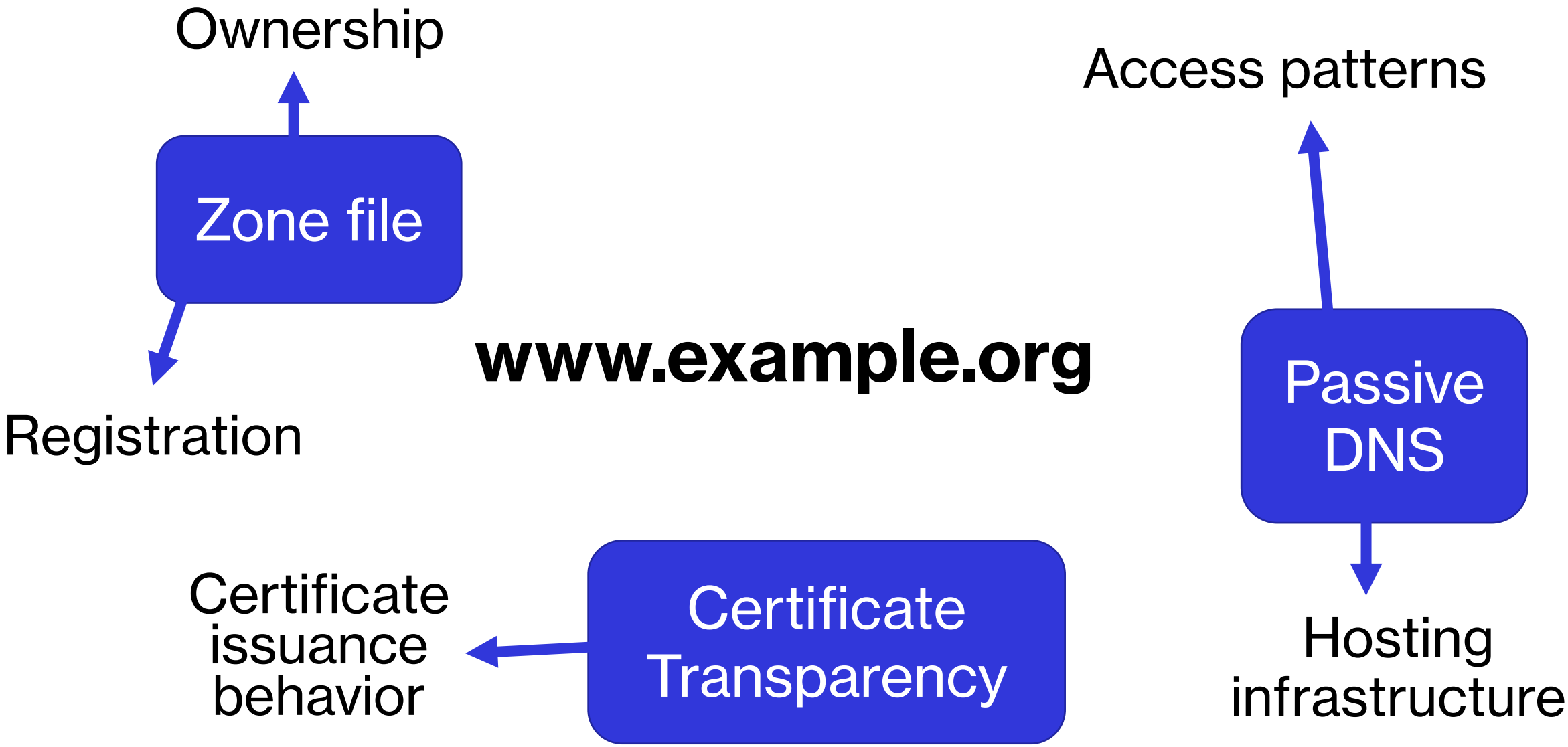
AALBORG
UNIVERSITY

NordSec 2021

# Gollector: Measuring Domain Name Dark Matter from Different Vantage Points

Kaspar Hageman, René Rydhof Hansen, and Jens Myrup Pedersen

# www.example.org


Ownership
Zone file
Registration
www.example.org
Access patterns
Passive DNS
Hosting infrastructure
Certificate issuance behavior
Certificate Transparency

