

# Counterfighting Counterfeit

## Detecting and taking down fraudulent webshops at a ccTLD

**Thymen Wabeke**, Giovane C. M. Moura, Nanneke Franken, and Cristian Hesselman  
{firstname}.{lastname}@sidn.nl

Published at PAM2020: [https://doi.org/10.1007/978-3-030-44081-7\\_10](https://doi.org/10.1007/978-3-030-44081-7_10)



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

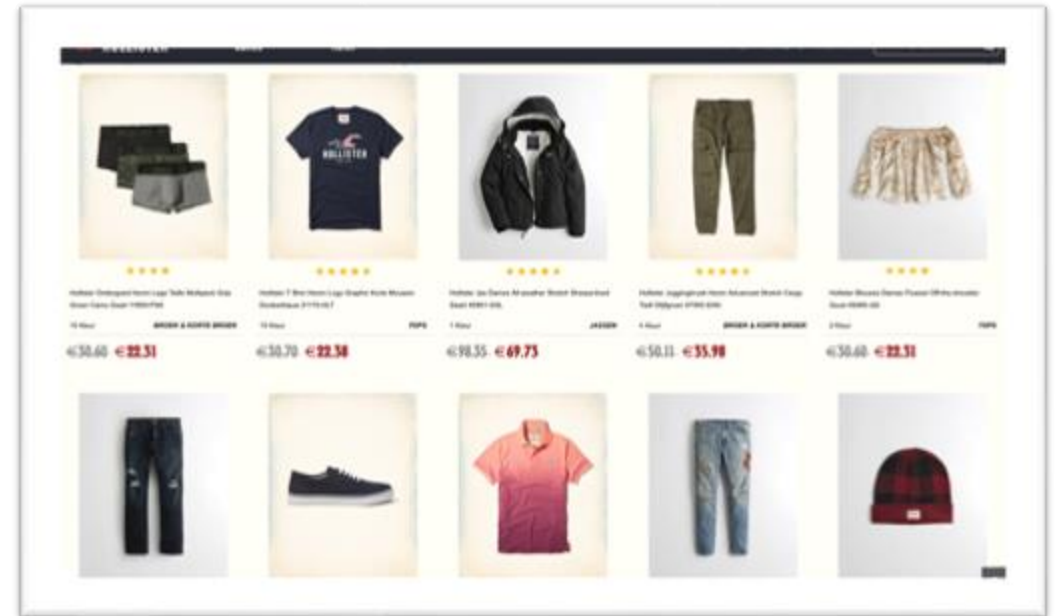
The main content area features a grid of ten clothing items, each with a product image, a star rating, a title, a description, and a price. The items are:

- Item 1:** Hollister Ondergoed Heren Lage Taille Multipack Grijs Groen Camo Zwart 11859-FNK. Rating: 4 stars. Price: €30.60 - €22.31. Category: BROEK & KORTE BROEK.
- Item 2:** Hollister T Shirt Heren Logo Graphic Korte Mouwen Donkerblauw 21170-HLT. Rating: 5 stars. Price: €30.70 - €22.38. Category: TOPS.
- Item 3:** Hollister Jas Dames All-weather Stretch Sherpa-lined Zwart 45801-GXL. Rating: 5 stars. Price: €98.35 - €69.73. Category: JASSEN.
- Item 4:** Hollister Joggingbroek Heren Advanced Stretch Cargo Twill Olijfgroen 97393-SXN. Rating: 5 stars. Price: €50.11 - €35.98. Category: BROEK & KORTE BROEK.
- Item 5:** Hollister Blouses Dames Fluwel Off-the-shoulder Goud 49289-JQI. Rating: 4 stars. Price: €30.60 - €22.31. Category: TOPS.
- 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.

# Counterfeit webshops scam, because users are unaware



VS



# SIDN's interest

- Consumer losses [1-4]
- Trust in Internet may decrease

## Perfect vantage point:

- List of *all* .nl-domains;
- Registration data and measurements.



# Results so far

- Detected thousands since 2016
- Protected users from being scammed
- 2 detection systems, 2 case studies
  - BrandCounter (2018 Q1-Q2)
  - FaDe (2019 Q1)

The image shows two overlapping screenshots. The top one is from the RIPE NCC website, displaying an article titled "Detecting and Taking Down Fraudulent Webshops at the .nl ccTLD". The bottom screenshot is from the Tweakers website, showing a news article with the headline "SIDN en .nl-registrars haalden vorig jaar 4340 fraudulente webwinkels offline". The article text includes: "SIDN heeft bekend gemaakt dat zij afgelopen jaar 4.340 fraudulente webwinkels offline hebben laten halen. Die tool heeft de afrekengegevens van de webwinkels verzameld. Webwinkels worden vervolgens gecontroleerd op overeenkomsten tussen de afrekengegevens en de daadwerkelijke bestellingen. In totaal heeft SIDN 4.340 webwinkels offline laten halen. Volgens de organisatie gaat het vaak om websites met erg lage prijs en gebrekkige vertalingen en zonder keurmerk." The article is dated 17 februari 2020 17:20.

