Joint Research: Phishing on .nl, .be, and .ie (+ academia)

Giovane C. M. Moura

SIDN Labs and TU Delft

2024-11-20
CENTR GA 72
Brussels, Belgium

Outline

$whoami

- Data Scientist at SIDN Labs
- Assistant Professor at TU Delft
- Research interests:
  - Intersection between operations and academia
- Active in both industry and academia

Research output example: RFC9199

Stream:	Independent Submission	
RFC:	9199	
Category:	Informational	
Published:	March 2022	
ISSN:	2070-1721	
Authors:	G. Moura *SIDN Labs/TU Delft*	W. Hardaker *USC/Information Sciences Institute*
	J. Heidemann *USC/Information Sciences Institute*	M. Davids *SIDN Labs*

RFC 9199
Considerations for Large Authoritative DNS Server Operators

- Academia (USC/ISI and TU Delft) and Industry Collaboration (SIDN Labs)
- 6 academic papers

Joint-study on phishing: .nl, .ie, .be and academia

Peer-reviewed paper, top security conference (10% accpt. rate)

ACM CCS 2024, Salt Lake City, USA

Characterizing and Mitigating Phishing Attacks at ccTLD Scale

Giovane C. M. Moura
SIDN Labs
Arnhem, The Netherlands
Delft University of Technology
Delft, The Netherlands

Thomas Daniels
DNS Belgium
Leuven, Belgium
KU Leuven
Department of Computer Science
Leuven, Belgium

Maarten Bosteels
DNS Belgium
Leuven, Belgium

Sebastian Castro
.IE Registry
Dublin, Ireland

Moritz Müller
SIDN Labs
Arnhem, The Netherlands
University of Twente
Enschede, The Netherlands

Thymen Wabeke
SIDN Labs
Arnhem, The Netherlands

Thijs van den Hout
SIDN Labs
Arnhem, The Netherlands

Maciej Korczyński
University of Grenoble Alps
Grenoble, France

Georgios Smaragdakis
Delft University of Technology
Delft, The Netherlands

Paper (PDF)

Collaboration Outcomes

1. What did we find?
2. How did we do it?
3. TLDs and Academia collaboration
4. How have we been profiting from it?

Outline

Phishing at three ccTLDs

1. First time 3 ccTLDs come together to analyze phishing:
   - The Netherlands' `.nl` (SIDN)
   - Ireland's `.ie` (.IE Registry)
   - Belgium's `.be` (DNS Belgium)
2. Longitudinal study (10 years)

Improving the state-of-the-art:

	Previous Works	Ours
Time	1 year	4–10 years
Companies	10	1233
Domains	1.4k	28.7k

Phishing at three ccTLDs

1. First time 3 ccTLDs come together to analyze phishing:
   - The Netherlands' `.nl` (SIDN)
   - Ireland's `.ie` (.IE Registry)
   - Belgium's `.be` (DNS Belgium)
2. Longitudinal study (10 years)

Improving the state-of-the-art:

	Previous Works	**Ours**
Time	1 year	4–10 years
Companies	10	1233
Domains	1.4k	28.7k

ccTLDs compared

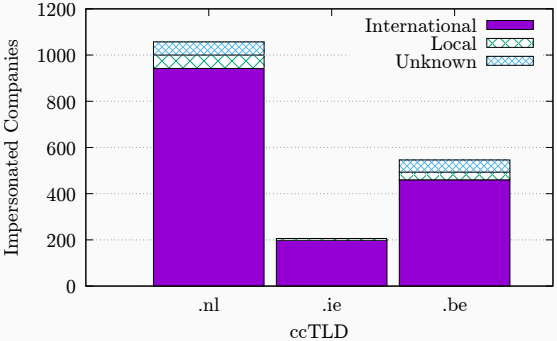
ccTLD	.nl	.ie	.be
# Domains	6.1M	330.1k	1.7M
Reg. Policy	Open	Restricted	Open
Country Population	17.5M	4.9M	11.5M

Table 1: ccTLDs overview.