**Different vantage point**
**=**
**Different view of the domain name space at a particular time**

# Different vantage point
# =
# Different view of the domain name space at a particular time

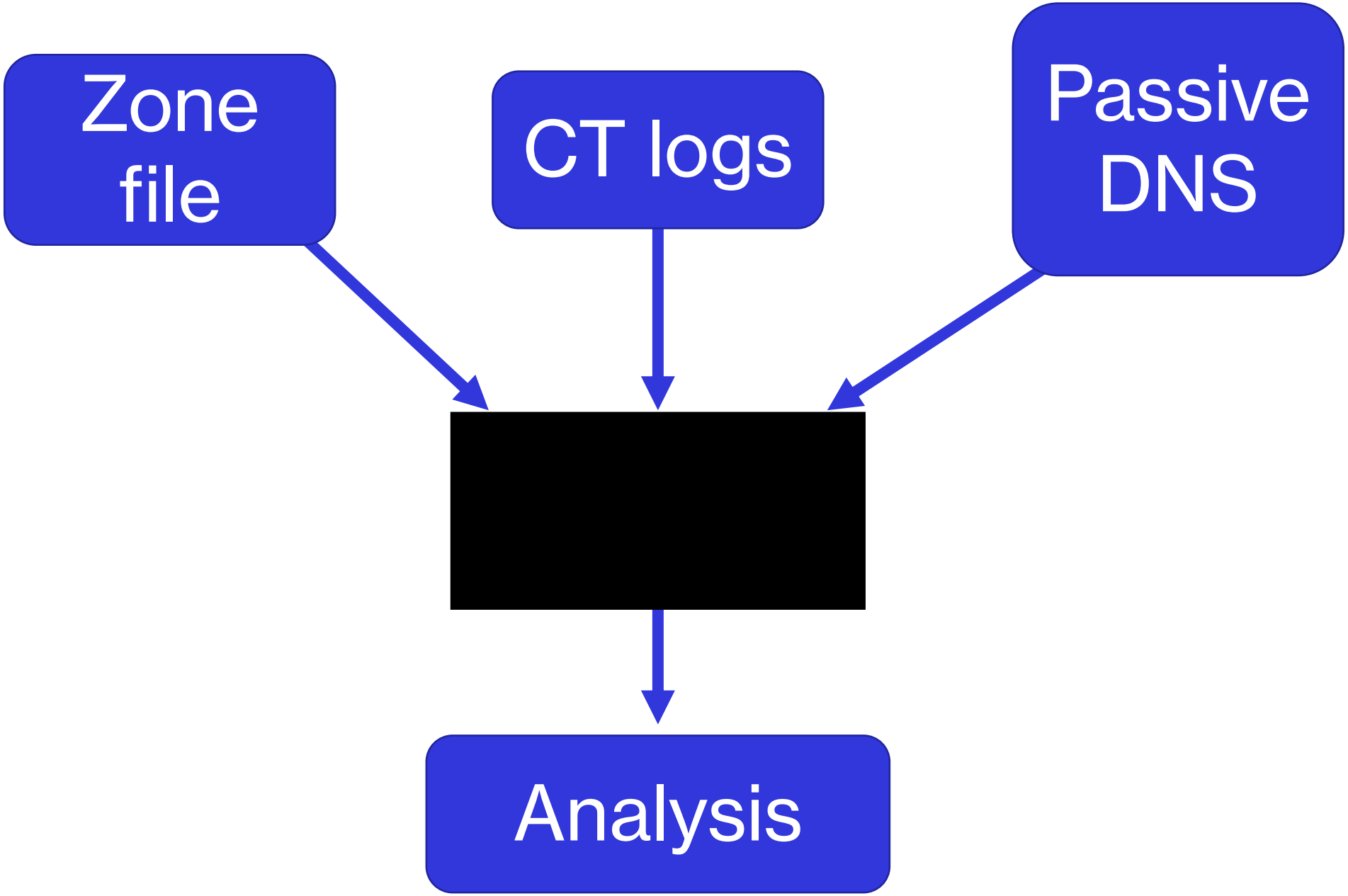
**CT log does not see domains which do no employ TLS**

