

# LEMMINGS: Preventing data leaks at TLD scale

Moritz Müller | ONE Conference 2023

October 3, 2023



# The (old) Problem

A deleted domain may lead to an accidental data leak

**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



Image for illustration.

**f** 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.

**| Just in**

- 08:47 Anderlec to go to f
- 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

presented by: **Certified Secure**


## security.nl

News Background Community

### News

#### Police leak sensitive emails through expired domain names

Friday 20 January 2017, 07:57 by **Redactie** , 20 reactions



Police sent emails containing sensitive information to expired domain names, ending up in the hands of a security expert who registered the domains. These are arrest warrants, reports from Safe Home and security plans, **BNR** reports .

The police were warned about this security risk two years ago. During the transition to the National Police, the police let all kinds of domain names expire, which can then be registered by others. Security expert Wouter Slotboom already warned about this in **February 2015** . However, the police would not have done anything with his report.

"I raised a problem and then presented a solution on a silver platter. When I found out after half a year that they hadn't done anything about it, I registered a few domain names myself to demonstrate that sensitive information was indeed sent via the mail comes in. I hope my point is clear now." It is not just about information that can be received. According to Gerrit van de Kamp of the ACP police union, information could also be requested via the domain names.

The National Police states that more than 3600 old domain names have now been registered by the police themselves, but a number have slipped through. Appropriate measures will be taken for this. The National Police also emphasizes that both citizens and police officers should only use the domain name politie.nl.

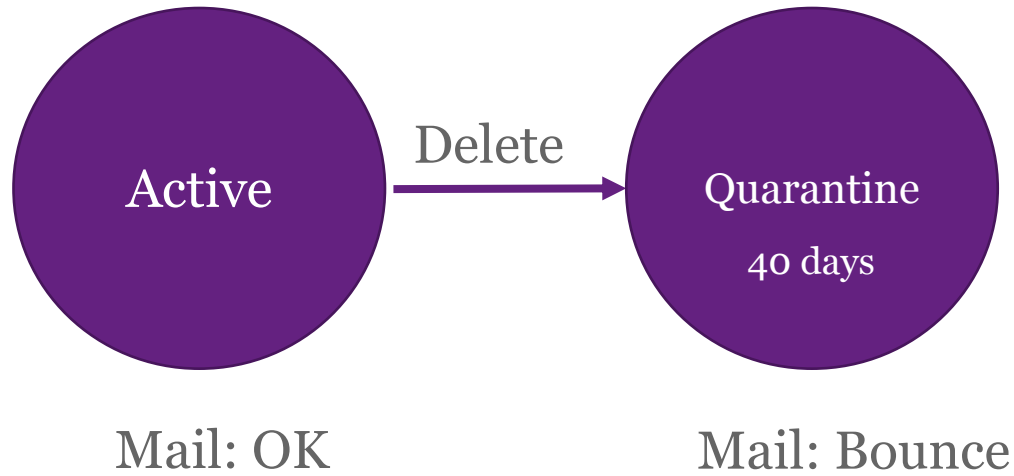
**Our goal:**  
Warn registrants about the potential  
danger of a deleted domain

