



Deleting Your Domain?

Preventing Data Leaks at TLD Scale



Who Are We?



Maarten Wullink
Research engineer
SIDN Labs



Moritz Müller
Research engineer
SIDN Labs

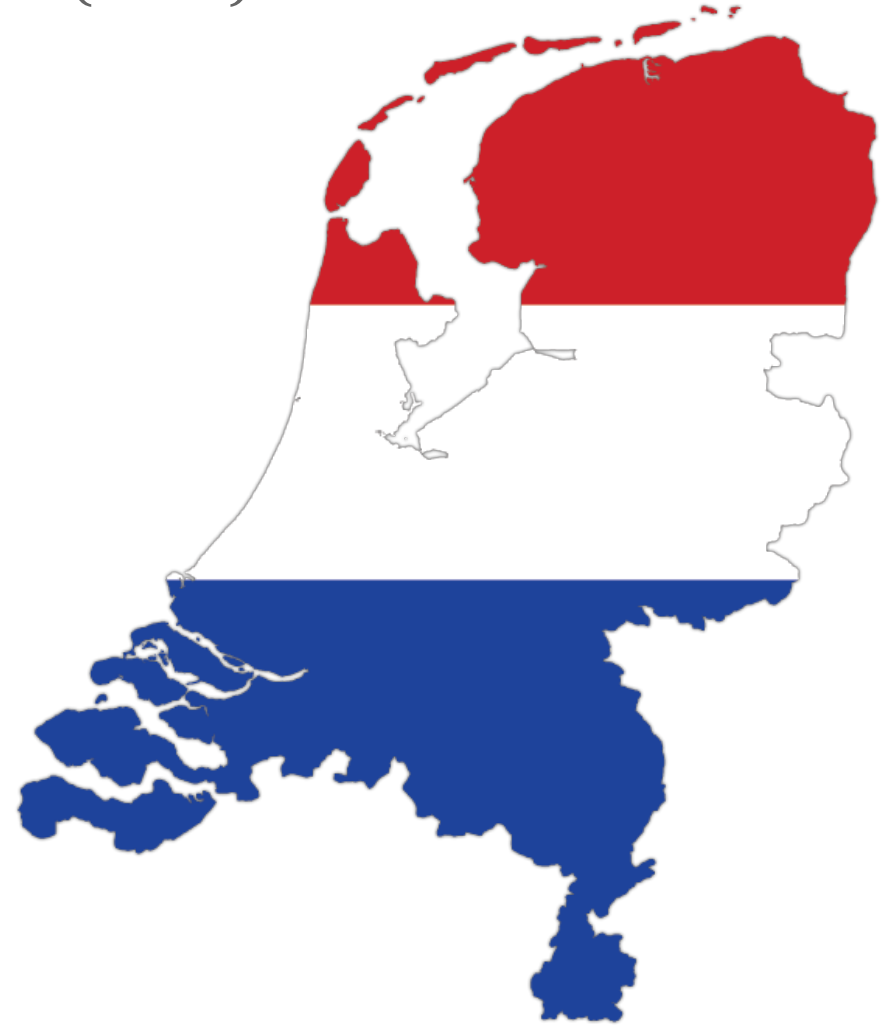
What to Expect?

- Introduction
- Data leaks and email
- LEMMINGS system
- DNS and email
- System functionality
- Results

About SIDN

Registry for the **.nl** country code top-level domain (cctld)

- **6.3** million **.nl** domains
- **61%** use DNSSEC
- Global Anycast DNS network



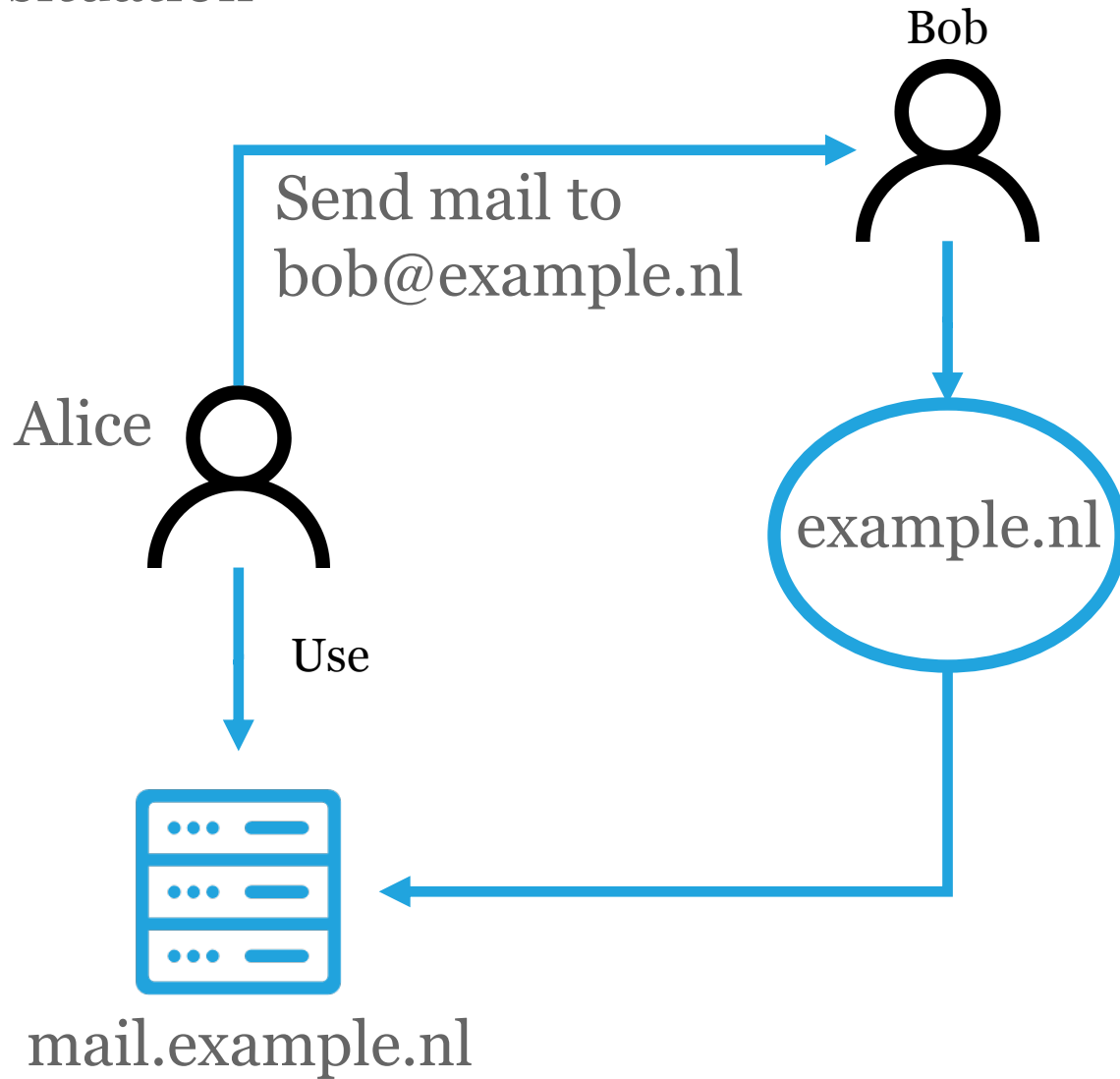
About SIDN Labs

Research arm of SIDN

- Applied technical research into the safety and stability of the Internet
- Main research themes
 - Domain name security
 - Infrastructure security
 - Emerging Internet technologies