**VS**

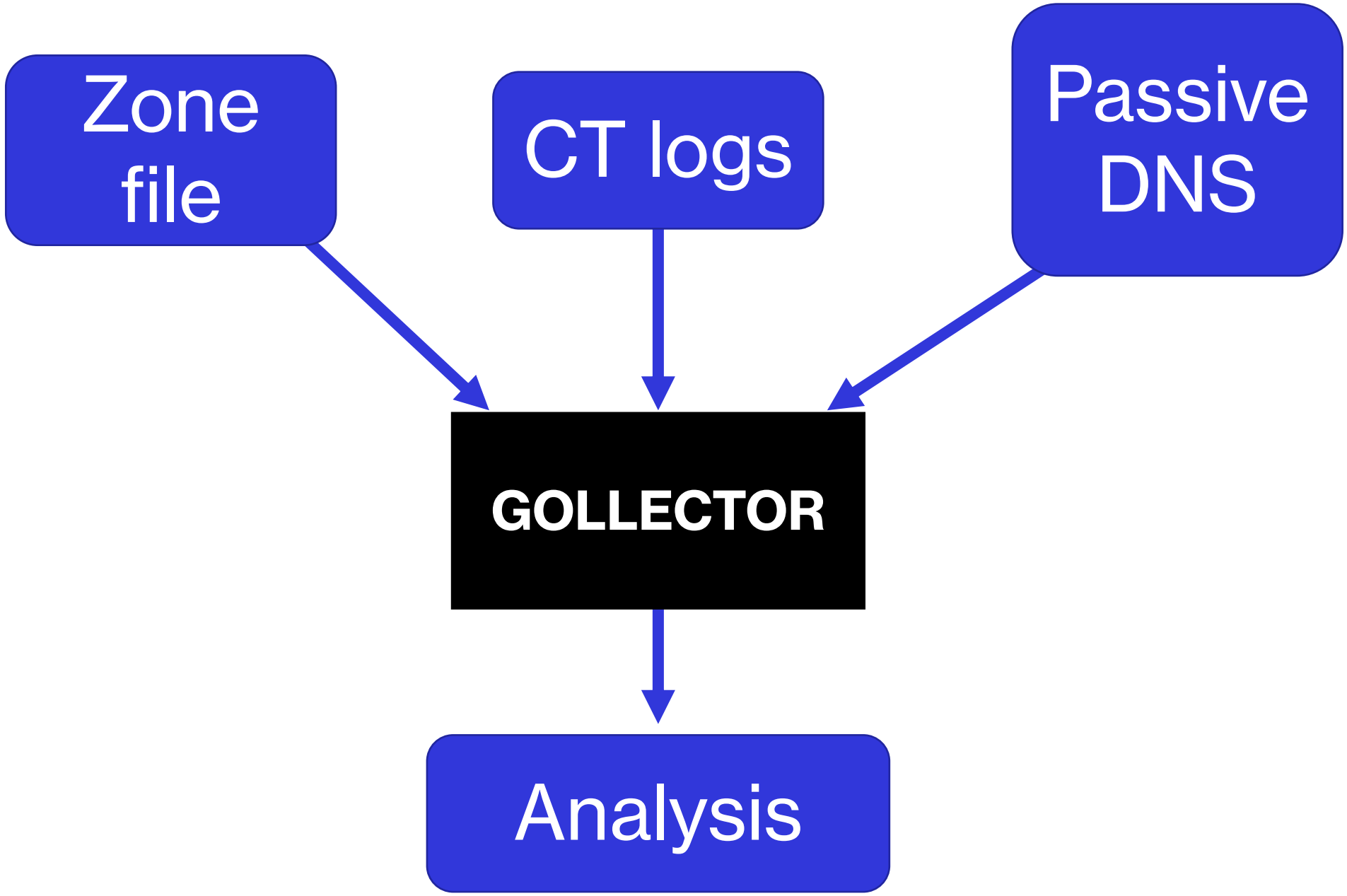
**Passive DNS from a Russian authoritative NS does not see non-Russian domains**

**Different vantage point**
**=**
**Different view of the domain name space at a particular time**

Certain parts of the domain name space are hidden for a given vantage point
**Domain name dark matter**




