industry and academia



Governance and Consolidation: end of privacy

Using data for the public good

Counterfeited webshops

Logo Misuse

Wrapping up

Back in 2016 ... strange websites

- We stumbled on these websites while looking for phishing
- They were rather *odd*
- We had many questions:
 1. does anyone even *buy* from them?
 2. what is their *business model*?
 3. how many they were (on .nl)?
 4. what can we do about it?

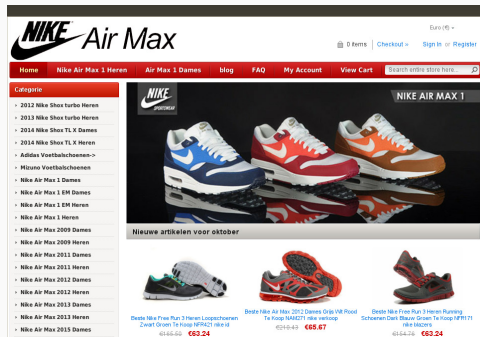


Figure 1: Screenshot of 2016 .nl website

Does anyone even buy from them?

- Yes, they were
- Scam: getting fake or no product
- Dealing with financial losses



Figure 2: NOS news (2018)

OK, so what to do about it

- SIDN is a Internet registry, not police
- But we have a mission to make the .nl zone safer for users
- And we were sitting on the data
- Ethical dilemma:
 - Turn the blind eye OR
 - Do something about it
- We talked to our lawyers
- We need to conform to our mandate and EU and NL laws

We decided to go ahead and measure it

OK, so what to do about it

- SIDN is a Internet registry, not police
- But we have a mission to make the .nl zone safer for users
- And we were sitting on the data
- Ethical dilemma:
 - Turn the blind eye OR
 - Do something about it
- We talked to our lawyers
- We need to conform to our mandate and EU and NL laws

We decided to go ahead and measure it

OK, so what to do about it

- SIDN is a Internet registry, not police
- But we have a mission to make the .nl zone safer for users
- And we were sitting on the data
- Ethical dilemma:
 - Turn the blind eye OR
 - Do something about it
- We talked to our lawyers
- We need to conform to our mandate and EU and NL laws

We decided to go ahead and measure it

What is their *business model*?

- Counterfeit (fake) industry is **huge**: books, computers, shoes, bags
 - EU borders seizures 2016: 670 million EUR
 - US 2017: US\$ 1.2 Billion
- Luxury goods have a massive demand



If you buy a fake from the street, you know it

- but not online
- so we got involved

What is their *business model*?

- The business model goes like this:
 1. Consumer demand [4]
 2. Manufacturing in China [1]
 3. These webshops connect both of them
- It's not only a .nl problem:
 - .de, .be, .com, and many others have the same issue
- We are dealing with *pros* here

How many were on the .nl zone?

- Back to 2016: we stumbled on them
- We realized they all share a similar pattern:

1. long `html <title>` tags

```
1 <title>Vans Schoenen On Sale 70% OFF | Geen  
   verzendkosten</title>
```

2. tags listing many brands (Nike, Reebok, Gucci, you name it..)

- **Question: Why this tactic?**

- Search Engine optimization → more clicks, more money [5]

How many were on the .nl zone?

- Back to 2016: we stumbled on them
- We realized they all share a similar pattern:

1. long `html <title>` tags

```
1 <title>Vans Schoenen On Sale 70% OFF | Geen  
   verzendkosten</title>
```

2. tags listing many brands (Nike, Reebok, Gucci, you name it..)

- **Question: Why this tactic?**

- Search Engine optimization → more clicks, more money [5]

How many were on the .nl zone?

- Back to 2016: we stumbled on them
- We realized they all share a similar pattern:

1. long `html <title>` tags

```
1 <title>Vans Schoenen On Sale 70% OFF | Geen  
   verzendkosten</title>
```

2. tags listing many brands (Nike, Reebok, Gucci, you name it..)