# Domain Life Cycle

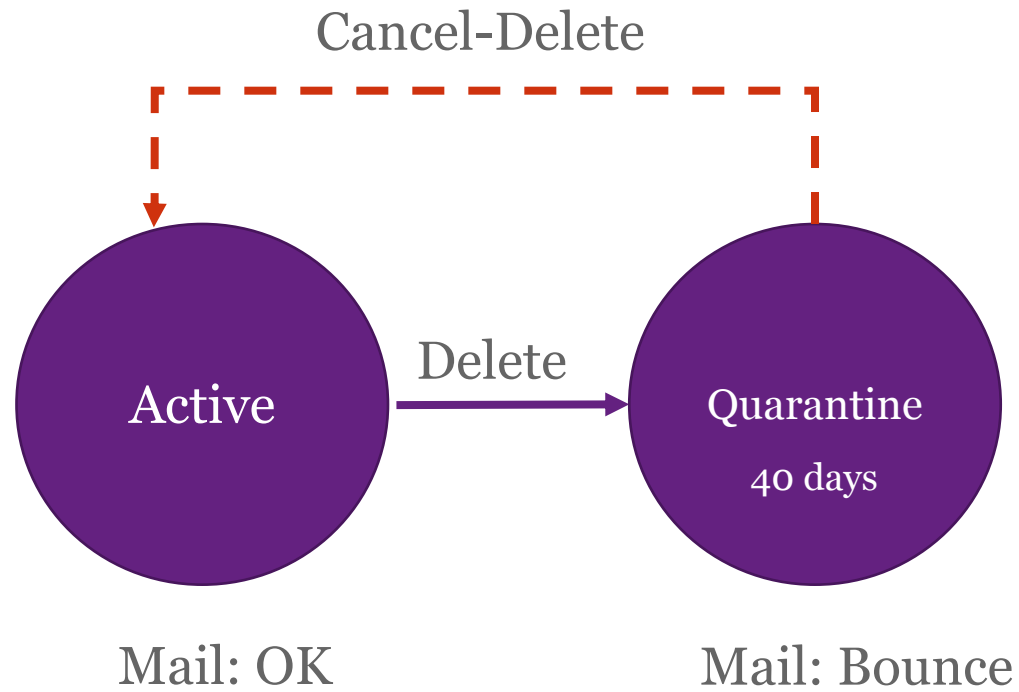


Mail: OK

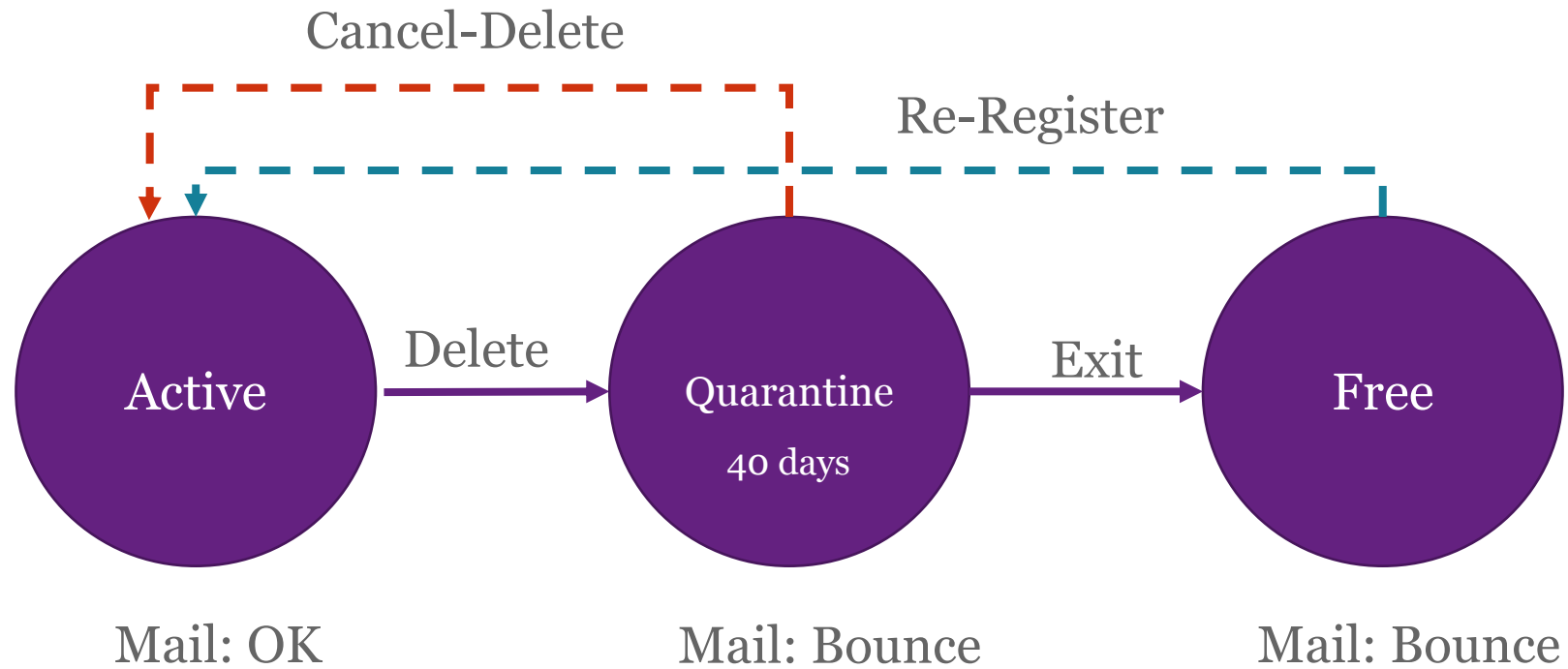
# Domain Life Cycle



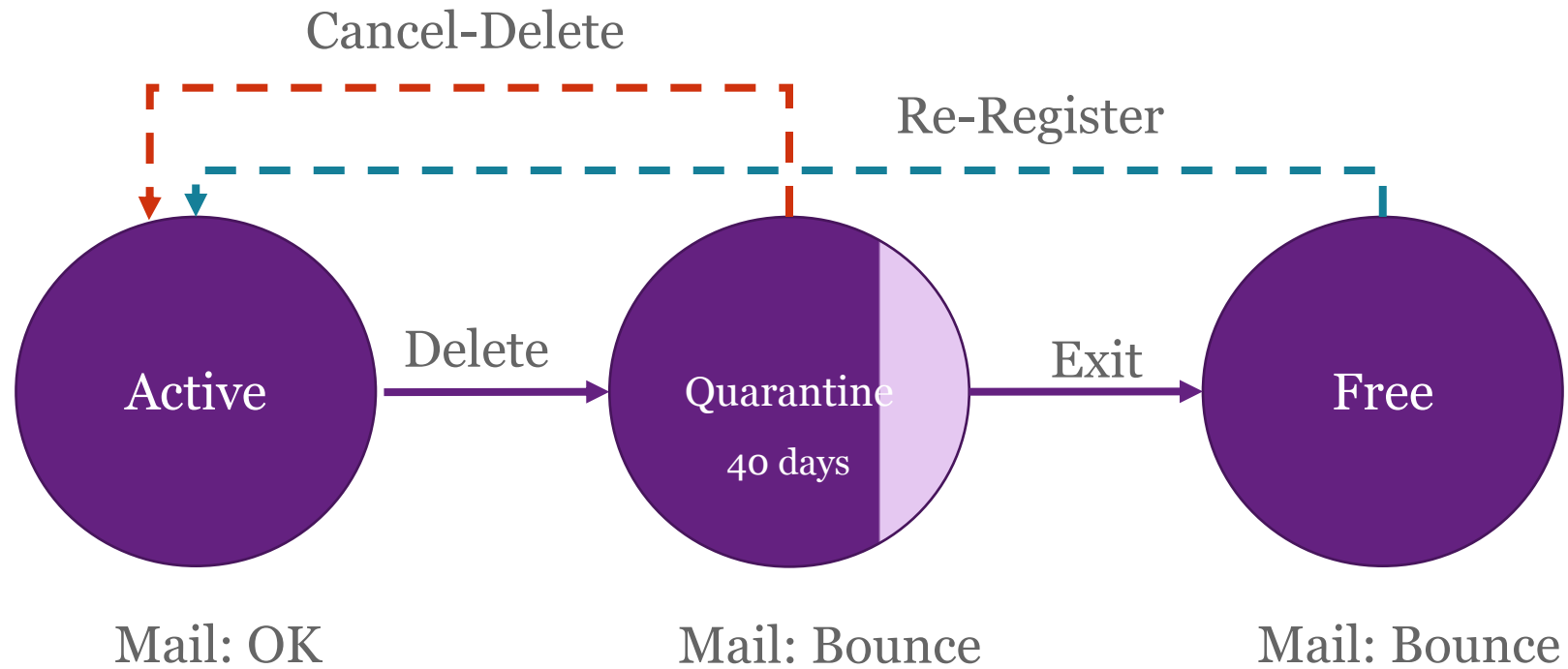
# Domain Life Cycle



# Domain Life Cycle



# Domain Life Cycle



LEMMINGS



Day 1 - 30



# Introducing: LEMMINGS

(deLetEd doMain MaIl warNinG System)

- A system for detecting legitimate mail transactions by analysing DNS data