K. Hageman et al., Gollector: Measuring Domain Name Dark Matter from Different Vantage Points

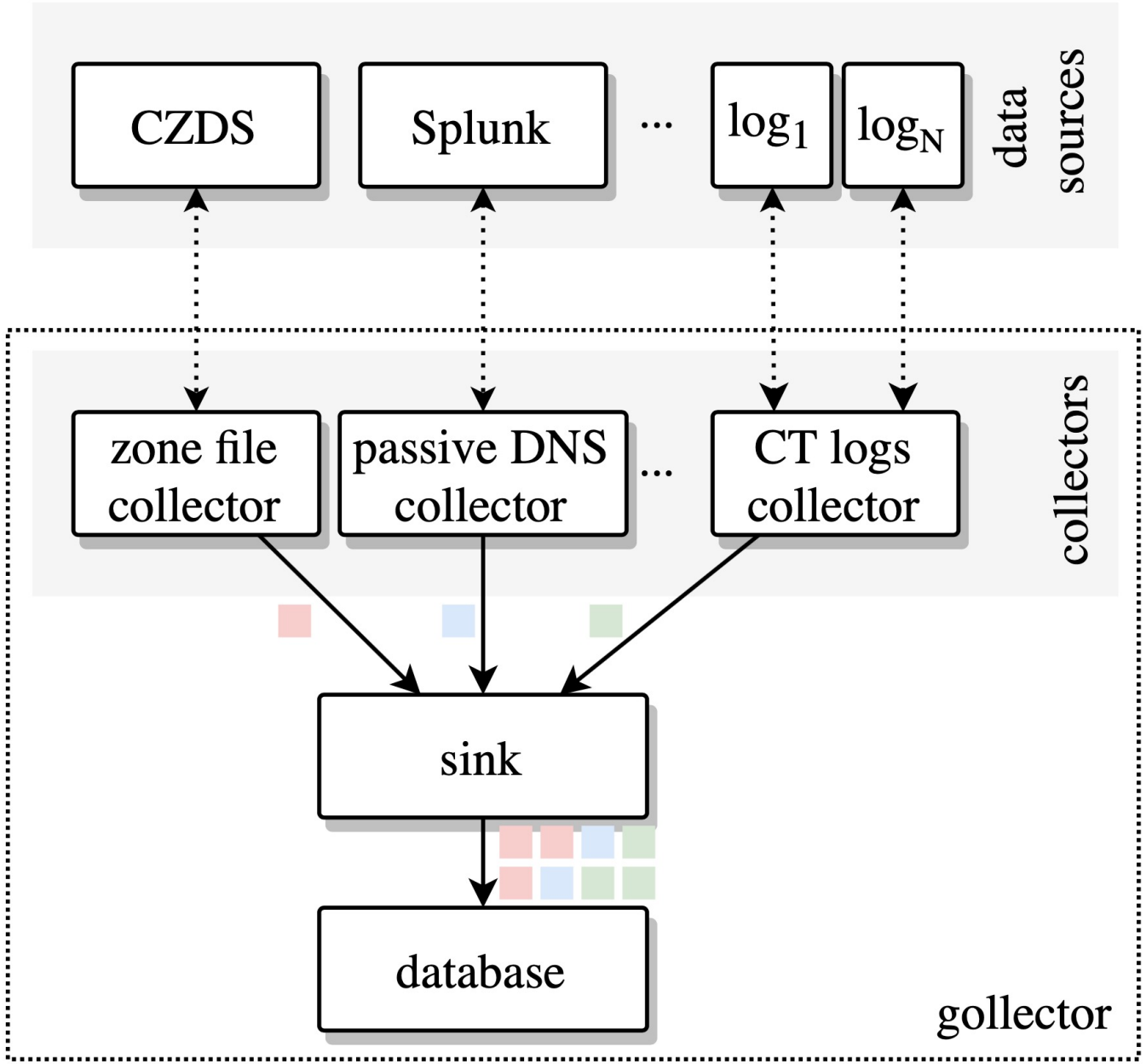



K. Hageman et al., Gollector: Measuring Domain Name Dark Matter from Different Vantage Points

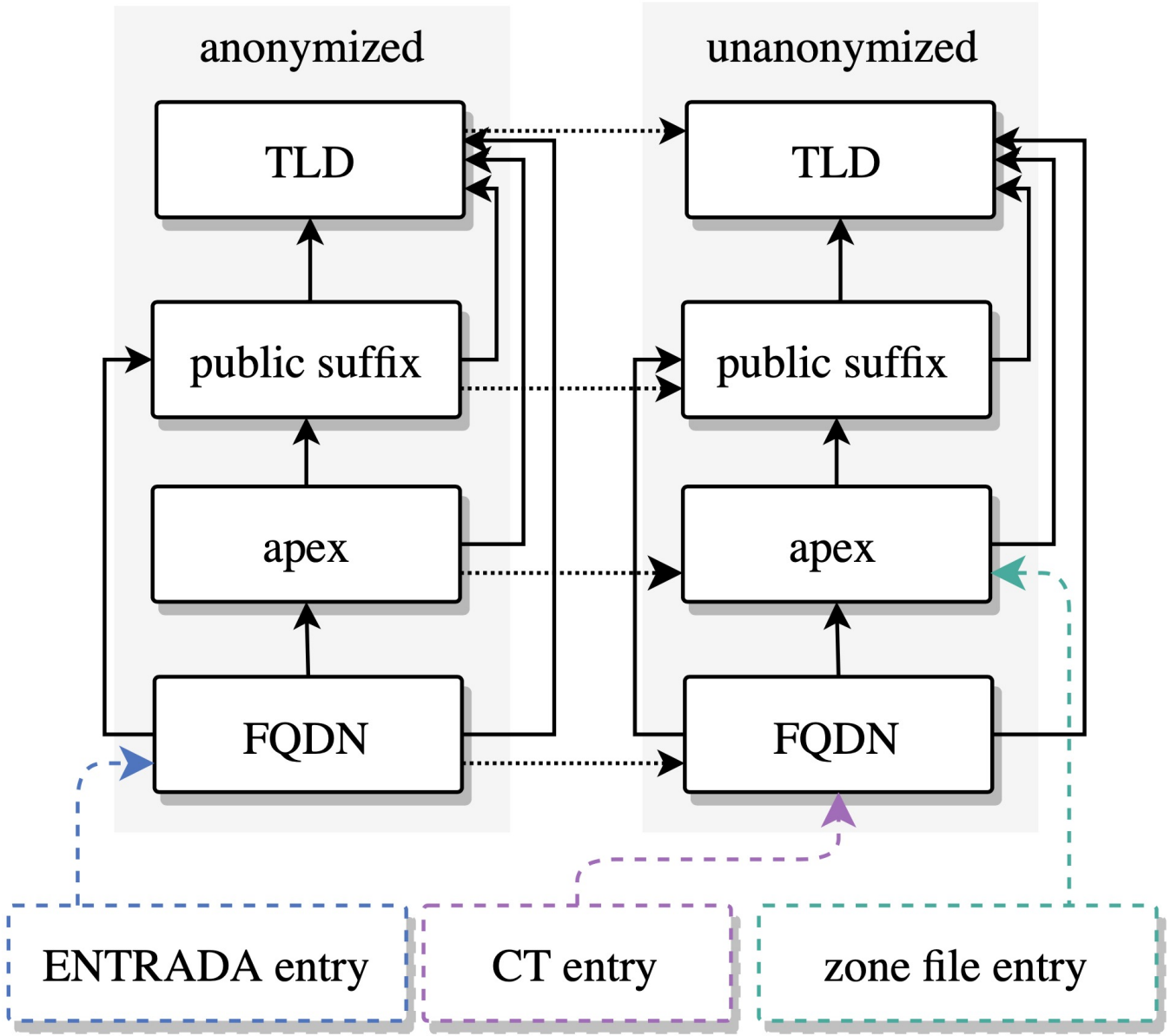
# Contributions

**#1** **DESIGN OF GOLLECTOR**

**#2** **USE CASES OF DOMAIN NAME DARK MATTER**




K. Hageman et al., Gollector: Measuring Domain Name Dark Matter from Different Vantage Points




K. Hageman et al., Gollector: Measuring Domain Name Dark Matter from Different Vantage Points

# Use cases

Early detection of domain names

Split horizon and data leakage

Subdomain enumeration

# Use cases

Early detection of domain names

Split horizon and data leakage

Subdomain enumeration

**Data collection: 3 weeks of data, 4 vantage points**

# Early detection of domain names

*‘Can we leverage non-zone file data for identifying domain registration early?’*

