

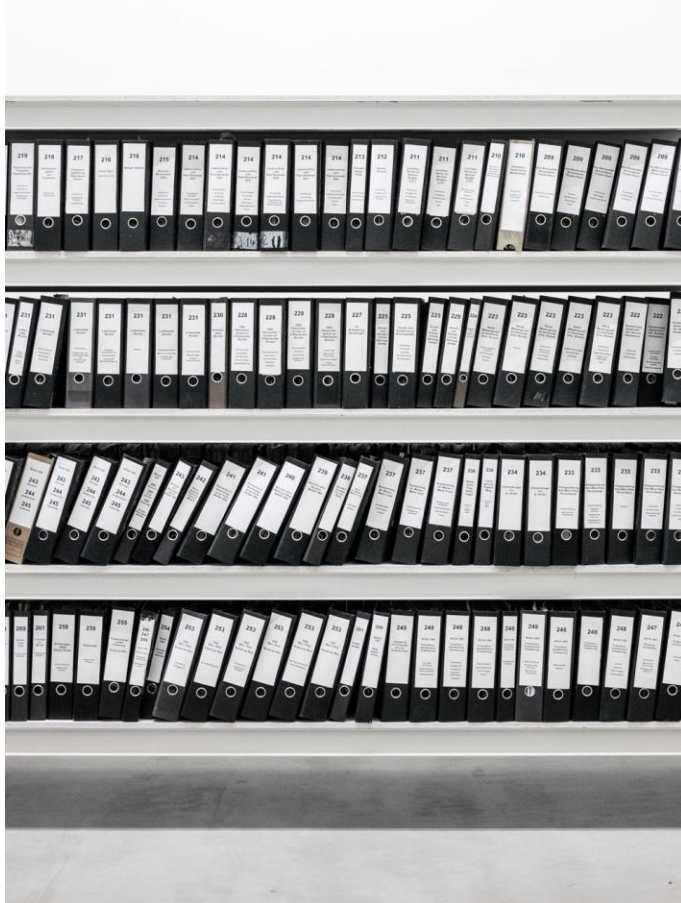
Using machine learning to boost internet and DNS security

Thymen Wabeke

September 23, 2021



SIDN: operator of the .nl TLD



Registration of domain names
6.2M .nl-domains



Lookup IP address of a domain name
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 & IoT security, trusted future internet infrastructures

Create ML applications that boost internet security

- Goal: apply ML to increase security of the Internet and DNS
 - Many large datasets are available
 - Manually extracting patterns is difficult
- 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)

Applying ML in a responsible way

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

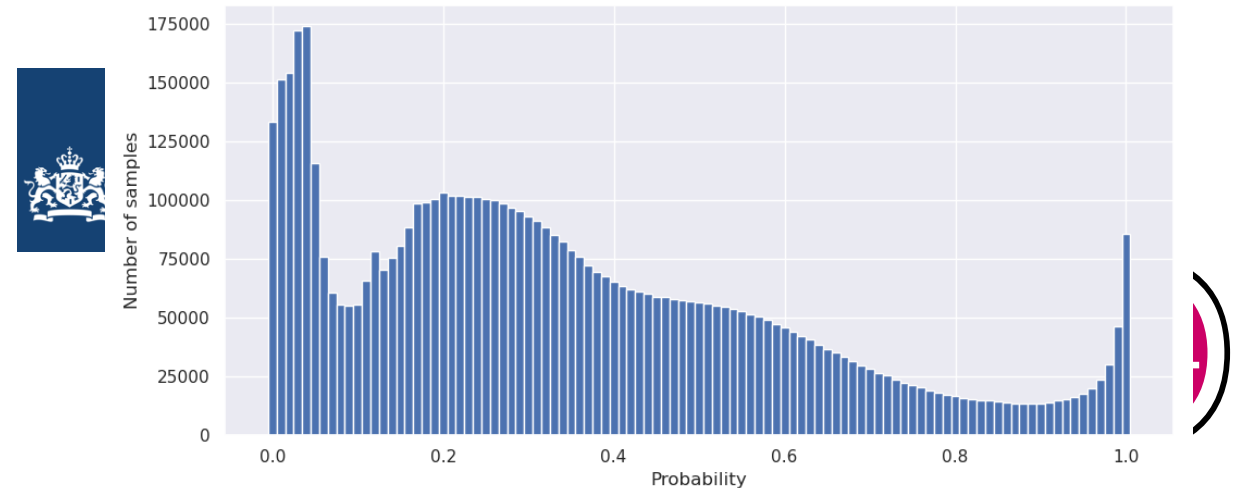
Radboud University



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

REALTIME REGISTER

Response Distribution Chart



UNIVERSITY OF TWENTE



for a large array of goods, such as pharmaceu-



are among the s from the con-

Thursday 10 June 2021

Article by: Thijs van den Heuvel, Thymen Wabeke, Cristian Hesselman

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



Remainder of presentation



nederlandwebshop.nl

The screenshot displays the Hollister website interface. At the top, there is a navigation bar with the Hollister logo, 'Dames' and 'Heren' category links, and utility links for 'Inloggen', 'Register', and 'Winkelwagen'. A search bar is located on the right side of the navigation bar.