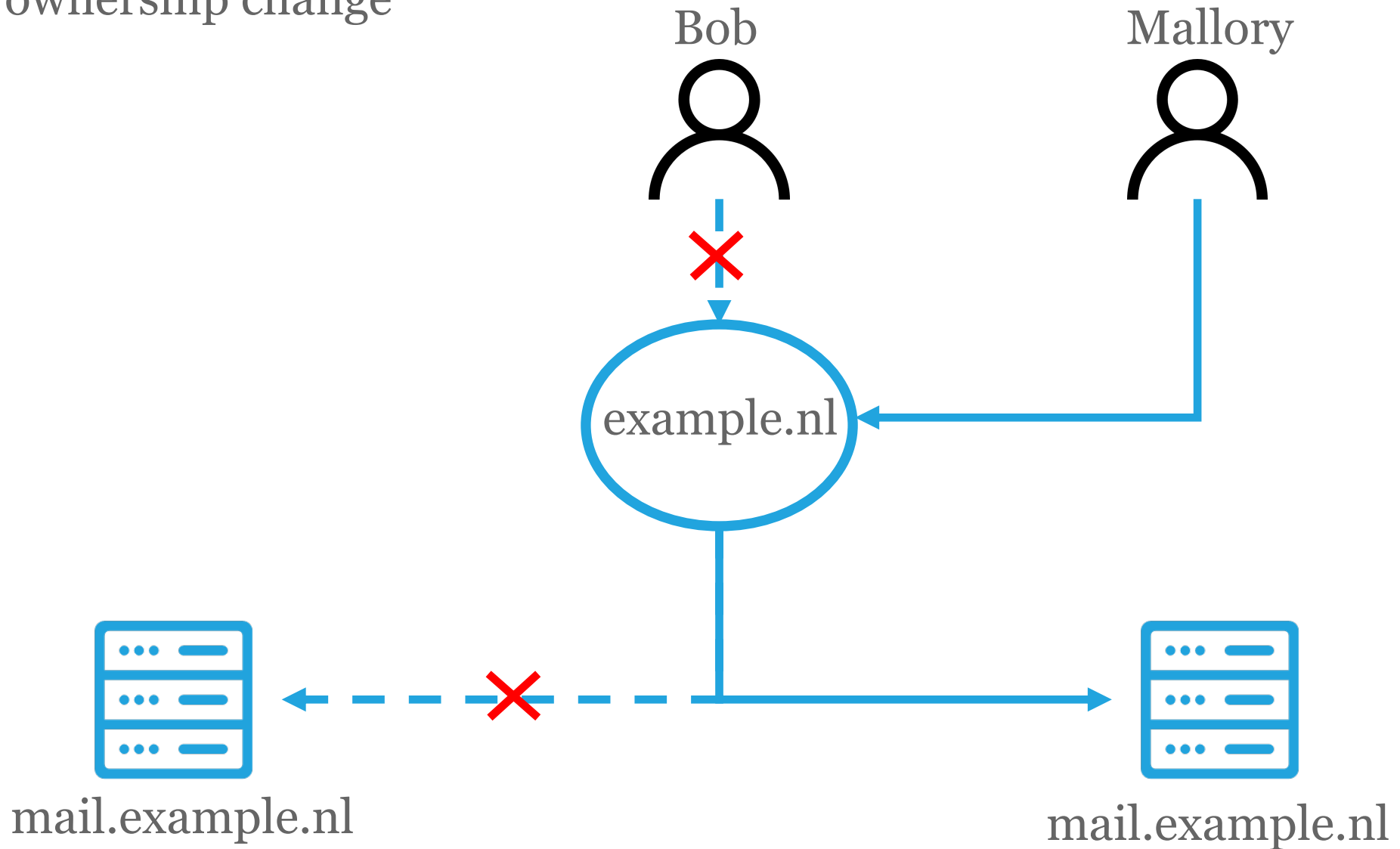
Data Leaks and Email

Normal situation



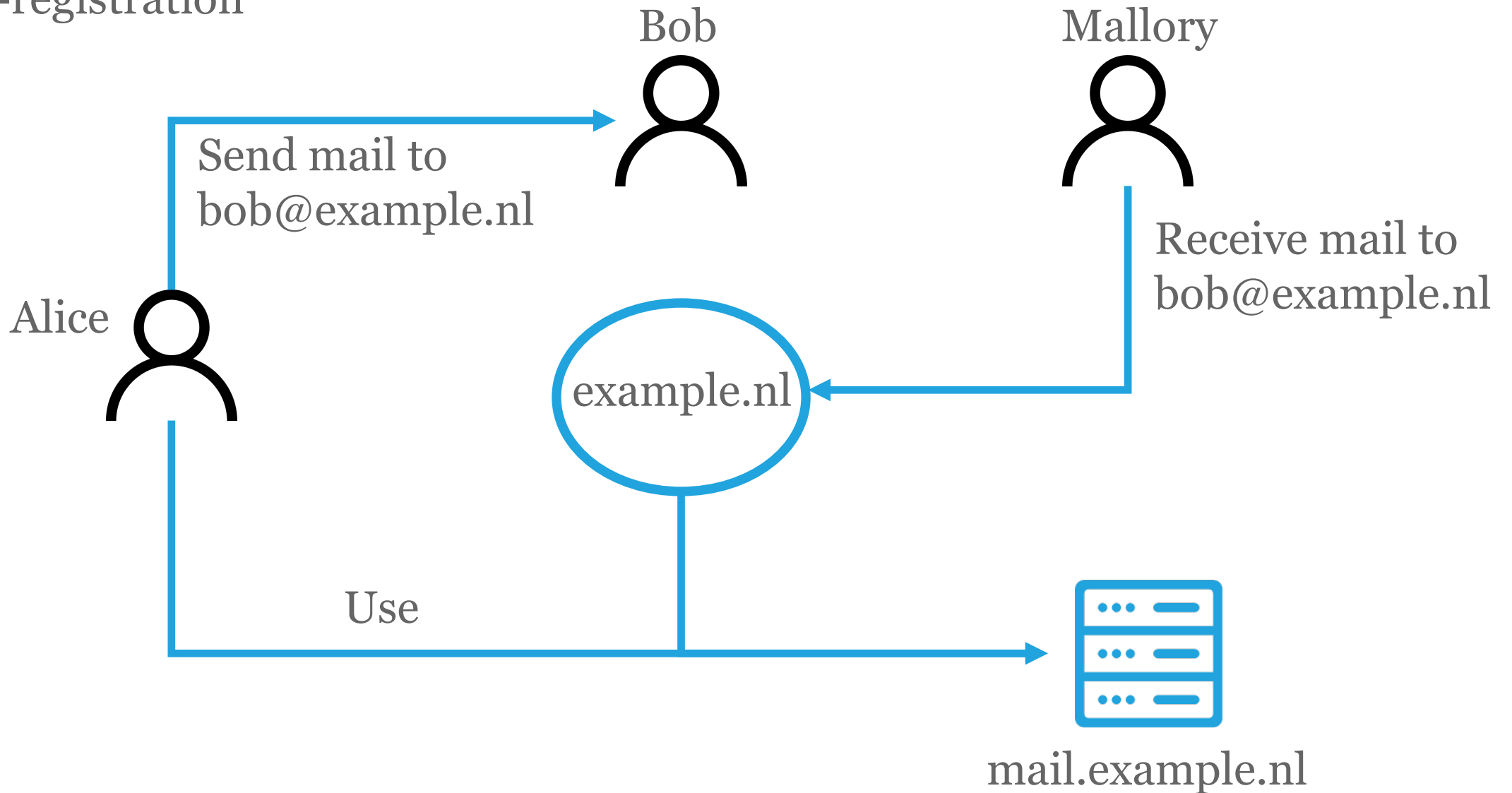
Data Leaks and Email

Domain ownership change



Data Leaks and Email

After re-registration



Data Leaks - Example

EXCLUSIVE

Major data breach at Jeugdriagg: medical records of vulnerable children leaked



By means of Daniel Verlaan

October 1, 2020 12:56 PM · Modified October 1, 2020 1:45 PM



Just in

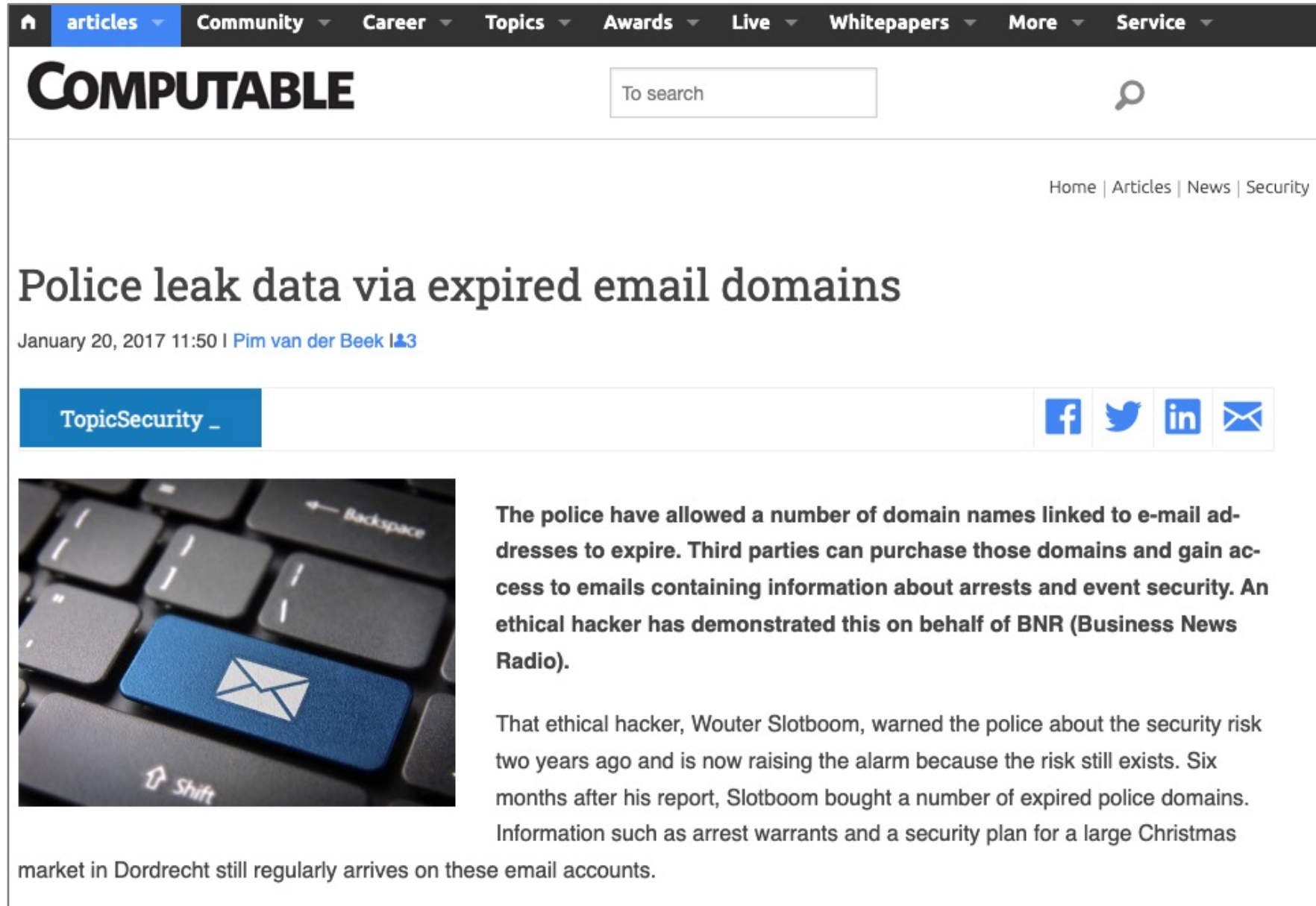
- 08:47 Anderlec to go to P
- 08:41 Greek tra ranking a
- 08:19 Eloise er with Prid
- 08:17 Barbie is director t dollars
- 08:11 Lost Rob surgery t



Due to an error at Jeugdriagg, the files of children with often serious psychological problems have been leaked. Despite efforts by Minister Hugo de Jonge to better secure healthcare institutions, hardly anything seems to have changed in a year and a half.