K. Hageman et al., Gollector: Measuring Domain Name Dark Matter from Different Vantage Points

# Early detection of domain names

Table 4: Detection of newly registered domain names for non-zone files vantage points. The results for both the full set of TLDs and the `.dk` zone only are shown.

| | Absolute | | | Percentual | | |
|---|---|---|---|---|---|---|
| | CT | Passive | ENTRADA | CT | Passive | ENTRADA |
| **All TLDs** | | | | | | |
| Overall | 971,318 | 533 | 46,628 | 23.6% | 0.01% | 1.1% |
| Before | 568,436 | 216 | 25,713 | 13.8% | 0.01% | 0.62% |
| Within 7 days | 325,277 | 169 | 4688 | 7.9% | 0.00% | 0.11% |
| **`.dk` only** | | | | | | |
| Overall | 16,476 | 63 | 46,495 | 34.9% | 0.13% | 98.5% |
| Before | 0 | 0 | 25,673 | 0.00% | 0.00% | 54.4% |
| Within 7 days | 639 | 3 | 4,601 | 1.35% | 0.01% | 9.74% |

K. Hageman et al., Gollector: Measuring Domain Name Dark Matter from Different Vantage Points

# Early detection of domain names

Table 4: Detection of newly registered domain names for non-zone files vantage points. The results for both the full set of TLDs and the `.dk` zone only are shown.

| | Absolute | | | Percentual | | |
|---|---|---|---|---|---|---|
| | CT | Passive | ENTRADA | CT | Passive | ENTRADA |
| **All TLDs** | | | | | | |
| Overall | 971,318 | 533 | 46,628 | 23.6% | 0.01% | 1.1% |
| Before | 568,436 | 216 | 25,713 | 13.8% | 0.01% | 0.62% |
| Within 7 days | 325,277 | 169 | 4688 | 7.9% | 0.00% | 0.11% |
| **.dk only** | | | | | | |
| Overall | 16,476 | 63 | 46,495 | 34.9% | 0.13% | 98.5% |
| Before | 0 | 0 | 25,673 | 0.00% | 0.00% | 54.4% |
| Within 7 days | 639 | 3 | 4,601 | 1.35% | 0.01% | 9.74% |

Absolute domain registrations detected in CT dataset **before zone files**

As percentage of **ground truth** (zone files)

K. Hageman et al., Gollector: Measuring Domain Name Dark Matter from Different Vantage Points

# Early detection of domain names

Table 4: Detection of newly registered domain names for non-zone files vantage points. The results for both the full set of TLDs and the `.dk` zone only are shown.

| | Absolute | | | Percentual | | |
|---|---|---|---|---|---|---|
| | CT | Passive | ENTRADA | CT | Passive | ENTRADA |
| **All TLDs** | | | | | | |
| Overall | 971,318 | 533 | 46,628 | 23.6% | 0.01% | 1.1% |
| Before | 568,436 | 216 | 25,713 | 13.8% | 0.01% | 0.62% |
| Within 7 days | 325,277 | 169 | 4688 | 7.9% | 0.00% | 0.11% |
| **.dk only** | | | | | | |
| Overall | 16,476 | 63 | 46,495 | 34.9% | 0.13% | 98.5% |
| Before | 0 | 0 | 25,673 | 0.00% | 0.00% | 54.4% |
| Within 7 days | 639 | 3 | 4,601 | 1.35% | 0.01% | 9.74% |

K. Hageman et al., Gollector: Measuring Domain Name Dark Matter from Different Vantage Points

# Split horizon and data leakage

*‘How much domain name information is leaked outside our university network?’*

K. Hageman et al., Gollector: Measuring Domain Name Dark Matter from Different Vantage Points

# Split horizon and data leakage

**Passive DNS from network resolver**

**VS**

**Passive DNS from authoritative name server**

K. Hageman et al., Gollector: Measuring Domain Name Dark Matter from Different Vantage Points

# Split horizon and data leakage

Table 5: The 10 apex domains with the most observed unique FQDNs in the passive DNS dataset collected from the university network.