- Restricted registration: check Irish ID, passport, or business in Ireland
- Open registration ( ): anyone can register a domain

Do they target mostly national companies?

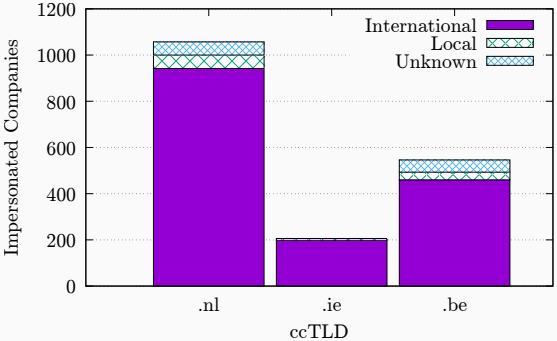
- Citizens have trust in their ccTLDs
  - Govs use it
- Do attackers exploit this trust for phishing?



- Most impersonated companies are **International**
- So most attackers **do not seem to care which TLD they use**.
  - Is it really so?

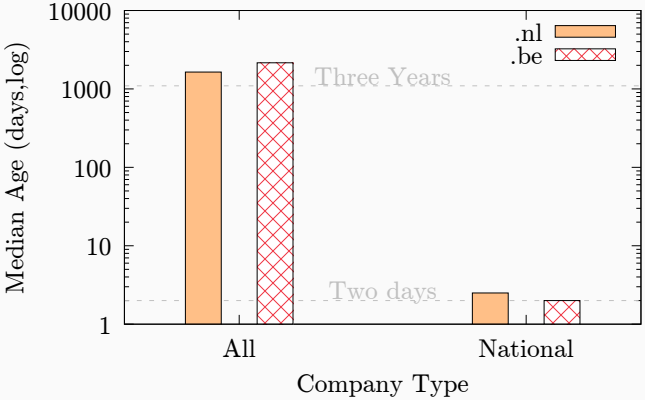
Do they target mostly national companies?

- Citizens have trust in their ccTLDs
  - Govs use it
- Do attackers exploit this trust for phishing?



- Most impersonated companies are **International**
- So most attackers **do not seem to care which TLD they use**.
  - Is it really so?

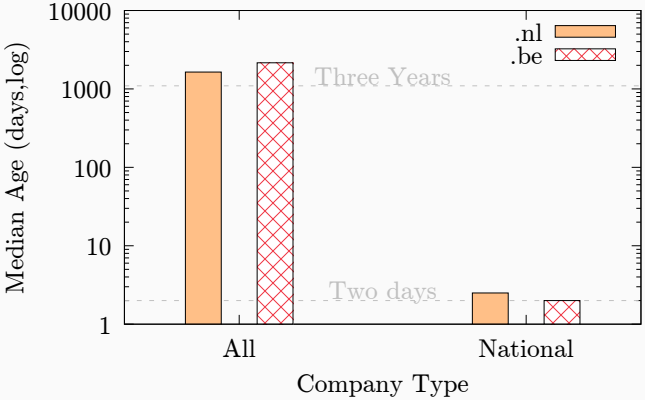
National companies vs international companies



We see a pattern:

1. **International** companies impersonated with old domains
2. **National** companies impersonated with new domains

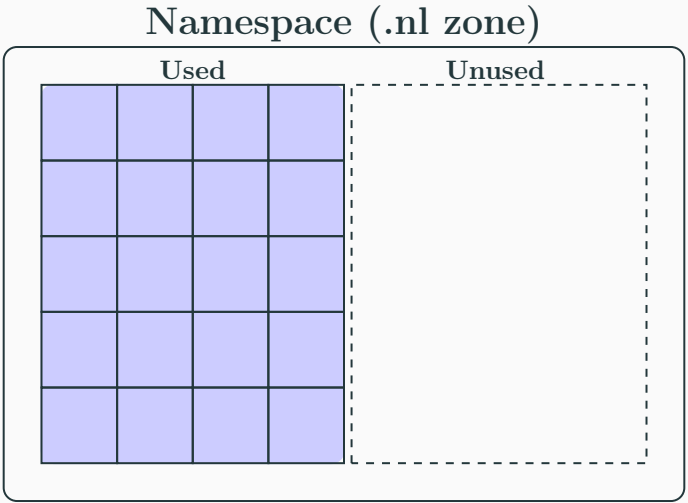
National companies vs international companies