The main content area features a grid of ten clothing items, each with a product image, a star rating, a title, a description, and a price. The items are:

- Item 1:** Hollister Ondergoed Heren Lage Taille Multipack Grijs Groen Camo Zwart 11859-FNK. 15 Kleur. BROEK & KORTE BROEK. Price: €30.60 - €22.31. Rating: 4 stars.
- Item 2:** Hollister T Shirt Heren Logo Graphic Korte Mouwen Donkerblauw 21170-HLT. 15 Kleur. TOPS. Price: €30.70 - €22.38. Rating: 5 stars.
- Item 3:** Hollister Jas Dames All-weather Stretch Sherpa-lined Zwart 45801-GXL. 1 Kleur. JASSEN. Price: €98.35 - €69.73. Rating: 5 stars.
- Item 4:** Hollister Joggingbroek Heren Advanced Stretch Cargo Twill Olijfgroen 97393-SXN. 4 Kleur. BROEK & KORTE BROEK. Price: €50.11 - €35.98. Rating: 5 stars.
- Item 5:** Hollister Blouses Dames Fluessel Off-the-shoulder Goud 49289-JQI. 2 Kleur. TOPS. Price: €30.60 - €22.31. Rating: 4 stars.
- Item 6:** (Image of dark jeans). Rating: 4 stars.
- Item 7:** (Image of a black sneaker). Rating: 4 stars.
- Item 8:** (Image of a red polo shirt). Rating: 4 stars.
- Item 9:** (Image of blue jeans). Rating: 4 stars.
- Item 10:** (Image of a red and black beanie). Rating: 4 stars.

SIDN's interest

- Consumer losses
- Trust in Internet may decrease

Perfect vantage point:

- List of *all* .nl-domains, including:
 - DRS: Registration data
 - DMAP: DNS, TLS, SMTP and HTTP crawls
 - ENTRADA: traffic observed by authoritative name servers



Main results

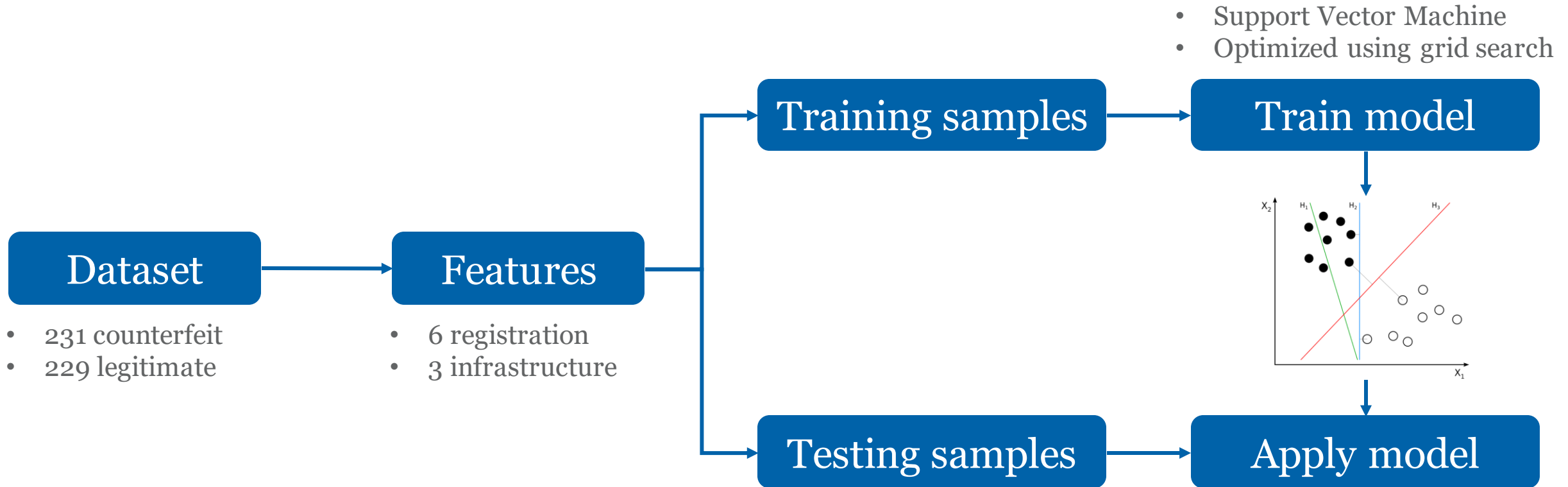
- Detected thousands since 2016
- Protected users from being scammed
- PAM2020 paper: 2 detection systems, 2 case studies
 - BrandCounter (2018 Q1-2)
 - **FaDe** (2019 Q1)



Fake Detector (FaDe)

- Collaboration with ICS Cards, a credit card issuer
- ICS Cards provided 231 shops involved in scams
- Classification model based on supervised machine learning