- Approach:
  - Analyse DNS MX queries for deleted domains.
  - Send alert to former registrant, when following is true:
    - Indication of legitimate email transactions
    - Domain not yet exited quarantine-period (1)

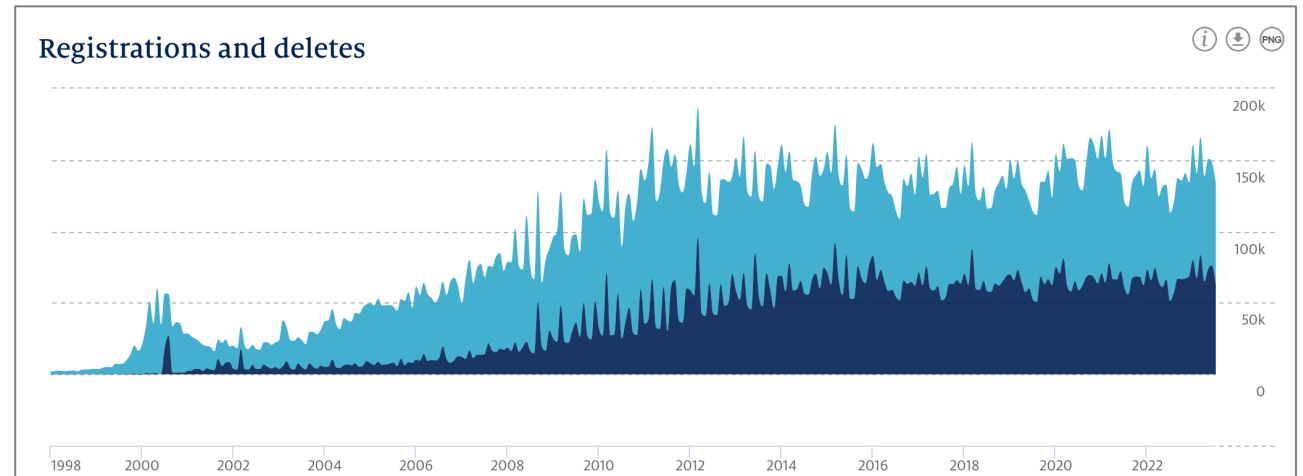
[1] Only the former registrant is allowed to remove the domain from quarantine

