LogoMotive

Detecting Logos on Websites to Identify Online Scams - A TLD Case Study

van den Hout, Thijs; Wabeke, Thymen; Moura, Giovane C.M.; Hesselman, Cristian

DOI
10.1007/978-3-030-98785-5_1

Publication date
2022

Document Version
Final published version

Published in
Passive and Active Measurement - 23rd International Conference, PAM 2022, Proceedings

LogoMotive: Detecting Logos on Websites to Identify Online Scams - A TLD Case Study

Thijs van den Hout[1(✉)], Thymen Wabeke[1], Giovane C. M. Moura[1,2], and Cristian Hesselman[1,3]

[1] SIDN Labs, Arnhem, The Netherlands
{thijs.vandenhout,thymen.wabeke,giovane.moura,cristian.hesselman}@sidn.nl
[2] TU Delft, Delft, The Netherlands
[3] University of Twente, Enschede, The Netherlands

Abstract. Logos give a website a familiar feel and promote trust. Scammers take advantage of that by using well-known organizations' logos on malicious websites. Unsuspecting Internet users see these logos and think they are looking at a government website or legitimate webshop, when it is a phishing site, a counterfeit webshop, or a site set up to spread misinformation. We present the largest logo detection study on websites to date. We analyze 6.2M domain names from the Netherlands' country-code top-level domain `.nl`, in two case studies to detect logo misuse for two organizations: the Dutch national government and *Thuiswinkel Waarborg*, an organization that issues certified webshop trust marks. We show how we can detect phishing, spear phishing, dormant phishing attacks, and brand misuse. To that end, we developed `LogoMotive`, an application that crawls domain names, generates screenshots, and detects logos using supervised machine learning. `LogoMotive` is operational in the `.nl` registry, and it is generalizable to detect any other logo in any DNS zone to help identify abuse.

1 Introduction

A logo is a critical element of the visual identity of an organization. They influence people's perception of an organization [47], and help people to identify the associated company quickly. In the *real world*, organizations are typically very keen to protect their corporate identity (and their logo's use), by using brand protection methods [61].

The same is true for unauthorized *online* use of corporate logos: phishing attacks, for example, often attempt to impersonate organizations and use their logos both in e-mails and on webpages [1,29] while counterfeit luxury goods webshops perform trademark infringement by misusing the original brand's logos [58,59].

Besides phishing and trademark infringement, logos can also be misused in *government impersonation scams* [13,15,16], in which fraudsters attempt to

O. Hohlfeld et al. (Eds.): PAM 2022, LNCS 13210, pp. 3–29, 2022.
https://doi.org/10.1007/978-3-030-98785-5_1

impersonate governments to perform a series of crimes: "extortion, tax fraud, social security fraud, asking for donations, lenient punishment, waiver of fines, and so on" [16]. Government impersonation ultimately undermines the government's own authority to enforce laws and policies [16].

This paper focuses on identifying various types of online abuse and scams that rely on logo misuse. We do so by detecting logos on all websites in the `.nl` zone and continuously monitoring newly registered domain names, providing brand owners' abuse analysts with a complete overview of their logo's use online. With this point of view, various forms of abuse can be detected, including phishing, spear phishing, trademark infringement, misinformation, and more. To that end, we present `LogoMotive` (Sect. 2), an application that employs deep-learning for logo detection on websites' screenshots. `LogoMotive` crawls a domain, generates screenshots, detects logos in these screenshots and provides analysts with a web dashboard for annotation. Our system is designed to have operational impact, which means that we want to prevent `LogoMotive` from making autonomous decisions about domain names – ultimately protecting domains from being mislabeled and their potential consequences, such as being suspended or removed from the DNS zone. Therefore we decided to follow the human-in-the-loop principle [38]. This means we leave the assessment of whether a website abuses a logo to human analysts and do not automatically classify websites.

We present two cases studies, in which we apply `LogoMotive` to the 6.2M domains present in the `.nl` DNS zone – the country-code top-level domain (ccTLD) of the Netherlands. As such, ours is the largest research on logo detection on websites to date (the largest study before us analyzed 350k websites [29]).

In the first case study, we partner with the Dutch national government to detect government impersonation scams. In the second case study, we team up with *Thuiswinkel Waarborg*, a widely recognized trust mark certificate issuer for webshops in the Netherlands, to identify false claims of trust mark membership. We detected over 10k domain names containing the logo in both studies, which were all annotated manually by human abuse analysts at the respective organizations.

We make the following contributions: first, we show that logo detection is a powerful method to detect phishing, spear phishing, and potential phishing attacks in government impersonation scams (Sect. 3.1): we detect 168 instances of government logo misuse, 6 were active phishing domains, which attempted to commit online identity theft and bank credential theft, targeting the citizens of the Netherlands. These phishing websites were removed from the `.nl` zone after the usual legal due diligence.

`LogoMotive` is a powerful tool because it detects abuse usually missed by the traditional blocklist or HTML-based detection methods. Furthermore, it is more broadly applicable than finding phishing attacks only; it provides a complete overview of a logo's (mis)use in a DNS zone. Optionally, users can choose to monitor only newly registered websites with `LogoMotive`, which greatly reduces the amount of manual work over analyzing the entire `.nl` zone. In our experiments, we found that most malicious use of logos is found in recently registered domain names.

Our second contribution is to document the presence of "dormant" phishing websites (Sect. 3.2): government typo-squatted domains, which employ HTTP redirects [14] to forward users to the *legitimate* government website. While seemly innocuous, these websites might do so to leverage search engine optimization to increase the number of visitors [59], and could, at their will, replace the HTTP redirect by an actual phishing website, potentially compromising their visitors. Typo-squad detection systems are insufficient since they rely on a predetermined list of domain names on which variants are based. We found 9 cases of dormant phishing websites and 2 cases of dormant spear phishing attacks (targeted at very specific government agencies). Some of these typo-squat domains had MX records [34] – which specify e-mail servers – indicating that phishing e-mails can be sent from these very suspicious domains. Worse, these malicious domains websites would *redirect* users to the *legitimate* government website, which could give a false sense of legitimacy to these suspicious domains, increasing the chances of spear phishing success.

Our third contribution is to show that logo detection can be successfully used to detect fake claims of trust mark certification (Sect. 4): we detect 208 domain names leading to webshops that falsely claimed to be certified by the trust mark organization by displaying their logo, thereby misleading consumers. The trust mark organization requested these websites to remove the logo.

`LogoMotive`, our tool, is operational and has been active in the `.nl` zone for the last 8 months for both use cases here presented. We show operational impact by removing phishing websites before users can be compromised and displaying its broad applicability in finding online logo abuse. `LogoMotive` can be applied to any DNS zone and easily trained to support different logos. Hence, we make `LogoMotive`'s source code available upon request for academic purposes on https://logomotive.sidnlabs.nl, and actively promote deployment by peer registries such that `LogoMotive` can be used to find abuse in other DNS zones besides `.nl`.

2 LogoMotive

Next we present `LogoMotive`, the application that we have developed to perform logo recognition on websites. It has three main modules, as shown in the lower part of Fig. 1: *Crawler*, which takes a list of domain names as input and generates screenshots from their webpages, *Logo Detector*, which applies a deep learning algorithm to detect logos on those screenshots, and the *Dashboard*, which is used by abuse analysts who are responsible for labeling the results.

`LogoMotive` detects the presence of logos on websites, but it is not designed to automatically determine if the logo is used legitimately or not – for that, we rely on manual validation. Analysts evaluate each domain name on which their respective logo was detected. Our case studies in Sect. 3 and Sect. 4 show how analysts evaluated more than 20k domain names from the `.nl` zone.

We also add other requirements for `LogoMotive`: It must be accurate enough to limit the number of false positives and stay manageable by human analysts.

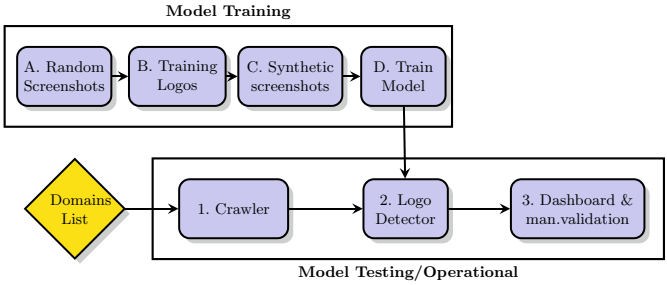

Fig. 1. `LogoMotive` architecture: model training (upper part) and operational part (lower part).