- **Question: Why this tactic?**

- Search Engine optimization → more clicks, more money [5]

Our measurements

1. Get all .nl domain names (5.8M)
 - private data
2. Scrape their websites (if they have)
 - We used DMap [6]
3. We deployed “state-of-the art” ML to detect
 - simply count the number of brands on `<title>`

```
<title>Vans Schoenen On Sale 70% OFF | Geen verzendkosten</title>
```

- if `> 5`, then flag it
- (we precompiled a list of brands and discount words)

Our measurements

1. Get all .nl domain names (5.8M)
 - private data
2. Scrape their websites (if they have)
 - We used DMap [6]
3. We deployed “state-of-the art” ML to detect
 - simply count the number of brands on `<title>`

```
<title>Vans Schoenen On Sale 70% OFF | Geen verzendkosten</title>
```

- if `> 5`, then flag it
- (we precompiled a list of brands and discount words)

Our measurements

1. Get all .nl domain names (5.8M)
 - private data
2. Scrape their websites (if they have)
 - We used DMap [6]
3. We deployed “state-of-the art” ML to detect
 - simply count the number of brands on `<title>`

```
<title>Vans Schoenen On Sale 70% OFF | Geen verzendkosten</title>
```

- if `> 5`, then flag it
- (we precompiled a list of brands and discount words)

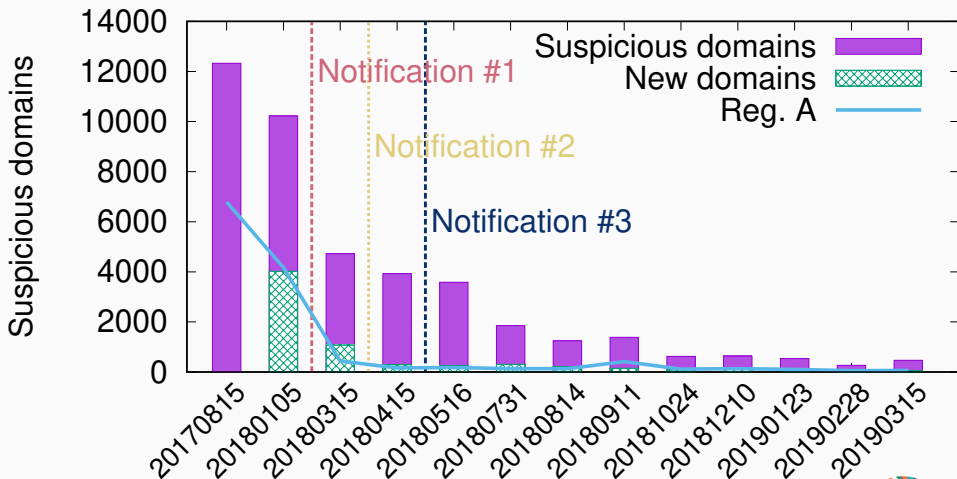
Our measurements

1. Get all .nl domain names (5.8M)
 - private data
2. Scrape their websites (if they have)
 - We used DMap [6]
3. We deployed “state-of-the art” ML to detect
 - simply count the number of brands on `<title>`

```
1 <title>Vans Schoenen On Sale 70% OFF | Geen  
   verzendkosten</title>
```

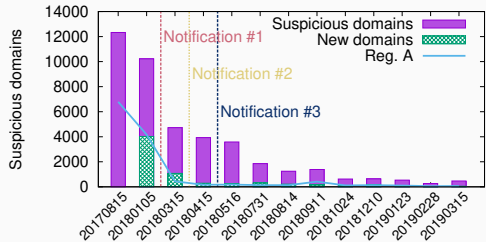
- if > 5, then flag it
- (we precompiled a list of brands and discount words)

What did we find?



How to take them down it?

- We could not take them down
- But there was a way to validate them:
 1. Notify a registrar that registered the domain
 2. Ask them to verify the ID of the registrant
 3. If it fails, then they can suspend the domain



1. How come does this even work?

- This is to show they suffered little pressure

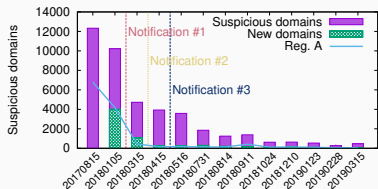
2. Why so many of these webshops?