# Challenges

- Filtering out noisy mail; (marketing, spam, social media)
- Explaining the security risk to a registrant
- Prevent the alert from looking like a scam/spam message
- Not every registrant can be contacted via e-mail
  - Privacy proxies and inzone mail addresses
- Prevent registrant questions flooding registrar supportdesk

# Data Sources

- Registration database
- Daily DNS queries for .nl
  - Total ~4 billion
  - MX ~180 million
- Web crawler data
  - ~6.1 million .nl domains,
  - Web content-type (business, ecommerce ...)
  - Use of email addresses on website
- Abuse feeds (Spamhaus, APWG), Sinkhole for filtering



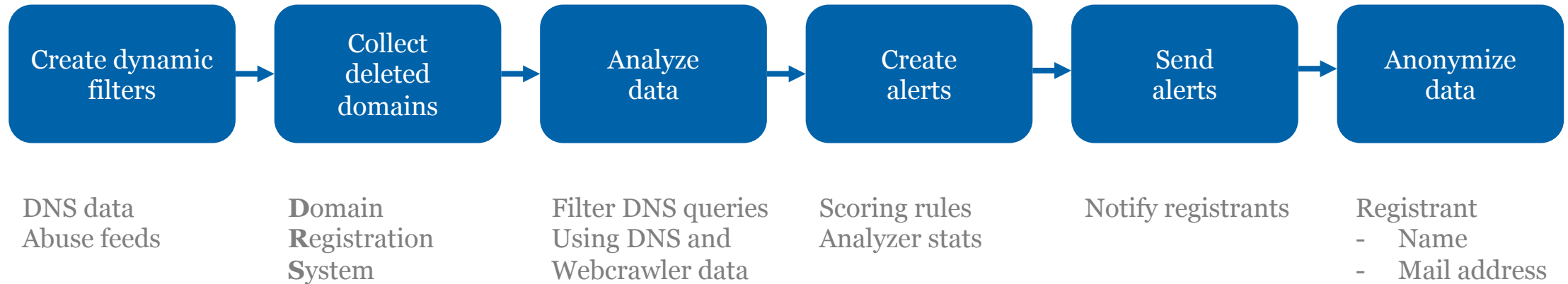
Registrations (light blue) and deletes (dark blue)

<https://stats.sidnlabs.nl/en/registration.html#registrations%20and%20deletes>

# Privacy Considerations

- We do not capture or analyse mail content
- PII information (registrant email address and name) are deleted after quarantine period
- No trackers in email alert sent to registrant

# System Process

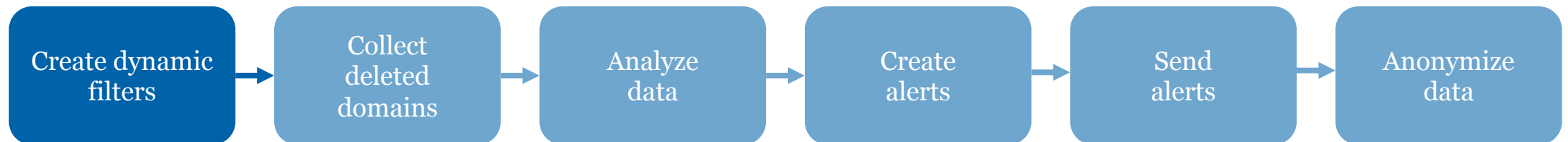


# Filters

Multiple filters used to remove "noise": Spam, marketing, social media

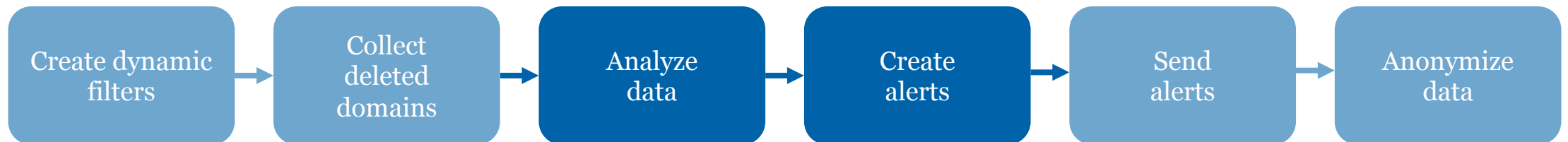
Filters based on DNS request attributes, for example:

- *ASN*: filter queries from autonomous system, e.g. mail marketing company, Facebook
- *High Nxdomain*: DNS resolvers having a high ratio of NXDOMAIN
- *Newly Seen IPs*: IP addresses of resolvers that have not been seen before