It should also be scalable, *i.e.,* able to be used to analyze DNS zones of different sizes in reasonable time, and adaptable, meaning it should be possible to detect new logos with ease.

2.1 LogoMotive Modules

Crawler: As shown in Fig. 1, the Crawler module takes a list of domain names as input and crawls the homepage of their websites (if available), following any redirects. For each website, it generates 1024×1024 pixels screenshots: two semi-overlapping from the page header towards the footer and, if not yet captured, two semi-overlapping screenshots from the footer (bottom) of the page up. We focus on the header and footer of websites, where logos are typically placed, reducing the search space for screenshots, especially on very long webpages. In our experiments, the crawler module generates 2.7 screenshots per website on average; in 65%, the first two screenshots of the header already cover all the homepage's display area.

To implement this module, we use Selenium hub [53], an automated, programmable browser that allows running multiple browser sessions in parallel. We run this module in Docker containers to easily deploy our crawler and up or down scale the number of browser nodes. Instead of downloading all image resources, we take screenshots to make sure we analyze the webpages as a regular user would see it. Furthermore, images might be hidden in CSS or SVG paths, and logos could be embedded in a larger image, thereby escaping detection.

Performance: our code is parallelized, and with 15 browser nodes, we are able to crawl 5.2 domain names per second (19k/hour) on a 12-CPU 64 GB memory machine.

Logo Detector: To detect logos on the screenshots of websites, we employ YOLO [45], which is a supervised machine learning (deep learning) algorithm designed to perform object detection.

Why YOLO: there is a large number of object detection algorithms that could be used for image recognition, such as Single Shot Multibox Detector [30], Fast(er)

Fig. 2. YOLO detects the government logo on a website.

R-CNN [46], or static feature-extraction and matching models using for example SIFT [31], SURF [3], or ORB [49]. We chose YOLO because recent comparison studies have shown YOLO outperforms other deep learning models in terms of inference speed, and in many cases also accuracy [51,54]. It is also easy and relatively fast to train, which is important for the scalability of `LogoMotive`. Because we do not change any significant details in the training or deployment of the detection model, we trust the existing comparison studies in making our choice. In our experiments, we found that the static detection methods using feature description matching perform worse than YOLO since it is not accelerated by a GPU. Also, it struggles with detection if multiple logos are shown on a page, and requires to be applied separately for each logo class we wish to detect.

How YOLO Works: YOLO is a one-stage detection model, which means it is trained to do bounding box regression and classification at the same time, making it faster than two-stage counterparts such as Faster R-CNN [54]. YOLO is a supervised machine learning model which is trained on images that contain the to-be-detected objects (logos in our case), and a list of coordinates and dimensions that describe bounding boxes around the objects, with their corresponding class labels.

At inference time, YOLO first divides the input image into multiple grids of a varying number of cells, each of which predicts several bounding boxes, as well as the class label and confidence scores of each detection. We show an example in Fig. 2. We crawl a random website and YOLO detects the logo on it. The output is the corner coordinates of the bounding boxes that describe where logos are found on the image, together with the logo class, and confidence scores (ranging from 0 to 1), which indicate how certain the model is of its detections.

Overlapping detections are filtered out by a process called nonmaximal suppression, which discards the results with lower confidences, keeping only the detections with the highest confidence.

Performance: In our setup (12-CPU/64 GB RAM machine), we use an Nvidia GeForce RTX 2080Ti for training YOLO, which is designed to leverage GPUs.

Table 1. Datasets used for `LogoMotive` training and validation.

	Value
Crawled random domains	25,000
Screenshots generated	64,893
Synthetic training samples	100,000
training set	95,000
validation set	5,000

In this setup, we can evaluate 50 screenshots per second. We use the YOLOv5 open-source python implementation by Ultralytics [57].

Scalability: YOLO can successfully be trained to detect many classes at once, e.g. the Objects365 dataset [52], which contains 365 object classes. This indicates YOLO will not be a bottleneck in the scalability of `LogoMotive`, as more logos are to be detected.

Dashboard: The last component is a web dashboard, on which analysts manually evaluate the logo detection results. It lists the domain names, the screenshots on which logos were detected, and metadata such as registration information to aid in the labeling. It allows the analysts to classify each detection, and label the use of logos on the websites as malicious or legitimate. This can later be used to follow up on the results with the appropriate measures. We host a dashboard for each class of logos we detect, meaning only relevant results are shown on each dashboard. We include a screenshot of the analysis pop-up on the dashboard in Appendix Sect. A.

Scalability: Abuse analysts at the brand owner's organization analyze the webpages on which their logo was found. This means that as `LogoMotive` is scaled to detect more logos, the workload of any particular analyst does not increase. The dashboard is a dockerized web application that can easily be scaled up.

2.2 Model Training

As a supervised learning algorithm, YOLO requires *labeled* data to be trained, so it can learn to recognize logos.

Generating labeled datasets for the training of object detection models is a very time-intensive task when performed manually. To avoid that, we generate a synthetic training dataset, a common practice in the training of various object detection models [12,56]. We generate the synthetic training datasets by (i) crawling 25k random `.nl` domain names with our crawler module, resulting in 64k screenshots, and (ii) overlaying the logos we wish to detect at random locations on these screenshots.

We randomly augment the logos by changing aspects such as scale, opacity, color, blur, occlusion, and others, such that the model becomes robust against the various appearances of logos on websites. Additionally, the augmentations make

the detector robust against simple adversarial attacks. We create 100k training samples with this process, which is sufficient to train our model to convergence, meaning the mean average precision does not increase further. We use 95% (95k) as training samples, and the remaining 5% for validation (5k), see Table 1. The validation set is used during training to monitor the generalizability of the model and to spot issues such as model overfitting. We generate 100K training samples in a little over 30 min using our method. This allows us to train the model on any logos we want to detect with minimal effort.

YOLO can be trained to detect multiple logos at once, so we do not need to train a separate detection model for each logo. We generate the synthetic training data with all the logos we wish to detect, resulting in a single dataset with screenshots that each contain one or more logos that should be detected. When more logos are to be detected, we can simply regenerate the training data including the new logos, and retrain the model, making this process scalable in practice.

The detection model was trained in 50 iterations over the whole training set, using the Adam optimizer [24], after which we found the model converged.

2.3 Model Tuning

YOLO assigns a confidence score to each logo it detects, as shown in Fig. 2. We can choose a confidence threshold to reduce the number of false positives. Detections with a confidence score lower than the confidence threshold are discarded. Tweaking the confidence threshold thus changes the trade-off between the model's precision and recall evaluation metrics. Precision is the fraction of detected domain names that indeed display the logo. Recall is the fraction of domain names that display the logo that we successfully detected. A high confidence threshold means we discard more detections, which leads to lower recall, but a higher precision.

The results of the logo detection module are manually analyzed and labeled on the web dashboard (Sect. A) by abuse analysts. Manual annotation of these results is a time-intensive task, so we would like to limit the number of false positive-samples the analysts must go through. Experimentally we found that a confidence threshold of 0.8 results in a precision of 90%, which is sufficient to still allow analysts to manually classify the results without overwhelming them. Given this confidence threshold, `LogoMotive` still finds logos that are visually altered by for example changing colors, stretching, and changing details. Logos that are altered beyond the point of recognition will not instill trust in the visitors and are therefore not a threat in the scope of this research.

2.4 Model Evaluation

In practice, we cannot determine the recall performance of our model in the entire `.nl` zone, because we do not have the ground truth of all `.nl`-websites. To evaluate the recall of our model, we generated a test set with the two logos of

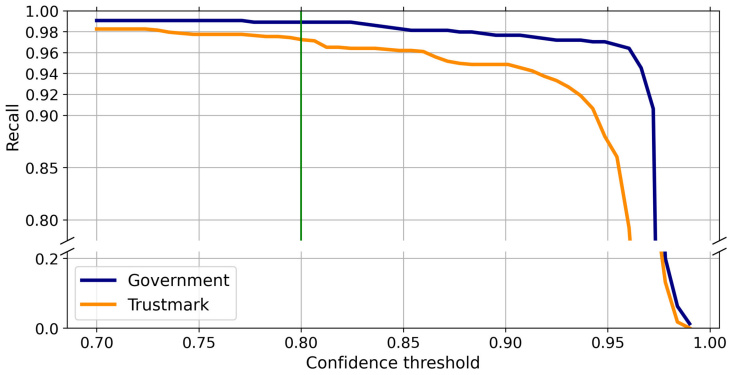

Fig. 3. Recall performance of LogoMotive at confidence thresholds. The vertical line denotes our chosen threshold.

the institutions we collaborate with: the logo of the Dutch national government, and the logo of *Thuiswinkel Waarborg*.

