NLnet *Labs*

Who is “really” running email?

Willem Toorop @ MAT Working Group

RIPE 91

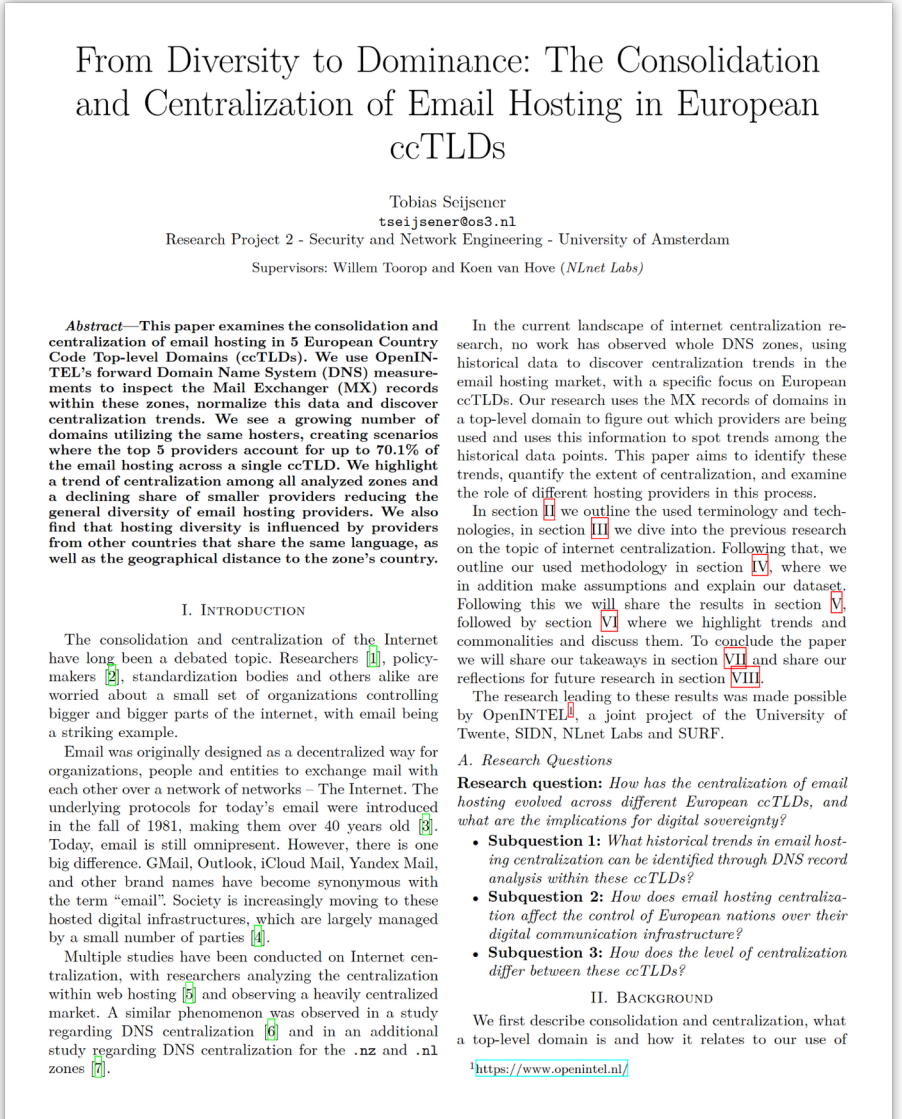
Goals

- Better expose the consolidation and centralization of email hosting
- Showcase vantage point of DNS resolvers, and their value for insightful data
- Collaborating and having fun

WHO IS “REALLY” RUNNING EMAIL?

Motivation

- Tobias Seijsener
  - study historical trends
  - using OpenINTEL data
    `.ch` `.ee` `.fr` `.se` `.sk`
  - national sovereignty



From Diversity to Dominance: The Consolidation and Centralization of Email Hosting in European ccTLDs

Tobias Seijsener
tseijsener@os3.nl
Research Project 2 - Security and Network Engineering - University of Amsterdam
Supervisors: Willem Toorop and Koen van Hove (*NLnet Labs*)

***Abstract*—This paper examines the consolidation and centralization of email hosting in 5 European Country Code Top-level Domains (ccTLDs). We use OpenINTEL's forward Domain Name System (DNS) measurements to inspect the Mail Exchanger (MX) records within these zones, normalize this data and discover centralization trends. We see a growing number of domains utilizing the same hosters, creating scenarios where the top 5 providers account for up to 70.1% of the email hosting across a single ccTLD. We highlight a trend of centralization among all analyzed zones and a declining share of smaller providers reducing the general diversity of email hosting providers. We also find that hosting diversity is influenced by providers from other countries that share the same language, as well as the geographical distance to the zone's country.**

I. Introduction

The consolidation and centralization of the Internet have long been a debated topic. Researchers [1], policy-makers [2], standardization bodies and others alike are worried about a small set of organizations controlling bigger and bigger parts of the internet, with email being a striking example.

Email was originally designed as a decentralized way for organizations, people and entities to exchange mail with each other over a network of networks – The Internet. The underlying protocols for today's email were introduced in the fall of 1981, making them over 40 years old [3]. Today, email is still omnipresent. However, there is one big difference. GMail, Outlook, iCloud Mail, Yandex Mail, and other brand names have become synonymous with the term “email”. Society is increasingly moving to these hosted digital infrastructures, which are largely managed by a small number of parties [4].

Multiple studies have been conducted on Internet centralization, with researchers analyzing the centralization within web hosting [5] and observing a heavily centralized market. A similar phenomenon was observed in a study regarding DNS centralization [6] and in an additional study regarding DNS centralization for the `.nz` and `.nl` zones [7].

In the current landscape of internet centralization research, no work has observed whole DNS zones, using historical data to discover centralization trends in the email hosting market, with a specific focus on European ccTLDs. Our research uses the MX records of domains in a top-level domain to figure out which providers are being used and uses this information to spot trends among the historical data points. This paper aims to identify these trends, quantify the extent of centralization, and examine the role of different hosting providers in this process.