Data Leaks - Example



The screenshot shows the top navigation bar of the 'COMPUTABLE' website with links for articles, Community, Career, Topics, Awards, Live, Whitepapers, More, and Service. Below the navigation is the site logo and a search bar. The article title is 'Police leak data via expired email domains', dated January 20, 2017, by Pim van der Beek. A blue tag 'TopicSecurity _' is visible. Social media sharing icons for Facebook, Twitter, LinkedIn, and Email are present. The main text describes how the police allowed expired email domains to be purchased by third parties, leading to data leaks. An ethical hacker, Wouter Slotboom, is mentioned as having warned the police about this security risk.


Home | Articles | News | Security

Police leak data via expired email domains

January 20, 2017 11:50 | Pim van der Beek 13

TopicSecurity _

[f](#) [t](#) [in](#) [✉](#)



The police have allowed a number of domain names linked to e-mail addresses to expire. Third parties can purchase those domains and gain access to emails containing information about arrests and event security. An ethical hacker has demonstrated this on behalf of BNR (Business News Radio).

That ethical hacker, Wouter Slotboom, warned the police about the security risk two years ago and is now raising the alarm because the risk still exists. Six months after his report, Slotboom bought a number of expired police domains. Information such as arrest warrants and a security plan for a large Christmas market in Dordrecht still regularly arrives on these email accounts.

LEMMINGS

DeLetEd doMain MaIl warNinG System

A system for detecting mail traffic
involving deleted domains and alerting
registrants



LEMMINGS – Get a bad Rep

Wait, LEMMINGS are those little suicidal maniacs, right?

NO, this is a MYTH!



This is what really happened!



LEMMINGS

- Method:
 - Analyse DNS queries for all deleted domains
 - Combine with web crawler and domain abuse data
- Alert the former registrant, when following is true
 - Indication domain is used for email
 - Domain has not yet exited quarantine period



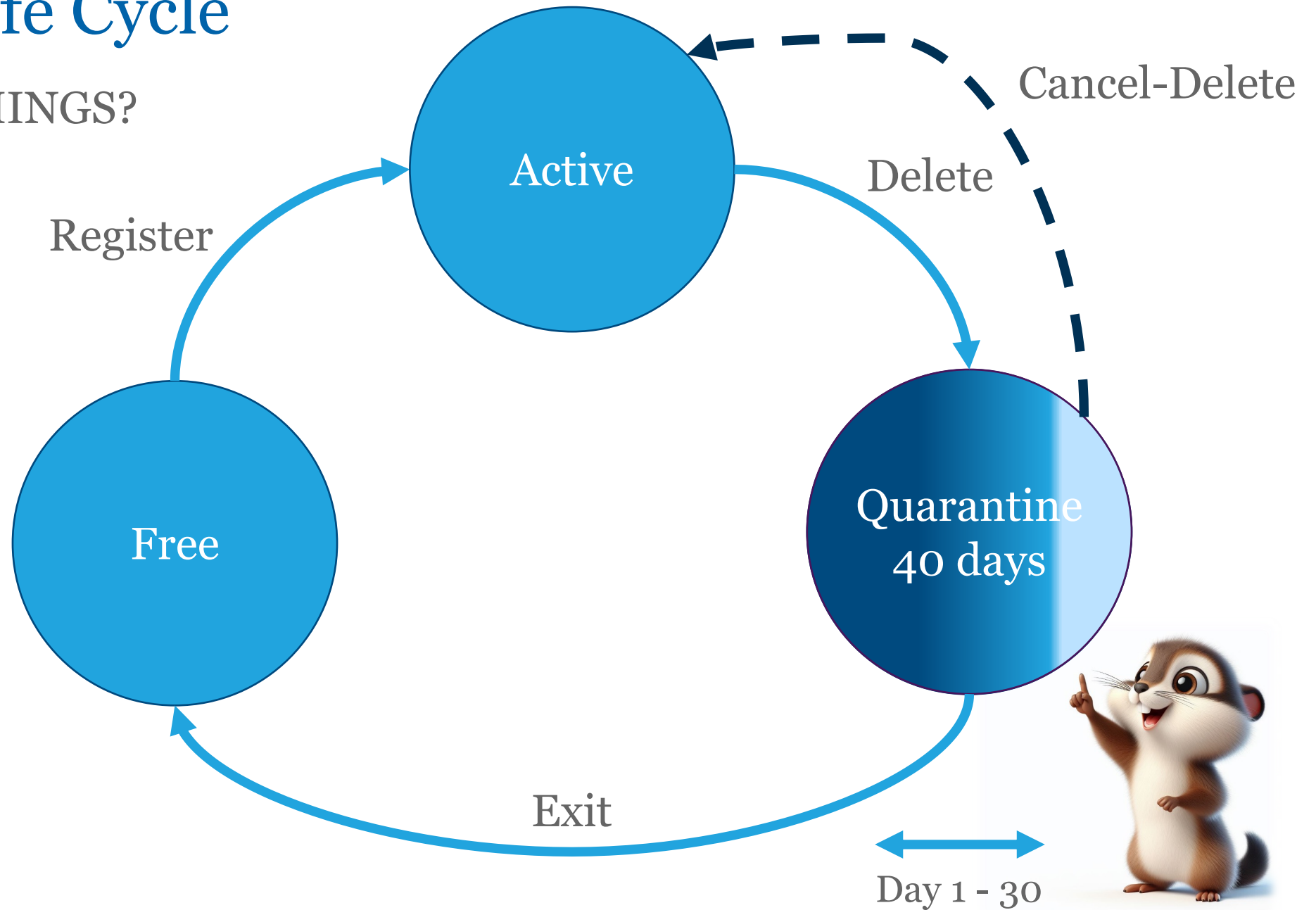
Privacy Considerations

- Not capturing or analysing actual mail content
- No trackers in mail alert to registrants
- Removing PII data after process is completed
- Published privacy policy



Domain Life Cycle

Where is LEMMINGS?