- Q1: How many counterfeit webshops?
- Q2: How to take counterfeit shops offline?
- Q3: How do counterfeiters operate?

# BrandCounter

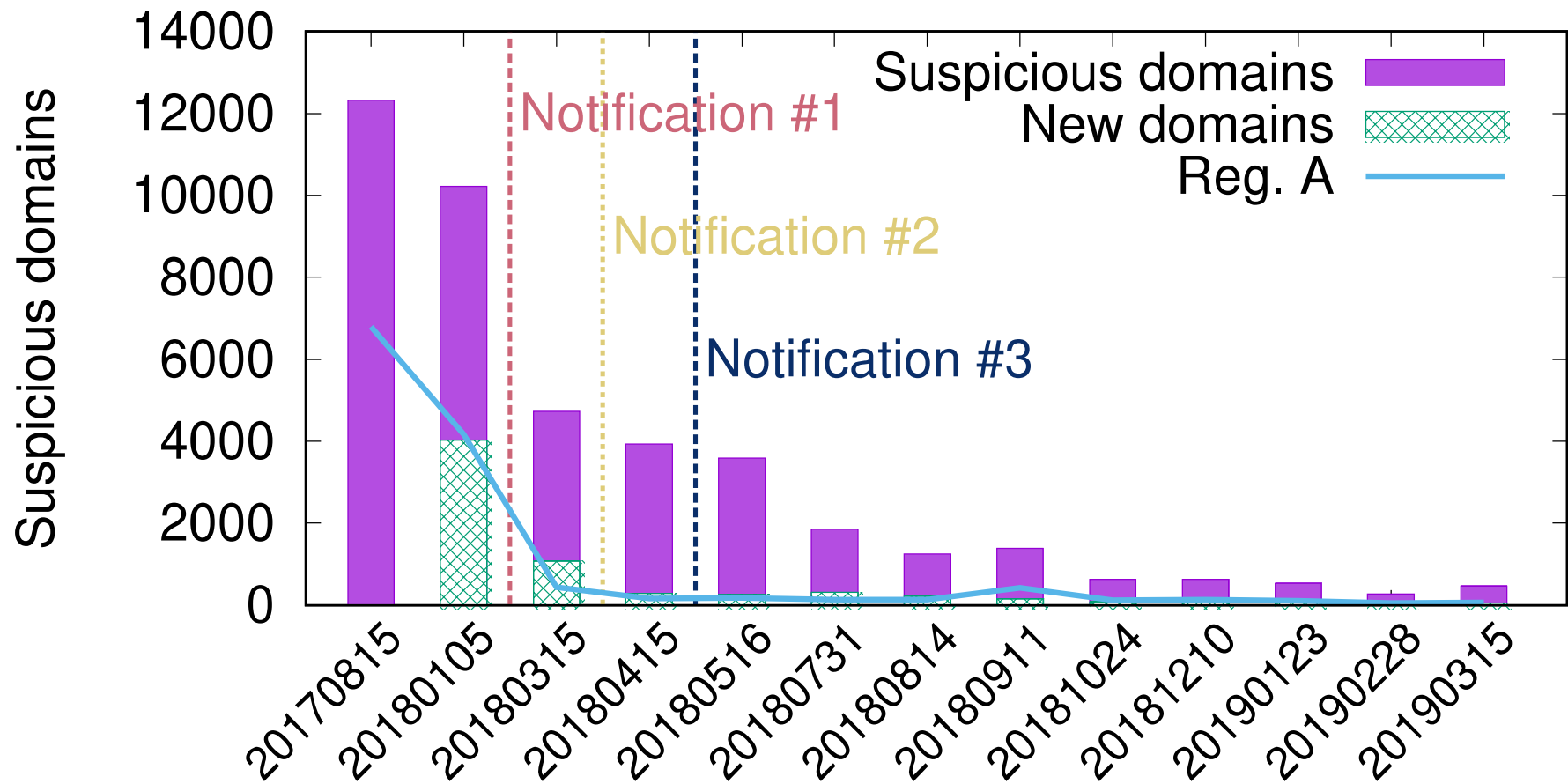
## Observation:

- Long `html <title>` tags listing brands (Nike, Reebok, Gucci, etc.)
- Helps rank high on search engines (SEO) [5]

## Method:

- Create a list with 1100 brands and discount words
- Count suspicious words in the `html <title>` of .nl-websites
- >5 words (arbitrary), mark as suspicious