In section II we outline the used terminology and technologies, in section III we dive into the previous research on the topic of internet centralization. Following that, we outline our used methodology in section IV, where we in addition make assumptions and explain our dataset. Following this we will share the results in section V, followed by section VI where we highlight trends and commonalities and discuss them. To conclude the paper we will share our takeaways in section VII and share our reflections for future research in section VIII.

The research leading to these results was made possible by OpenINTEL[1], a joint project of the University of Twente, SIDN, NLnet Labs and SURF.

A. Research Questions

Research question: *How has the centralization of email hosting evolved across different European ccTLDs, and what are the implications for digital sovereignty?*

- **Subquestion 1:** *What historical trends in email hosting centralization can be identified through DNS record analysis within these ccTLDs?*
- **Subquestion 2:** *How does email hosting centralization affect the control of European nations over their digital communication infrastructure?*
- **Subquestion 3:** *How does the level of centralization differ between these ccTLDs?*

II. Background

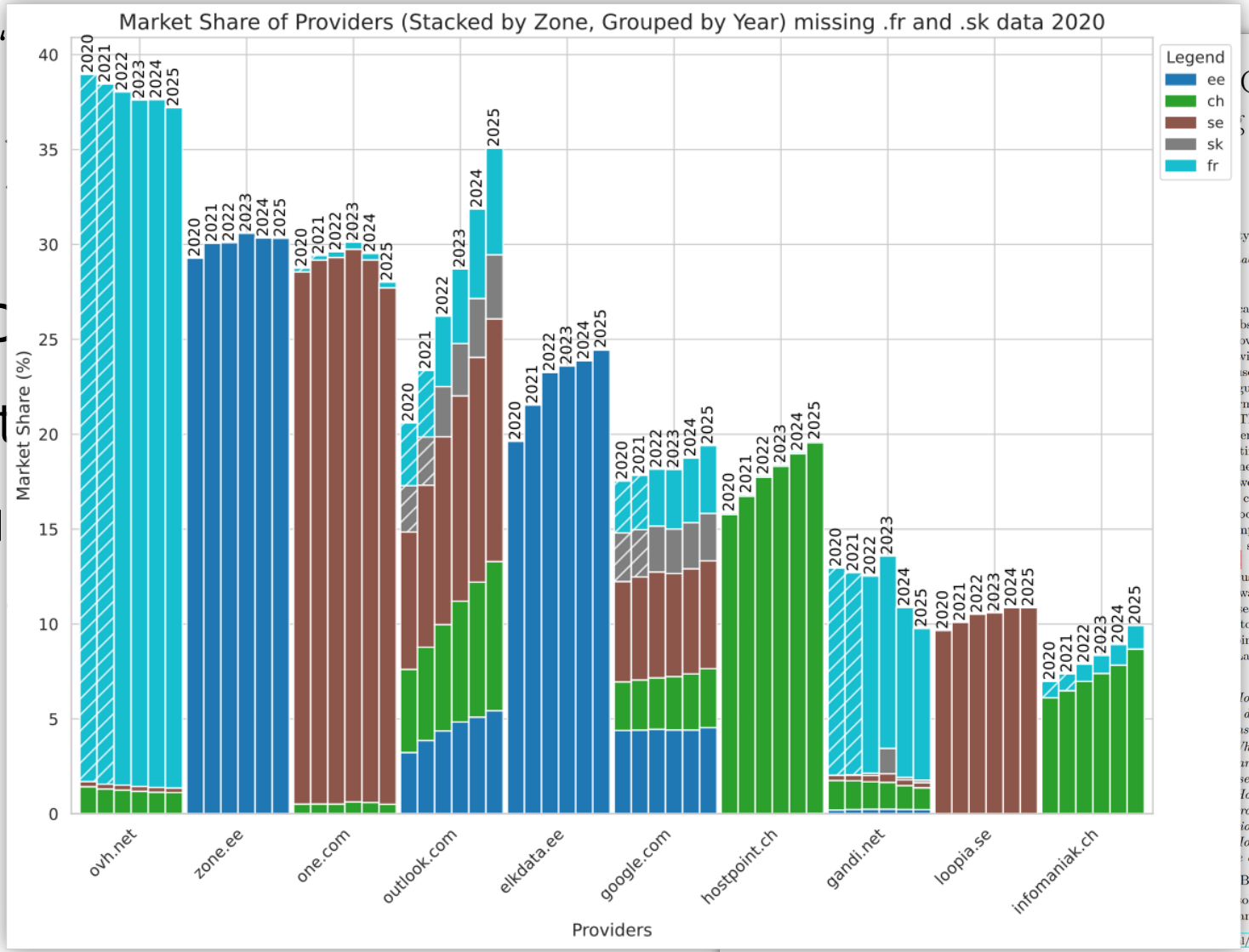
We first describe consolidation and centralization, what a top-level domain is and how it relates to our use of

[1] https://www.openintel.nl/

WHO IS "

Mot

- Tob
  - st
  - u
  -



Consolidation
in European

y of Amsterdam
abs)

ape of internet centralization re-
bserved whole DNS zones, using
over centralization trends in the
vith a specific focus on European
ses the MX records of domains in
gure out which providers are being
rmation to spot trends among the
This paper aims to identify these
ent of centralization, and examine
ting providers in this process.
he used terminology and tech-
we dive into the previous research
centralization. Following that, we
odology in section IV, where we
nptions and explain our dataset.
share the results in section V,
where we highlight trends and
uss them. To conclude the paper
ways in section VII and share our
earch in section VIII.
to these results was made possible
int project of the University of
abs and SURF.

low has the centralization of email different European ccTLDs, and s for digital sovereignty?
What historical trends in email hosting n be identified through DNS record se ccTLDs?
low does email hosting centralization rol of European nations over their ion infrastructure?
low does the level of centralization ccTLDs?

BACKGROUND

olidation and centralization, what nd how it relates to our use of

Motivation

> “ V. Results
>
> In each ccTLD studied, the top provider is consistently a local one from the respective country.