DNS and Email

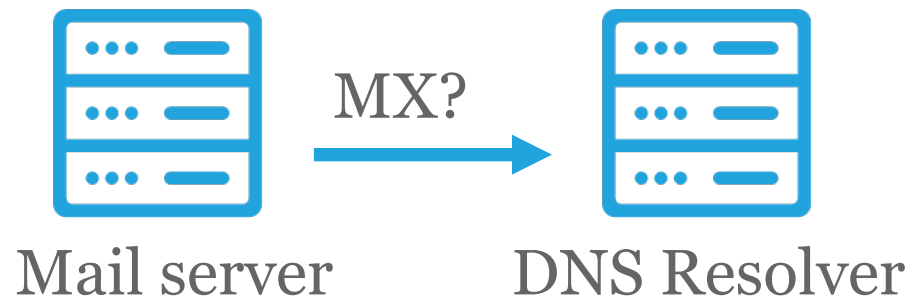
Delivering mail for bob@example.nl



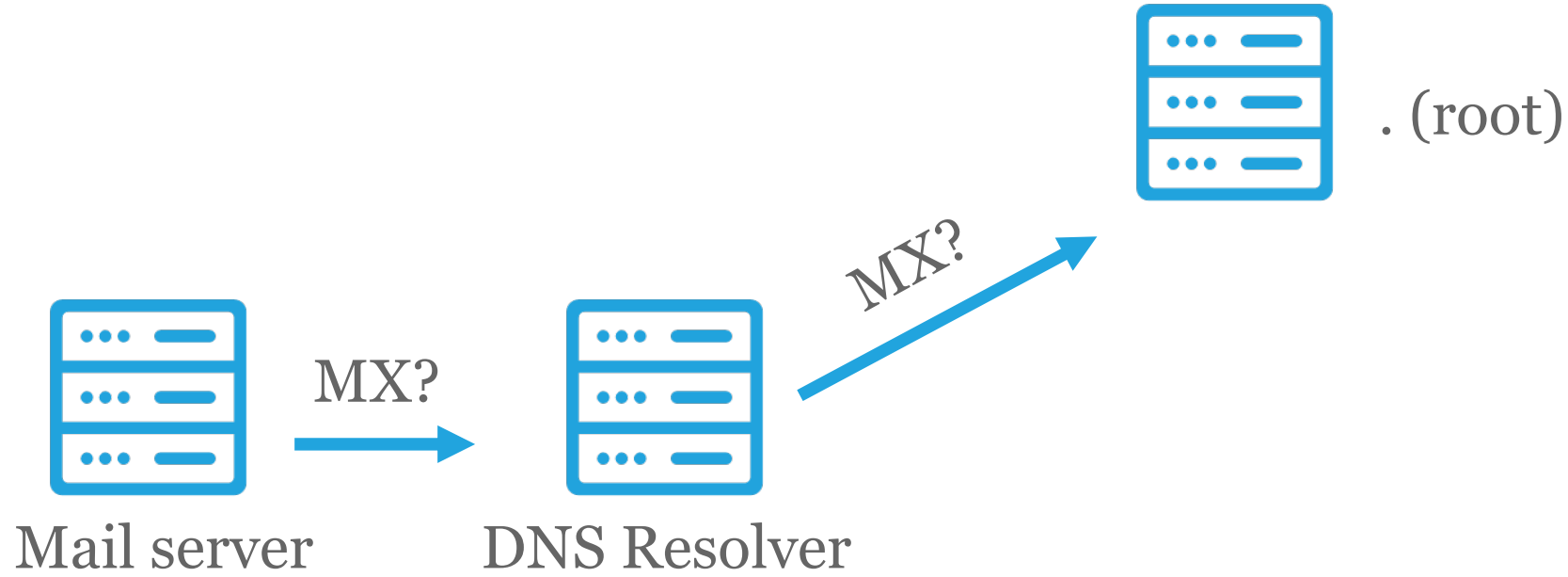
Mail server



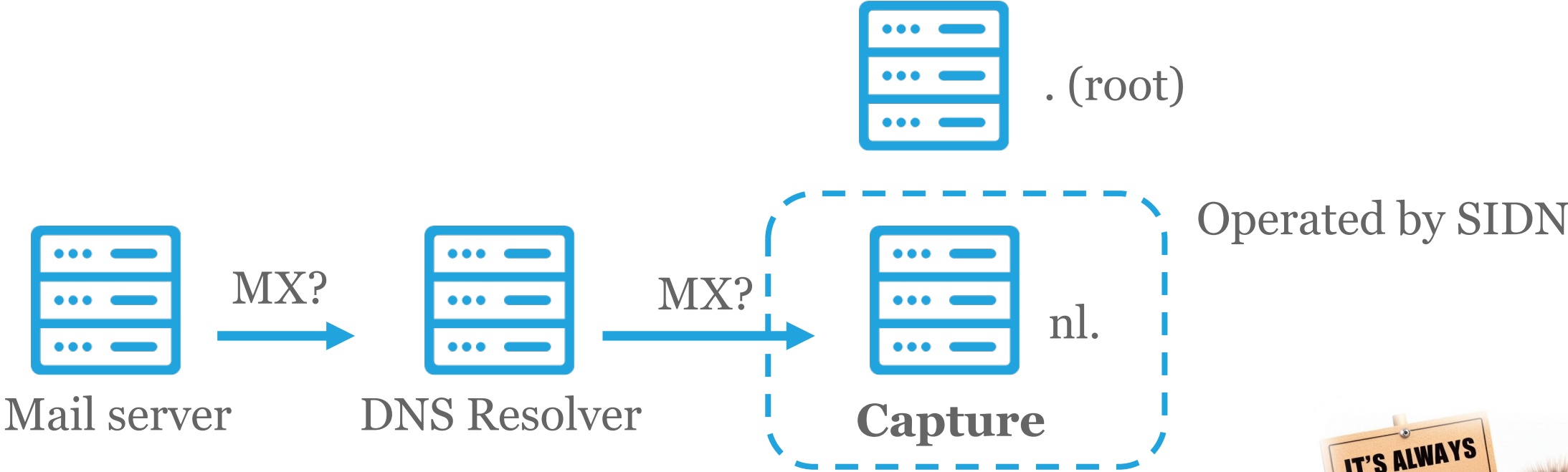
DNS and Email



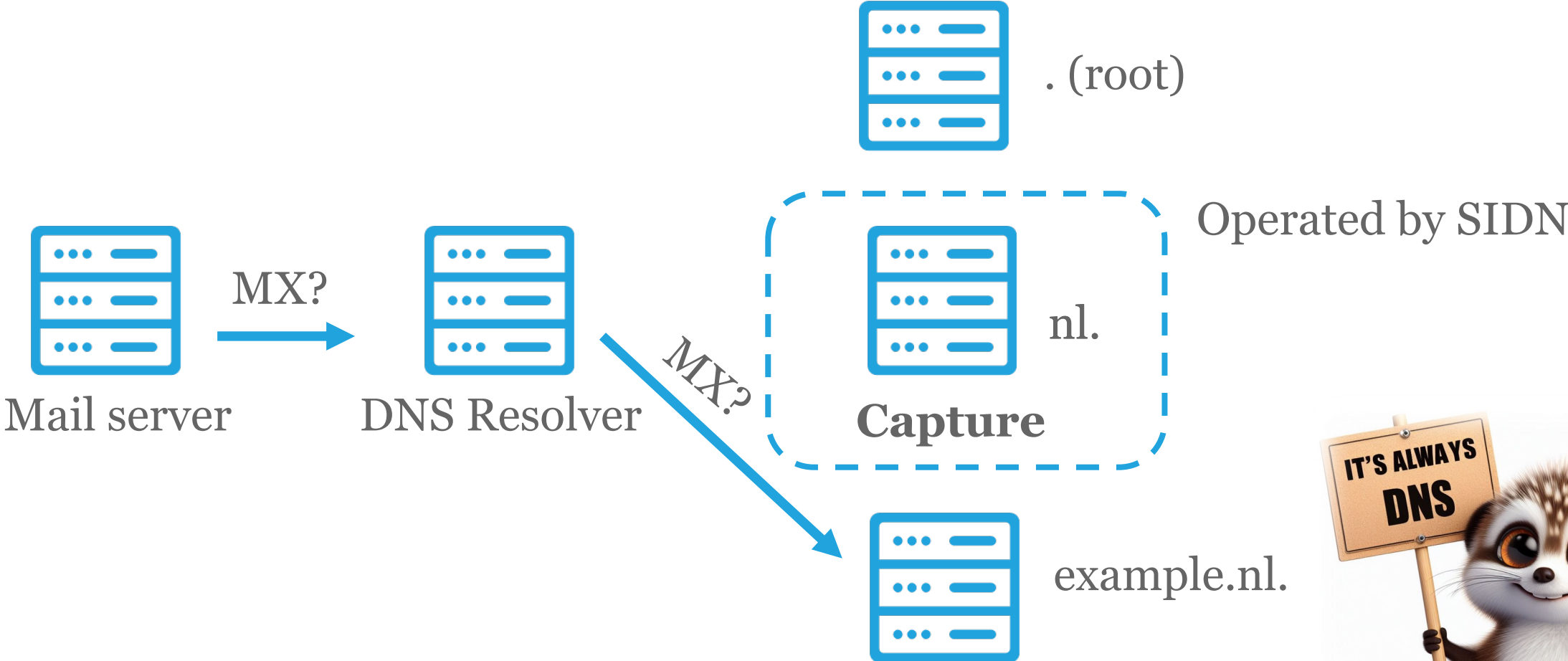
DNS and Email



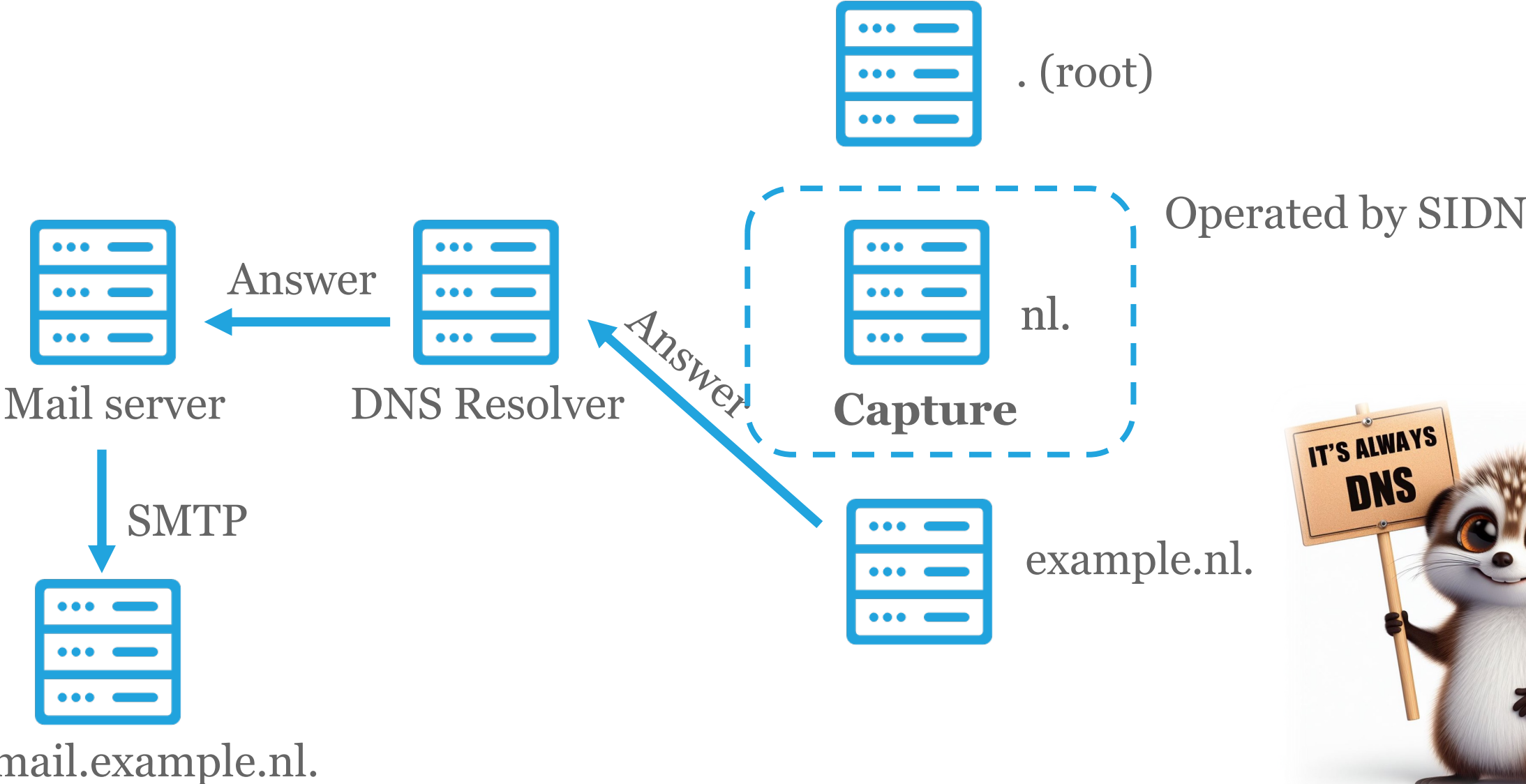
DNS and Email



DNS and Email



DNS and Email



Challenges

- Analysing large volume of DNS data
- Filtering noise (marketing, spam mail)
- Explaining the security risk to registrants