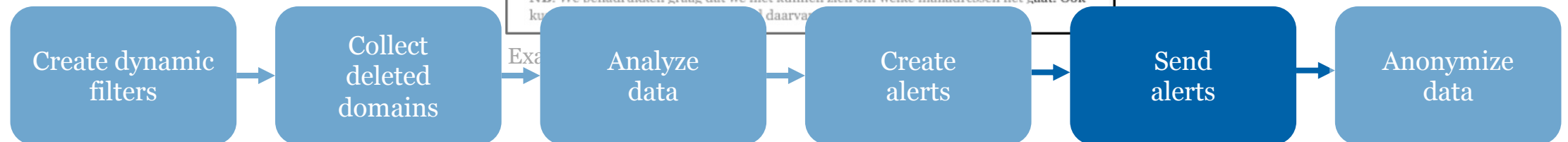


# Alerts Rules

- Risk categories
  - Based on 10-day average of daily MX queries (after filtering)
    - $\leq 5$  low
    - $> 5$  en  $\leq 10$  medium
    - $> 10$  high
- Special conditions
  - If keyword or business activity match, then risk is high
  - If email address is found on website, then at least medium



# Alert Message





# Alert Message

Message is sent on day 30 of 40 day quarantine period

- Designed in collaboration with registrars and registrants
- Explain about the risk and offer suggested actions
  - Inform contacts
  - Re-register domain

## Alert modes

- Direct to registrant,
- To the registrar, which then forwards the alert
- Registrar opt-out, no alerts are sent



Belangrijke informatie over je opgezegde domeinnaam



**Belangrijke informatie over je opgezegde domeinnaam zovintage.nl**

Er is mogelijk nog mailverkeer naar de domeinnaam

*An English version of this e-mail can be found at [www.sidn.nl](http://www.sidn.nl)*

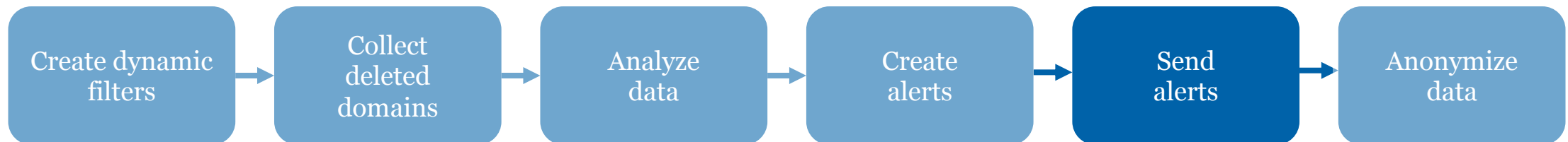
Dit is een bericht van SIDN, wij beheren het .nl-domein en ook de domeinnaam zovintage.nl. Je hebt deze domeinnaam opgezegd op 2021-07-28. Met het opheffen van zovintage.nl vervallen ook alle daaraan gekoppelde e-mailadressen. We sturen je dit bericht, omdat er waarschijnlijk nog gemaïld wordt naar een of meerdere e-mailadressen die gekoppeld waren aan de opgezegde domeinnaam. Hier schuilt een risico in. We vertellen je er graag meer over.

**Wat is er aan de hand?**

Als beheerder van de .nl-domeinnamen kunnen we zien dat er waarschijnlijk nog gemaïld wordt naar mailadressen die gekoppeld waren aan de opgezegde domeinnaam. Het zou dus kunnen zijn dat er nog voor jou of je organisatie bestemde mail naar oude mailadressen wordt gestuurd. Vanaf 2021-09-06 kan de domeinnaam door een ander geregistreerd worden. Het risico bestaat dat deze nieuwe houder toegang krijgt tot de voor jou bestemde mails die dan nog verstuurd worden. Daarmee kan er mogelijk voor jou of je organisatie bestemde persoonlijke en/of gevoelige informatie in handen van derden komen, met alle gevolgen van dien.

**NB.** We benadrukken graag dat we niet kunnen zien om welke mailadressen het gaat. Ook kunnen wij de mails zelf en de inhoud daarvan niet zien.