> “ VII. Conclusion
>
> The degree of centralization and providers involved are less non-European than hypothesised

> “ VIII. Future Work
>
> What this method fails to take into consideration is the actual traffic to these MX servers

From Diversity to Dominance: The Consolidation and Centralization of Email Hosting in European ccTLDs

Tobias Seijsener
tseijsener@os3.nl
Research Project 2 - Security and Network Engineering - University of Amsterdam
Supervisors: Willem Toorop and Koen van Hove (*NLnet Labs*)

***Abstract*—This paper examines the consolidation and centralization of email hosting in 5 European Country Code Top-level Domains (ccTLDs). We use OpenINTEL's forward Domain Name System (DNS) measurements to inspect the Mail Exchanger (MX) records within these zones, normalize this data and discover centralization trends. We see a growing number of domains utilizing the same hosters, creating scenarios where the top 5 providers account for up to 70.1% of the email hosting across a single ccTLD. We highlight a trend of centralization among all analyzed zones and a declining share of smaller providers reducing the general diversity of email hosting providers. We also find that hosting diversity is influenced by providers from other countries that share the same language, as well as the geographical distance to the zone's country.**

I. Introduction

The consolidation and centralization of the Internet have long been a debated topic. Researchers [1], policy-makers [2], standardization bodies and others alike are worried about a small set of organizations controlling bigger and bigger parts of the internet, with email being a striking example.

Email was originally designed as a decentralized way for organizations, people and entities to exchange mail with each other over a network of networks – The Internet. The underlying protocols for today's email were introduced in the fall of 1981, making them over 40 years old [3]. Today, email is still omnipresent. However, there is one big difference. GMail, Outlook, iCloud Mail, Yandex Mail, and other brand names have become synonymous with the term “email”. Society is increasingly moving to these hosted digital infrastructures, which are largely managed by a small number of parties [4].

Multiple studies have been conducted on Internet centralization, with researchers analyzing the centralization within web hosting [5] and observing a heavily centralized market. A similar phenomenon was observed in a study regarding DNS centralization [6] and in an additional study regarding DNS centralization for the `.nz` and `.nl` zones [7].

In the current landscape of internet centralization research, no work has observed whole DNS zones, using historical data to discover centralization trends in the email hosting market, with a specific focus on European ccTLDs. Our research uses the MX records of domains in a top-level domain to figure out which providers are being used and uses this information to spot trends among the historical data points. This paper aims to identify these trends, quantify the extent of centralization, and examine the role of different hosting providers in this process.

In section II we outline the used terminology and technologies, in section III we dive into the previous research on the topic of internet centralization. Following that, we outline our used methodology in section IV, where we in addition make assumptions and explain our dataset. Following this we will share the results in section V, followed by section VI where we highlight trends and commonalities and discuss them. To conclude the paper we will share our takeaways in section VII and share our reflections for future research in section VIII.

The research leading to these results was made possible by OpenINTEL[1], a joint project of the University of Twente, SIDN, NLnet Labs and SURF.

A. Research Questions

Research question: *How has the centralization of email hosting evolved across different European ccTLDs, and what are the implications for digital sovereignty?*

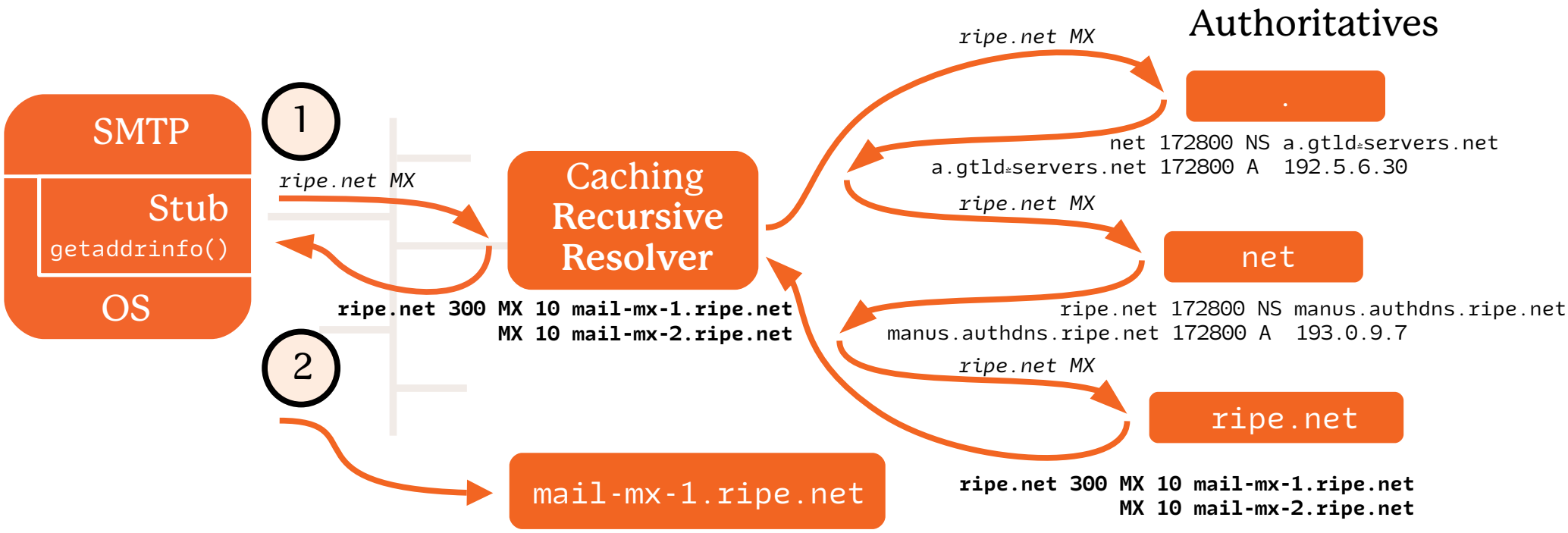
- **Subquestion 1:** *What historical trends in email hosting centralization can be identified through DNS record analysis within these ccTLDs?*
- **Subquestion 2:** *How does email hosting centralization affect the control of European nations over their digital communication infrastructure?*
- **Subquestion 3:** *How does the level of centralization differ between these ccTLDs?*

II. Background

We first describe consolidation and centralization, what a top-level domain is and how it relates to our use of