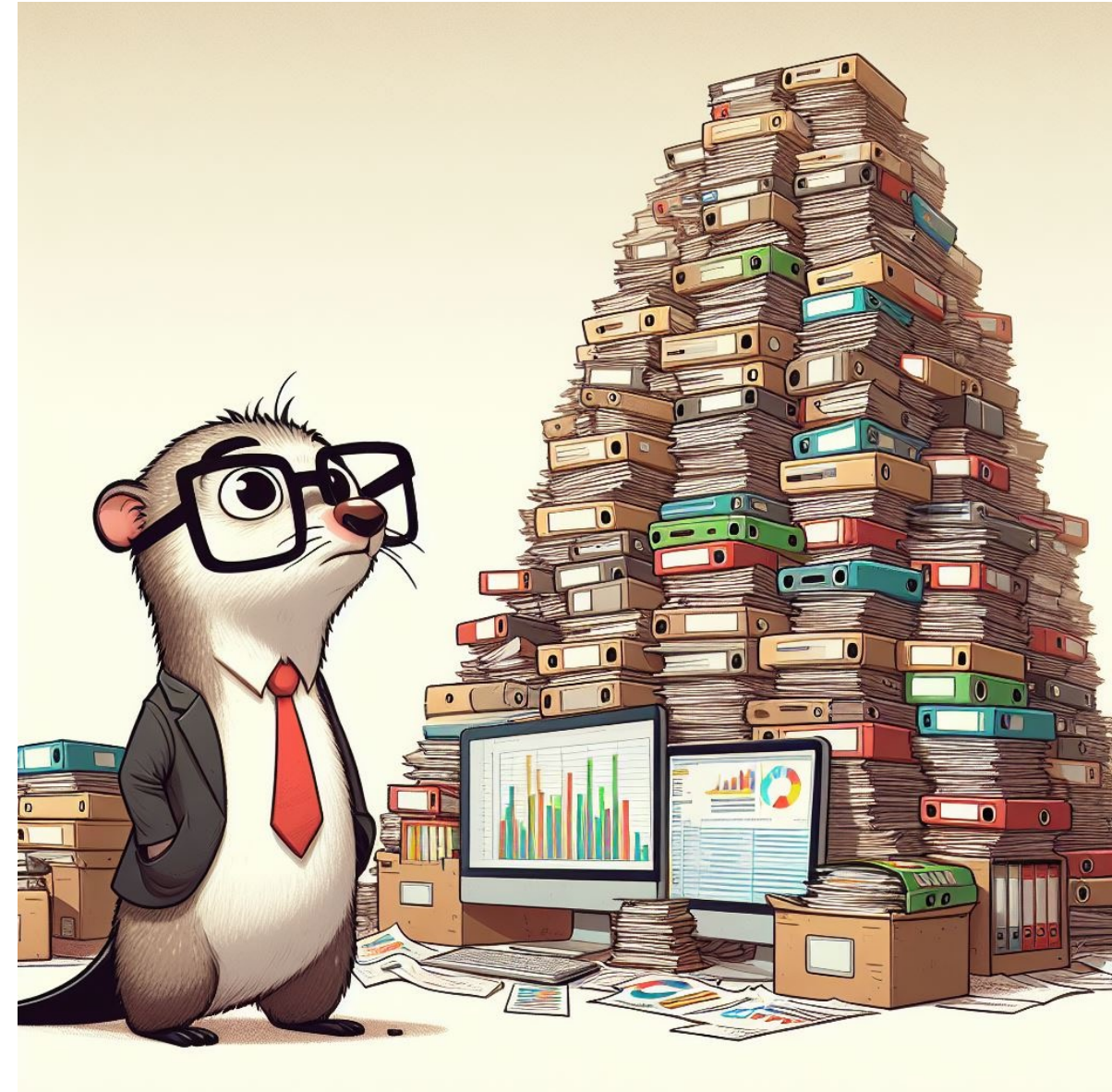
Challenges

- Try not to make the alert look like a spam message
- Registrant contact email address is not reachable
 - Privacy proxy
 - In-zone email address



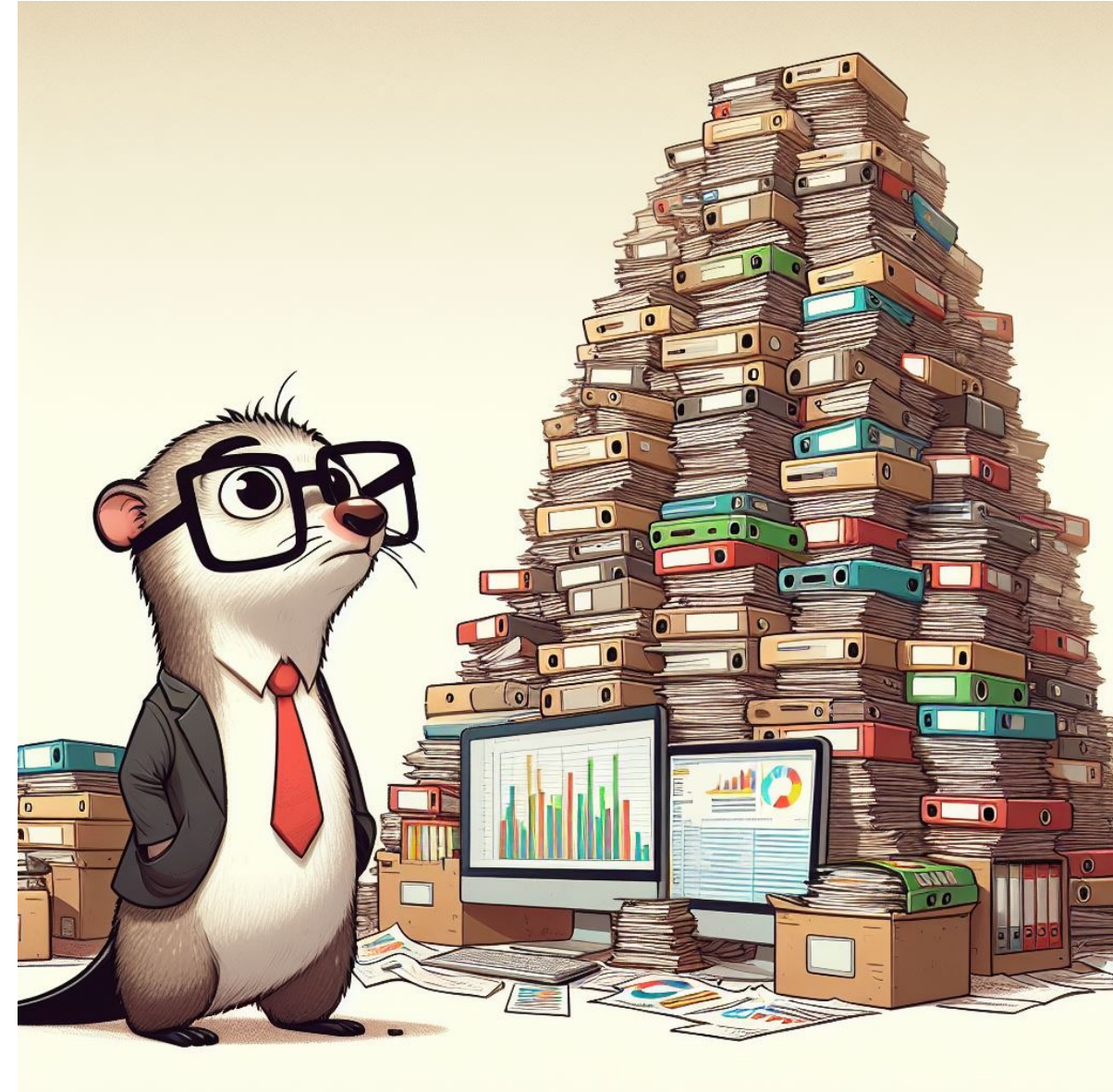
Data Sources

- DNS queries for 6.3 million .nl domains
 - ~4 billion daily
 - ~180 million email related