Dataset Generation: We used our crawler module to take screenshots of 1271 domain names from the government's website portfolio, 1300 domain names from the trust mark's member list, and 1300 random domain names from the `.nl`-zone, which generally do not include either logo. We manually annotated the screenshots generated by the crawler module to determine which logos are shown on the crawled websites. Our manual annotation resulted in a test set with 635 domain names showing a website with the Dutch national government logo, 962 with the *Thuiswinkel Waarborg* logo, and 2096 showing neither logo.

We then applied our logo detection model on this test set to compute the recall of our algorithm at various confidence thresholds. The results are shown in Fig. 3.

Using our default confidence threshold of 0.8, we obtain precision scores of 0.986 and 0.983 and recall scores of 0.989 and 0.968 for the government logo and trust mark logo respectively. This indicates that the model misses very few logos in the `.nl` zone. The small difference in recall between logos represents the difficulty of detecting a particular logo. Generally speaking, the more distinct features a logo contains, the easier it is to detect.

Note that the precision in this set is higher than it would be in the entire `.nl` zone, because this evaluation set contains a larger fraction of screenshots with a logo. The precision in practice is around 90% in the entire `.nl` zone at a confidence threshold of 80%, according to the manual annotation of the results in our use cases.

3 Government Impersonation Case Study

After training our model to detect logos, we apply it to detect Dutch national government impersonation scams in the `.nl` zone, which is the primary TLD

used by the national government. We apply `LogoMotive` in two modes. In the first mode (Sect. 3.1), we evaluate monthly snapshots of the entire zone. To detect short-lived scams, we also apply the model in the live mode, in which we evaluate every newly registered domain name (Sect. 3.2).

3.1 Full Zone Evaluation

Data Collection: Using the Crawler module (Sect. 2.1), we obtain screenshots from websites in the entire `.nl` zone. For this case study, we report 5 full passes on the `.nl` zone (covering March to July 2021). Table 2 shows our datasets.

Table 2. Datasets for government impersonation case study (2021).

	March	April	May	June	July
Domains	6.02M	6.18M	6.19M	6.20M	6.20M
Domains without websites	3.75M	3.53M	3.28M	3.30M	3.56M
Domains with websites	2.27M	2.65M	2.91M	2.90M	2.64M
unchanged websites	–	750K	744K	985K	873K
changed/new websites	–	1.90M	2.17M	1.92M	1.77M
Domains processed	2.27M	1.90M	2.17M	1.92M	1.77M

Reducing Search Space: The `.nl` zone has over 6.2M domain names. `.nl` crawls its entire zone monthly using DMAP [62]; another crawler tool that collects metadata of `.nl`-websites. We use this metadata to reduce our search space by removing domain names that do not host a webpage, show an empty page or give HTTP errors. As shown in Table 2, this allows us to go from 6.02M domain names to 2.27M domains on which we generate screenshots, for March. For April through July, we can further reduce the number of domain names the Crawler visits by excluding websites that have not changed compared to the previous visit, which we identify by analyzing the hash of the webpage.

Results: The "Full-Zone" column in Table 3 shows the results (we explain the "Newly-Registered" column in Sect. 3.2). In total, `LogoMotive` detected 12.8K domain names, 11.7K of which indeed displayed the government logo (91% precision). Given `LogoMotive` only detects the presence (or absence) of the government logo, we need to rely on manual inspection of these domains to determine if they are malicious or not. Abuse analysts at the Dutch national government manually went through the 12.8K results and categorized all of them. The analysts deal with domain name abuse daily and by working for the government they are in the best position to determine whether or not their logo is used maliciously or legitimately. Due to the time-intensive nature of the annotation work, each result is labeled by one analyst, which restricts us from comparing results between analysts.

Table 3. Manual validation results for government impersonation case 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%)

The most critical category covers government impersonation. We found 151 domains in this category, of which 3 were phishing domains. Two of them were phishing websites mimicking the National Tax Authority (Belastingdienst) of the Netherlands, and the other one was a phishing website related to the national online authentication system (DigiD). The average age of these domains was 15 d upon detection. This shows that `LogoMotive` can help to pick up scams that have not been reported to blocklists yet. Shorter-lived attacks may be detected quicker by continuously scanning the zone. This is why we look specifically at new registrations in Sect. 3.2.

The government impersonation category also contains 73 domain names that are a potential threat. This includes domains that return an HTTP redirect to a legitimate government domain but are registered by a third party who has no connection with the government. This includes suspicious names, such as a domain containing the terms 'vaccination' and 'appointment' which redirected to the official government website (`coronatest.nl`), the website on which Dutch citizens can plan an appointment for a COVID-19 test.

Domain Name Popularity: we estimate, indirectly, how popular these government impersonation scams are compared to legitimate government websites. To do that, we use Authoritative DNS server logs. Authoritative DNS servers are a type of DNS server that knows the contents of a zone from memory [22]. DNS resolvers, such as Google Public DNS [17] and the default configurations in user devices, ultimately ask the authoritative servers for records in their zones.

As the `.nl` operator, we have access to historical authoritative DNS traffic of two of the three authoritative servers. We collect this data and store it in ENTRADA [63], our open-source Hadoop-based database from which we can query this data. Although we do not receive every DNS query because of caching [35,36], this still provides an indication of how popular the domains are.

We compare the average number of daily queries for the 73 potentially malicious domain names and the 952 legitimate government domain names, as

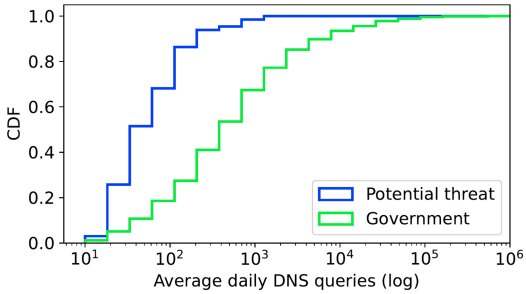

Fig. 4. CDF of average daily DNS queries seen at `.nl` authoritative nameservers.

observed by our authoritative name servers in the 7 days before annotation. (We choose one week given the known weekly and diurnal patterns of Internet traffic [44]).

Figure 4 shows the results for average daily DNS queries across these two groups in a cumulative distribution function (CDF). The fraction of domain names on the y-axis receives at most the number of DNS queries on the x-axis. It shows that potentially malicious domains receive fewer queries ($\mu = 132.8, \sigma = 239.0$) than government domains ($\mu = 5566.2, \sigma = 30724.2$). However, there are some potentially malicious domains that are popular. For instance, the suspicious COVID-19-related domain name receives 1% of the queries received by the official `coronatest.nl`. Although 1% may seem small, it still represents a large number of users given the popularity of the official domain name during the pandemic.

Domain Registration Data: We also manually inspected the registration data of domain names that are a potential threat. As the registry for `.nl`, we store the registrant information when a domain is registered. This data appears valid and legitimate for most of these domains, similar to what we see in normal registrations.

Domain Age: we show in Fig. 5 the CDF of the domain age per group. We see that the potential threat domain names are newer than government domains: 80% of the government domains are at least 2 years old, whereas only 20% of the potential threat domains are two years old.

This raises the question of whether attackers build a reputation with the seemingly legitimate domain name, and later on, launch an attack. Attackers may, at any time, direct users to a scam or send e-mails that appear to be from the government. We observe that 49% of the suspicious domain names published MX records on the day of labeling and could potentially send malicious e-mails that appear to be sent from a government domain.

75 domain names fall into the category *other* under government impersonation. This category comprises dubious applications of the government logo, such as false testimonials or endorsements, visually altered logos, and satirical websites. Analysts discussed these cases with the government's communication department who in some cases asked the webmaster to remove the logo or be

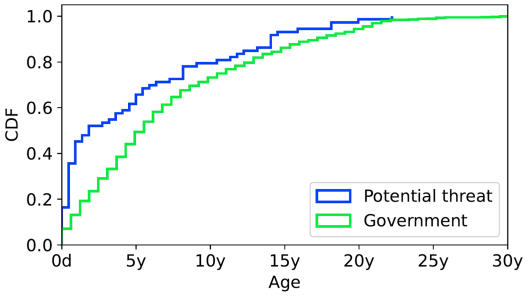

Fig. 5. CDF of domain age.

more clear in the relation with the government. Notably, `LogoMotive` could even detect highly visually altered variants of the government logo.