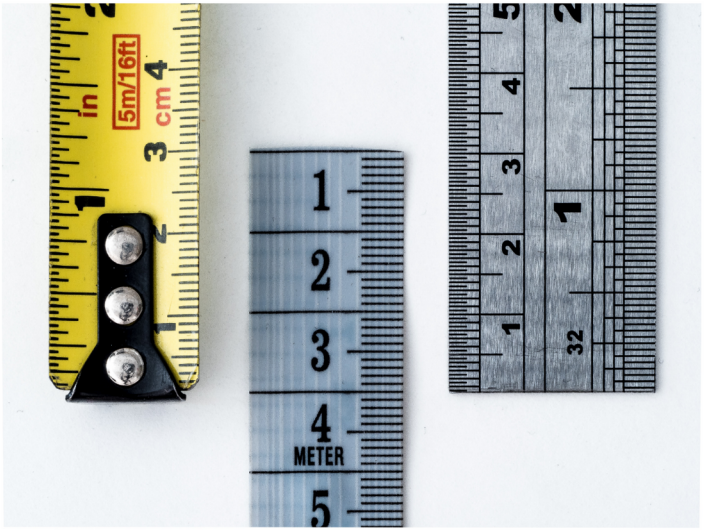
[1] https://www.openintel.nl/

Method


SMTP
Stub
getaddrinfo()
OS
1
ripe.net MX
ripe.net 300 MX 10 mail-mx-1.ripe.net
MX 10 mail-mx-2.ripe.net
2
Caching Recursive Resolver
mail-mx-1.ripe.net
Authoritatives
ripe.net MX
.
net 172800 NS a.gtld-servers.net
a.gtld-servers.net 172800 A 192.5.6.30
ripe.net MX
net
ripe.net 172800 NS manus.authdns.ripe.net
manus.authdns.ripe.net 172800 A 193.0.9.7
ripe.net MX
ripe.net
ripe.net 300 MX 10 mail-mx-1.ripe.net
MX 10 mail-mx-2.ripe.net

Method

> *“The total number of signed domain names or the number of validating resolvers give a distorted view, if widely-used domain names are not protected and popular resolvers are not validating. We therefore propose a metrics focused on the number of transactions protected with DNSSEC.*



DNSSEC Deployment Metrics Research

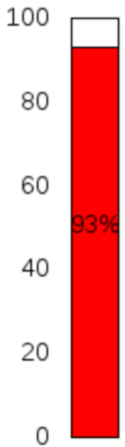
2022/08/08

MORITZ MÜLLER, SIDN Labs
JELTE JANSEN, SIDN Labs
MARCO DAVIDS, SIDN Labs
WILLEM TOOROP, NLnet Labs

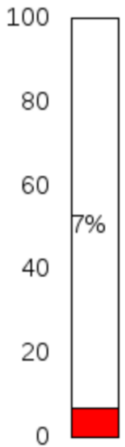
DNSSEC Deployment Report

Sun Oct 19 02:56:58 PDT 2025
Total TLDs: 1439 / Signed TLDs in root: 1346 / Recently added: et. (09/20/2025)

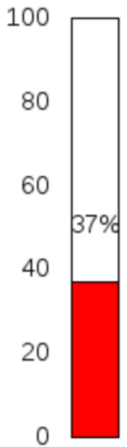
% of TLDs signed in root



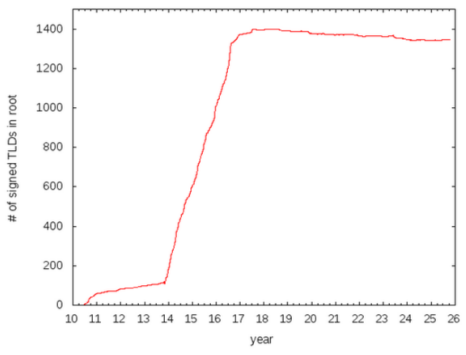
Approx % of these 2LDs signed



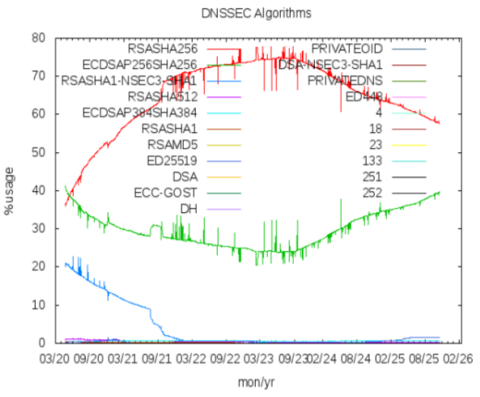
Approx* % Users Validating



Trend



Algorithm Trends



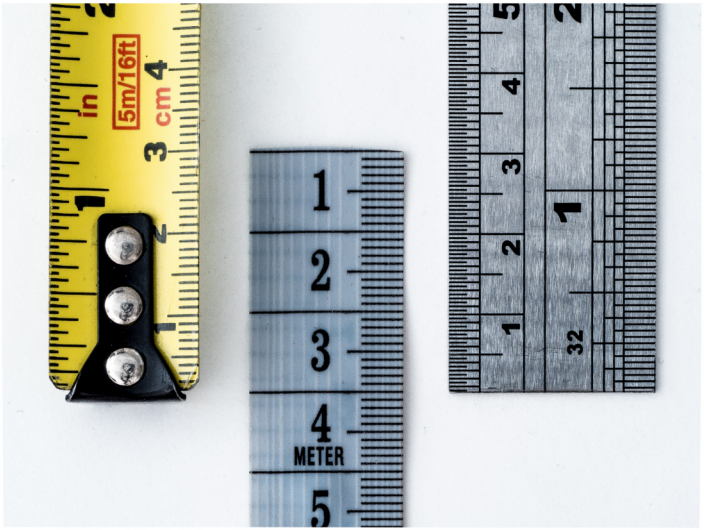
ccTLD Map

*From http://stats.labs.apnic.net/dnssec Some tools: http://www.co.tt keys/algs DEWS% DEWS

20200420-Took an hour out of C19 downtime to add code to pull CZDS. You will see more TLDs with AlgNo:Count. HLL still used for unavailable. % 2LD estimate unchanged. more