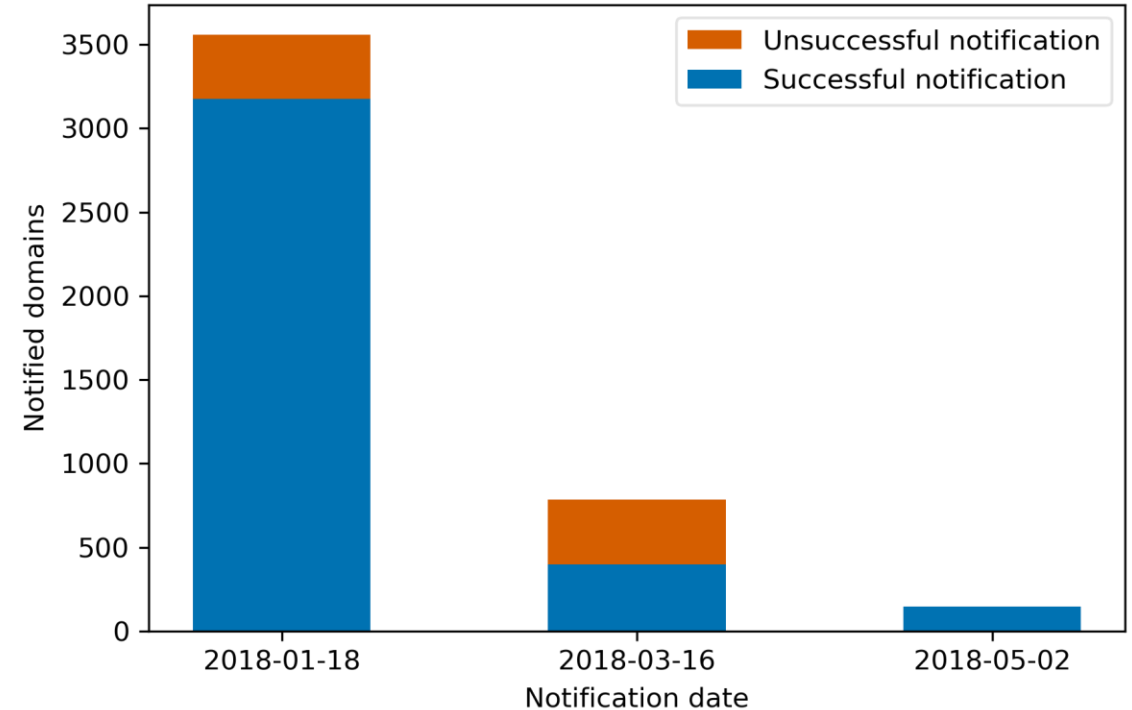


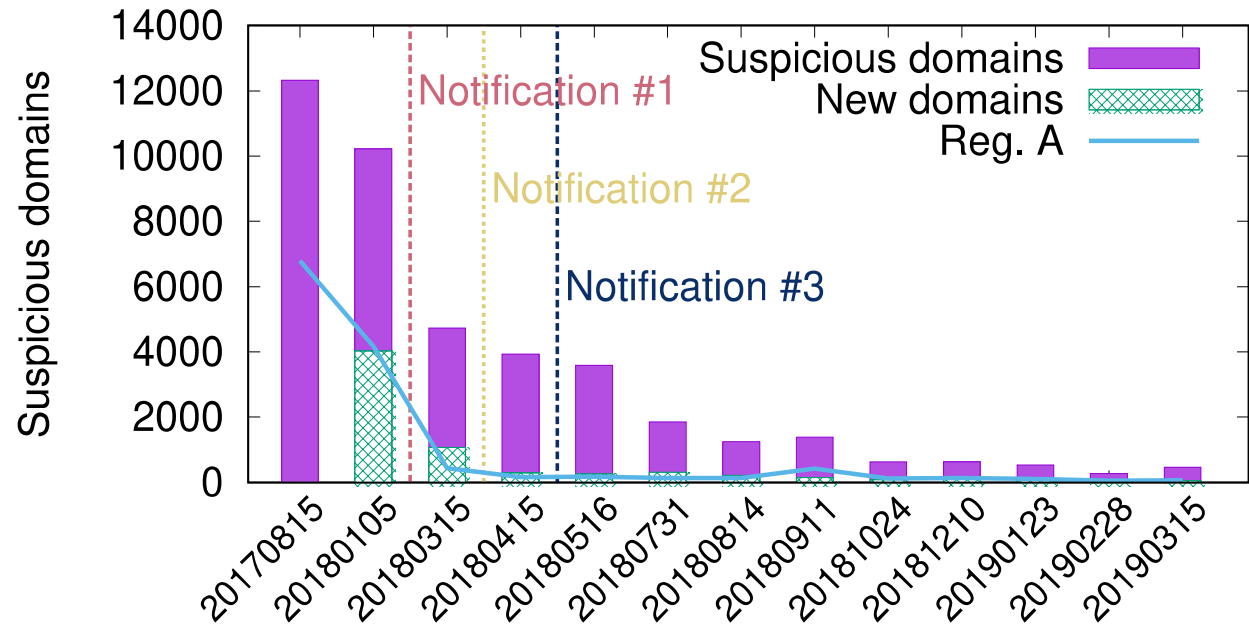




# Registrar A notification

- We (SIDN) have limited possibilities to take down domains directly
  - 42.3% registered with *Registrar A*
- Sent 4107 notifications to *Registrar A*
  - 3708 took down (90.31%)