Most domains on which the logo was found fall into the *benign* category, *i.e.,* websites that are not directly related to the government but make use of the logo legitimately. For example, these include news websites and companies using the government logo in a testimonial.

The last category includes government websites. We detected 636 known government websites but also discovered 316 legitimate government domain names that were not listed in the website portfolio of the Dutch national government. 109 of these discovered domains could be added to the portfolio directly and 207 cases are still pending investigation at the time of writing. Asking the analysts about this, we found that the reason for this is that different branches of government such as ministries, government agencies, and regulators often use their own registrar instead of the national government registrar. This is against the government policy for registering domains and has several drawbacks. As opposed to existing abuse detection methods, `LogoMotive` can find unknown domain names that are not malicious.

Domains outside the government portfolio are a latent risk, for example, because they can expire and be re-registered by a third party, *e.g.,* a domain drop-catcher [26], who are specialized in registering the domain names within seconds after they become available. This has happened before and led to data breaches at the police and a health organization in the Netherlands [40,41]. This resulted in sensitive information such as arrest warrants and the health records of thousands of children reaching journalists.

Security Standards Adoption: the Dutch national government has committed to adopting security protocols on their domain names, and monitors the adoption rate of the websites in their domain name portfolio. These security protocols include DNSSEC [2], which adds authenticity and integrity to DNS using cryptographic algorithms, and DMARC [25], which adds authenticity to mail servers and provides domain owners with a way of advertising their legitimate e-mail servers using DNS. Domains *not* in the government portfolio cannot be monitored on the adoption of these standards, and may therefore unknowingly be insecure.

Table 4. DNSSEC and DMARC adoption in Dutch national government domain names.

	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%)

We evaluate DMARC and DNSSEC adoption in these two categories of domains using our crawler DMAP [62]. `.nl`-websites. Table 4 shows the results (our crawler missed data for 7 domains that are not in the portfolio). We see that the adoption of DNSSEC and DMARC is prevalent in the monitored government websites. The unmonitored domains have lower adoption rates for both DNSSEC and DMARC, meaning that for those domain names users are not protected against DNS and e-mail spoofing.

Summary: `LogoMotive`'s evaluation of the entire `.nl` zone allowed to detect 3 phishing domains, 73 domains that could become malicious, and 75 other embodiments of government impersonation. An unexpected finding was that we also discovered legitimate domain names that communicate on behalf of the government, but were not included in the government's portfolio, which could become a security threat. `LogoMotive` helps in detecting such threats.

3.2 Live Registration Monitoring

Given phishing domains tend to be short-lived [4], we applied `LogoMotive` to domains right after their registration, to detect potential scams faster.

Data Collection: For this case study, we continuously process every domain added to the `.nl` zone for two months (Aug. 15th to Oct. 15th, 2021). For each domain added to the zone, we ran the `LogoMotive` pipeline every 3 h for 15 days after the registration date We only generate *new* screenshots if the page contents changed. In total, we analyzed 134.4k domains, and 44.4k eventually had a webpage.

Results: We apply `LogoMotive` to detect the government logos on these 44.4k domains, and found 53 domains with the government logo on their website. Similar to the full zone scans, the government analysts also validated these results.

The "Newly-Registered" column in Table 3 shows the results. First, there were no false positives – so every detected domain indeed displayed the logo of the government. We then use the same categories from Sect. 3.1 to further classify these domains. 33 domains (62%) were labeled as benign and 4 (7.55%) as government domains, but 16 (30.19%) domains were government impersonation scams.

Phishing Domains: from the 16 impersonation scams domains, we found 3 phishing websites. Their target group comprised all citizens of the Netherlands. Two impersonated the National Tax Authority – they presented the website visitors with a tax penalty warning, and request the users to proceed with the payment. The other phishing domain attempt to obtain citizens' online identification credentials, by impersonating the national online authentication system (DigiD). We found that 2 of these 3 domains were present in Netcraft's [32] blocklist, a popular phishing URL provider, which collects phishing URLs from reports by volunteers that run their toolbar. This suggests that `LogoMotive` complements existing techniques. Upon detection and evaluation, these domains were ultimately removed from the `.nl` zone, after the required legal procedures.

Potential Threats: further, the analysts classified 9 domains as potential threats. They all follow the same pattern: they typo-squat a legitimate government domain, but instead of directing web users to a malicious scam page, they redirect users to the *legitimate* government website, using HTTP redirects. We expect this is a strategy to build up a domain reputation with users, and, once popular enough, use the domain name to host a phishing website. Three of these domains also published MX records, which means they could potentially be used to send e-mails that appear to be sent by the national government. These domains were also removed after the evaluation.

Dormant Spear Phishing: 2 of the potential threats are likely dormant spear phishing attacks. 1 of these redirected to a specialized branch of the government that is likely not known by the general public. This domain name also published MX records which pointed to a mail server that is often used in shady activities, according to our abuse analysts. The other dormant spear phishing attack redirected to a service that is only intended for government employees. These domain names could become a serious threat because compromising a national-level agency could have severe implications. For example, the United States has documented cases of spear phishing against various government agencies [9]. Given that spear phishing is harder to detect, they tend to not appear on lists like Netcraft – indicating that our method is complementary to existing techniques. These malicious domain names were removed from the zone.

Summary: `LogoMotive`'s live monitoring of the `.nl` zone allowed us to detect 3 phishing domains and 9 potential threats. 2 of these threats were targeted at very specialized branches of the Dutch national government. `LogoMotive` can find scams that have not yet been reported to blocklists.

4 Trustmark Abuse Case Study

Background: *Thuiswinkel Waarborg* is an organization that certifies online webshops to show visitors which shops are secure, trustworthy, and honest. They evaluate whether webshops meet certain legal, security and financial stability requirements, for example, the shop must offer lawful return policies and pay after delivery options.

The *Thuiswinkel Waarborg* logo is widely recognized in the Netherlands, by more than 90% of the population [33]. The consumers association in the Netherlands (*Consumentenbond*) also recommends this trust mark [10]. This suggests that consumers are more likely to trust a webshop having the *Thuiswinkel Waarborg* logo on it, and therefore, be more likely to shop on these certified webshops. As a consequence, online shops have an incentive to obtain the trust mark legitimately, or to *abuse* it.

Data Collection. As in Sect. 3.1, we apply `LogoMotive` to evaluate the entire `.nl` zone from 2021-06-24 until 2021-09-27 and detect webpages that contain *Thuiswinkel Waarborg*'s logo. We use the member list of *Thuiswinkel Waarborg* to automatically label domain names specified by members as benign. The certified members are required to include the *Thuiswinkel Waarborg* logo on their websites, which should link to a page on the *Thuiswinkel Waarborg* website where the certificate details can be verified. For example, for which shop *Thuiswinkel Waarborg* has issued the certificate, and which of the shop's domain names may use it. We check whether the webpages we detect publish a hyperlink to a valid *Thuiswinkel Waarborg* certificate and if the detected domain matches with the certificate's domain we also automatically label it as benign.

Logo Detection and Validation. Table 5 shows `LogoMotive` results: it found 10,669 domain names with the logo of *Thuiswinkel Waarborg*. To validate our results, we shared them with analysts at *Thuiswinkel Waarborg* using `LogoMotive`'s dashboard, who manually labeled the domain names from 2021-09-23 until 2021-12-16.

We also add a column with unique URLs, because we observe that webshops often register multiple domain names, where one is used as the primary domain and the others serve an HTTP redirect forwarding users to the primary domain. This strategy of using multiple outlets is probably done for search engine optimization (SEO) and/or marketing purposes.

From the set of 10,669 domains names, 5582 (52.32%) were automatically labeled as benign because they belong to certified webshops and show the correct certificate, so they did not require manual annotation. 83 domains (0.78%) did not show the trust mark at the time of inspection. The remaining 10,586 (99.22%) indeed showed the trust mark logo.

The *Thuiswinkel Waarborg* analysts then classified each domain with the trust mark into subcategories. The majority of the domains fall in the *benign* category, *i.e.,* certified webshops and a few other domains, for instance, the website of *Thuiswinkel Waarborg* itself and those of events they organized.