Example alert message



# Alerts sent

Running LEMMINGS for 10-month period

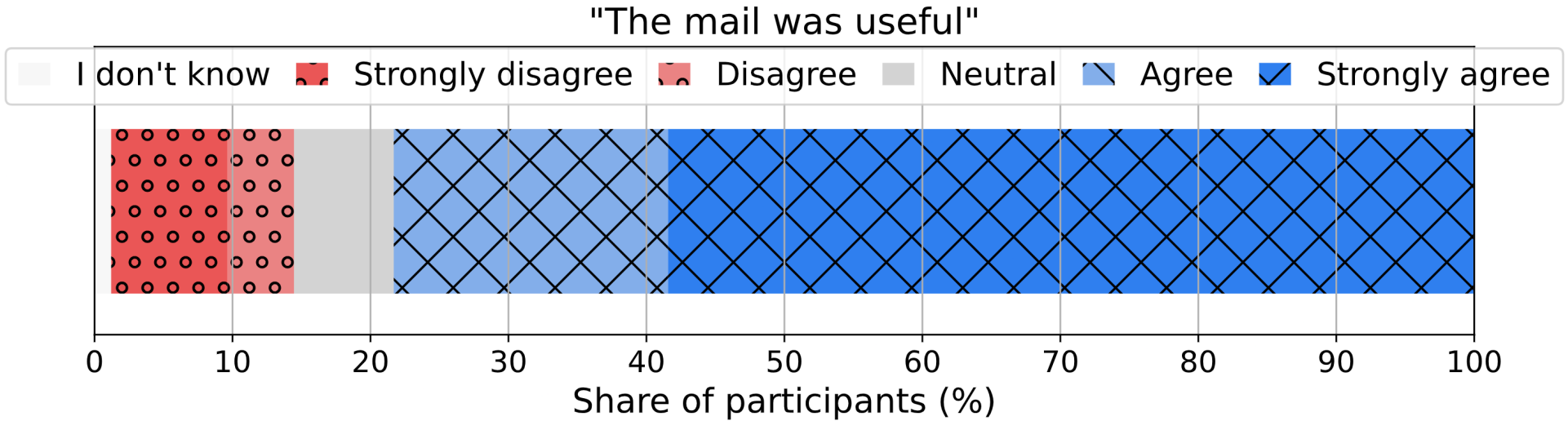
- 587.778 domains have been deleted
- Filtering removed 75% of MX queries
  - Average daily MX queries for domain drops from 4.7 to 1.2
- 54.410 alerts have been sent (9.2%)

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

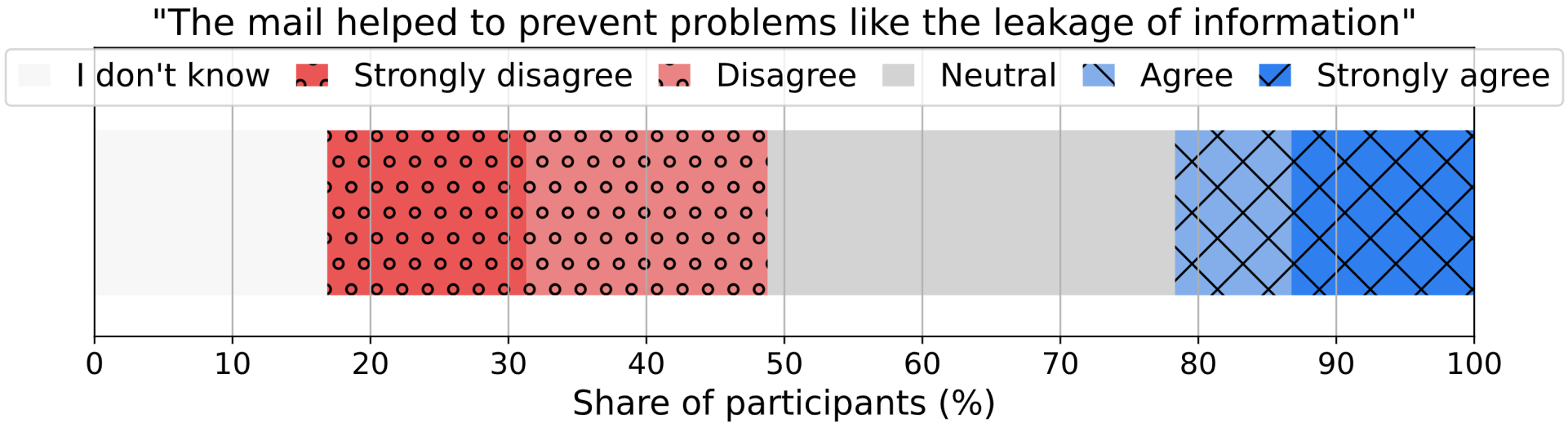
# How effective is LEMMINGS?

- Impossible to directly measure prevented data leaks
- Proxies:
  - Survey
  - Cancel-delete request as a proxy for preventing potential data leaks