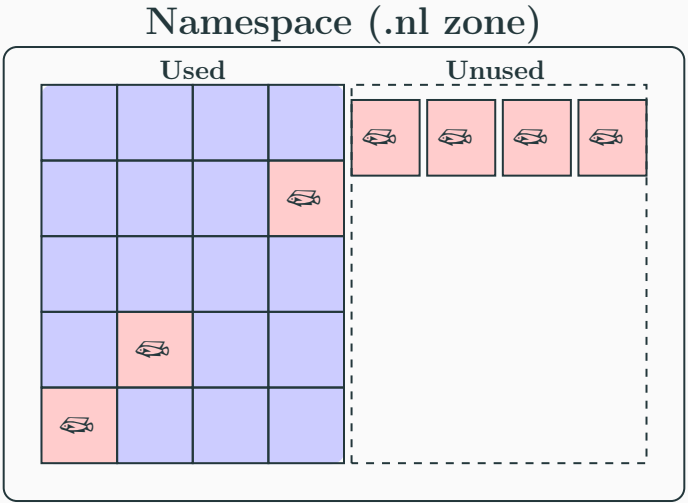
We see a pattern:

1. **International** companies impersonated with old domains
2. **National** companies impersonated with new domains

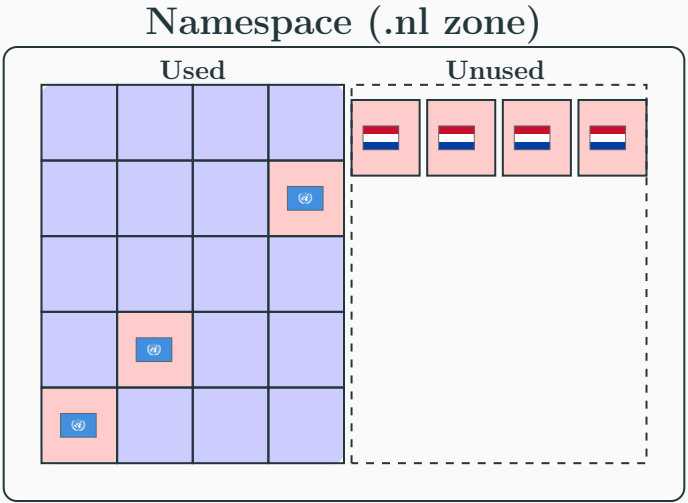
Two attack strategies



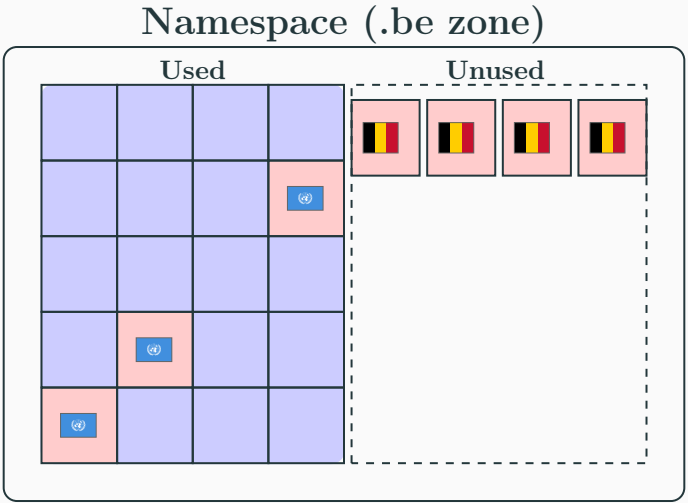
Two attack strategies



Two attack strategies



Same for .be



Top 10 impersonated companies (`.nl` zone)