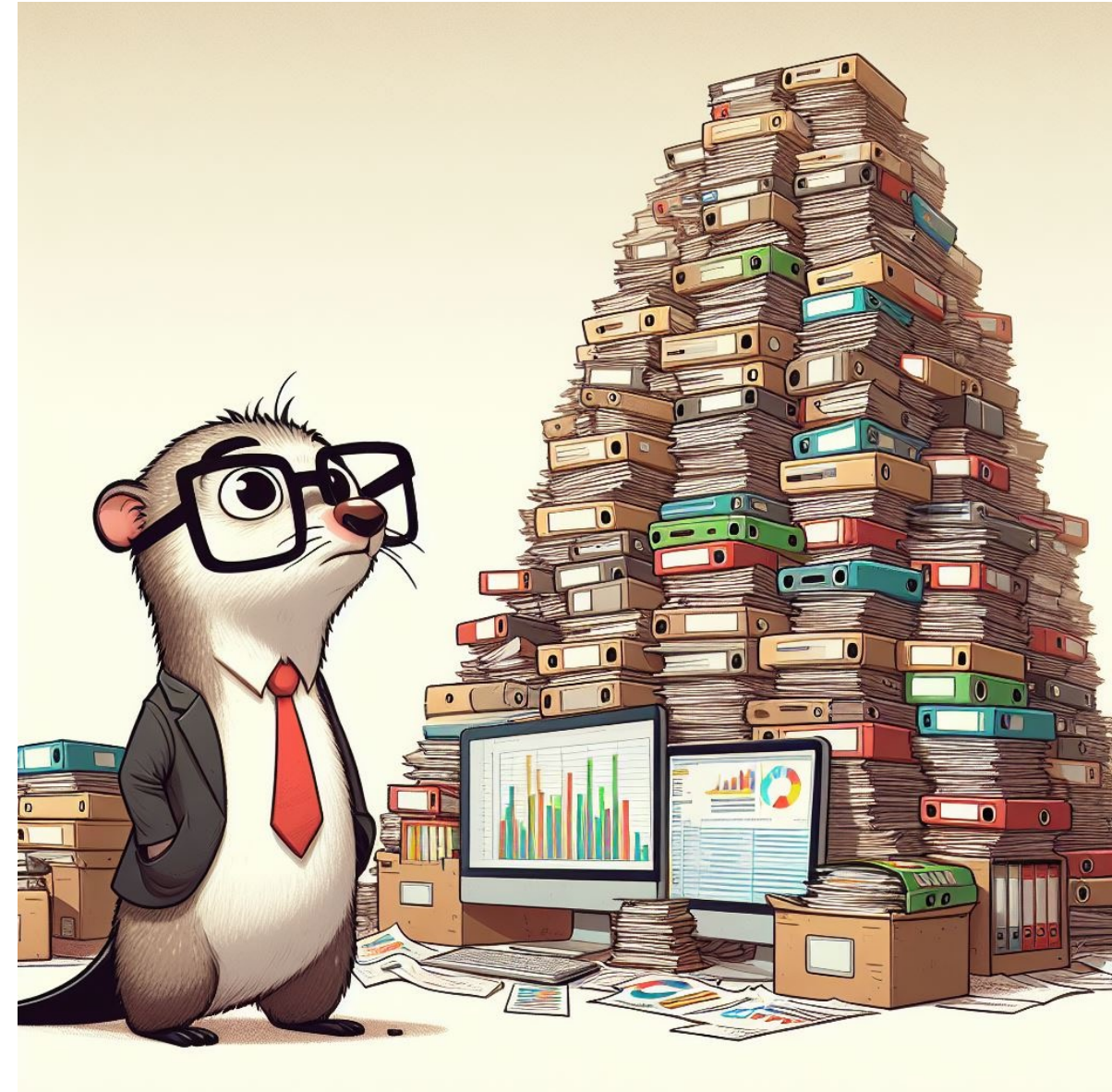
Data Sources

- Web crawler data
 - 6.3 million .nl domains
- Attributes
 - Web content-type
 - NACE code
 - Email addresses



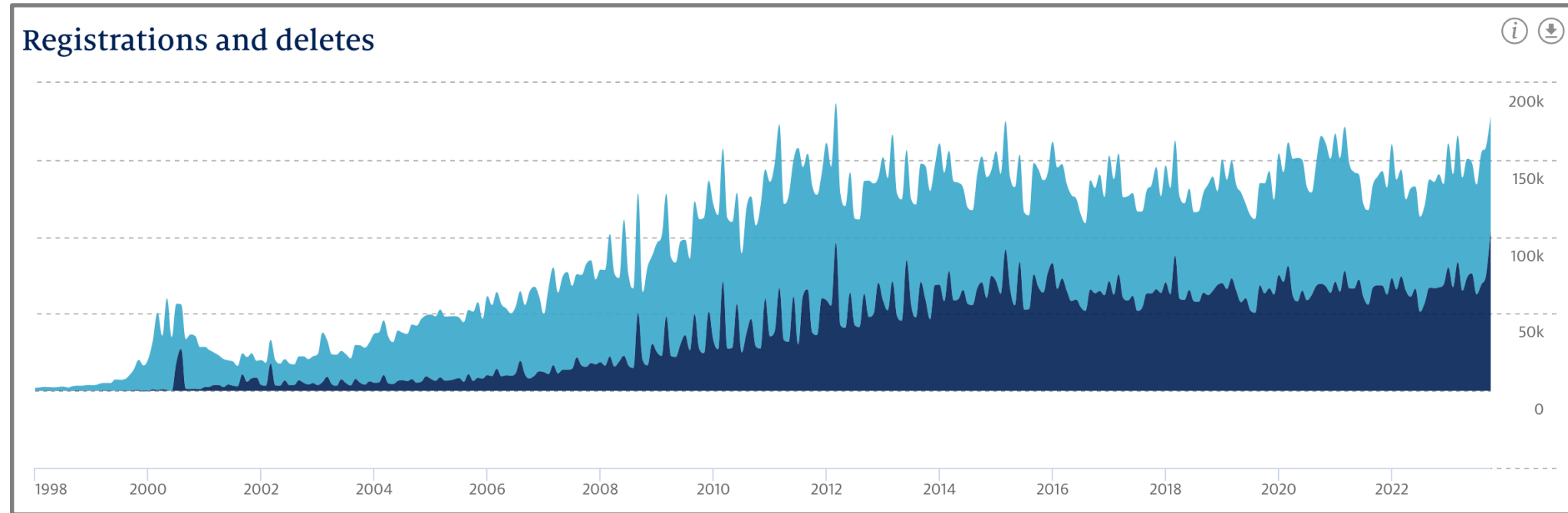
Data Sources

- Abuse feeds (Spamhaus, APWG)
 - SPAM senders
- Sinkhole (botnet C&C domains)
 - Botnet client IPs



Data Sources

Domain registration database



Registrations (light blue) and deletes (dark blue)

<https://stats.sidnlabs.nl>



Data Platform



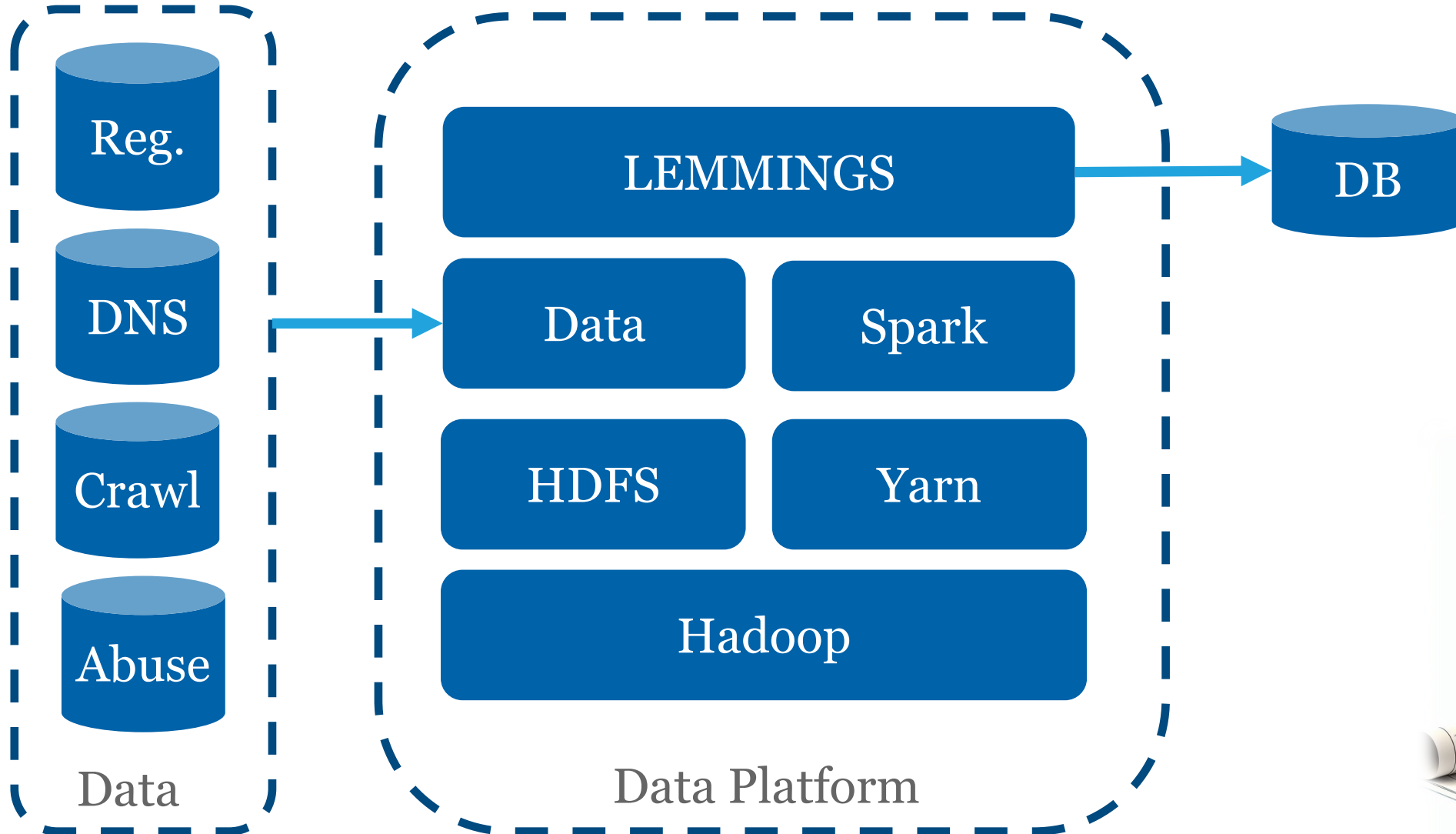
SEE
MORE



<https://entrada.sidnlabs.nl>



Architecture



Workflow



Filters

- Use of multiple filters for removing noise
 - Spam, marketing related
- Based on attributes such as
 - IP address
 - ASN
 - Web content-type



Filters

Filter types

- Static
 - Manually maintained
- Dynamic
 - Automatically generated



Static Filters

Static and manually maintained lists

- **AS Number:** e.g. mail marketing company networks
- **Country:** High volume SPAM countries
- **IP Address:** e.g. other researchers



Static Filters – ASN Example

```
[  
  {  
    "asn" : [47205],  
    "name" : "MAILERLITE",  
    "reason" : "Email marketing"  
  },  
  {  
    "asn" : [396479],  
    "name" : "MailGun",  
    "reason" : "Bulk Email"  
  },  
]
```