| Apex domain | Unique FQDN count | % |
|---|---|---|
| aau.dk | 3,829,837 | 63% |
| googlesyndication.com | 344,058 | 6% |
| technicolor.net | 61,151 | 1.01% |
| cedexis-radar.net | 44,771 | 0.74% |
| sophosxl.net | 39,297 | 0.65% |
| bbsyd.net | 36,758 | 0.61% |
| office.com | 30,215 | 0.50% |
| emnet.dk | 23,540 | 0.39% |
| obelnet.dk | 22,909 | 0.38% |
| webspeed.dk | 21,569 | 0.36% |

K. Hageman et al., Gollector: Measuring Domain Name Dark Matter from Different Vantage Points

# Split horizon and data leakage

Table 5: The 10 apex domains with the most observed unique FQDNs in the passive DNS dataset collected from the university network.

| Apex domain | Unique FQDN count | % |
|---|---|---|
| aau.dk | 3,829,837 | 63% |
| googlesyndication.com | 344,058 | 6% |
| technicolor.net | 61,151 | 1.01% |
| cedexis-radar.net | 44,771 | 0.74% |
| sophosxl.net | 39,297 | 0.65% |
| bbsyd.net | 36,758 | 0.61% |
| office.com | 30,215 | 0.50% |
| emnet.dk | 23,540 | 0.39% |
| obelnet.dk | 22,909 | 0.38% |
| webspeed.dk | 21,569 | 0.36% |

18,499 of these seen outside university network

2,813 both seen in- and outside university network

2,300 non-common subdomains

K. Hageman et al., Gollector: Measuring Domain Name Dark Matter from Different Vantage Points

# Split horizon and data leakage

Table 5: The 10 apex domains with the most observed unique FQDNs in the passive DNS dataset collected from the university network.

| Apex domain | Unique FQDN count | % |
|---|---|---|
| aau.dk | 3,829,837 | 63% |
| googlesyndication.com | 344,058 | 6% |
| technicolor.net | 61,151 | 1.01% |
| cedexis-radar.net | 44,771 | 0.74% |
| sophosxl.net | 39,297 | 0.65% |
| bbsyd.net | 36,758 | 0.61% |
| office.com | 30,215 | 0.50% |
| emnet.dk | 23,540 | 0.39% |
| obelnet.dk | 22,909 | 0.38% |
| webspeed.dk | 21,569 | 0.36% |

18,499 of these seen outside university network

2,813 both seen in- and outside university network

2,300 non-common subdomains !

K. Hageman et al., Gollector: Measuring Domain Name Dark Matter from Different Vantage Points

# Subdomain enumeration

*‘Can we use our dataset to intelligently find subdomains under a given apex domain?’*

K. Hageman et al., Gollector: Measuring Domain Name Dark Matter from Different Vantage Points

# Subdomain enumeration

1. Find subdomains that are seen under the same apex domain (graph)
2. Find cliques of graph
3. Generate candidate FQDNs based on these cliques

# Subdomain enumeration

Table 6: Examples of cliques

| Description | Subdomain count | Apex count | Subdomains |
|---|---|---|---|
| High-entropy subdomains | 237 | 2 | `adfqjkxr`, `aeovrpvk`, `anhpfctcxzcp`, `asqzcggxiy`, `bdzvxofezaejku`, ... |
| Email servers | 5 | 34,249 | `imap`, `xwa`, `xas`, `pop`, `smtp` |
| Western language-related subdomains | 7 | 26,730 | `en`, `es`, `fr`, `pt`, `it`, `ru`, `de` |
| More language-related subdomains | 6 | 3,764 | `ko`, `zh`, `cs`, `nl`, `ar`, `ja` |
| Content deliver network | 9 | 5,197 | `cdn-1`, `cdn-3`, `cdn-2`, `cdn-5`, `cdn-7`, ... |

**59% of candidates exist**

K. Hageman et al., Gollector: Measuring Domain Name Dark Matter from Different Vantage Points

# That’s all folks!

**Gollector** is available at
https://github.com/aau-network-security/gollector/

Our experiments show the utility of the tool,
what’s next?