Method

> *Ideally, we recommend collecting this metric directly on recursive resolvers*

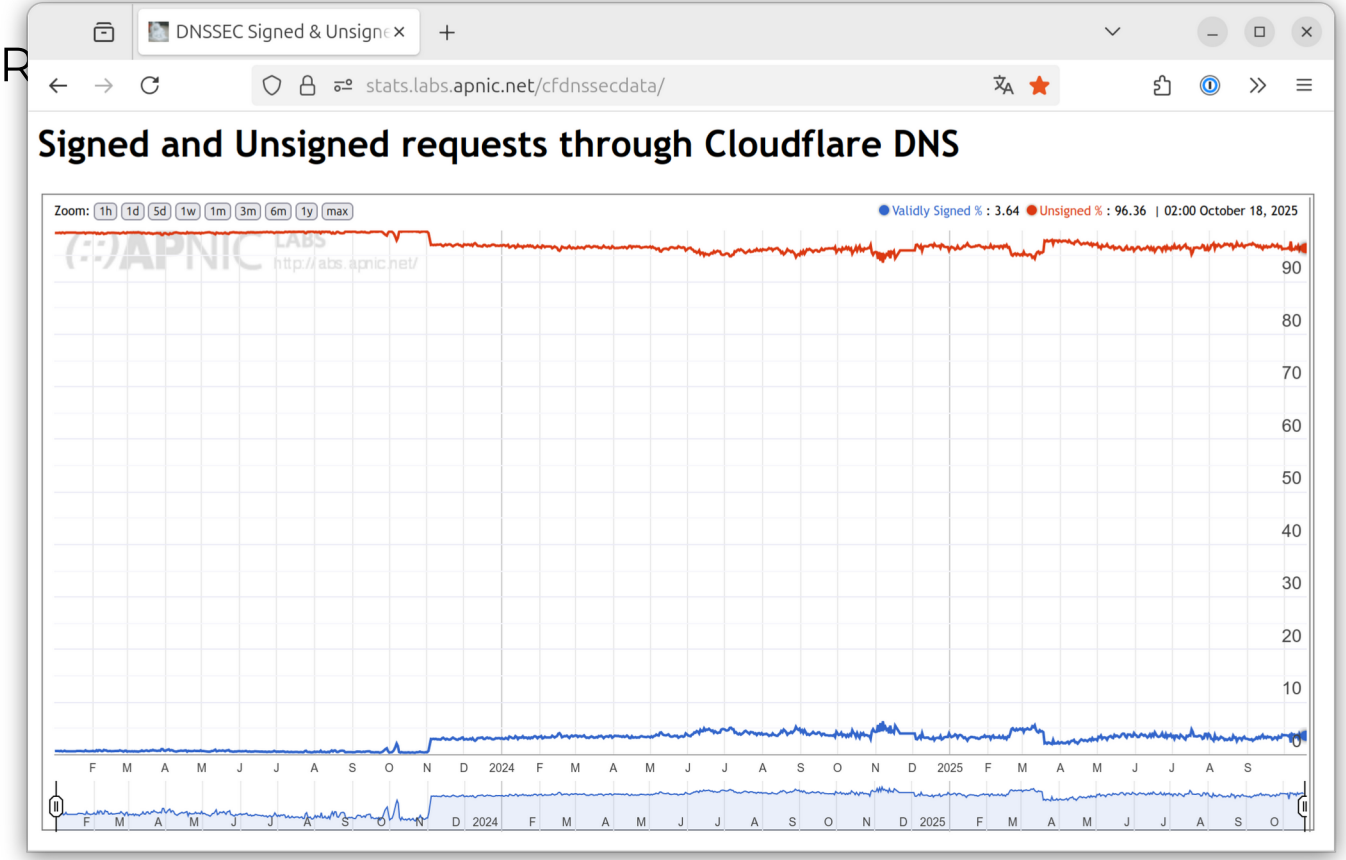


DNSSEC Deployment Metrics Research

2022/08/08

MORITZ MÜLLER, SIDN Labs
JELTE JANSEN, SIDN Labs
MARCO DAVIDS, SIDN Labs
WILLEM TOOROP, NLnet Labs

Method



- 3.64%
- APNIC measures this using 1.1.1.1 resolver
- Presented @ OARC 41 by Joao Damas

Method

- quad9
- How Quad9 Handles Your Data | 2025

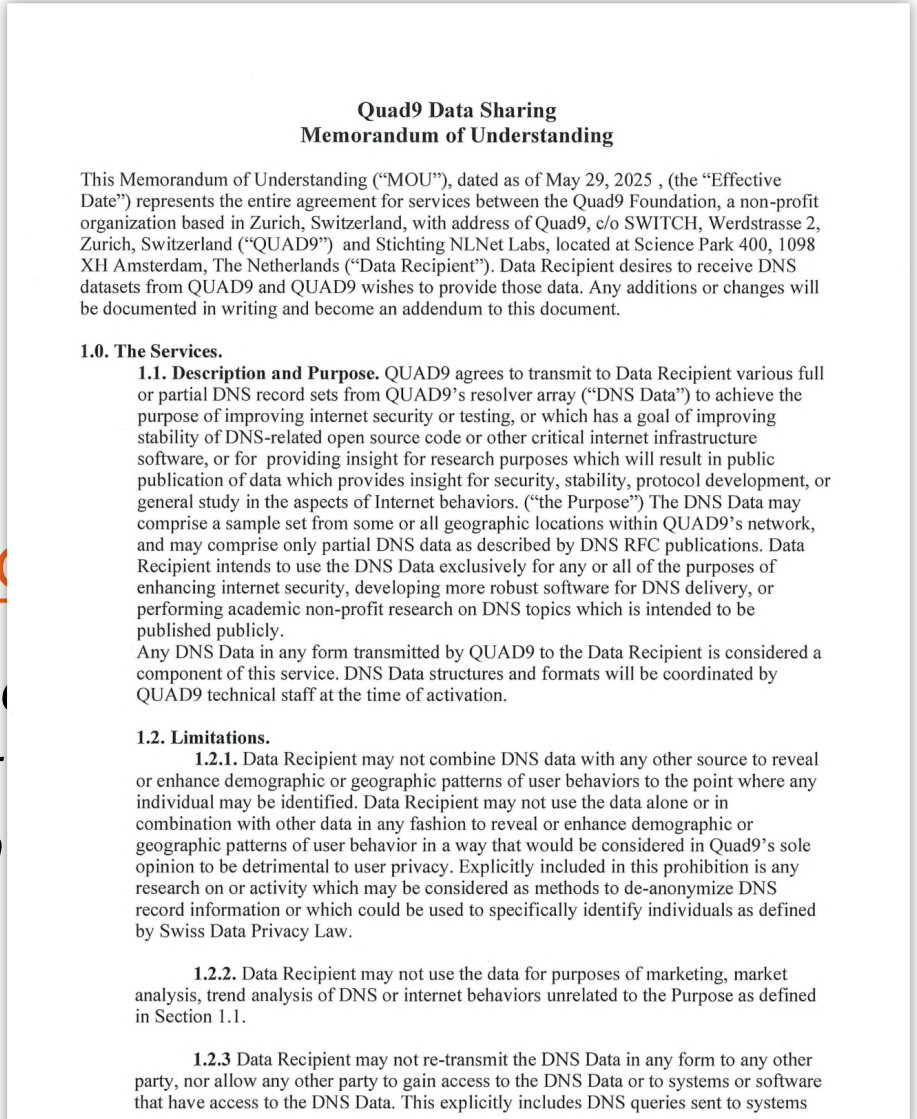
> *“We do share other aggregated data (again, no personal data) with some partners or researchers for the express purposes of improving security and performance of the DNS.*