# Registrant Survey (1)



# Registrant Survey (2)



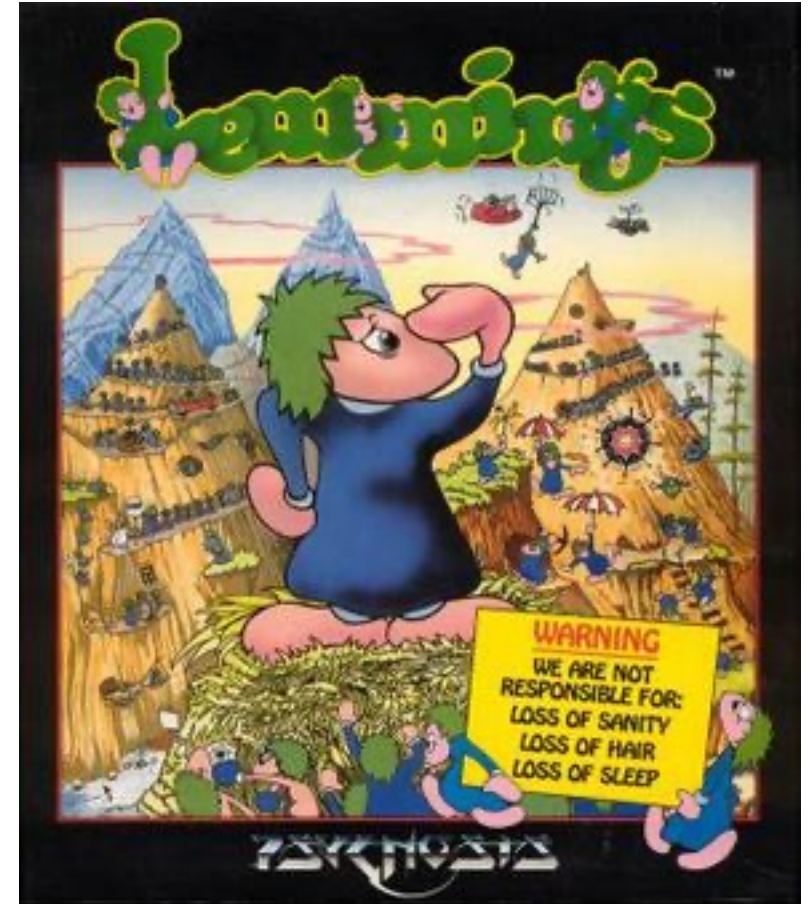
# How effective is LEMMINGS?

- Cancel-delete request as a proxy for preventing potential data leaks
- Calculate cancel-delete request baseline for the 12-month period before using LEMMINGS
  - 0.13% over 627.285 domains
- Compare to cancel-delete ratio seen during LEMMINGS period

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

# Conclusion

- Difficult to measure the effect of LEMMINGS
- Reaction (mostly) positive
  - Low number of registrant/registrar complains about the alerts
  - Positive response Dutch internet community



Source:  
<https://en.wikipedia.org/wiki/File:Lemmings-BoxScan.jpg>

 SIDN.nl

 @SIDN

 SIDN

# Questions?

[www.sidnlabs.nl](http://www.sidnlabs.nl) | [stats.sidnlabs.nl](http://stats.sidnlabs.nl)