The second category is trust mark abuse. This category contains 208 domains of webshops that have the *Thuiswinkel Waarborg* trust mark, while they are not a member. These shops are unlikely to meet the requirements that *Thuiswinkel Waarborg* members must meet and therefore pose a risk to consumers who are likely not aware of this deception.

Table 5. Manual validation results for trust mark abuse case study.

Label	Domains	Unique-URLs
Total	10669	3890
Without trust mark	83 (0.78%)	64 (1.65%)
With trust mark	10586 (99.22%)	3826 (98.35%)
Benign	10324 (96.77%)	3691 (94.88%)
Trustmark abuse	208 (1.95%)	106 (2.72%)
Discovered	54 (0.51%)	29 (0.75%)

Action Taken: Thuiswinkel Waarborg contacted the companies behind the abusive domains with the request to remove the trust mark. *Thuiswinkel Waarborg* may take legal actions if this request is not responded to. At the time of writing, 104 of the 208 (50%) of domains that abused the *Thuiswinkel Waarborg* trust mark have removed it from their website.

Domains Profile: We manually analyzed a sample of the 208 domains, and most of them seem to be legitimate shops, with rich and well-designed websites, and some even mention a valid Chamber of Commerce number, which in the Netherlands indicates that it is an existing business.

Domain Age: Next we look into the average age of the domain names in both groups. Figure 6 shows that, for both groups, the domains are relatively old: half of the domains are at least 11 years old (benign) and the trust mark abuse are least 6 years old. That is very different from phishing, in which domains tend to be short-lived (Sect. 3.2).

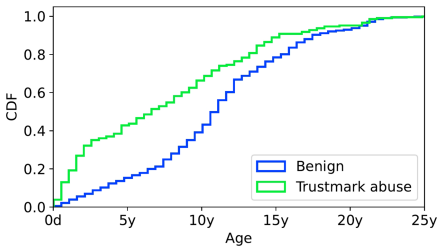

Fig. 6. Age of benign and trust mark abuse webshops.

Reasons for Trust Mark Misuse: feedback from *Thuiswinkel Waarborg* analysts confirms that webshops that misuse the trust mark are not necessarily malicious. They hypothesize that these webshops misuse the logo to improve sales, and avoid obtaining their own certificates from *Thuiswinkel Waarborg* either due to the costs involved and/or the legal and financial requirements they have to meet to obtain the certification.

Domains Popularity: We indirectly measure the popularity of these domains by analyzing incoming DNS queries for the `.nl` authoritative DNS server, as we did in Sect. 3.1. For each domain name, we compute the number of average daily queries and unique IP addresses of resolvers we observed one week before the annotation date. While the number of queries and resolvers do not correspond to the number of unique visitors (due to caching at DNS resolvers), it indicates how popular a domain name is.

Figure 7 shows the average number of daily DNS queries for the 151 trust mark abuse and 9659 benign domains, and Fig. 8 shows the average daily number of resolvers. Differently from the government impersonation case (Sect. 3.1), we see that both classes of domains have a similar number of queries *and* resolvers. The fraction of domains receiving over 1k queries per day is in the same range (14.0% for benign and 12.5% for trust mark misuse).

In addition, 3.6% (367 domains) of the benign domains receive more than 10k queries compared to 1.4% (3 domains) with trust mark misuse. This shows that domains that misuse the trademark are also very popular (not necessarily because of the trademark misuse), which is different from the government impersonation case.

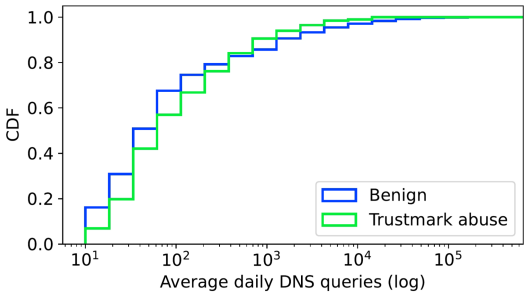

Fig. 7. Cumulative distribution function on the average number of daily DNS queries.

The last category contains 54 domain names of *Thuiswinkel Waarborg* members that were previously unknown to *Thuiswinkel Waarborg*. Analysts found that these domains belong to a certified *Thuiswinkel Waarborg* member, but the domain itself was not specified by the member. Similar to the discovered domains of the Dutch national government, these domains are a lurking risk, because *Thuiswinkel Waarborg* was unable to monitor whether those domains comply with the rules and standards imposed by the organization.

Summary: `LogoMotive` evaluation of the entire `.nl` zone allowed to detect 208 domains that abuse the *Thuiswinkel Waarborg* trust mark. We show that these domains have a long life cycle (6.8 years) and attract as many visitors as certified *Thuiswinkel Waarborg* webshops.

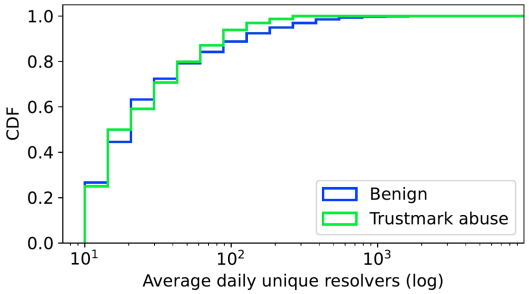

Fig. 8. Cumulative distribution function on the average number of daily unique resolvers.

5 Related Work

Phishing Detection: phishing detection is a very active research field *e.g.,* [4,11, 29,43]. Most previous research focus on textual features such as URLs [27,39,50], or HTML content [28,43,48]. We, on the other hand, rely on visual features on a page, namely logos, to detect phishing and other malicious content involving logo abuse.

Single feature: we design `LogoMotive` to detect logos, and rely on human validation to determine if there is abuse or not. Therefore, we did not explicitly design it to detect phishing or other types of malicious websites. `LogoMotive` is a broadly applicable tool for any organization aiming to protect its brand from online abuse. Given we are interested only in logo detection, we refrain from using textual features, such as registrant details, to make no assumption about the usage of logos. We can in future research combine the output from `LogoMotive` to phishing detection systems.

Visual features have been used for the detection of phishing in previous research. [66] Uses the global visual similarity of a target page to a suspicious page, combined with the existence of the target logo, to detect phishing sites. [1] detect phishing websites by comparing websites to a profile of trusted websites' appearances, which is created using text features, as well as logos which are detected using SIFT, a method for image matching. They apply their method on 1000 phishing sites from the Phishtank dataset and 200 legitimate sites from the Alexa top sites. [60] also uses SIFT image matching to find logos on pages and use a browser plugin to warn users if a logo is found on an unauthorized page.

Logo Detection: logo detection is a subset of object detection, for which various methods exist and have been applied in previous related research. Deterministic feature extraction and matching methods using SIFT [31] and SURF [3] have been employed to detect logos on websites by [1,60]. [5] use features of Histogram of Oriented Gradients to obtain a visual representation of phishing target brand logos. We found these methods to perform inadequately for our purpose, and

instead use a machine learning object detection method. Recently, deep learning object detection models are used to detect logos on websites. [29,64] use Faster R-CNN to detect the identity logo and input text fields [29]. Our approach applies YOLO version 5 to detect logos on screenshots of `.nl` webpages, because of its fast inference speed.

Phishing Detection with Logos: logo recognition has been previously used to detect phishing websites. [7,8,29,65] all detect logos on webpages and attempt to match the logo's brand identity to an organization. Next, they use this information to determine if a website is a phishing site by comparing the domain name with the organization's known URLs, for example, if the Amazon logo is found, whether the domain name matches amazon.com or not. Their underlying assumption is that the relationship between the logo and domain name is exclusive, implying that logos may only be used legitimately on a select set of domain names. We, however, show in Sect. 3 that the government logo can legitimately be used on domain names that do not belong to the government. Another example is credit card company logos, which are often placed on e-commerce websites, which simply use these payment services. In contrast, our work does not assume any relationship between a logo and a company in the detection process: it simply detect logos, and it is up to to the brand owners to further label the classification results. In certain cases, we can prioritize results or automatically label a subset of the results, for example, if a logo may only be used on specific websites.