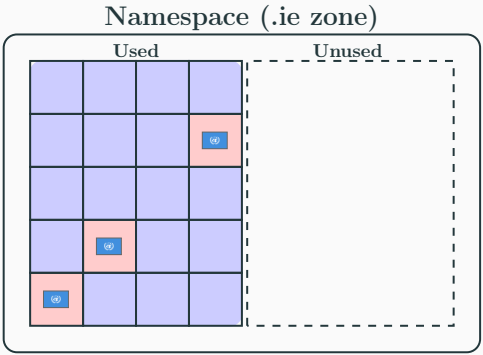
Rank	Company	Domains	Median Age (days)
1	Microsoft	2,319	2,251
2	PayPal	2,134	1,751
3	ING	1,815	1
4	ICS	1,410	2
5	Apple	1,276	1,775
6	ABN AMRO	1,259	1
7	Google	1,236	1,416
8	Rabobank	1,222	1
9	Webmail Users	1,054	2,247
10	Netflix	756	1,653

Top 10 impersonated companies in phishing attacks on the `.nl` zone ( ).

But what about Ireland?

Only two new phishing domains

- `.ie` = restricted registration policy
- Restricted policy prevents part of the phishing attacks
  - But cannot prevent compromised domain names



Implications of this finding

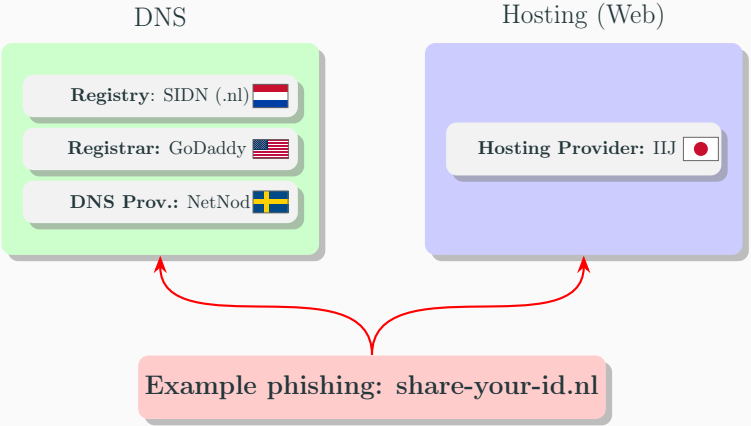
1. Most phishing research focus on new domains
    - call for action to investigate compromised domains
2. Policy: restricted registration is effective against malicious new domain names
    - but most phishing is from compromised
3. Following research:
    - why make these websites vulnerable?
    - what is the role of hosting providers and registrars?
    - can we identify patterns to try to remediate it?
    - what about other abuse types, as malware?

Finding 2: Impact of mitigation policies

- Phishing mitigation ***is not*** a single event
- Different parties can mitigate it independently
  - registrant (example.nl) → Registrar (GoDaddy) → Registry (SIDN)

Finding 2: Impact of mitigation policies

- Phishing mitigation ***is not*** a single event
- Different parties can mitigate it independently
  - registrant (example.nl) → Registrar (GoDaddy) → Registry (SIDN)



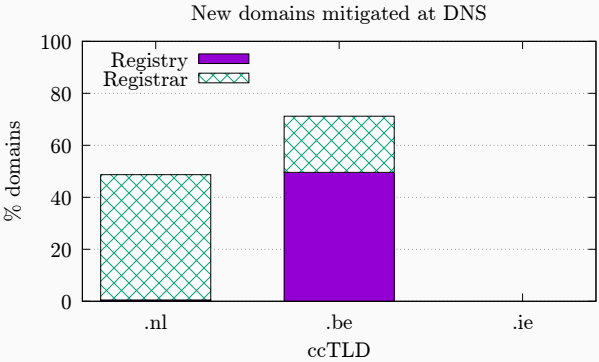
ccTLD Mitigation Policy

- ccTLDs can perform 3 operations at the DNS level
- Each of them have its own policy (§B in [4])

	.nl	.ie	.be
Suspend domain	✓ After 66h	✓ After 30 days	✓ASAP
Delete domain	✓	✓After two weeks	✓
Change NS records	–	–	✓