Dynamic Filters

Automatically generated each day

- **High Nxdomain:** DNS resolvers showing a high ratio of NXDOMAIN
- **Newly Seen:** IP addresses of resolvers that have not been seen before
- **Suspicious**
 - Open resolvers
 - Sinkhole clients
 - Abuse feeds (Spamhaus, APWG)



Analysing Data

Generate statistics per domain per day

- # of queries
- # of queries after filtering
- Results for each filter
- Unique # of ASNs, IPs and countries



Alerts

Use the generated daily statistics

- Determine if a domain needs to be alerted
 - Simple rule-based system



Alert Rules

Distinct risk categories

Based on 10-day average of daily DNS queries (after filtering)

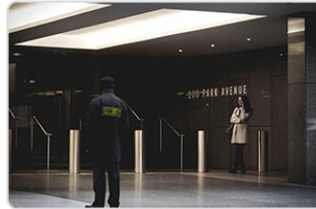
- Low: ≤ 5
- Medium: $5 < \leq 10$
- High > 10



Alert Rules

Special conditions

- Is NACE code match?
 - then risk is high



O - PUBLIC ADMINISTRATION AND DEFENCE; COMPULSORY SOCIAL SECURITY



P - EDUCATION



Q - HUMAN HEALTH AND SOCIAL WORK ACTIVITIES



Alert Rules

Special conditions

- Is keyword match?
 - then risk is high

```
{  
  "match": [  
    "legal",  
    "lawyer",  
    "government",  
    "doctor",  
    "pediatrician",  
    "dentist",  
    "healthcare",  
    "medical"  
  ]  
}
```



Alert Rules

Special conditions

- Did the web crawler find email address linked to domain?
 - then risk at least medium



Alert Message

- Designed in collaboration with registrars and registrants
- Sent on day 30 of 40 day quarantine period
- Registrant has 10 days to take action



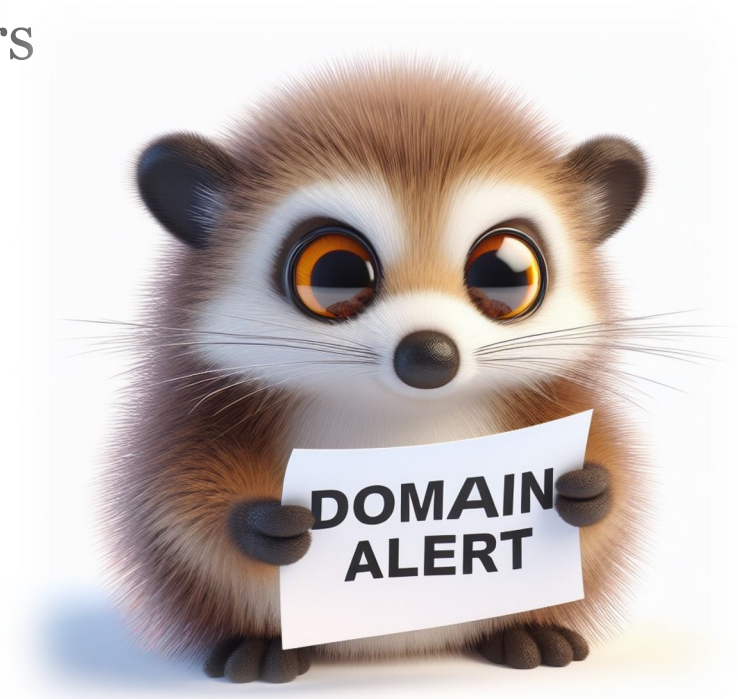
Alert Message

- Designed to explain the risk and suggest actions, e.g.
 - Informing contacts
 - Restoring the domain
- Multiple alert modes
 - To registrant
 - To registrar, who then forwards the alert
 - Registrar opt-out, no alerts are sent



Alert Message

- Domain grouping
 - Prevent sending multiple alerts to same registrar
- Weekly registrar digest
 - Gives registrar overview of alerts sent to their customers
- Support for sector related CERTs
 - E.g. send batch alerts to CERT for healthcare sector



Alert Message - Example

Belangrijke informatie over je opgeheven domeinnaam



**Belangrijke informatie over je
opgeheven domeinnaam
mariethereseheijnen.nl**

Er is mogelijk nog mailverkeer naar de domeinnaam

An English version of this e-mail can be found at www.sidn.nl

Dit is een bericht van SIDN, wij beheren het .nl-domein en ook de domeinnaam mariethereseheijnen.nl. Je hebt deze domeinnaam opgezegd op 2023-05-06. Met het opheffen van mariethereseheijnen.nl vervallen ook alle daaraan gekoppelde e-mailadressen. We sturen je dit bericht, omdat er waarschijnlijk nog gemaïld wordt naar een of meerdere e-mailadressen die gekoppeld waren aan de opgeheven domeinnaam. Hier schuilt een risico in. We vertellen je er graag meer over.

Alert Message – Daily Overview

LEMMINGS alert statistics

Datum	2023-12-02 00:00:00
Domains	671
Opt out	130
Direct-mail	0
Registars	62
Resellers	34
High risk	60
Medium risk	84
Low risk	527

Overview of alerted domains

Domain	Ok (avg)	Risk	Available	Keyword match	SBI section	Web	Registrar	Reseller	Opt-out	Direct-mail
<u>leviabbink.nl</u>	2.32	low	2023-12-12	False	-	False	1API GmbH	-	False	False
<u>nanningbakker.nl</u>	19.03	high	2023-12-12	False	-	False	Amen Nederland B.V.	-	False	False
<u>professional-backoffice.nl</u>	4.55	low	2023-12-12	False	N	False	Antagonist B.V.	-	False	False
<u>jobdistrict.nl</u>	3.06	low	2023-12-12	False	-	False	Antagonist B.V.	-	False	False
<u>dickrenses.nl</u>	3.68	low	2023-12-12	False	-	False	Antagonist B.V.	-	False	False
<u>bijstandsboete.nl</u>	4.06	low	2023-12-12	False	L	False	Antagonist B.V.	-	False	False

Anonymisation

PII information is deleted after a domain exits the 40-day quarantine period

- Registrant
 - Identifier
 - Name
 - Email address



2 Pilots

Do registrants understand the warning?



2 Pilots

Registrars worried about an increase in support calls



Alerts sent

After running LEMMINGS for 10-month period

- **587.778** deleted domains analysed
- Filtering removed 75% of mail related DNS queries
 - The average daily number of queries for a domain dropped from **4.7 to 1.2**
- **54.410** alerts have been sent to registrants



Alerts sent

54.410 alerts have been sent (9.2% of deleted domains)

Risk category	Alerts	Percentage
Low	44.701	82.15%
Medium	8.080	14.85%
High	4.639	8.53%



DNS Query Filters

Effectiveness per filter

Filter name	Removed	Percentage
ASN	47.177.603	44.4%
High Nxdomain	38.125.287	35.7%
Time	18.675.125	17.6%
Newly Seen IP	16.759.368	15.8%
Spamhaus	4.228.491	4.0%
Country	3.733.279	3.5%
Resolver Stability	2.675.135	2.5%
IP Address	2.552.351	2.4%

Measuring the Effect

- Not possible to directly measure the number of prevented data leaks
- Using a proxy:
 - Cancel-delete request as a proxy for prevented data leaks
 - Registrant survey



Cancel-delete Proxy

- Cancel-delete as proxy for prevented data leaks
- Cancel-delete baseline for the 12-month period before using LEMMINGS
 - **0.13%** of **627.285** deleted domains received cancel-delete

Cancel-delete Proxy

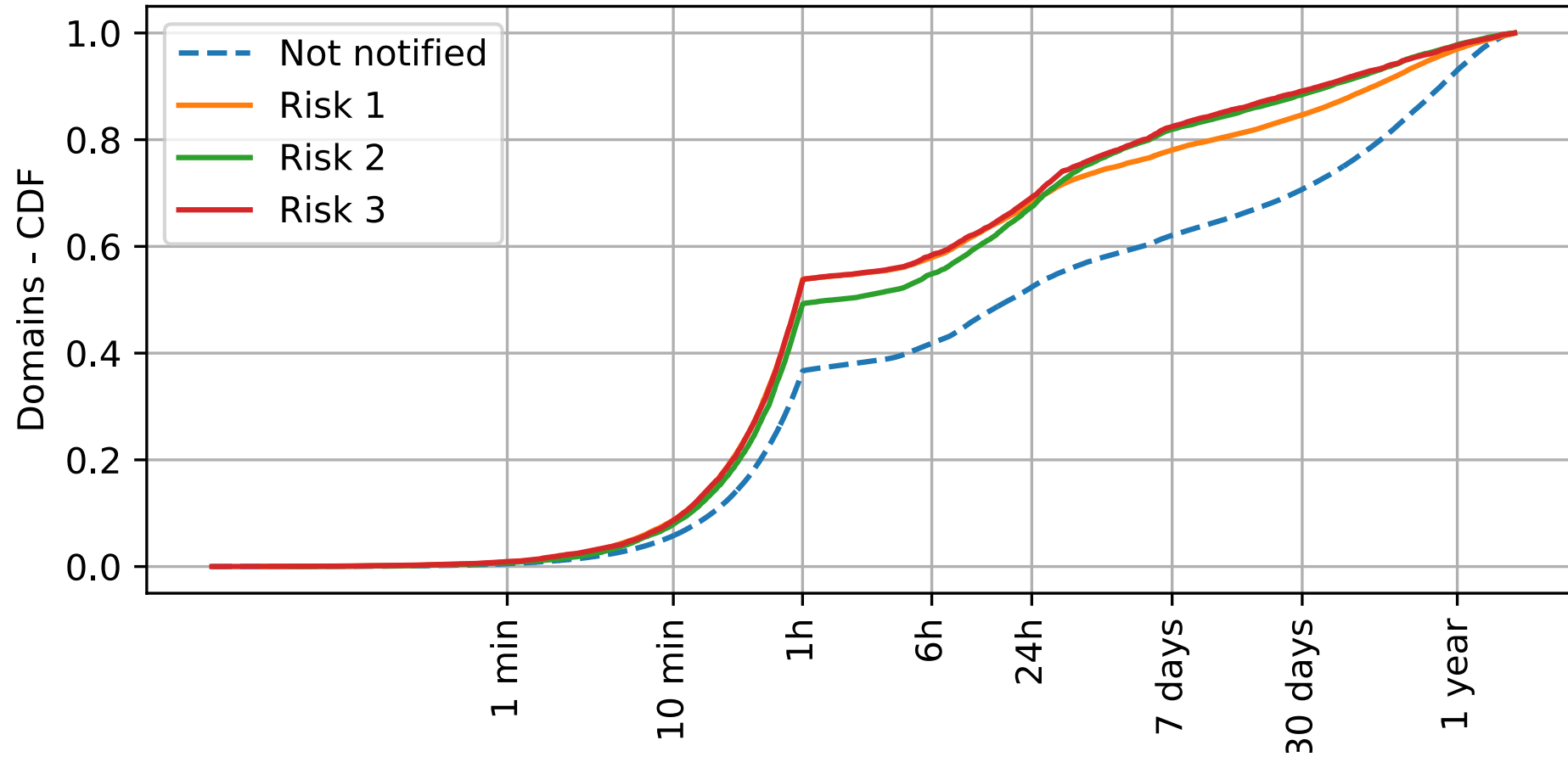
LEMMINGS cancel-delete ratio vs. baseline (**0.13%**)