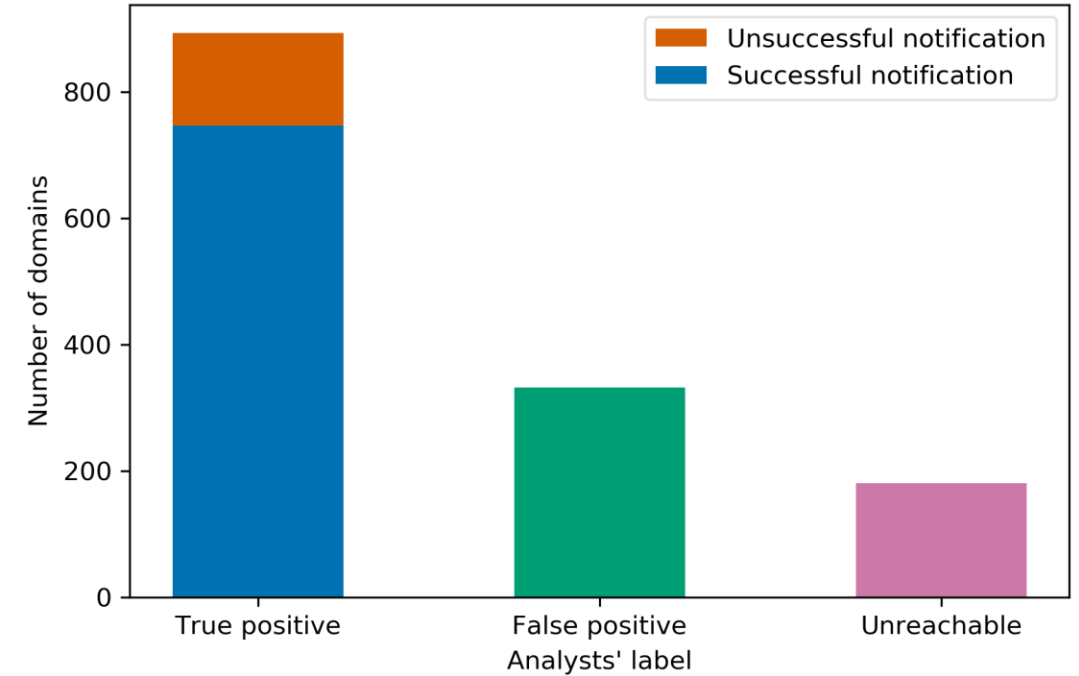




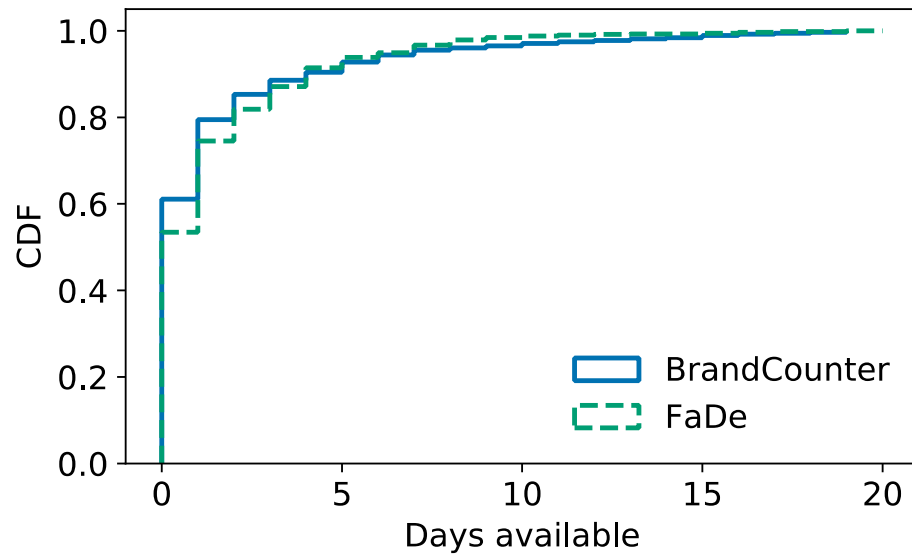
Samples	Precision	Recall
Train (cross-validation)	0.98	0.97
Test	1.0	1.0

FaDe notification

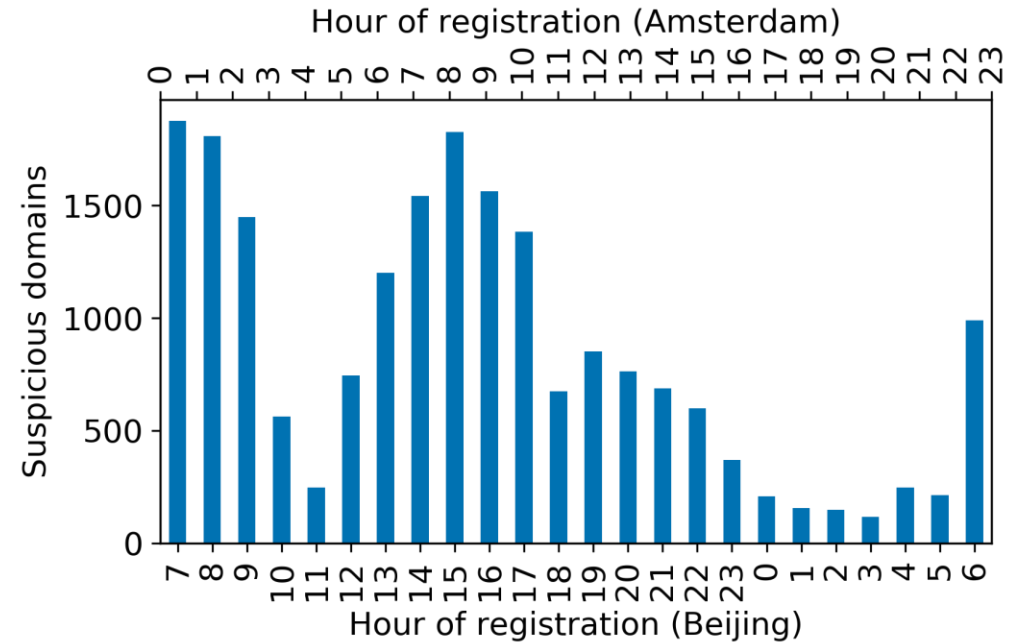
- Applied model to 30k .nl-domains
 - 1407 classified as suspicious
 - 894 true positives (73%)
- Sent 894 notifications to registrars
 - 747 taken down (84%)