- it's unlikely there are that many counterfeiters
- *Domains are cheap and disposable*
- automation heavily used
- 10 down does not even make a difference

3. Why 6K were registered with only one registrar?

- API for automatic registration & good price

Take downs were effective, in partnership with our registrars



- Later they changed strategy, we had a new system
- See [PAM2020 \[3\]](#)

Lessons

1. How come does this even work?

- This is to show they suffered little pressure

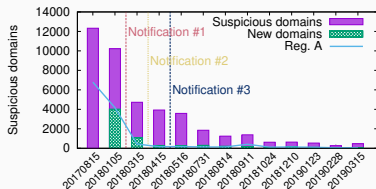
2. Why so many of these webshops?

- it's unlikely there are that many counterfeiters
- *Domains are cheap and disposable*
- automation heavily used
- 10 down does not even make a difference

3. Why 6K were registered with only one registrar?

- API for automatic registration & good price

Take downs were effective, in partnership with our registrars



- Later they changed strategy, we had a new system
- See [PAM2020 \[3\]](#)

Lessons

1. How come does this even work?

- This is to show they suffered little pressure

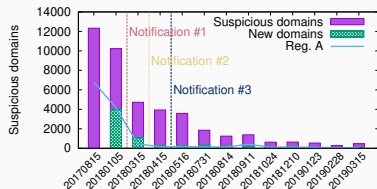
2. Why so many of these webshops?

- it's unlikely there are that many counterfeiters
- *Domains are cheap and disposable*
- automation heavily used
- 10 down does not even make a difference

3. Why 6K were registered with only one registrar?

- API for automatic registration & good price

Take downs were effective, in partnership with our registrars



- Later they changed strategy, we had a new system
- See [PAM2020 \[3\]](#)

Lessons

1. How come does this even work?

- This is to show they suffered little pressure

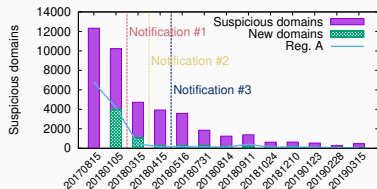
2. Why so many of these webshops?

- it's unlikely there are that many counterfeiters
- *Domains are cheap and disposable*
- automation heavily used
- 10 down does not even make a difference

3. Why 6K were registered with only one registrar?

- API for automatic registration & good price

Take downs were effective, in partnership with our registrars



- Later they changed strategy, we had a new system
- See [PAM2020 \[3\]](#)

1. How come does this even work?

- This is to show they suffered little pressure

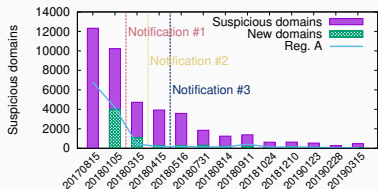
2. Why so many of these webshops?

- it's unlikely there are that many counterfeiters
- *Domains are cheap and disposable*
- automation heavily used
- 10 down does not even make a difference

3. Why 6K were registered with only one registrar?

- API for automatic registration & good price

Take downs were effective, in partnership with our registrars

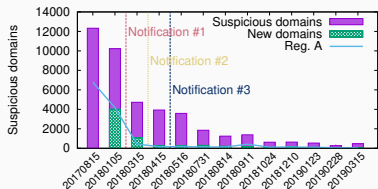


- Later they changed strategy, we had a new system
- See [PAM2020 \[3\]](#)

Lessons

1. How come does this even work?
 - This is to show they suffered little pressure
2. Why so many of these webshops?
 - it's unlikely there are that many counterfeiters
 - *Domains are cheap and disposable*
 - automation heavily used
 - 10 down does not even make a difference
3. Why 6K were registered with only one registrar?
 - API for automatic registration & good price

Take downs were effective, in partnership with our registrars



- Later they changed strategy, we had a new system
- See [PAM2020 \[3\]](#)

Lessons

1. How come does this even work?

- This is to show they suffered little pressure

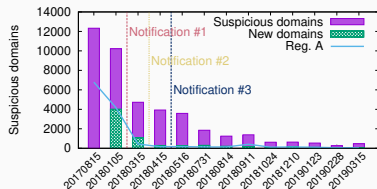
2. Why so many of these webshops?