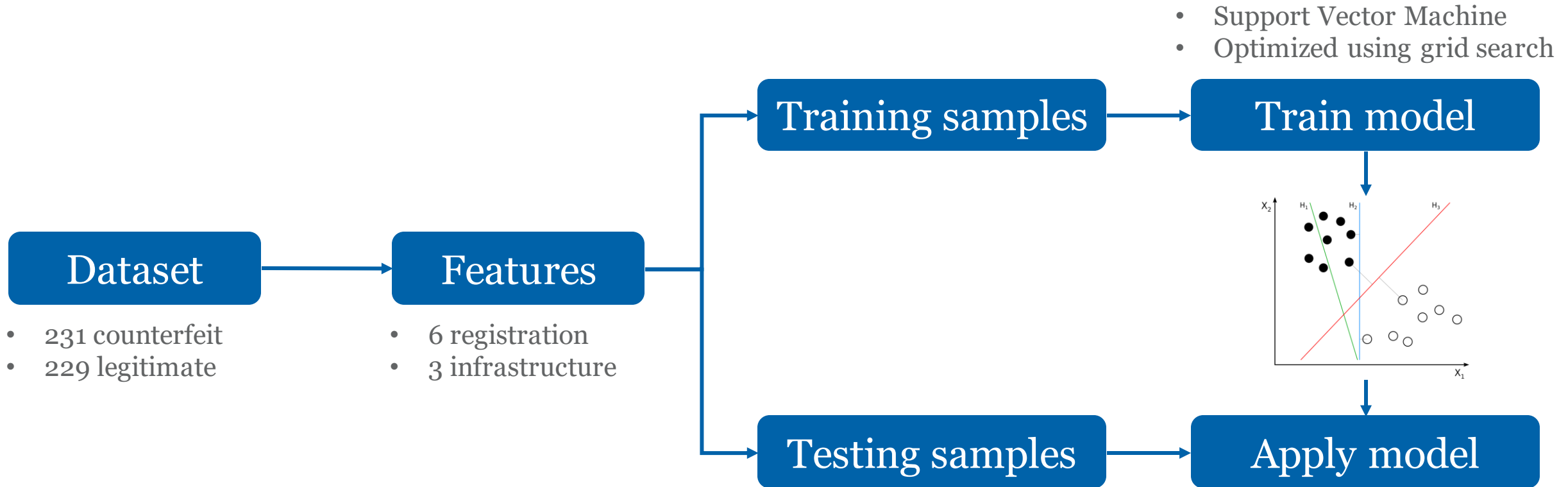
Have counterfeiters given up?  
 Learned to avoid BrandCounter?

# Fake Detector (FaDe)

- Not dependent of page titles
- Not biased towards SIDN's perspective

## **Solution:**

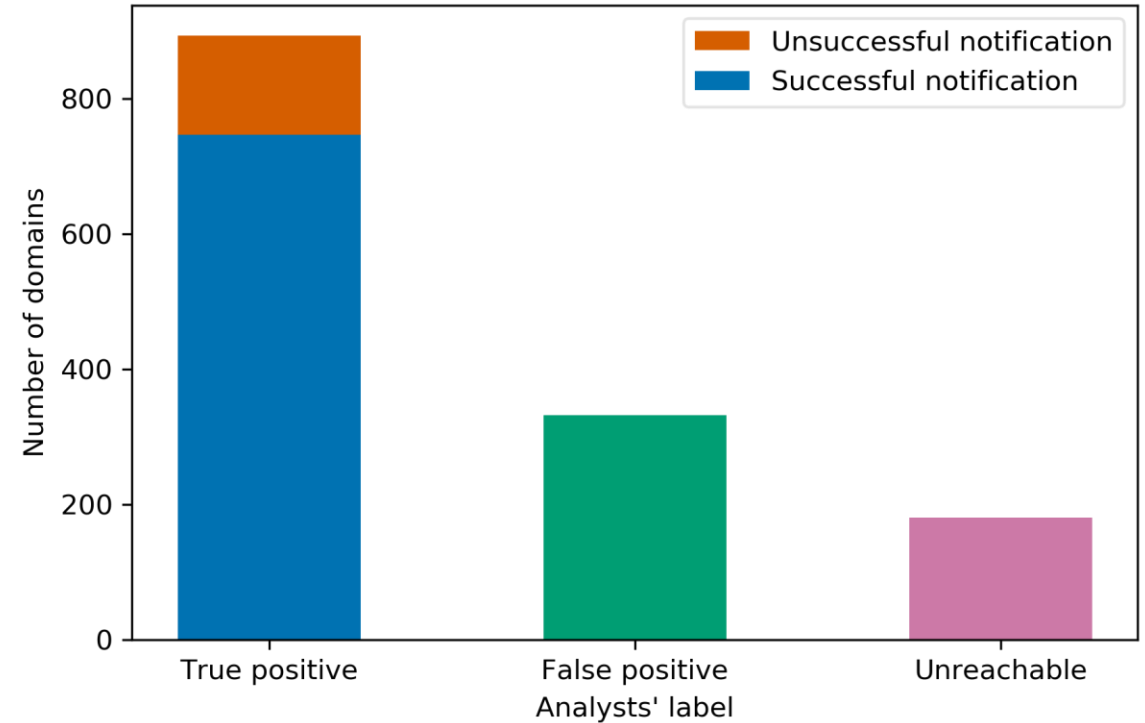
- Collaborate with ICS, a credit card issuer in The Netherlands
- ICS provided 231 counterfeit shops involved in scams
- Used supervised machine learning to train a classification model