How do counterfeiters operate?



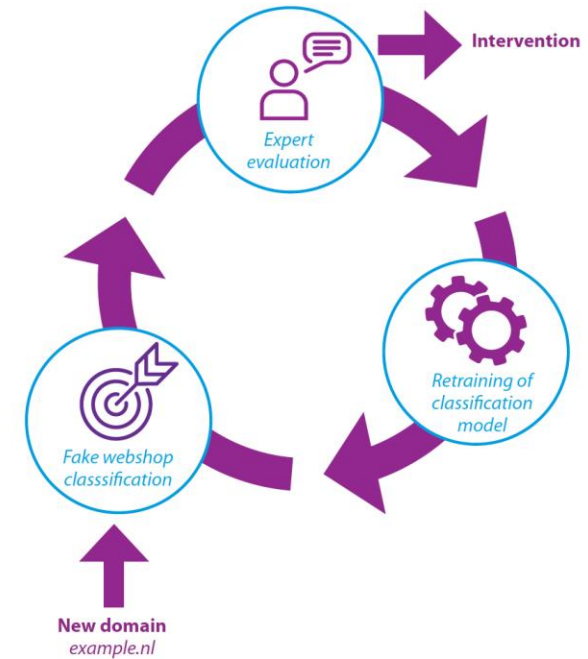
Days between domain expiration and re-registration



Time of domain registration

Lessons learned

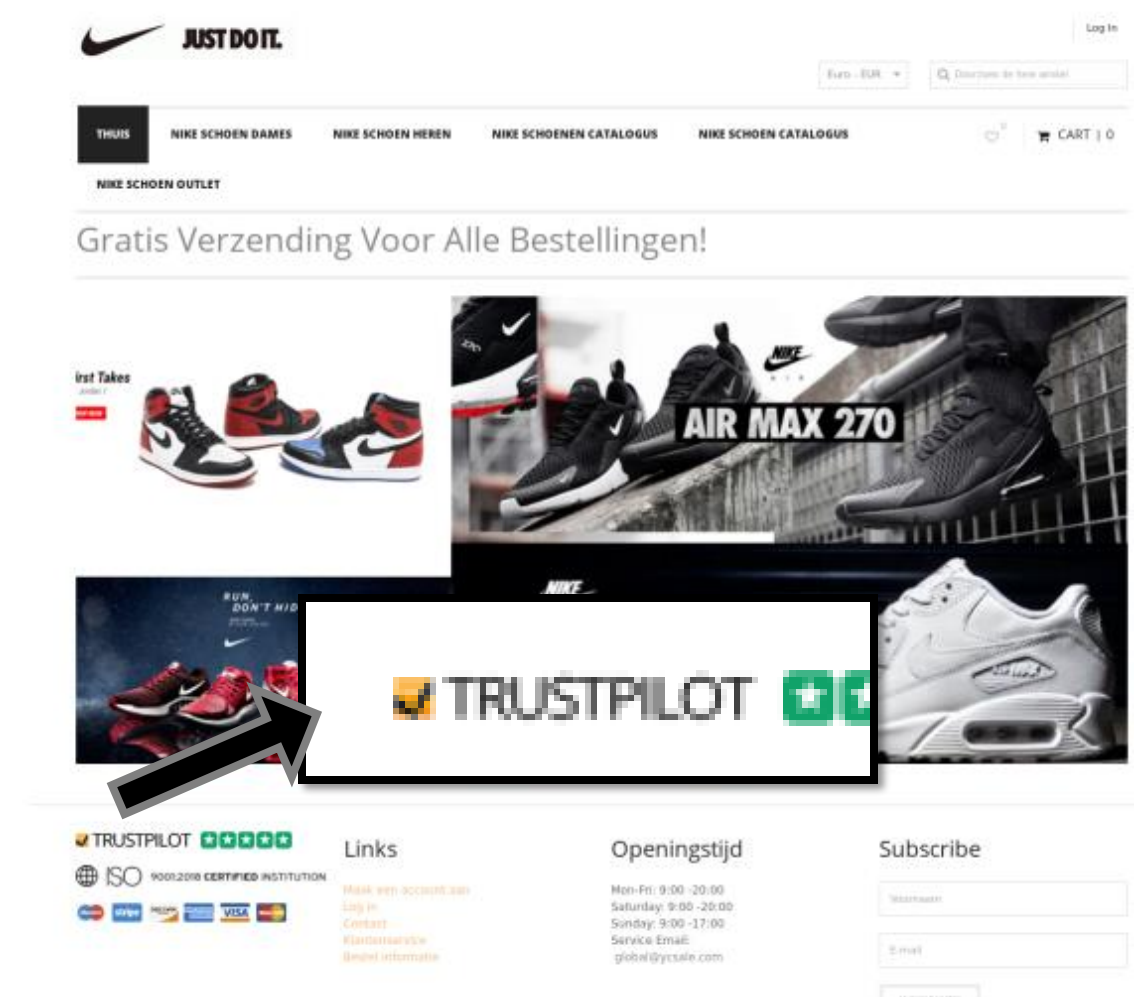
- Registrars and ICS collaboration was key
- Detectors are simple yet effective
 - Registries have perfect vantage point
 - Suggests little pressure
- It's an ever-going whack-a-mole game
 - Monitor features and evaluate model regularly
 - Fewer takedowns = fewer scams?