- it's unlikely there are that many counterfeiters
- *Domains are cheap and disposable*
- automation heavily used
- 10 down does not even make a difference

3. Why 6K were registered with only one registrar?

- API for automatic registration & good price

Take downs were effective, in partnership with our registrars



- Later they changed strategy, we had a new system
- See **PAM2020** [3]

Governance and Consolidation: end of privacy

Using data for the public good

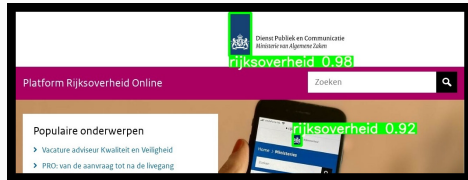
Counterfeited webshops

Logo Misuse

Wrapping up

From text to logo detection: LogoMotive

- My colleagues did a study evaluating misuse of Dutch government logo
- It became a **brand protection** service
- See **PAM2022** [2] paper



Detecting logos misuse with ML

The screenshot shows a website footer with various logos. Green boxes highlight detected logos with their confidence scores: 'rijksoverheid' (0.9), 'rijksoverheid' (0.98), 'sidn' (0.97), and 'thuiswinkel' (0.95). A red arrow points from the text 'Detected logos with confidence' to these highlighted logos.

Pagina's

- Home
- Problemen
- Vragen
- Nieuws

Volg ons

- Facebook
- Twitter
- Instagram
- YouTube
- Vimeo

Privacyverklaring | Cookieverklaring | Responsible disclosure | Disclaimer | Digitoegankelijkheid

Een initiatief van:

- rijksoverheid 0.9
- rijksoverheid 0.98
- Ministerie van Economische Zaken en Klimaat
- Nationaal Cyber Security Centrum
- BCP
- Platform voor de Innovatie Samenleving

Mede mogelijk gemaakt door:

- kpn
- vodafone
- gogo
- sidn 0.97
- 5BN
- Google
- Microsoft
- POLITIE
- thuiswinkel 0.95
- NLdigital
- thuiswinkel.org
- ACM
- ConsuWijzer
- Co-financed by the European Union
- Connecting Europe Facility
- veilig internetten.nl

Detected logos with confidence

The screenshot shows a login page with a detected logo 'rijksoverheid' (0.98) highlighted in green. A red arrow points from the text 'Detected logos with confidence' to this logo.

EN | NL

Inloggen bij DigiD Online

Hoe wilt u inloggen?

- Met de DigiD app
De makkelijkste manier om veilig in te loggen
- Met een sms-controle
- Met mijn identiteitskaart

Annuleren

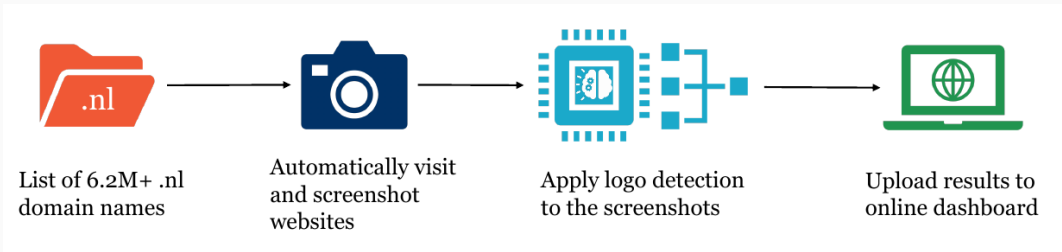
Kunt u niet verder? Download dan de DigiD app (opent in een nieuw venster) of activeer de sms-controle (opent in een nieuw venster)

Nog geen DigiD? Vraag uw DigiD aan

Vraag en antwoord

- Ik ben mijn gebruikersnaam vergeten

How does LogoMotive work?