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 took down (84%)



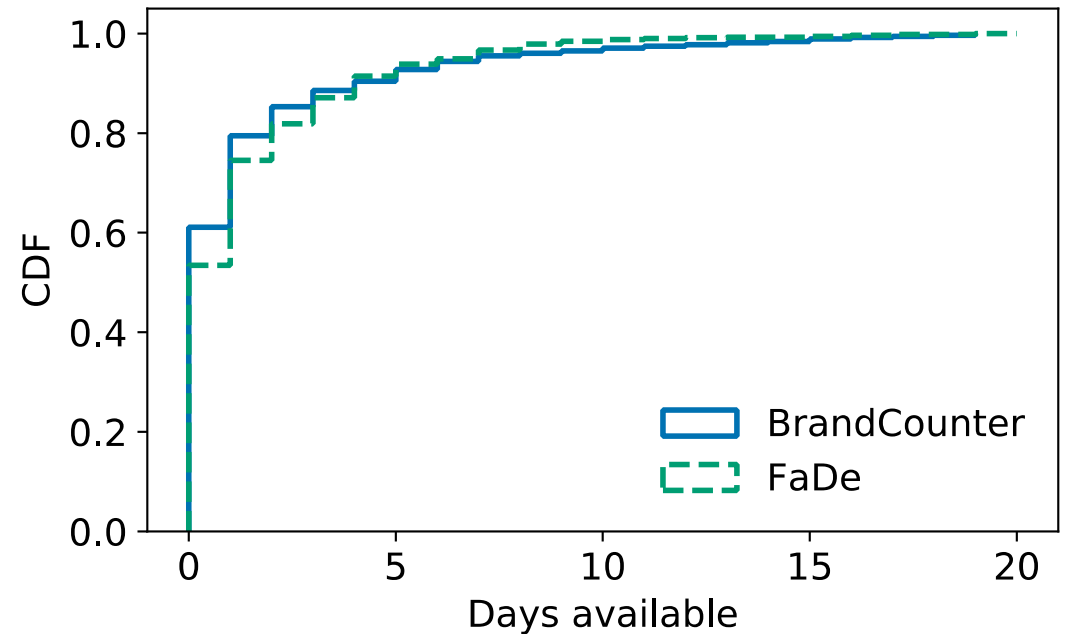
# How do counterfeiters operate?



Photo by JESHOOOTS.COM on Unsplash

# Production farm of shops: automation

- Mostly cheap registrars that offer APIs
- Similar yet different website templates
- 80% is a re-registered domain
  - Majority re-registered immediately
  - Benefit from “residual reputation” [6]

