

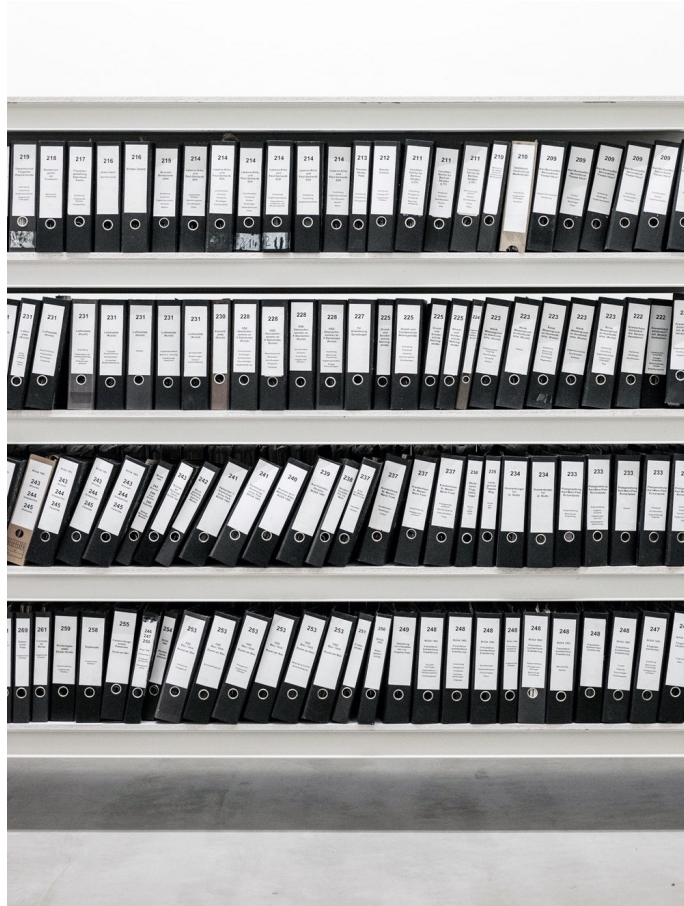
# Using machine learning to boost internet and DNS security