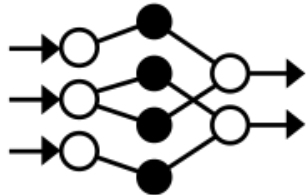
Year	Taken down
2018	~12,000
2019	4,340
2020	481

Number of counterfeit webshops taken down

Malicious websites use well-known organizations' logos



Help analyst find suspect websites using logo detection



1. Crawl and take screenshots

- Should be efficient: skip duplicates
- Using Selenium

2. Object detection algorithm

- Should be flexible: no manual labeling
- Based on YOLOv5

3. Annotation dashboard

- Should be easy-to-use and generic

Automatic training data generation



Jeugdhulp in Gezinsvormen. Opgroeien en opvoeden 'thuis' in de buurt. Netwerken van gezinsvormers zullen de hulp aan jeugd en gezinnen lokaal organiseren.

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

Beleid & Onderzoek
Agenda -
Onderzoek Gezinsvormen -
Naar een nieuwe jeugdhulp

Over Gezinsvormen
Werken en leven in een gezinsvorm -

Minder actief?
Ik ben nu veel minder actief met 'gezinsvormen'. 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 netwerkbijeenkomst 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 3 februari was te

Random screenshot

Jeugdhulp in Gezinsvormen. Opgroeien en opvoeden 'thuis' in de buurt. Netwerken van gezinsvormers zullen de hulp aan jeugd en gezinnen lokaal organiseren.

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

Beleid & Onderzoek
Agenda -
Onderzoek Gezinsvormen -
Naar een nieuwe jeugdhulp

Over Gezinsvormen
Werken en leven in een gezinsvorm -

Minder actief?
Ik ben nu veel minder actief met 'gezinsvormen'. 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 netwerkbijeenkomst 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 3 februari was te

Resulting datapoint

Filter

Label

All



Status

All



Filter

Bulk update

Label

Select



Status

Select



Update

Logo's found

Show 10 entries

 Select All

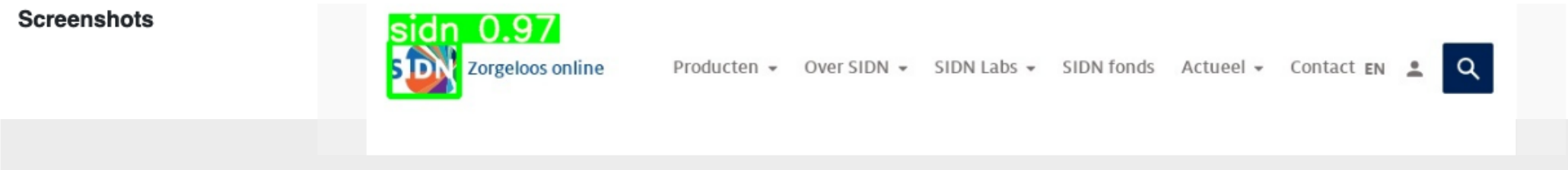
Search:

Domain name	↑↓ Screenshot date	↑↓ Registrar	↑↓ Registrant	↑↓ Registered on	↑↓ Label	↑↓ Status	↑↓
sidn.nl	2021-08-12 11:11	Stichting Interne...	Stichting Interne...	1999-11-18	-	Open	Annotate
domainregistry.nl	2021-08-12 11:11	Stichting Interne...	Stichting Interne...	2001-03-27	-	Open	Annotate
dnsops.nl	2021-08-12 11:11	Stichting Interne...	Stichting Interne...	2007-09-26	-	Open	Annotate
dnssec.nl	2021-08-12 01:11	Stichting Interne...	Stichting Interne...	2001-03-26	-	Open	Annotate
6miljoen.nl	2021-08-11 22:08	Stichting Interne...	Stichting Interne...	2020-06-12	-	Open	Annotate
sidnlabs.nl	2021-08-11 10:18	Stichting Interne...	Stichting Interne...	2010-05-10	-	Open	Annotate
abuse204.nl	2021-08-11 01:39	Stichting Interne...	Stichting Interne...	2014-08-11	Correct use	Afgehandeld	Annotate
otic.nl	2021-07-16 09:47	Stichting Interne...	Stichting Interne...	2010-04-01	Correct use	Afgehandeld	Annotate

Logo found on sidn.nl



Screenshot date	12-08-2021 11:11
Page also found on	dnsops.nl , schijtbakkes.nl , domainregistry.nl , vdstbv.nl
Registrant	Stichting Internet Domeinregistratie Nederland
Registrar	Stichting Internet Domeinregistratie Nederland 2
Registration date	18-11-1999 00:00



Comment

Clear label

Previous

Comment...