WHO IS “REALLY” RUNNING EMAIL?

Method

- quad9
- How Quad9 Handles Y

“*We do share other aggregate personal data) with some par for the express purposes of im performance of the DNS.*



Quad9 Data Sharing
Memorandum of Understanding

This Memorandum of Understanding (“MOU”), dated as of May 29, 2025 , (the “Effective Date”) represents the entire agreement for services between the Quad9 Foundation, a non-profit organization based in Zurich, Switzerland, with address of Quad9, c/o SWITCH, Werdstrasse 2, Zurich, Switzerland (“QUAD9”) and Stichting NLNet Labs, located at Science Park 400, 1098 XH Amsterdam, The Netherlands (“Data Recipient”). Data Recipient desires to receive DNS datasets from QUAD9 and QUAD9 wishes to provide those data. Any additions or changes will be documented in writing and become an addendum to this document.

1.0. The Services.

1.1. Description and Purpose. QUAD9 agrees to transmit to Data Recipient various full or partial DNS record sets from QUAD9’s resolver array (“DNS Data”) to achieve the purpose of improving internet security or testing, or which has a goal of improving stability of DNS-related open source code or other critical internet infrastructure software, or for providing insight for research purposes which will result in public publication of data which provides insight for security, stability, protocol development, or general study in the aspects of Internet behaviors. (“the Purpose”) The DNS Data may comprise a sample set from some or all geographic locations within QUAD9’s network, and may comprise only partial DNS data as described by DNS RFC publications. Data Recipient intends to use the DNS Data exclusively for any or all of the purposes of enhancing internet security, developing more robust software for DNS delivery, or performing academic non-profit research on DNS topics which is intended to be published publicly.
Any DNS Data in any form transmitted by QUAD9 to the Data Recipient is considered a component of this service. DNS Data structures and formats will be coordinated by QUAD9 technical staff at the time of activation.

1.2. Limitations.

1.2.1. Data Recipient may not combine DNS data with any other source to reveal or enhance demographic or geographic patterns of user behaviors to the point where any individual may be identified. Data Recipient may not use the data alone or in combination with other data in any fashion to reveal or enhance demographic or geographic patterns of user behavior in a way that would be considered in Quad9’s sole opinion to be detrimental to user privacy. Explicitly included in this prohibition is any research on or activity which may be considered as methods to de-anonymize DNS record information or which could be used to specifically identify individuals as defined by Swiss Data Privacy Law.

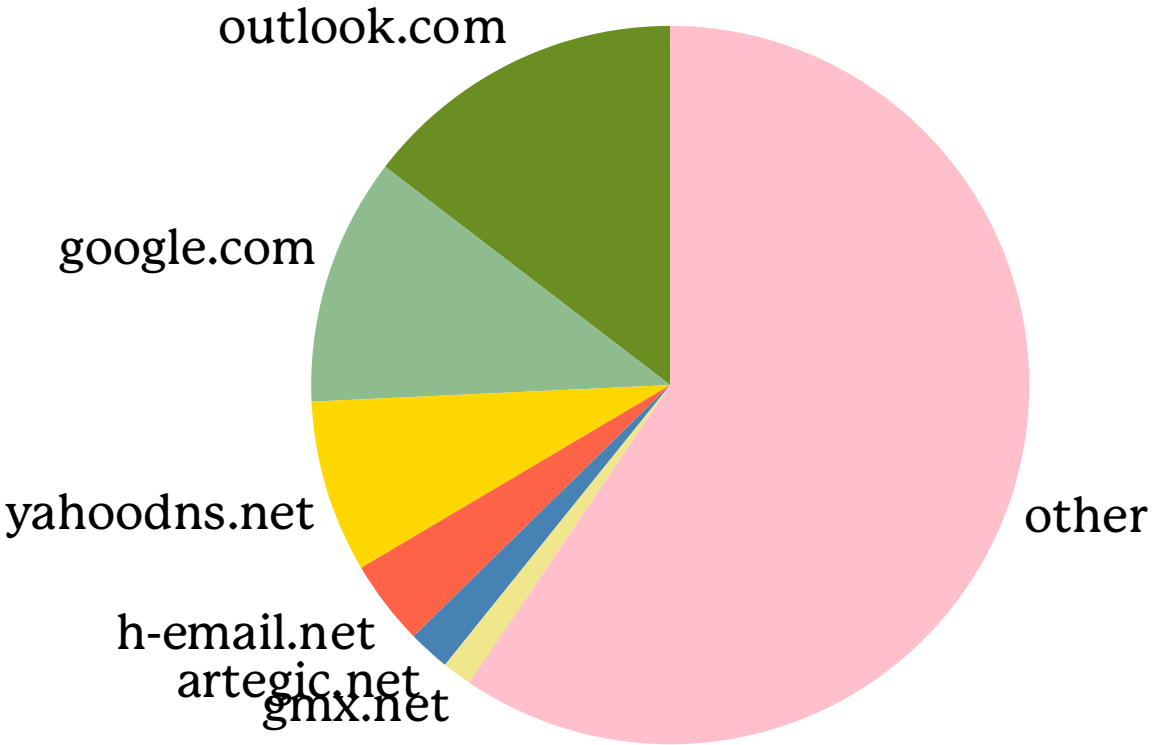
1.2.2. Data Recipient may not use the data for purposes of marketing, market analysis, trend analysis of DNS or internet behaviors unrelated to the Purpose as defined in Section 1.1.