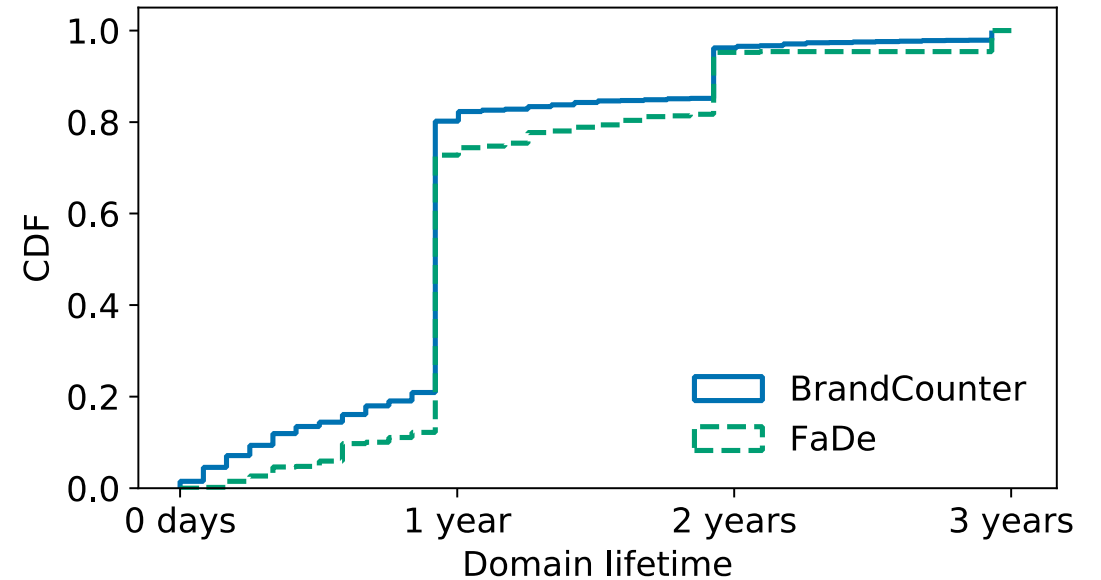


*Days in between domain expiration and re- registration.*



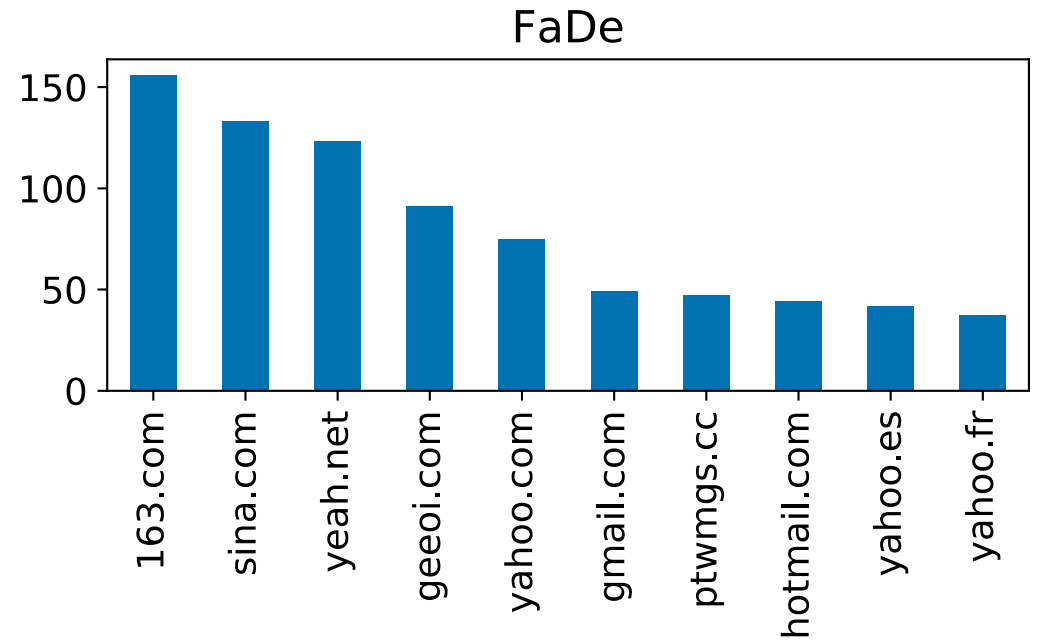
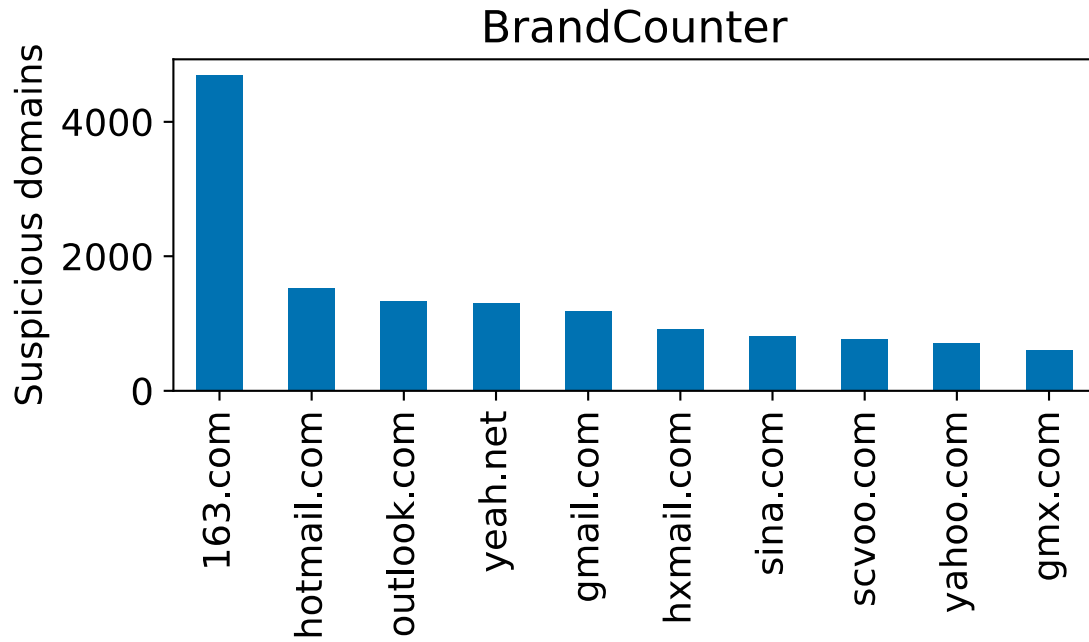
# Spreading changes: domains cheap and disposable

- Domain names do not match content
- Spelling mistakes, translation errors
- Domains have short lifetimes