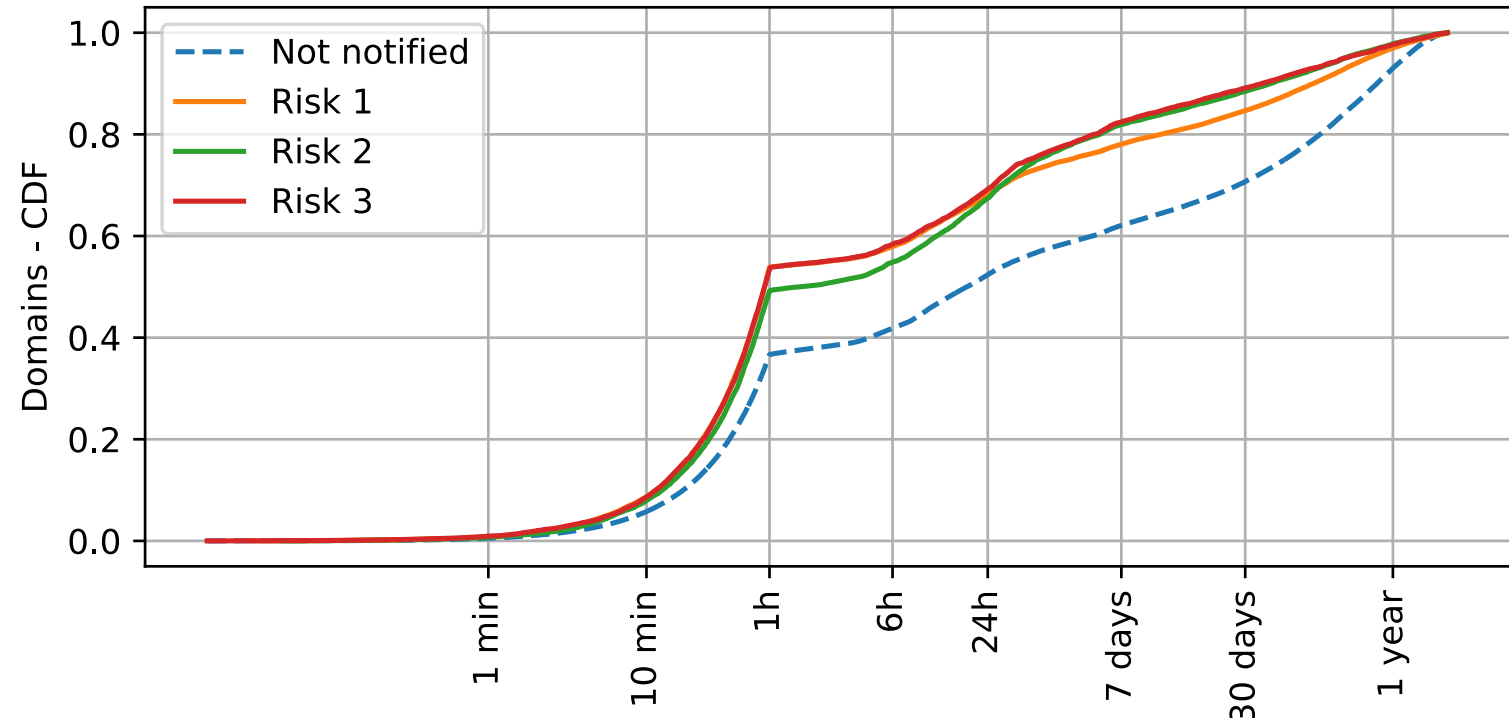


Additional slides

# MX Query Filters

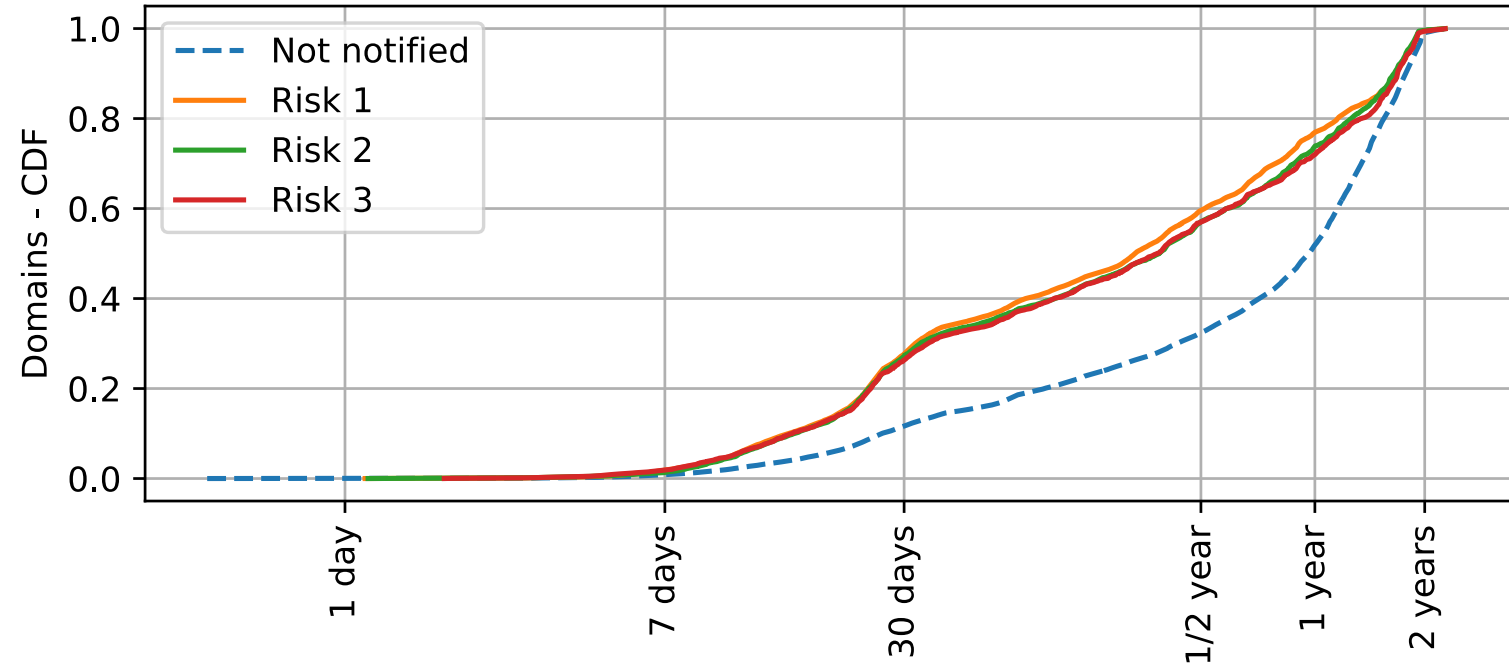
Filter name	Queries removed	Percentage
ASN	47.177.603	44.4%
High Nxdomain	38.125.287	35.7%
Time	18.675.125	17.6%
Newly Seen IP	16.759.368	15.8%
Spamhaus	4.228.491	4.0%
Country	3.733.279	3.5%
Resolver Stability	2.675.135	2.5%
IP Address	2.552.351	2.4%
Open Resolver	742.604	0.7%
No Mail	731.991	0.7%
Sinkhole	14.166	0.0%
APWG	1.650	0.0%

# Do we reach the right domain names?



Warned domain names are re-registered faster

# Do we reach the right domain names?



Warned domain names have a new mail server faster