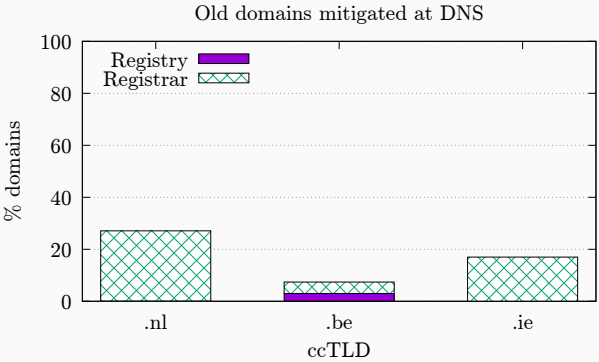
Table 2: ccTLDs phishing detection and mitigation procedure.

DNS mitigation and ccTLD policy: new domains



- `.be` suspend new domains ASAP
- `.nl` notifies registrars, hosting who take action
- Rest is mitigated at Web level

Phishing Mitigation at DNS: Old Domains



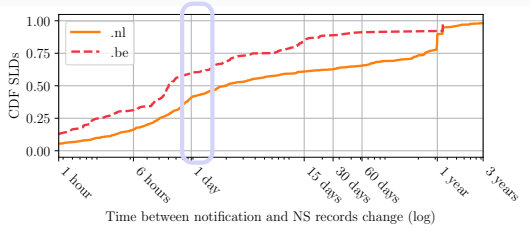
- Most old domains are compromised
  - Web mitigation is preferred
- Exceptions: aged domains

DNS vs Web Mitigation speed

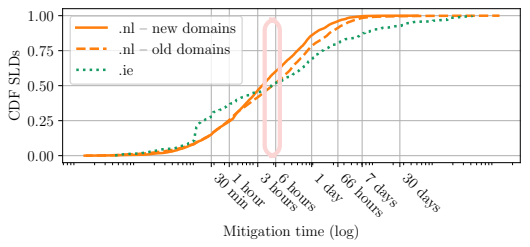
Web mitigation is faster than DNS mitigation

DNS: 50–60% first 24h

Web: 50–60% first 6h



(a) DNS mitigation: Domain suspension



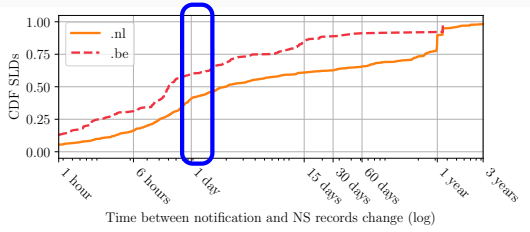
(b) Web mitigation

DNS vs Web Mitigation speed

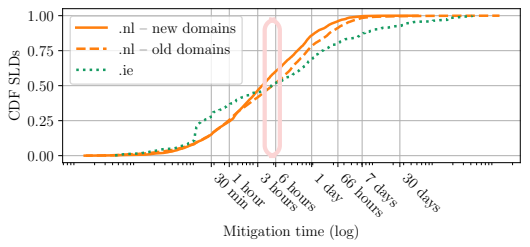
Web mitigation is faster than DNS mitigation

DNS: 50–60% first 24h

Web: 50–60% first 6h



(c) DNS mitigation: Domain suspension



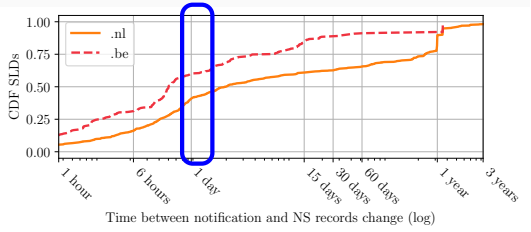
(d) Web mitigation

DNS vs Web Mitigation speed

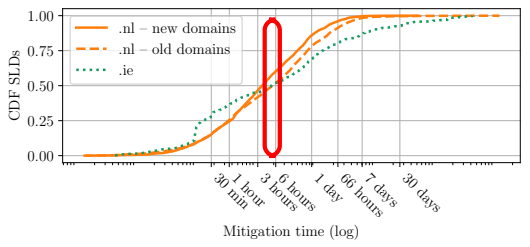
Web mitigation is faster than DNS mitigation