Generating training datasets

- We've used **Yolo** for image recognition
- It requires labeled data
- So we've generated it

	Value
Screenshots generated	64,893
Synthetic training samples	100,000
training set	95,000
validation set	5,000

Table 1: Datasets used for training and validation.

Generating training datasets

- We've used **Yolo** for image recognition
- It requires labeled data
- So we've generated it

	Value
Screenshots generated	64,893
Synthetic training samples	100,000
training set	95,000
validation set	5,000

Table 1: Datasets used for training and validation.

Generating training datasets



Jagthulp in Groenvenen. Opnieuw en spreiden 'blau' in de haart. Netwerken van groenvenen willen de hulp aan jagd en groenvenen lokal organiseren.

Actuele berichten (home)
Lokale Netwerken -
Groenvenen -
Publicaties & Blogs

Over Groenvenen
Werken en leven in een groenvenen
Naar een nieuwe jagthulp

Minder actief?
Ik ben nu veel minder actief met 'groenvenen'. De site hou ik nog wel in de lucht. Uiteraard wil ik...

Bijeenkomsten om landelijke en regionale pleegzorgontwikkelingen met elkaar te verbinden
In mei en juni 2019 organiseert de NVP vier bijeenkomsten voor pleegouders, verspreid over Nederland. Op deze bijeenkomsten horen we...

Bijeenkomst voor pleeg- en gezinshuis-ouders Zeist, De Bilt, Bunnik, Utrechtse Heuvelrug en Wijk bij Duurstede
Op 17 april organiseert de regio Zuid-Oost Utrecht een netwerkbijsamenkomst voor pleeg- en gezinshuisouders uit de gemeenten Zeist, De Bilt...

Minister wil intensivering Actieplan Pleegzorg
Om de dagelijkse praktijk van pleeggezinnen te verbeteren, wil minister Hugo de Jonge een intensivering van het Actieplan Pleegzorg. Dat...

Versterk pleeggezinnen
In de uitzending van De Monitor van zondag 5 februari was te

Random screenshot



Jagthulp in Groenvenen. Opnieuw en spreiden 'blau' in de haart. Netwerken van groenvenen willen de hulp aan jagd en groenvenen lokal organiseren.

Actuele berichten (home)
Lokale Netwerken -
Groenvenen -
Publicaties & Blogs

Over Groenvenen
Werken en leven in een groenvenen
Naar een nieuwe jagthulp

Minder actief?
Ik ben nu veel minder actief met 'groenvenen'. De site hou ik nog wel in de lucht. Uiteraard wil ik...

Bijeenkomsten om landelijke en regionale pleegzorgontwikkelingen met elkaar te verbinden
In mei en juni 2019 organiseert de NVP vier bijeenkomsten voor pleegouders, verspreid over Nederland. Op deze bijeenkomsten horen we...

Bijeenkomst voor pleeg- en gezinshuis-ouders Zeist, De Bilt, Bunnik, Utrechtse Heuvelrug en Wijk bij Duurstede
Op 17 april organiseert de regio Zuid-Oost Utrecht een netwerkbijsamenkomst voor pleeg- en gezinshuisouders uit de gemeenten Zeist, De Bilt...

Minister wil intensivering Actieplan Pleegzorg
Om de dagelijkse praktijk van pleeggezinnen te verbeteren, wil minister Hugo de Jonge een intensivering van het Actieplan Pleegzorg. Dat...

Versterk pleeggezinnen
In de uitzending van De Monitor van zondag 5 februari was te