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 we use logo detection to boost internet security?

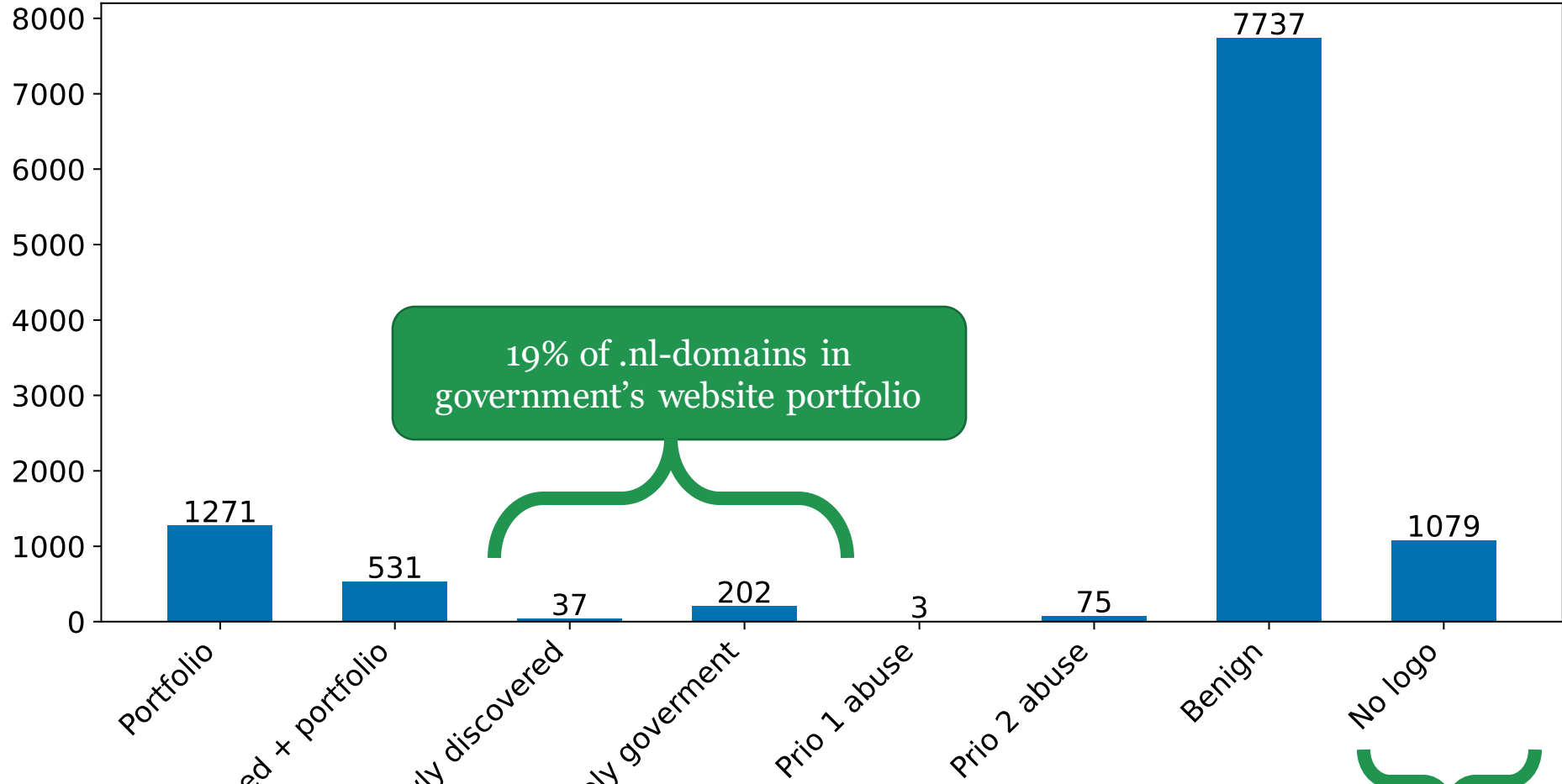
Apply to 6.2M .nl-domains and evaluate with  thuiswinkel waarborg

- Detect malicious webshop by matching results with member list
- Out of scope for today 😊

Apply to 6.2M .nl-domains and evaluate with  Rijksoverheid

- Detect abusive domains
- Discover unknown government domains

Number of domains per label (total = 10935)



19% of .nl-domains in government's website portfolio

42% of .nl-domains in government's website portfolio

Precision = .89



Example detected abusive webpages

Prio 1

Belastingdienst
rijksoverheid 0.99

Menu

Belastingaangifte

Bij een controle omtrent uw aangifte hebben wij kunnen zien dat u in aanmerking komt voor een belastingteruggave. U zult een teruggave ontvangen van: **€1161,51**, met kenmerk BTG593025.