1.2.3 Data Recipient may not re-transmit the DNS Data in any form to any other party, nor allow any other party to gain access to the DNS Data or to systems or software that have access to the DNS Data. This explicitly includes DNS queries sent to systems

Method

```
{ "geoip"        : { "country" : { "iso_code" : "NL" } },
, "pop_code"     : "ams"
, "requestData"  : { "time" : 17608714679789908O
                   , "timePrecision" : "ns" }
, "responseData" :
  { "answers" :
    [ { "class"      : "IN"
      , "domainName" : "*****"
      , "rData"      : "10 *****.mail.protection.outlook.com."
      , "recordType" : "MX"
      , "ttl"        : 60 }, … ]
, …
, "sample_rate"  : 20 }
```

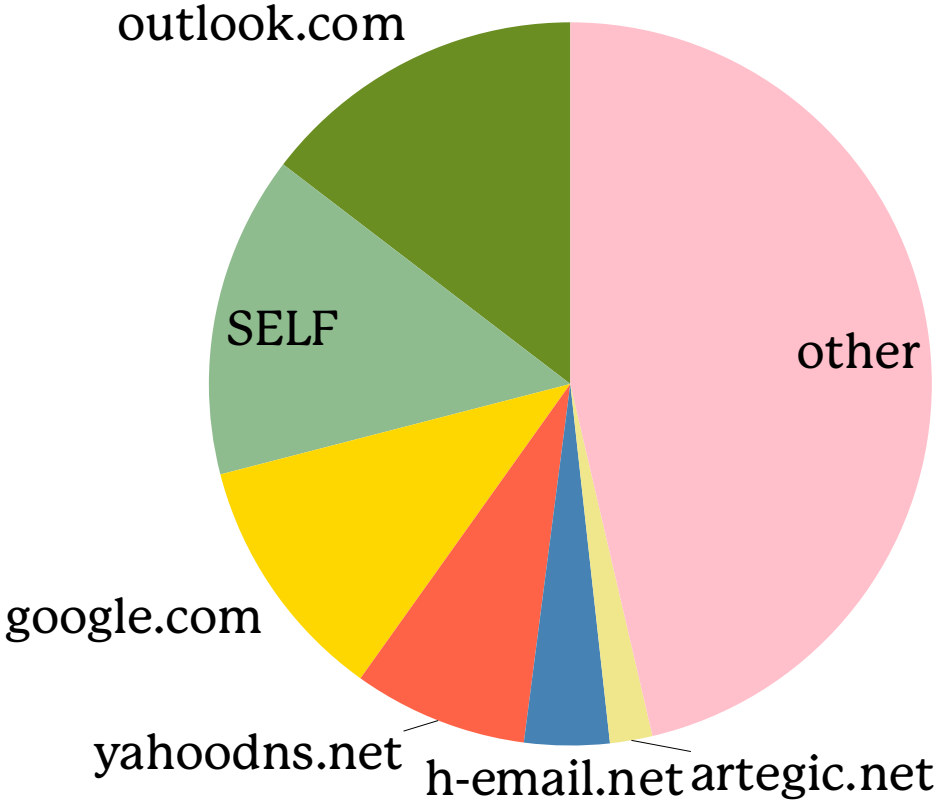
Top mostly used MX hosts



outlook.com	14.6%
google.com	11.1%
yahoodns.net	7.8%
h-email.net	3.8%
artegic.net	1.9%
gmx.net	1.3%
other	59.5%

RESULTS

Top mostly used MX hosts

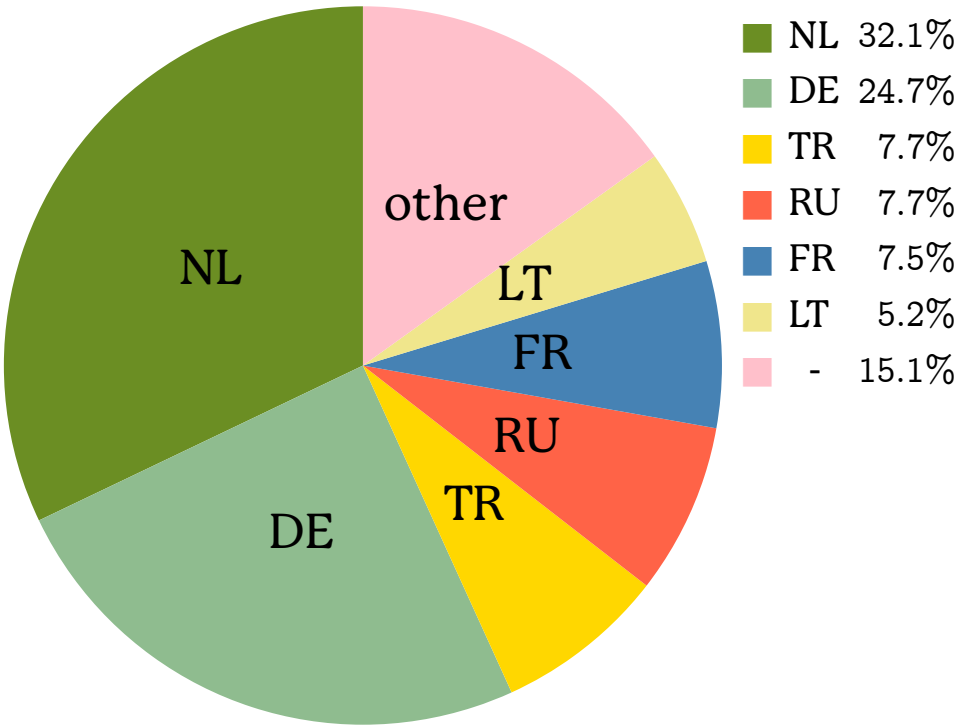


outlook.com	14.6%
SELF	14.4%
google.com	11.1%
yahoodns.net	7.8%
h-email.net	3.8%
artegic.net	1.9%
other	46.3%