DNS: 50–60% first 24h

Web: 50–60% first 6h



(e) DNS mitigation: Domain suspension



(f) Web mitigation

Implications of this finding

1. Phishing mitigation is a multi-party process
    - DNS provider, registrars, registries, hosting, upstream
2. Web mitigation (both .nl and .ie) is faster than DNS mitigation
    - but most phishing is from compromised domains
3. Follow-up research:
    - how can we reduce uptimes?

Outline

How did we do it?

- Started as a project with TU Delft
- Then we invited .be and .ie:
  - we knew them from previous collaborations
  - we need to compare results with other TLDs
- We set up an information collaboration:
  - Same goals
  - No contracts
  - No NDAs
  - No redtape
- It became an Academia/Industry collaboration

How did we do it?

- Datasets were never shared
  - Only aggregated results and figs
- Each registry run the same code locally
- Most issues resolved on `gitlab`
  - few calls (3?)
- We are planning a second study with more registries
  - Please consider joining!

main / phishing-lifecycle / paper / code / netcraft / netcraft_with_sql.ipynb

Find file | Blame | History | Permalink

adding all row
Giovane Moura authored 8 months ago

f2ecca4c

netcraft_with_sql.ipynb 232.89 KiB

Blame | Edit | Replace | Delete

In [226]:

```
tld = "ie"
anonimsed_tld = 'anonEU2'

netcraft_src = "ie/Takedown_export_2023-08-25_17_10_52.csv.gz"
dest_folder = "ie"

# We assume you have a parquet file (~/data/new_registrations.parquet) that contains your registration history

# Ideally the file should have these two columns:
# * domainname_ascii  : it's just the domainname if you don't do IDN (else it should be the a_label)
# * reg_date_utc      : the timestamp of the registration in UTC

# If you have a CSV file instead it's probably enough to change 'parquet' into 'csv' in the next statement.

registration_history_source = 'ie-registration-info.parquet'

# If your file has different columns, you can change them here:

registration_history_domain_name_col = 'Domain'
registration_history_reg_date_col = 'Registration'

# Set open_generated_pdfs to True if you want to see the pdf's after they are generated
open_generated_pdfs = True
```

Outline

TLDs and Academia collaboration

	TLDs	Academia
1	Reduce dark data	Gain access to private data (indirectly)
2	Scrutinize registration and mitigation policies	Advance the state of the art
3	Compare with other TLDs	Address real-world problems
4	Access academic networks	Connect with domain experts and industry networks
5	Visibility, reputation boost, and contribute to the community	

Visibility so far:

- Presentations: ACM CCS 2024, CENTR Tech (FRA), RIPE 89, DNS-OARC
- Blog posts: RIPE, SIDN Labs, APNIC, TU Delft

TLDs and Academia collaboration

	TLDs	Academia
1	Reduce dark data	Gain access to private data (indirectly)
2	Scrutinize registration and mitigation policies	Advance the state of the art
3	Compare with other TLDs	Address real-world problems
4	Access academic networks	Connect with domain experts and industry networks
5	Visibility, reputation boost, and contribute to the community	

Visibility so far:

- Presentations: ACM CCS 2024, CENTR Tech (FRA), RIPE 89, DNS-OARC
- Blog posts: RIPE, SIDN Labs, APNIC, TU Delft

TLDs and Academia collaboration

	TLDs	Academia
1	Reduce dark data	Gain access to private data (indirectly)
2	Scrutinize registration and mitigation policies	Advance the state of the art
3	Compare with other TLDs	Address real-world problems
4	Access academic networks	Connect with domain experts and industry networks
5	Visibility, reputation boost, and contribute to the community	

Visibility so far:

- Presentations: ACM CCS 2024, CENTR Tech (FRA), RIPE 89, DNS-OARC
- Blog posts: RIPE, SIDN Labs, APNIC, TU Delft