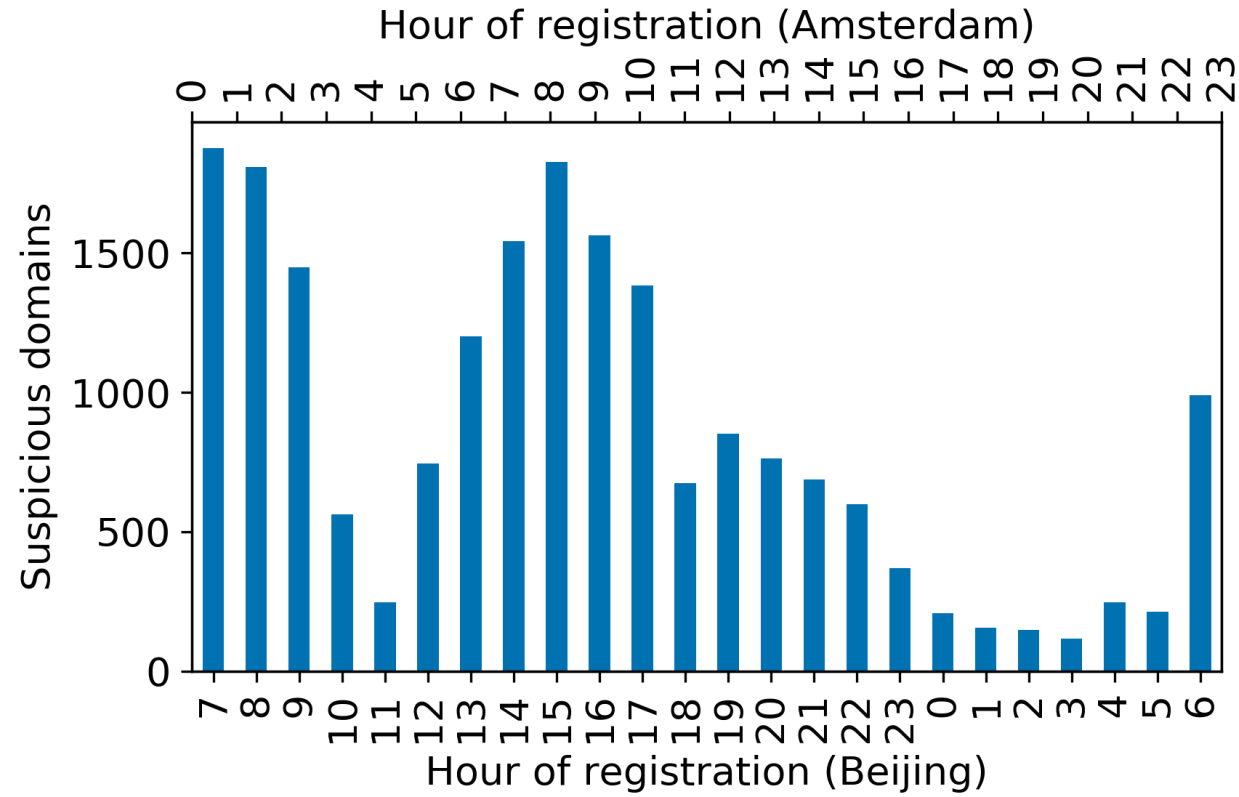


*Most domains not renewed after 1 year— the registration period.*

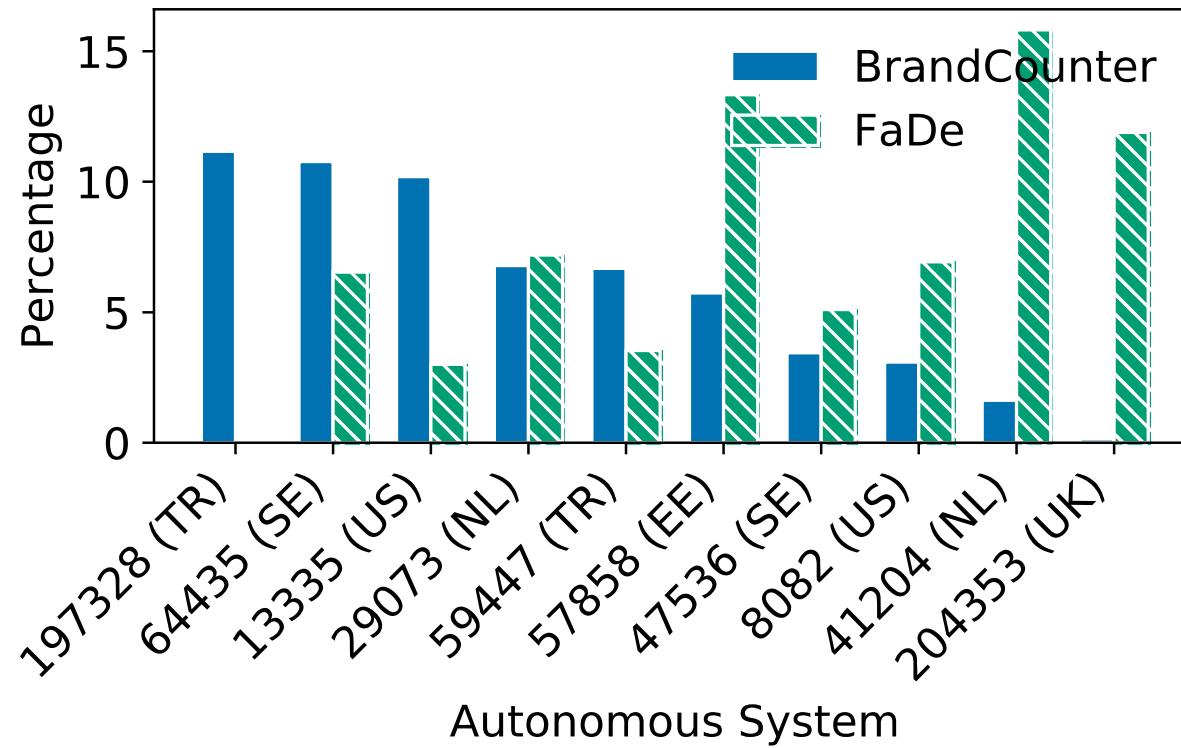
# Registrations from China



# Registrations from China



# Not hosted in China

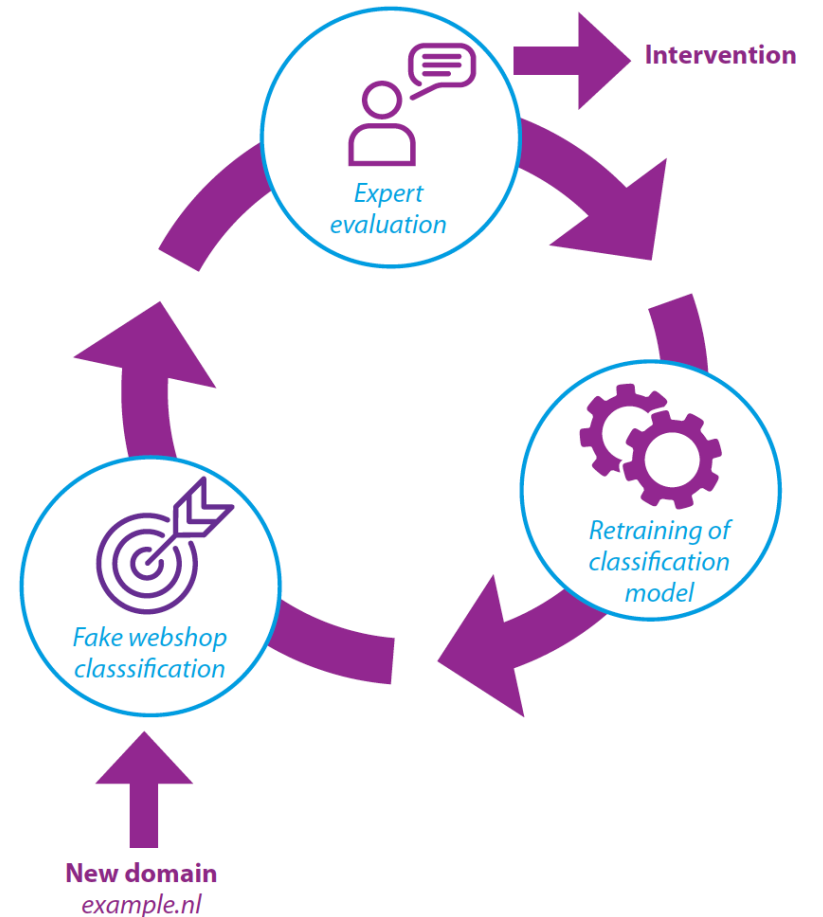


# We helped to take down 4455 counterfeit webshops



# 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 wack-a-mole game
  - Monitor features and evaluation constantly
  - We already have a new system in place





# References

1. RTL Nieuws: Dit jaar al 307 nep-webwinkels oine gehaald door politie (in Dutch) (Dec 12 2018), <https://www.rtlnieuws.nl/geld-en-werk/artikel/4520646/dit-jaar-al-307-nep-webwinkels-offline-gehaald-door-politie>
2. NOS: Consumenten voor 5 miljoen euro opgelicht via nepwinkels op sociale media (in Dutch) (Dec 12 2018), <https://nos.nl/artikel/2258095-consumenten-voor-5-miljoen-euro-opgelicht-via-nepwinkels-op-sociale-media.html>
3. NOS: Waar komen al die nep-webshops toch vandaan? (in Dutch) (May 5 2018), <https://nos.nl/artikel/2230087-waar-komen-al-die-nep-webshops-toch-vandaan.html>
4. Peter Hornung: Gef□alschte Sneaker von der FDP? (In German). <https://www.tagesschau.de/wirtschaft/fakeshops-plagiate-sneaker-china-101.html> (2019)
5. Wang, D.Y., Der, M., Karami, M., Saul, L., McCoy, D., Savage, S., Voelker, G.M.: Search + seizure: The effectiveness of interventions on seo campaigns. In: Proceedings of the 2014 Conference on Internet Measurement Conference. pp. 359--372. IMC '14, ACM, New York, NY, USA (2014). <https://doi.org/10.1145/2663716.2663738>
6. Lever, C., Walls, R., Nadji, Y., Dagon, D., McDaniel, P., Antonakakis, M.: Domainz: 28 registrations later measuring the exploitation of residual trust in domains. In: 2016 IEEE Symposium on Security and Privacy (SP). pp. 691{706 (May 2016). <https://doi.org/10.1109/SP.2016.47>