Thymen Wabeke | Ofcom

8 December 2021



# SIDN: operator of the .nl ccTLD



Registration of domain names  
6.2M .nl-domains

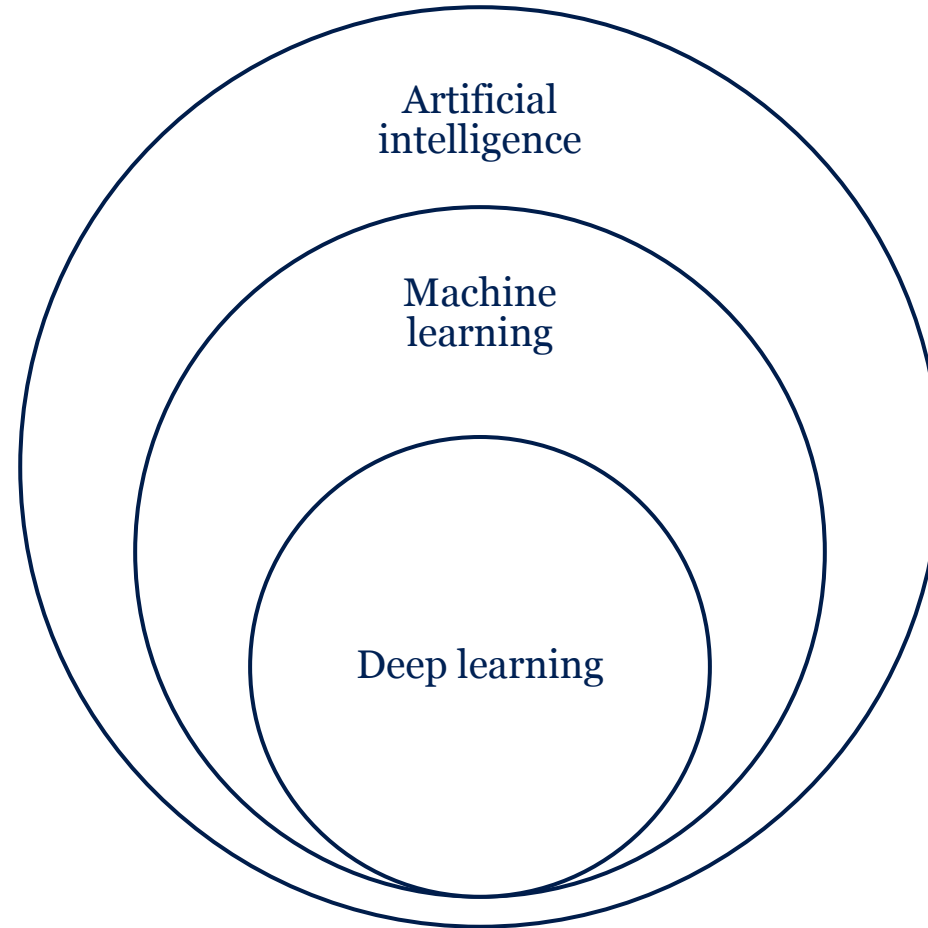


Publish domain names via DNS  
2.5B DNS queries/day

# SIDN Labs: research team

- Goal: increase the trustworthiness (security, stability, resilience, and transparency) of our society's internet infrastructure, for .nl and the Netherlands in particular
- Strategies:
  - Applied research (measurements, design, prototyping, evaluation)
  - Make results publicly available and useful for various target groups
  - Work with universities, infrastructure operators, and other labs
- Three research areas: network security, **domain name security**, trusted future internet infrastructures

# Machine learning in perspective





# Research agenda

- Apply ML to increase security of the Internet and DNS
- Approach: explore and integrate promising algorithms, papers and tools
  - Innovating *with* ML, not innovation *of* ML
- Target group: DNS actors (registries, registrars and DNS operators)

# Research topics



RQ1: How can we get even better at proactive abuse detection?



RQ2: How can we train shared abuse models without exchanging data?



RQ3: How can we use ML to improve our anycast infrastructure monitoring and management?

# Applying ML in a responsible way

- Human-in-the-loop
- Simple and interpretable models
- Collaborate and publish
- Monitor performance

**Radboud University**

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

Thymen Wabeke<sup>1</sup>, Giovane C. M. Moura<sup>1,3</sup>, Nanneke Franken<sup>2</sup>, and Cristian Hesselman<sup>1,4</sup>

<sup>1</sup> SIDN Labs, Arnhem, The Netherlands  
<sup>2</sup> SIDN, Arnhem, The Netherlands

**REALTIME REGISTER**

**Centr**

**DEAL**

**UNIVERSITY OF TWENTE**

**SIDN LABS**

Government of the Netherlands

Using logo detection technology to identify malicious .nl websites

LogoMotive helps the fight against internet crime by flagging up unauthorised logo use

Thursday 10 June 2021  
Article by: Thijs van den Hout, Thymen Wabeke, Cristian Hesselman

The original blog is in Dutch. This is the English translation.

**ICCS** VISA MasterCard  
INTERNATIONAL CARD SERVICES

**SIDN LABS**

# Remainder of presentation



**HOLLISTER** Dames Heren Inloggen Register (0) Omschrijving

<p>★★★★★</p> <p>Hollister Undergoed Heren Lage Taille Multipack Grijs Groen Camo Zwart 11859-FNX</p> <p>15 Kleur <b>BROEK &amp; KORTE BROEK</b></p> <p><del>€30.60</del> <b>€22.31</b></p>	<p>★★★★★</p> <p>Hollister T Shirt Heren Logo Graphic Korte Mouwen Donkerblauw 21170-HLT</p> <p>15 Kleur <b>TOPS</b></p> <p><del>€30.70</del> <b>€22.38</b></p>	<p>★★★★★</p> <p>Hollister Jas Dames All-weather Stretch Sherpa-lined Zwart 45801-GXL</p> <p>1 Kleur <b>JASSEN</b></p> <p><del>€98.35</del> <b>€69.73</b></p>	<p>★★★★★</p> <p>Hollister Joggingbroek Heren Advanced Stretch Cargo Twill Olijfgroen 97393-SXN</p> <p>4 Kleur <b>BROEK &amp; KORTE BROEK</b></p> <p><del>€50.11</del> <b>€35.98</b></p>	<p>★★★★★</p> <p>Hollister Blouses Dames Fluweel Off-the-shoulder Goud 49289-JQI</p> <p>2 Kleur <b>TOPS</b></p> <p><del>€30.60</del> <b>€22.31</b></p>
<p>★★★★★</p>	<p>★★★★★</p>	<p>★★★★★</p>	<p>★★★★★</p>	<p>★★★★★</p>