TLDs and Academia collaboration

	TLDs	Academia
1	Reduce dark data	Gain access to private data (indirectly)
2	Scrutinize registration and mitigation policies	Advance the state of the art
3	Compare with other TLDs	Address real-world problems
4	Access academic networks	Connect with domain experts and industry networks
5	Visibility, reputation boost, and contribute to the community	

Visibility so far:

- Presentations: ACM CCS 2024, CENTR Tech (FRA), RIPE 89, DNS-OARC
- Blog posts: RIPE, SIDN Labs, APNIC, TU Delft

TLDs and Academia collaboration

	TLDs	Academia
1	Reduce dark data	Gain access to private data (indirectly)
2	Scrutinize registration and mitigation policies	Advance the state of the art
3	Compare with other TLDs	Address real-world problems
4	Access academic networks	Connect with domain experts and industry networks
5	Visibility, reputation boost, and contribute to the community	

Visibility so far:

- Presentations: ACM CCS 2024, CENTR Tech (FRA), RIPE 89, DNS-OARC
- Blog posts: RIPE, SIDN Labs, APNIC, TU Delft

TLDs and Academia collaboration

	TLDs	Academia
1	Reduce dark data	Gain access to private data (indirectly)
2	Scrutinize registration and mitigation policies	Advance the state of the art
3	Compare with other TLDs	Address real-world problems
4	Access academic networks	Connect with domain experts and industry networks
5	Visibility, reputation boost, and contribute to the community	

Visibility so far:

- Presentations: ACM CCS 2024, CENTR Tech (FRA), RIPE 89, DNS-OARC
- Blog posts: RIPE, SIDN Labs, APNIC, TU Delft

Since C* folks are in the room...

Should you start a research team in your TLD?

- It pays off
- It requires board support
- It requires research mindset
  - Academic mindset helps
  - gold standard: original Bell Labs
- Academic and industry collaboration are key

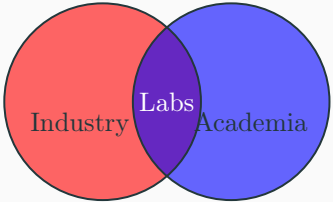


Figure 1: SIDN Labs research positioning.

We are working on a paper about it

SME-Academia Open Research Collaboration Models: a case study

Cristian Hesselman
SIDN Labs and University of Twente
The Netherlands
cristian.hesselman@sidn.nl

Giovane C. M. Moura
SIDN Labs and TU Delft
The Netherlands
giovane.moura@sidn.nl

Summary

Three EU ccTLDs on the largest phishing characterization study

1. Two main attacker types:
   - National companies → new domains
   - Intl' → old, compromised domains
2. Policy impact on mitigation:
   - `.ie`'s restricted registration prevents new phishing domains
   - `.be` registry does most of DNS mitigation.
   - `.nl`'s registrars do most of DNS mitigation
3. Academia and Industry Collaboration pays off

Real phishing victims in the Netherlands go on the record

Source: NOS.nl

References i

[1] US Federal Bureau of Investigation, Internet Crime Complaint Center.
Internet Crimer Report.
https://www.ic3.gov/Media/PDF/AnnualReport/2023_IC3Report.pdf, 2023.

[2] European Union Agency for Cybersecurity.
ENISA Threat Landscape 2023.
https://www.enisa.europa.eu/publications/enisa-threat-landscape-2023, 2023.

References ii

[3] European Union Agency for Cybersecurity.
Malware, Phishing, and Ransomware.
https://www.enisa.europa.eu/publications/enisa-threat-landscape-2023, 2024.

[4] Giovane C. M. Moura, Thomas Daniels, Maarten Bosteels, Sebastian Castro, Moritz Müller, Thymen Wabeke, Thijs van den Hout, Maciej Korczyński, and G. Smaragdakis.
***Characterizing and Mitigating Phishing Attacks at ccTLD Scale (extended)*, volume EWI-TR-2024-1.**

References iii

Delft University of Technology, Faculteit Elektrotechniek, Wiskunde en Informatica, 2024.