Verifieer vooraf uw gegevens en bankrekening om uw teruggave te ontvangen. Uw teruggave kan alleen op uw rekening worden bijgeschreven nadat de verificatie is voldaan. Na het verifiëren van uw gegevens, ontvangt u van ons een e-mail ter bevestiging.

! Wat te doen met belastingteruggave?
Omdat je je belastingteruggave meestal in de zomerperiode krijgt, ligt het voor veel mensen voor de hand het geld te besteden aan een vakantie. Maar je kunt met het geld natuurlijk ook extra aflossen op je hypotheek of andere leningen.

Om er zeker van te zijn dat de verificatie van uw account zal worden voldaan door de rechtmatige eigenaar, vragen wij u als rekeninghouder zich te identificeren. Voor het innen van uw belasting teruggave verwachten wij dat de naam van de rekeninghouder overeenkomt met de naam waarop de debetreffende teruggave op geregistreerd staat.

Verifieer mijn bankrekening met iDeal

Selecteer uw bank

Over de Belastingdienst

Prio 2

CORONA ANTIGEEN SNELTEST | RIVM GOEDGEKEURD

thuiswinkel waarborg

rijksoverheid 0.95
OVERHEID GOEDGEKEURD

Corona Antigeen sneltest | RIVM goedgekeurd | 20 stuks

€199,00

- Direct uit voorraad leverbaar
- FDA / MDR / IVDR / CE goedkeuring
- 1 werkdag levering
- Ophalen ook mogelijk
- Directe verzending met track en trace
- Voor 15:00 besteld! volgende werkdag in huis
- Verpakking 20 sneltesten per doos
- BTW vrij

rijksoverheid 0.93
OVERHEID GOEDGEKEURD

BESTEL CORONA-SNELTEST >

Corona Antigeen sneltest kopen | Om te zien of je op het moment corona hebt.

Deze antigeen sneltest geeft u het antwoord op de vraag: "Wel of Geen Corona?", binnen 15 minuten!

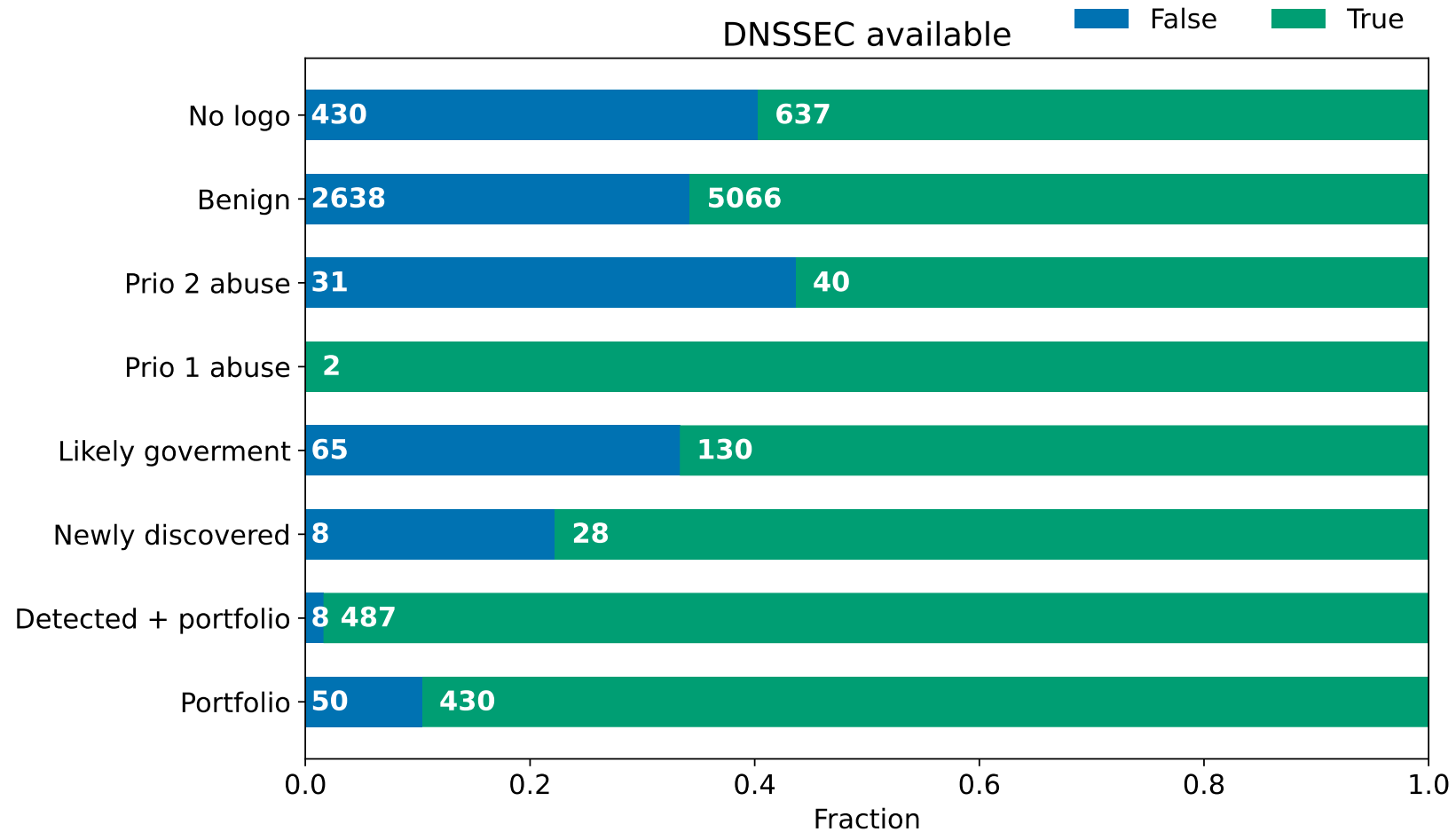
De antigeen test werkt heel nauwkeurig en snel en eenvoudig in gebruik. De test werkt met behulp van een wattenstaafje in de neusholte. Door het hoge gebruiksgemak kunnen veel mensen in korte tijd getest worden op Corona.