The most recent related study is Phishpedia [29]. This study aimed to detect phishing pages using logo detection, while our goal is much wider: we want to detect all websites containing a particular logo, which exposes many kinds of logo (mis)use, including phishing. This difference is reflected in multiple aspects. First, Phishpedia detects the most prominent logo (identity logo) using Faster-RCNN, while we detect all logos on a webpage using YOLO. Second, Phishpedia also identifies whether webpages contain text fields. This information is not relevant for our use case, because `LogoMotive` does not solely detect phishing websites. Third, Phishpedia compares the URL with a white-list of brand URLs. If a webpage contains a brand's identity logo, a text input field, and its URL is not white-listed, it is marked as phishing. In contrast, `LogoMotive` does not automatically make decisions about the websites it finds but facilitates abuse analysts in determining the motive with which their logo is used. Finally, Phishpedia is evaluated on a set of 30K phishing websites obtained from the OpenPhish database. `LogoMotive` is deployed in the `.nl` zone and is evaluated in two case studies, showing operational impact.

It is difficult to compare the accuracy of `LogoMotive` to any existing work, given that our proposed system employs the human-in-the-loop principle. `LogoMotive` is used by abuse analysts of various companies to help detect online abuse of their logo. Comparing it to phishing detection methods falls short of the broader applicability of `LogoMotive`. However, the phishing domain names we detected were not yet reported to blocklists, indicating that `LogoMotive` can be used to detect phishing websites more quickly than existing methods.

Dataset Size: [8] evaluate their results on a dataset consisting of 1140 webpages. [29] apply their method on the OpenPhish service for six months, in total accumulating 350K phishing URLs. After crawling and filtering out empty and legitimate pages, they end up with 29,496 phishing websites used in their evaluations. They use 29,951 benign webpages from the top-ranked Alexa dataset to evaluate their detection method. [7] evaluate their method on 400 phishing websites from the popular PhishTank database and 50 legitimate websites from Alexa. [65] use a dataset of 726 webpages containing both phishing and legit webpages. We evaluate our contributions on the `.nl` zone, consisting of over 6.2 million domain names. Moreover, `LogoMotive` is operational in the `.nl` DNS zone. In addition, we generate a ground-truth dataset of over 20K manually annotated domain names with the logos detected for our use cases. This dataset can be used for future research on abuse detection, including the automatic prioritization of `LogoMotive`'s results.

Recently, phishing kit detection has risen in popularity [4,11,42]. Phishing kits are purchasable, easily deployable phishing site templates. [4] study phishing in the wild by detecting the use of such phishing kits. They apply their method to the TLS Transparancy Logs Project [18] and find 1,363 phishing domains targeted at a Dutch audience. They also found that phishing sites are very short-lived, with a median up-time of merely 24 h. This supports our findings; we found few phishing domains when crawling the entire `.nl` DNS zone, but found more abuse when monitoring new domain name registrations live.

Spear Phishing Detection: Spear phishing detection has previously been based on e-mail and textual analysis [19,55]. Our system is not focused on spear phishing per se, but it has proved useful in detecting them and filling a void where phishing blocklists fail to cover.

Fake Webshop Detection: fake webshop detection and counterfeit luxury goods detection has been also done in the past [58,59]. In both cases, while logos were mentioned, logo detection was not employed. Both studies found a very large set of webpages selling counterfeit goods. Our case study with *Thuiswinkel Waarborg* (Sect. 4) differs from them because we find a smaller set of domains that misuse the trust mark, but they are legitimate webshops that have been active for many years but still mislead visitors.

6 Legal, Ethical, and Privacy Considerations

Legal Considerations: In this study, we first obtained legal permission from SIDN, the `.nl` registry, as well as from both institutions in the case studies. We established a data-sharing agreement between SIDN and both the Dutch national government and *Thuiswinkel Waarborg*, so the metadata associated with the detected domains could be shared through the dashboard (Sect. 2.1) for the analysts to evaluate and label the results. These agreements conform to both EU and Dutch [6,20] legislation. Because the `.nl` zonefile is not publicly available, we cannot share the list of domain names analyzed in the presented case studies.

Ethical Considerations: Object detection, which we employ for the detection of logos, is a field in artificial intelligence that raises several ethical, privacy, and legal concerns. We address them by training our model solely to detect the logos of organizations on the screenshots of public `.nl` websites. Therefore, it has no notion of other concepts or objects of which the automatic detection could raise ethical and legal issues, such as persons, faces, or race [21,23,37], and thus cannot be used for these purposes.

The goal of this study was to determine the feasibility of logo detection in detecting scams and trust mark misuse – both ultimately affect real users. By helping to detect and remove such scams, we help protect Internet users by preventing those scams to take place; our use case partners support this claim. We chose these two use cases because they directly affect real Internet users. Finally, `LogoMotive` operates based on the "human in the loop" principle [38]. This means it cannot make autonomous decisions about domain names like removing them from the zone, and always requires human input. We meet this requirement by presenting the results on a dashboard that helps human abuse analysts to assess suspect domain names.

Privacy Considerations: In this study, all data analysis and measurements were conducted by SIDN employees. Only the manual validation and annotation of logo detection results was carried out by analysts at the Dutch national government and *Thuiswinkel Waarborg*, and in that regard, we minimized the data shared with these organizations – restricting it only to screenshots and metadata of the domains on which logos were found. This way of working was approved by our Data Protection Officer. In addition, we have developed a publicly available data privacy framework [6] with our legal department that conforms to both EU and Dutch [6,20] legislation. This framework has been applied for this study and the resulting privacy policy is monitored by a privacy board that oversees SIDN Labs' research.

7 Conclusions and Future Work

Logos are widely used on websites, with both benign and malicious intentions. We proposed `LogoMotive`, a system that detects logos on `.nl`-websites and provides analysts with insights into their logo's (mis)use. `LogoMotive` outperforms existing logo detection methods with high recall values over 97% and its unique flexibility due to our automatic and dynamic training method. Our vantage point as manager of the `.nl` ccTLD zone allowed us to detect logos, and by extension phishing, trust mark misuse, and other malicious logo use, in 6.2M domain names.

We evaluate `LogoMotive` in two use cases. In a use case with the Dutch national government, we detect and annotate 11.6K domain names that display the government logo, In total, we found 6 phishing domains, 82 potential future threats

including dormant phishing attacks, and 80 cases of another misuse of the government logo. We also evaluated `LogoMotive` with a renowned Dutch certified webshop trust mark. Webshops require certification and must comply with strict rules and standards regarding security and consumer protection before they may display the logo on their website. Unauthorized use of the trust mark is, per definition, misleading visitors. `LogoMotive` found close to 10K domain names leading to 3253 unique websites that display the trust mark's logo. 151 of these webshops unjustly displayed the logo and received a cease and desist letter from *Thuiswinkel Waarborg*.

Our work has an operational impact: the government acted upon 168 embodiments of impersonation and 104 trust mark misuse websites removed their logos. It also allowed for the Dutch government to detect and include government websites that were registered outside the official regulations, and, this way, mitigate the risks associated with domain expiration and lack of security standards adoption – namely DNSSEC and DMARC.

`LogoMotive` has proven a useful system and will be deployed in production at SIDN. In future research, we could use the ground truth data resulting from the two use cases in further efforts to automatically detect malicious websites. We intend to combine its results with phishing-tailored detection systems. Furthermore, in the future, we can explore whether existing data sources allow us to prioritize the websites on which a logo is found based on their likeliness to be malicious. Automatic classification or prioritization of the logo detection results would require a use case-specific approach because the definition of misuse varies between logos. We currently focus on the home page of websites, because traversing websites' internal pages in the entire zone is not feasible. In future work, we can explore if it is feasible to traverse internal pages of recently registered domain names, where most abuse is found.

Acknowledgments. We thank very much the manual validation and annotation work carried by the anonymous analysts at the Dutch national government and *Thuiswinkel Waarborg*, for more than 10k domain names. We would also like to thank our colleagues at SIDN for reviewing and indirectly contributing to this study.

SIDN was partly funded by the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 830927 (https://cordis.europa.eu/project/id/830927). Project website: https://www.concordia-h2020.eu/.

A Appendix: LogoMotive Dashboard

(See Fig. 9).