Resulting datapoint

Evaluating the model

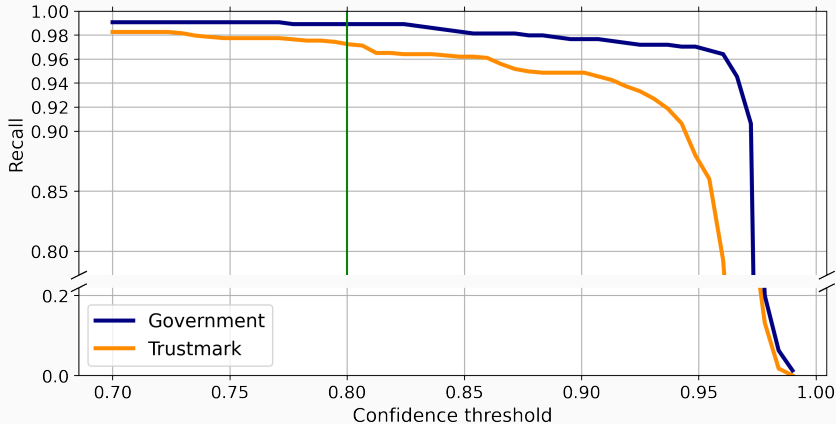


Figure 3: Recall performance of LogoMotive at confidence thresholds. The vertical line denotes our chosen threshold.

Results

Label	Full-Zone	Newly-Registered
Total	12862 (100.00%)	53
Without gov. logo (FP)	1164 (9.05%)	0 (0.00%)
With gov. logo (TP)	11698 (90.95%)	53 (100.0%)
Benign	10595 (82.37%)	32 (60.38%)
Government impersonation	151 (1.17%)	17 (32.09%)
Phishing	3 (0.02%)	3 (5.66%)
Potential threat	73 (0.57%)	9 (16.98%)
Other (false endorsements, satire, etc.)	75 (0.58%)	5 (9.43%)
Government domains	952 (7.40%)	4 (7.55%)
In portfolio	636 (4.94%)	2 (0.00%)
Not in portfolio	316 (2.46%)	2 (3.77%)
Added	109 (0.85%)	1 (1.89%)
Pending	207 (1.61%)	1 (1.89%)

Table 2: Manual validation results for government impersonation case study.

- See [PAM2022 \[2\]](#) paper for more details
- There was a second case study
- It became a brand protection service

Governance and Consolidation: end of privacy

Using data for the public good

Counterfeited webshops

Logo Misuse

Wrapping up

Conclusions

- We are living in the dark ages of privacy
- Internet consolidation is increasing
 - less competitiveness
 - the US and EU are moving slowly to fix it
- Data can still be used for the public good
 - business models that DO NOT rely on surveillance
- We showed two case studies
- Hope you folks help to address these issues

Slides: <https://tinyurl.com/sidn23>

Contact: giovane.moura@sidn.nl <https://sidnlabs.nl/en>



US Spy Satellite logo

[1] SCHMIDLE, N.

Inside the Knockoff-Tennis-Shoe Factory - The New York Times.

<http://www.nytimes.com/2010/08/22/magazine/22fake-t.html>, 2010.

[2] VAN DEN HOUT, T., WABEKE, T., MOURA, G. C. M., AND HESSELMAN, C.

Logomotive: detecting logos on websites to identify online scams - a tld case study.

In Passive and Active Measurement (2022).

[3] WABEKE, T., MOURA, G. C. M., FRANKEN, N., AND HESSELMAN, C.

Counterfighting Counterfeit: detecting and taking down fraudulent webshops at a ccTLD.

In Proceedings of the Passive and Active Measurement Workshop
(Eugene, OR, USA, 2020).

[4] WALL, D. S., AND LARGE, J.

Jailhouse frocks: Locating the public interest in policing counterfeit luxury fashion goods.

The British Journal of Criminology 50, 6 (2010), 1094–1116 –
<http://ssrn.com/abstract=1649773>.

- [5] WANG, D. Y., DER, M., KARAMI, M., SAUL, L., MCCOY, D., SAVAGE, S.,
AND VOELKER, G. M.

**Search + seizure: The effectiveness of interventions on seo
campaigns.**

In Proceedings of the 2014 Conference on Internet Measurement
Conference (New York, NY, USA, 2014), IMC '14, ACM, pp. 359–372.

[6] WULLINK, M., MOURA, G. C., AND HESSELMAN, C.

Dmap: Automating domain name ecosystem measurements and applications.

In Proceedings of the IEEE Network Traffic Monitoring and Analysis Conference (Vienna, Austria, June 2018), IEEE, pp. 1–8.