# SIDN's interest

- Consumer losses
- Trust in Internet may decrease

## Perfect vantage point:

- List of *all* .nl - domains
- Passive and active measurements

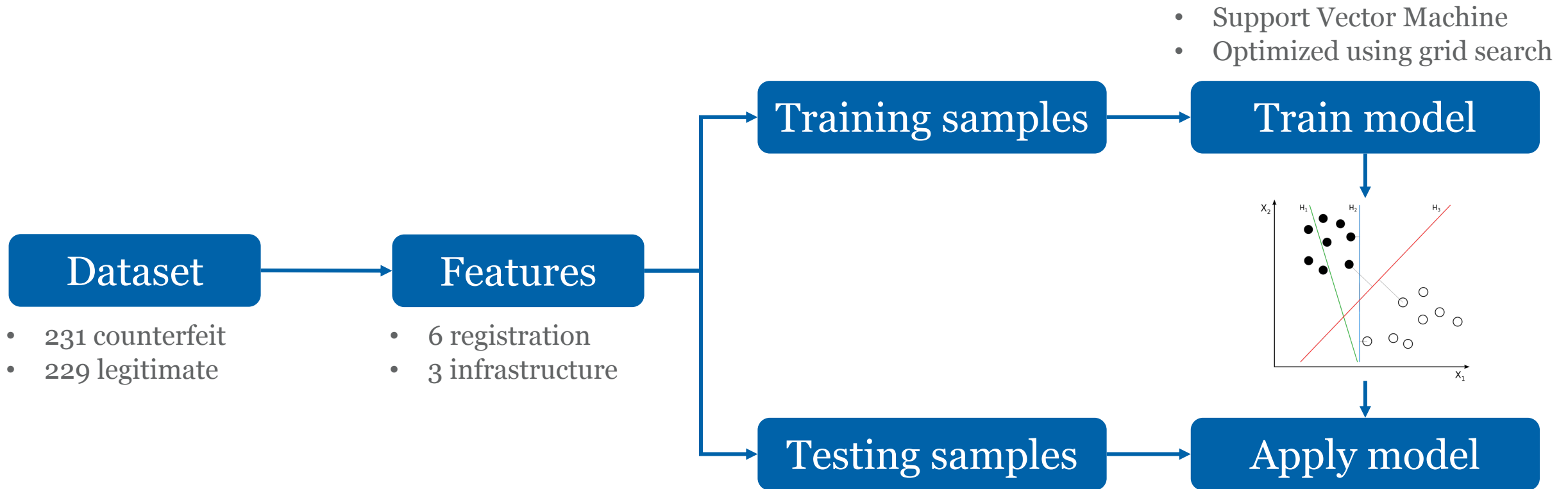


# Main results

- Detected thousands since 2016
- Protected users from being scammed
- PAM2020 paper:
  - BrandCounter (2018 Q1-2)
  - FaDe (2019 Q1)



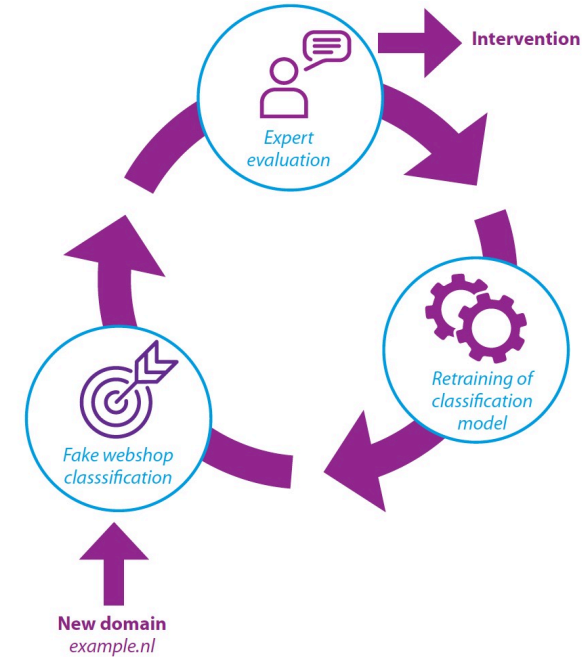




<b>Samples</b>	<b>Precision</b>	<b>Recall</b>
Train (cross-validation)	0.98	0.97
Test	1.0	1.0

# Lessons learned

- Registrar and ICS collaboration was key
- Detectors are simple yet effective
  - Registries have perfect vantage point
  - Suggests little pressure
- It's an ever-going whack-a-mole game
  - Monitor features and evaluate model regularly
  - Fewer takedowns = fewer scams?