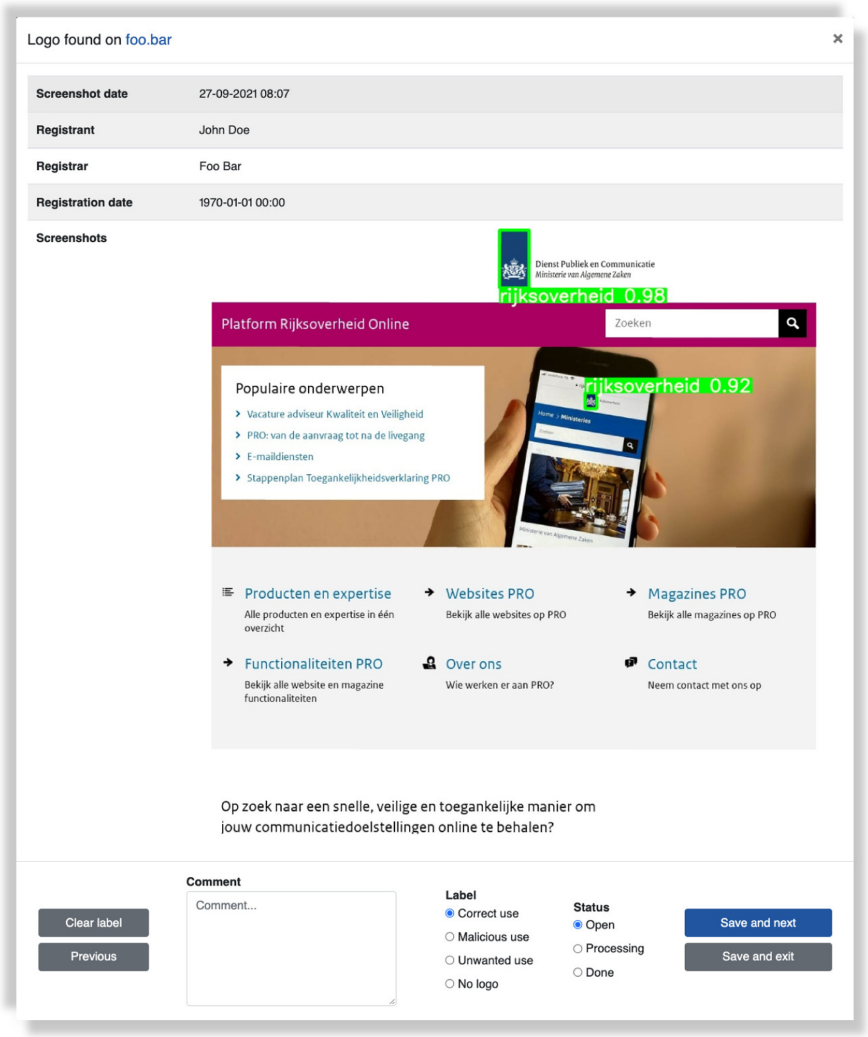

Fig. 9. Dashboard annotation pop-up screen

References

1. Afroz, S., Greenstadt, R.: PhishZoo: detecting phishing websites by looking at them. In: 2011 IEEE Fifth International Conference on Semantic Computing. IEEE, September 2011. https://doi.org/10.1109/icsc.2011.52

2. Arends, R., Austein, R., Larson, M., Massey, D., Rose, S.: DNS Security Introduction and Requirements. RFC 4033, IETF, March 2005. http://tools.ietf.org/rfc/rfc4033.txt
3. Bay, H., Ess, A., Tuytelaars, T., Gool, L.V.: Speeded-up robust features (SURF). Comput. Vis. Image Underst. **110**(3), 346–359 (2008). https://doi.org/10.1016/j.cviu.2007.09.014
4. Bijmans, H., Booij, T., Schwedersky, A., Nedgabat, A., van Wegberg, R.: Catching phishers by their bait: investigating the Dutch phishing landscape through phishing kit detection. In: USENIX Security 2021, pp. 3757–3774. USENIX Association, August 2021
5. Bozkir, A.S., Aydos, M.: LogoSENSE: a companion HOG based logo detection scheme for phishing web page and e-mail brand recognition. Comput. Secur. **95**, 101855 (2020). https://doi.org/10.1016/j.cose.2020.101855
6. Hesselman, C., Jansen, J., Wullink, M., Vink, K., Simon, M.: A privacy framework for DNS big data applications. Technical report, SIDN (2014). https://www.sidnlabs.nl/downloads/yBW6hBoaSZe4m6GJc_0b7w/2211058ab6330c7f3788141ea19d3db7/SIDN_Labs_Privacyraamwerk_Position_Paper_V1.4_ENG.pdf
7. Chang, E.H., Chiew, K.L., Sze, S.N., Tiong, W.K.: Phishing detection via identification of website identity. In: 2013 International Conference on IT Convergence and Security (ICITCS). IEEE, December 2013. https://doi.org/10.1109/icitcs.2013.6717870
8. Chiew, K.L., Chang, E.H., Sze, S.N., Tiong, W.K.: Utilisation of website logo for phishing detection. Comput. Secur. **54**, 16–26 (2015). https://doi.org/10.1016/j.cose.2015.07.006
9. CISA: Sophisticated Spearphishing Campaign Targets Government Organizations, IGOs, and NGOs, May 2021. https://us-cert.cisa.gov/ncas/alerts/aa21-148a
10. Consumentenbond: Keurmerken webwinkels: hoe betrouwbaar zijn ze? (2019). https://www.consumentenbond.nl/online-kopen/keurmerken-webwinkels. Accessed 20 Oct 2021
11. Cui, Q., Jourdan, G.-V., Bochmann, G.V., Onut, I.-V.: Proactive detection of phishing kit traffic. In: Sako, K., Tippenhauer, N.O. (eds.) ACNS 2021. LNCS, vol. 12727, pp. 257–286. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-78375-4_11
12. Eggert, C., Winschel, A., Lienhart, R.: On the benefit of synthetic data for company logo detection. In: Proceedings of the 23rd ACM International Conference on Multimedia. ACM, October 2015. https://doi.org/10.1145/2733373.2806407
13. FBI: FBI Warns Public to Beware of Government Impersonation Scams, April 2021. https://www.fbi.gov/contact-us/field-offices/boston/news/press-releases/fbi-warns-public-to-beware-of-government-impersonation-scams
14. Fielding, R., Reschke, J.: Hypertext Transfer Protocol (HTTP/1.1): Semantics and Content. RFC 7231, IETF, June 2014. http://tools.ietf.org/rfc/rfc7231.txt
15. FTC: How To Avoid a Government Impersonator Scam, April 2021. https://www.consumer.ftc.gov/articles/how-avoid-government-impersonator-scam
16. Goel, R.K.: Masquerading the government: drivers of government impersonation fraud. Public Finan. Rev. **49**(4), 548–572 (2021)
17. Google: Google Public DNS (2021). https://developers.google.com/speed/public-dns/
18. Google Inc.: Certificate transparency. https://certificate.transparency.dev/
19. Han, Y., Shen, Y.: Accurate spear phishing campaign attribution and early detection. In: Proceedings of the 31st Annual ACM Symposium on Applied Computing. ACM, April 2016. https://doi.org/10.1145/2851613.2851801