Risk category	Cancel-delete	Percentage	Increase
Low	237	0.53%	3.8x
Medium	38	0.84%	6.0x
High	50	1.08%	7.7x



Do we alert the correct domains?

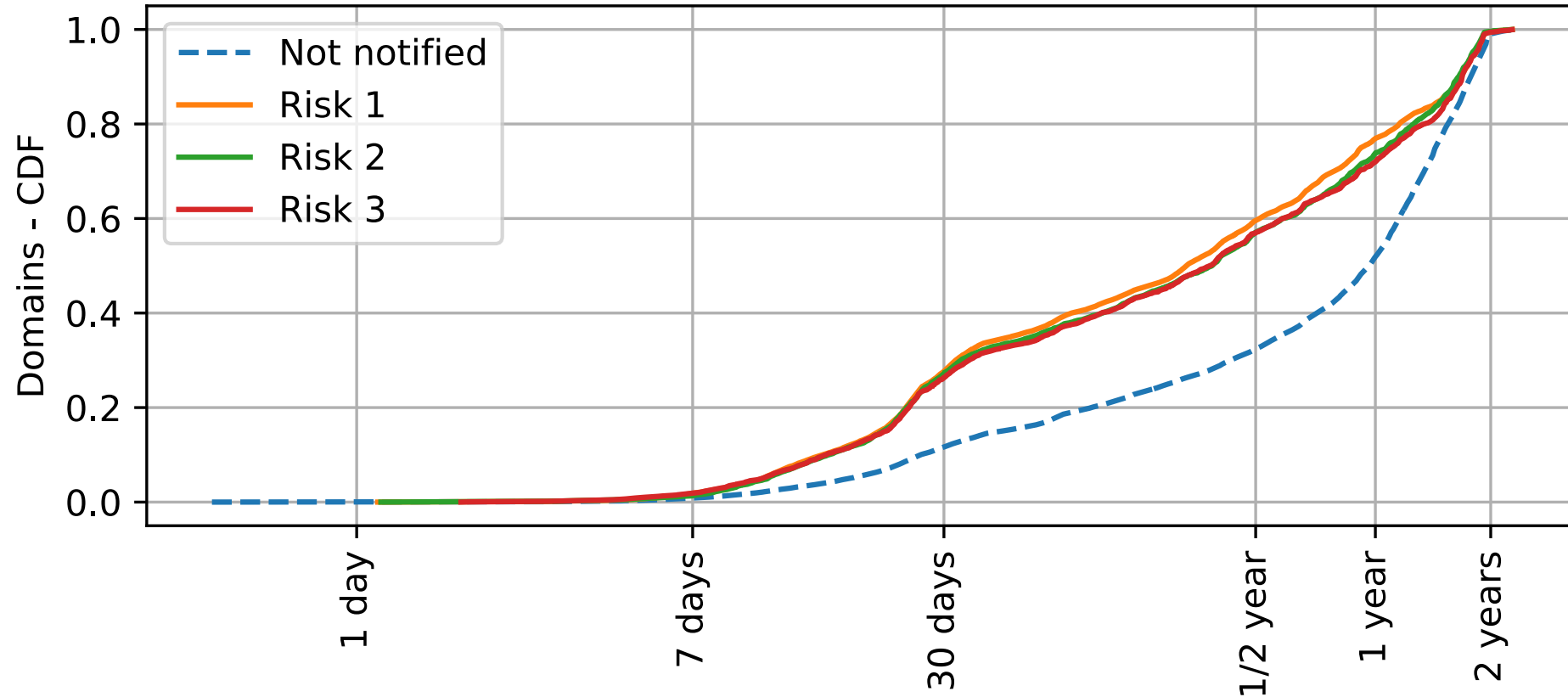
Re-registration time



Domains receiving an alert are re-registered more quickly

Do we alert the correct domains?

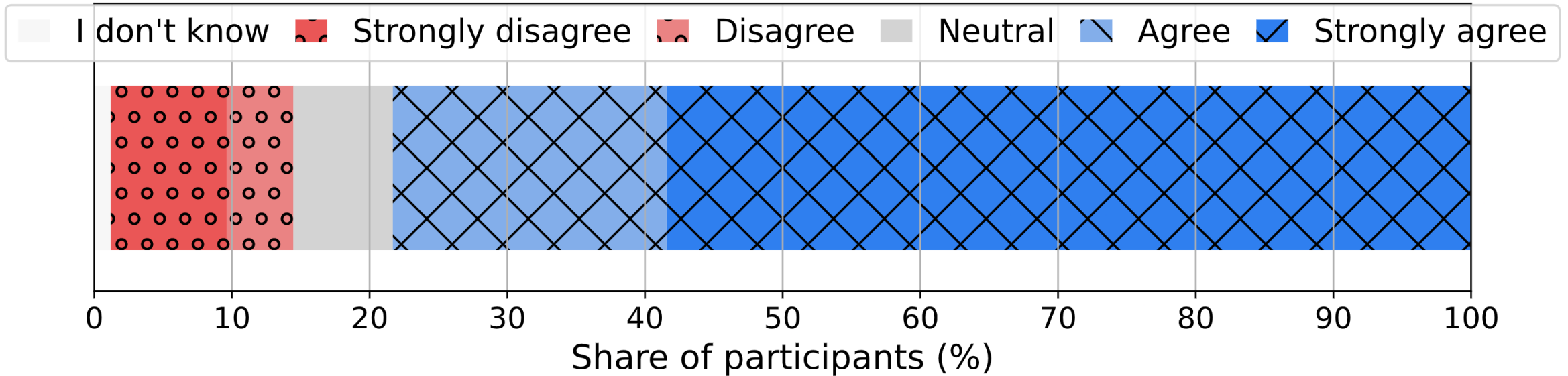
New mail server time



Alerted and re-registered domains have a new mail server more quickly

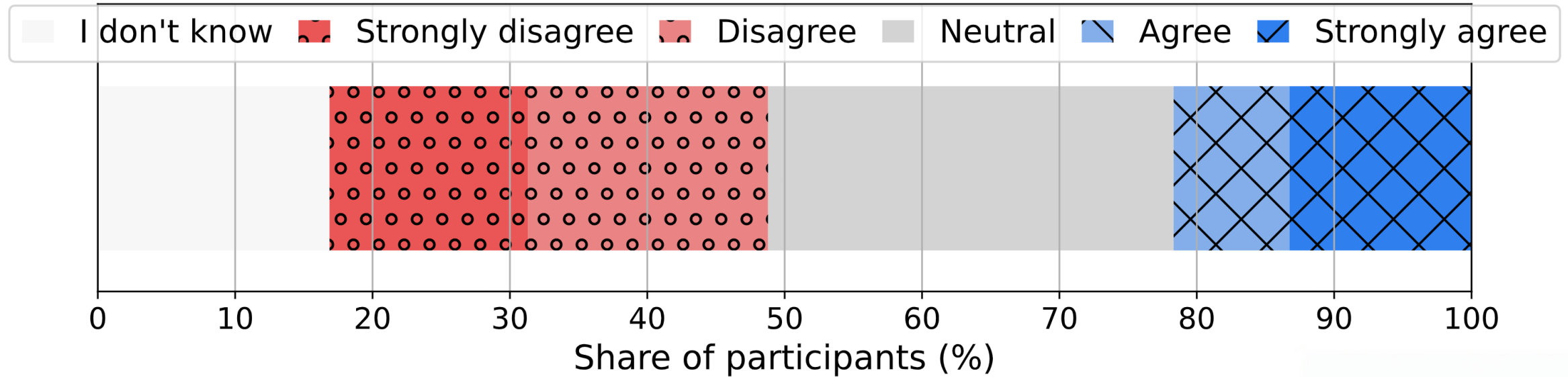
Registrant Survey

"The mail was useful"



Registrant Survey

"The mail helped to prevent problems like the leakage of information"



Future Work

- Analyse the impact of DNS Qname Minimisation
- Improve DNS filters



Takeaways

- Data leaks due to deleted domains are a real thing
- It's difficult to directly measure the effect of LEMMINGS
- Explaining the security risk is challenging
 - Low number of registrant and registrar questions



 SIDN.nl

 @SIDN

 SIDN

Questions?

www.sidnlabs.nl | stats.sidnlabs.nl