Year	Taken down
2018	~12,000
2019	4,340
2020	481

*Number of counterfeit webshops taken down*

# LogoMotive: finding malicious .nl-domains with logo detection

**Pagina's**

- Home
- Problemen
- Vragen
- Nieuws
- Video's
- Quizen
- Over ons

**Volg ons**

- Facebook
- Twitter
- Instagram
- YouTube
- Vimeo

Privacyverklaring   Cookieverklaring   Responsible disclosure   Disclaimer   Ditoegankelijkheid

Een initiatief van:

- rijksoverheid 0.9** / **rijksoverheid 0.98**  
Ministerie van Economische Zaken en Klimaat
- Nationaal Cyber Security Centrum  
Ministerie van Justitie en Veiligheid
- ECP  
Platform voor de InformatieSamenleving

Mede mogelijk gemaakt door:

- kpn   **vodafone**   Ziggo   Betaalvereniging Nederland
- sidn 0.97**   SIDN   T-Mobile   Google
- Microsoft   **POLITIE**   thuiswinkel **0.95**   thuiswinke.org   SENORWEB   medienet.nl   SIC
- NLdigital   FRAUDEHELPDESK.nl
- ACM   ConsuWijzer   Co-financed by the European Union  
Connecting Europe Facility
- veilig internetten.nl

EN | **NL**

**rijksoverheid 0.98**

Inloggen bij GGD 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

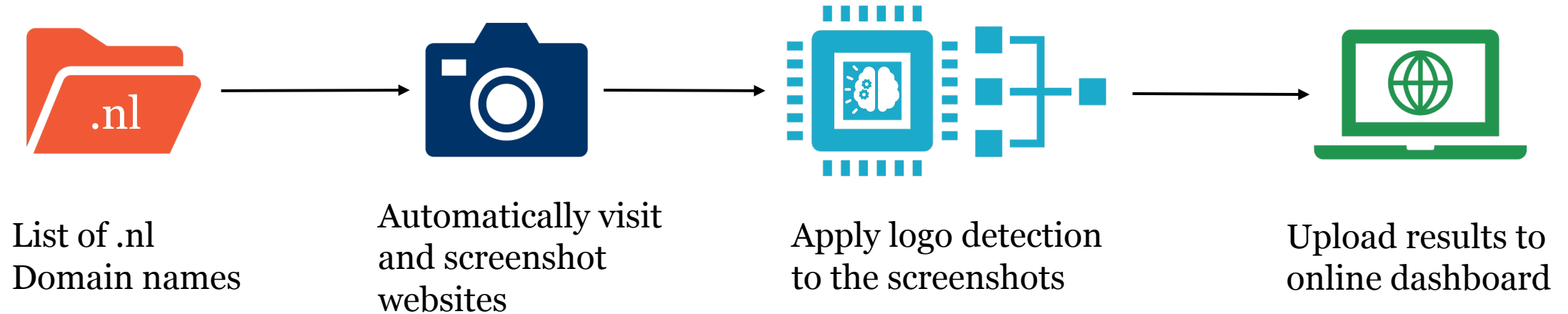
Kunt u niet verder? Download dan de DigiD app [opent in een nieuw venster] of activeer de sms-controle [opent in een nieuw venster]

Nog geen DigiD? Vraag uw DigiD aan

Vraag en antwoord

- Ik ben mijn gebruikersnaam vergeten

# How does LogoMotive work?



# Online dashboard

Logomotive - SIDN  
logomotive.sidnlabs.nl/sidn

SIDN LABS LOGOMOTIVE SIDN

Filter

Label: All  
Status: All  
Filter

Bulk update

Label: Select  
Status: Select  
Update

Logo's found

Show 10 entries  Select All Search:

Domain name	Screenshot date	Registrar	Registrant	Registered on	Label	Status	
laatdelinksliggen.nl	2021-10-04 04:07				-	Open	Annotate
hostinghero.nl	2021-09-27 09:43				-	Open	Annotate
webrestyle.nl	2021-09-27 09:43				-	Open	Annotate
studioactive.nl	2021-09-27 09:43				-	Open	Annotate
hostingu2.nl	2021-09-27 09:35				-	Open	Annotate
easyrek.nl	2021-09-27 09:28				-	Open	Annotate
houten-legbordstelling.nl	2021-09-27 09:28				-	Open	Annotate
coronakantoorinrichting.nl	2021-09-27 08:57				-	Open	Annotate
dutchantiddoscoalition.nl	2021-09-27 08:56				-	Open	Annotate
nomoredos.nl	2021-09-27 08:56				-	Open	Annotate

Showing 1 to 10 of 722 entries

Previous 1 2 3 4 5 ... 73 Next

Logo found on laatdelinksliggen.nl

Screenshot date: 04-10-2021 04:07

Page also found on: veiliginternetten.nl, digivaardigdigiveilig.nl, checkjeupdates.nl, maakhetniettemakkelijk.nl, doejouupdates.nl, digibewust.nl, cloudbewust.nl, doejeupdates.nl, digivaardigdigibewust.nl, jewachtwoord.nl, beschermjebedrijf.nl, internettenveilig.nl

Registrant: Stichting ECP

Registrar: team.blue nl B.V.