Importance of accurate domain portfolio

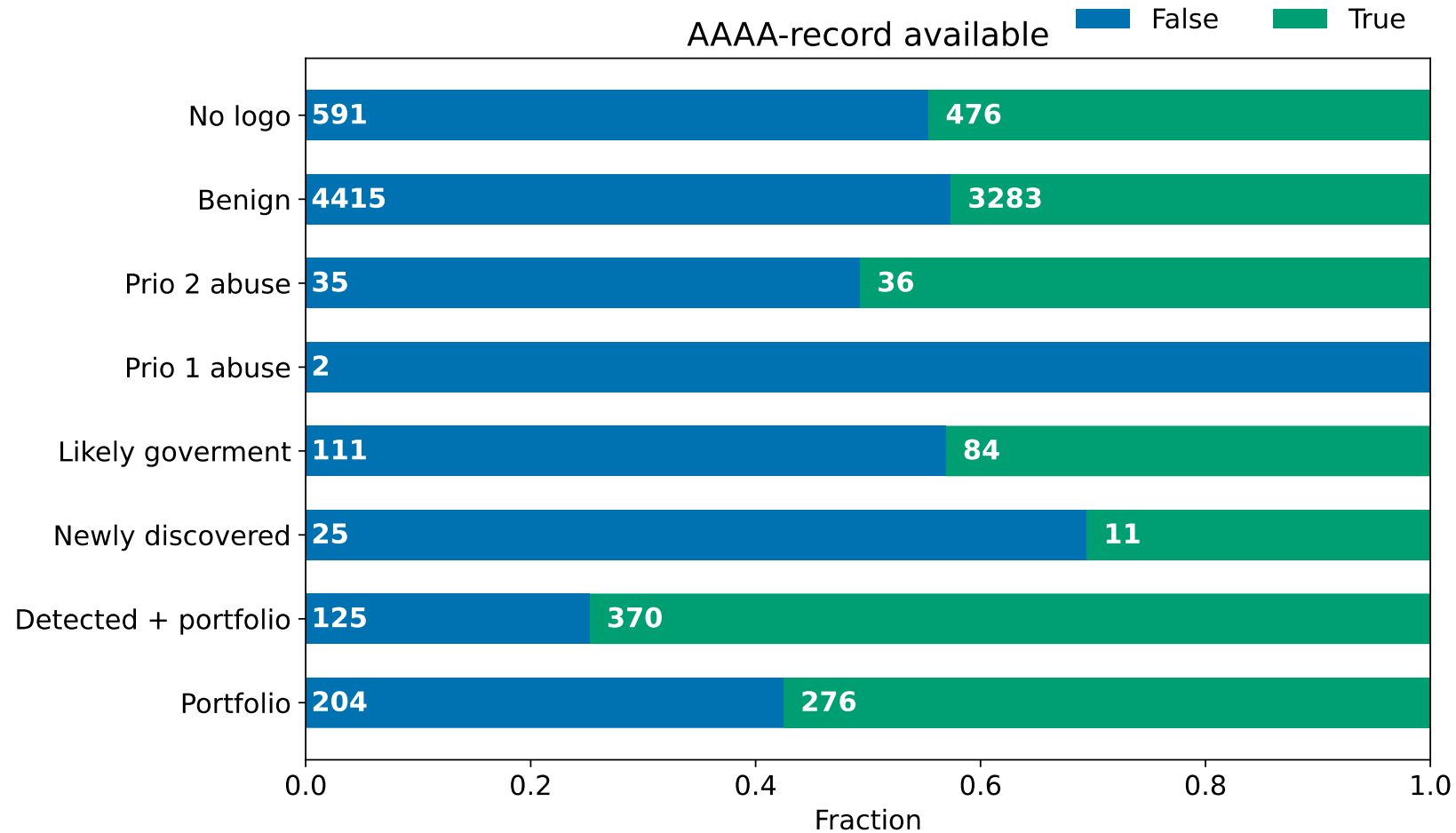
- Prevent leakage if domain got canceled by accident
- Full control over domain name and its DNS records
- Adopt and monitor security standards and technologies



DNSSEC adoption per label



IPv6 adoption per label



Lessons learned and future work

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

Next steps:

- Monitor new registrations
- Write academic paper and publish code

Volg ons

 SIDN.nl

 @SIDN

 SIDN

Q&A

www.sidnlabs.nl | stats.sidnlabs.nl

Thymen Wabeke
Research engineer
thymen.wabeke@sidn.nl