20. Hesselman, C., Moura, G.C., Schmidt, R.D.O., Toet, C.: Increasing DNS security and stability through a control plane for top-level domain operators. IEEE Commun. Mag. **55**(1), 197–203 (2017). https://doi.org/10.1109/mcom.2017.1600521cm
21. Hill, K.: The Secretive Company That Might End Privacy as We Know It, January 2020. https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html
22. Hoffman, P., Sullivan, A., Fujiwara, K.: DNS Terminology. RFC 8499, IETF, November 2018. http://tools.ietf.org/rfc/rfc8499.txt
23. Introna, L.D.: Disclosive ethics and information technology: disclosing facial recognition systems. Ethics Inf. Technol. **7**(2), 75–86 (2005). https://doi.org/10.1007/s10676-005-4583-2
24. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization (2017)
25. Kucherawy, M., Zwicky, E.: Domain-based Message Authentication, Reporting, and Conformance (DMARC). RFC 7489, IETF, March 2015. http://tools.ietf.org/rfc/rfc7489.txt
26. Lauinger, T., Buyukkayhan, A.S., Chaabane, A., Robertson, W., Kirda, E.: From deletion to re-registration in zero seconds. In: Proceedings of the Internet Measurement Conference 2018. ACM, October 2018. https://doi.org/10.1145/3278532.3278560
27. Le, A., Markopoulou, A., Faloutsos, M.: PhishDef: URL names say it all. In: 2011 Proceedings IEEE INFOCOM. IEEE, April 2011. https://doi.org/10.1109/infcom.2011.5934995
28. Li, Y., Yang, Z., Chen, X., Yuan, H., Liu, W.: A stacking model using URL and HTML features for phishing webpage detection. Futur. Gener. Comput. Syst. **94**, 27–39 (2019). https://doi.org/10.1016/j.future.2018.11.004
29. Lin, Y., et al.: Phishpedia: a hybrid deep learning based approach to visually identify phishing webpages. In: 30th USENIX Security Symposium (USENIX Security 2021) (2021)
30. Liu, W., et al.: SSD: single shot multibox detector. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9905, pp. 21–37. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46448-0_2
31. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004). https://doi.org/10.1023/b:visi.0000029664.99615.94
32. Netcraft Ltd.: Netcraft, 10 October 2021. https://www.netcraft.com/
33. Markt, A.C.: Onderzoek naar de kennis, houding en gedrag van consumenten ten aanzien van keurmerken (2016). https://web.archive.org/web/20180420203000/www.thuiswinkel.org/data/uploads/publication/ACM_en_GfK_onderzoek_keurmerken_2016.pdf. Accessed 20 Oct 2021
34. Mockapetris, P.: Domain names - implementation and specification. RFC 1035, IETF, November 1987. http://tools.ietf.org/rfc/rfc1035.txt
35. Moura, G.C.M., Heidemann, J., Müller, M., de O. Schmidt, R., Davids, M.: When the dike breaks. In: Proceedings of the Internet Measurement Conference 2018. ACM, October 2018. https://doi.org/10.1145/3278532.3278534
36. Moura, G.C.M., Heidemann, J., de O. Schmidt, R., Hardaker, W.: Cache me if you can. In: Proceedings of the Internet Measurement Conference. ACM, October 2019. https://doi.org/10.1145/3355369.3355568
37. Mozurl, P.: One Month, 500,000 Face Scans: How China Is Using A.I. to Profile a Minority, April 2019. https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html

38. Munro, R.: Human-in-the-Loop Machine Learning. Manning Publications, New York, October 2021
39. Nguyen, L.A.T., To, B.L., Nguyen, H.K., Nguyen, M.H.: A novel approach for phishing detection using URL-based heuristic. In: 2014 International Conference on Computing, Management and Telecommunications (ComManTel), pp. 298–303. IEEE (2014)
40. Nieuws, R.: Politiegeheimen op straat door verlopen mailadressen (2017). https://www.rtlnieuws.nl/nieuws/nederland/artikel/240411/politiegeheimen-op-straat-door-verlopen-mailadressen. Accessed 15 Oct 2021
41. Nieuws, R.: Groot datalek bij jeugdzorg: dossiers duizenden kwetsbare kinderen gelekt (2019). https://www.rtlnieuws.nl/tech/artikel/4672826/jeugdzorg-datalek-dossiers-kinderen-utrecht-email. Accessed 15 Oct 2021
42. Oest, A., Safei, Y., Doupe, A., Ahn, G.J., Wardman, B., Warner, G.: Inside a phisher's mind: understanding the anti-phishing ecosystem through phishing kit analysis. In: 2018 APWG Symposium on Electronic Crime Research (eCrime). IEEE, May 2018. https://doi.org/10.1109/ecrime.2018.8376206
43. Opara, C., Wei, B., Chen, Y.: HTMLPhish: enabling phishing web page detection by applying deep learning techniques on HTML analysis. In: 2020 International Joint Conference on Neural Networks (IJCNN). IEEE, July 2020. https://doi.org/10.1109/ijcnn48605.2020.9207707
44. Quan, L., Heidemann, J., Pradkin, Y.: When the internet sleeps. In: Proceedings of the 2014 Conference on Internet Measurement Conference. ACM, November 2014. https://doi.org/10.1145/2663716.2663721
45. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, June 2016. https://doi.org/10.1109/cvpr.2016.91
46. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. **39**(6), 1137–1149 (2017). https://doi.org/10.1109/tpami.2016.2577031
47. van Riel, C.B., van den Ban, A.: The added value of corporate logos - an empirical study. Eur. J. Mark. **35**(3/4), 428–440 (2001). https://doi.org/10.1108/03090560110382093
48. Roopak, S., Thomas, T.: A novel phishing page detection mechanism using HTML source code comparison and cosine similarity. In: 2014 Fourth International Conference on Advances in Computing and Communications. IEEE, August 2014. https://doi.org/10.1109/icacc.2014.47
49. Rublee, E., Rabaud, V., Konolige, K., Bradski, G.: ORB: an efficient alternative to SIFT or SURF. In: 2011 International Conference on Computer Vision. IEEE, November 2011. https://doi.org/10.1109/iccv.2011.6126544
50. Sahingoz, O.K., Buber, E., Demir, O., Diri, B.: Machine learning based phishing detection from URLs. Expert Syst. Appl. **117**, 345–357 (2019)
51. Sanchez, S.A., Romero, H.J., Morales, A.D.: A review: comparison of performance metrics of pretrained models for object detection using the TensorFlow framework. In: IOP Conference Series: Materials Science and Engineering, vol. 844, p. 012024, June 2020. https://doi.org/10.1088/1757-899x/844/1/012024
52. Shao, S., et al.: Objects365: a large-scale, high-quality dataset for object detection. In: 2019 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, October 2019. https://doi.org/10.1109/iccv.2019.00852
53. Software Freedom Conservancy: Selenium hub. https://hub.docker.com/r/selenium/hub/tags

54. Srivastava, S., Divekar, A.V., Anilkumar, C., Naik, I., Kulkarni, V., Pattabiraman, V.: Comparative analysis of deep learning image detection algorithms. J. Big Data **8**(1), 1–27 (2021). https://doi.org/10.1186/s40537-021-00434-w
55. Stringhini, G., Thonnard, O.: That ain't you: blocking spearphishing through behavioral modelling. In: Almgren, M., Gulisano, V., Maggi, F. (eds.) DIMVA 2015. LNCS, vol. 9148, pp. 78–97. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-20550-2_5
56. Su, H., Zhu, X., Gong, S.: Deep learning logo detection with data expansion by synthesising context. In: 2017 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, March 2017. https://doi.org/10.1109/wacv.2017.65
57. Ultralytics: Yolov5. https://github.com/ultralytics/yolov5
58. Wabeke, T., Moura, G.C.M., Franken, N., Hesselman, C.: Counterfighting counterfeit: detecting and taking down fraudulent webshops at a ccTLD. In: Sperotto, A., Dainotti, A., Stiller, B. (eds.) PAM 2020. LNCS, vol. 12048, pp. 158–174. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-44081-7_10
59. Wang, D.Y., et al.: Search + seizure. In: Proceedings of the 2014 Conference on Internet Measurement Conference. ACM, November 2014. https://doi.org/10.1145/2663716.2663738
60. Wang, G., et al.: Verilogo: proactive phishing detection via logo recognition. Department of Computer Science and Engineering, University of California (2011)
61. Wilson, J.M., Grammich, C.A.: Brand protection across the enterprise: toward a total-business solution. Bus. Horiz. **63**(3), 363–376 (2020). https://doi.org/10.1016/j.bushor.2020.02.002
62. Wullink, M., Moura, G.C.M., Hesselman, C.: DMAP: automating domain name ecosystem measurements and applications. In: 2018 Network Traffic Measurement and Analysis Conference (TMA). IEEE, June 2018. https://doi.org/10.23919/tma.2018.8506521
63. Wullink, M., Moura, G.C.M., Muller, M., Hesselman, C.: ENTRADA: a high-performance network traffic data streaming warehouse. In: NOMS 2016–2016 IEEE/IFIP Network Operations and Management Symposium. IEEE, April 2016. https://doi.org/10.1109/noms.2016.7502925
64. Yao, W., Ding, Y., Li, X.: Deep learning for phishing detection. In: ISPA/IUCC/BDCloud/SocialCom/SustainCom. IEEE, December 2018. https://doi.org/10.1109/bdcloud.2018.00099
65. Yao, W., Ding, Y., Li, X.: LogoPhish: a new two-dimensional code phishing attack detection method. In: ISPA/IUCC/BDCloud/SocialCom/SustainCom. IEEE, December 2018. https://doi.org/10.1109/bdcloud.2018.00045
66. Zhou, Y., Zhang, Y., Xiao, J., Wang, Y., Lin, W.: Visual similarity based anti-phishing with the combination of local and global features. In: 2014 IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications, pp. 189–196. IEEE (2014)