Registration date: 24-09-2021 00:00

Screenshots

Pagina's: Home, Problemen, Vragen, Nieuws, Video's, Quizen, Over ons

Volg ons: Facebook, Twitter, Instagram, YouTube, Vimeo

Een initiatief van: Ministerie van Economische Zaken en Klimaat, Nationaal Cyber Security Centrum, ECP

Mede mogelijk gemaakt door: kpn, vodafone, Ziggo, sidn 0.97, Google, Microsoft, POLITIE, NLdigital, FRAUDEBUDESK.nl, thuiswinkel.org, AGM ConsuWijzer, Co-financed by the European Union

Veilig internetten.nl

Comment: Comment...

Clear label, Previous

Label:  Correct use,  Incorrect use,  Geen logo

Status:  Open,  In behandeling,  Afgehandeld

Save and update all related domains, Save and next, Save and exit



# Can logo detection contribute to a safe .nl-zone?

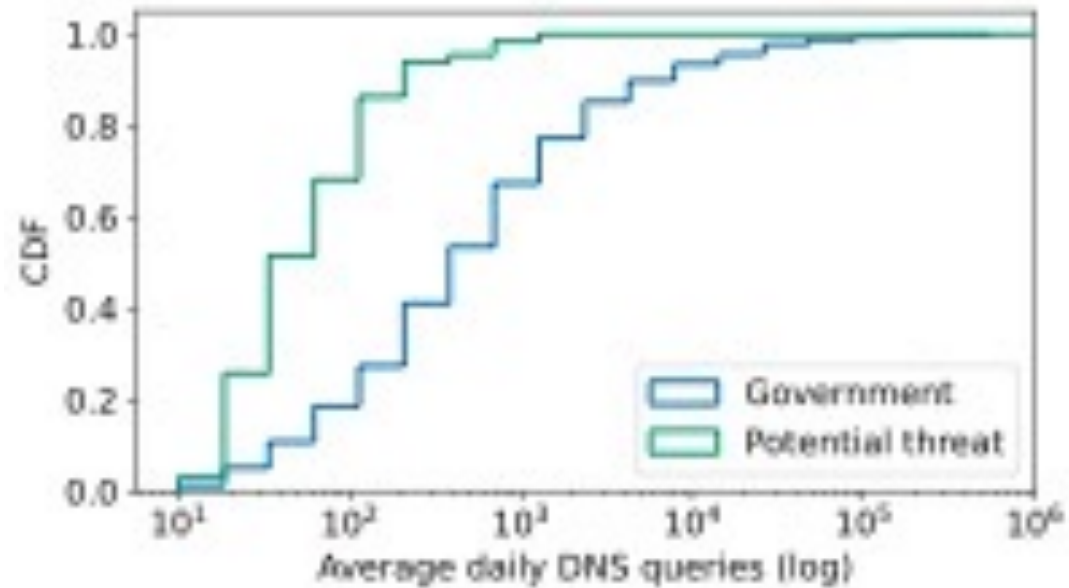
- Case study with Dutch national government
  - Goal: find government impersonation attacks
  - Apply to full zone (6.2M domains) and to new domains (2 month)
- Case study with Dutch e-commerce trustmark (Thuiswinkel.org)
  - Goal: find webshops that abuse the trustmark
  - Apply to full zone (6.2M domains)

# Manual validation results for government study

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%)



# DNS queries seen at .nl authoritative name servers



# Adoption rate of security standards

	Government Domains	
	In portfolio	Not in portfolio
Total		
with DNSSEC	623 (98%)	230 (74%)
without DNSSEC	13 (2%)	79 (26%)
with DMARC	584 (92%)	126 (41%)
without DMARC	52 (8%)	183 (59%)

# Lessons learned and future work

- Visual aspects like logo's help us to detect abuse
- Logo's also help to keep domain portfolio accurate
- Large gray area of unwanted, but not abusive content

## Next steps:

- Publish academic paper (under review)
- Share code with peers and university
- Integrate with '*SIDN BrandGuard*'

*Volg ons*

 SIDN.nl

 @SIDN

 SIDN

Q&A

[www.sidnlabs.nl](http://www.sidnlabs.nl) | [stats.sidnlabs.nl](http://stats.sidnlabs.nl)

Thymen Wabeke  
Research engineer  
[thymen.wabeke@sidn.nl](mailto:thymen.wabeke@sidn.nl)