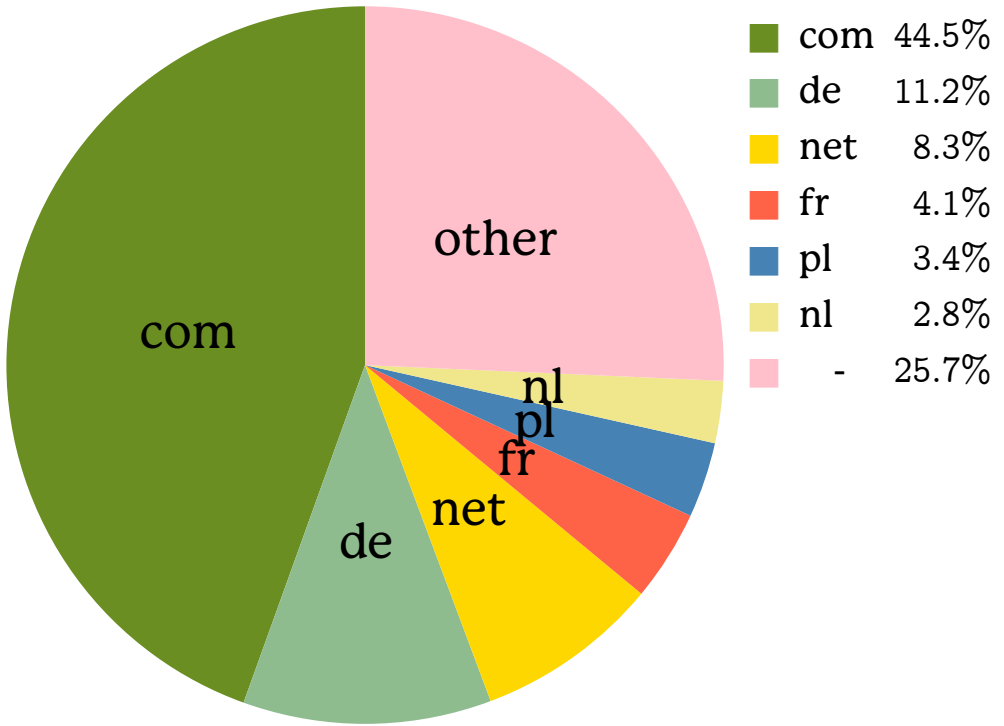
SELF	
saava.co.ke	0.6%
webcfs07.com	0.2%
icrtravel.com	0.2%
yingjian.cc	0.2%
focuspoint.tech	0.1%
andro.ovh	0.1%
zarzadca.biz	0.1%
laposte.net	0.1%
appleid.com	0.1%
openrainbow.net	0.1%
appway.biz	0.1%
qwer12.top	0.0%
other	13.0%

Top source and target localities

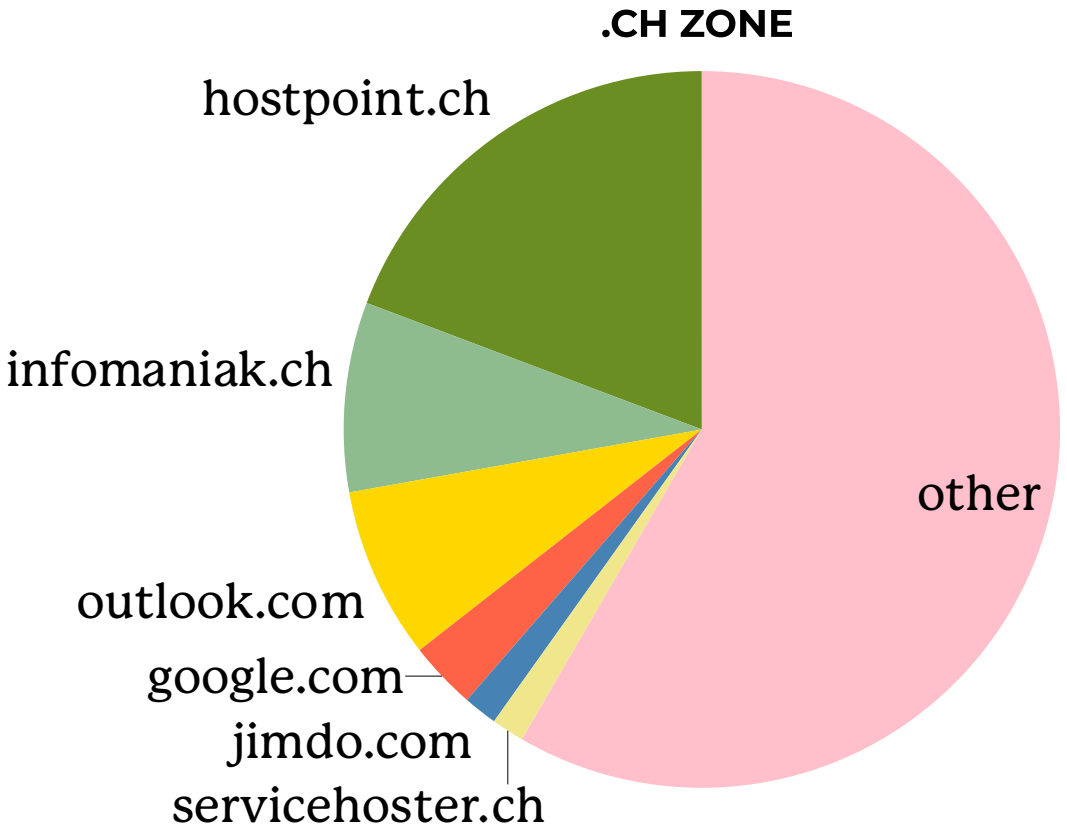
By GeoIP

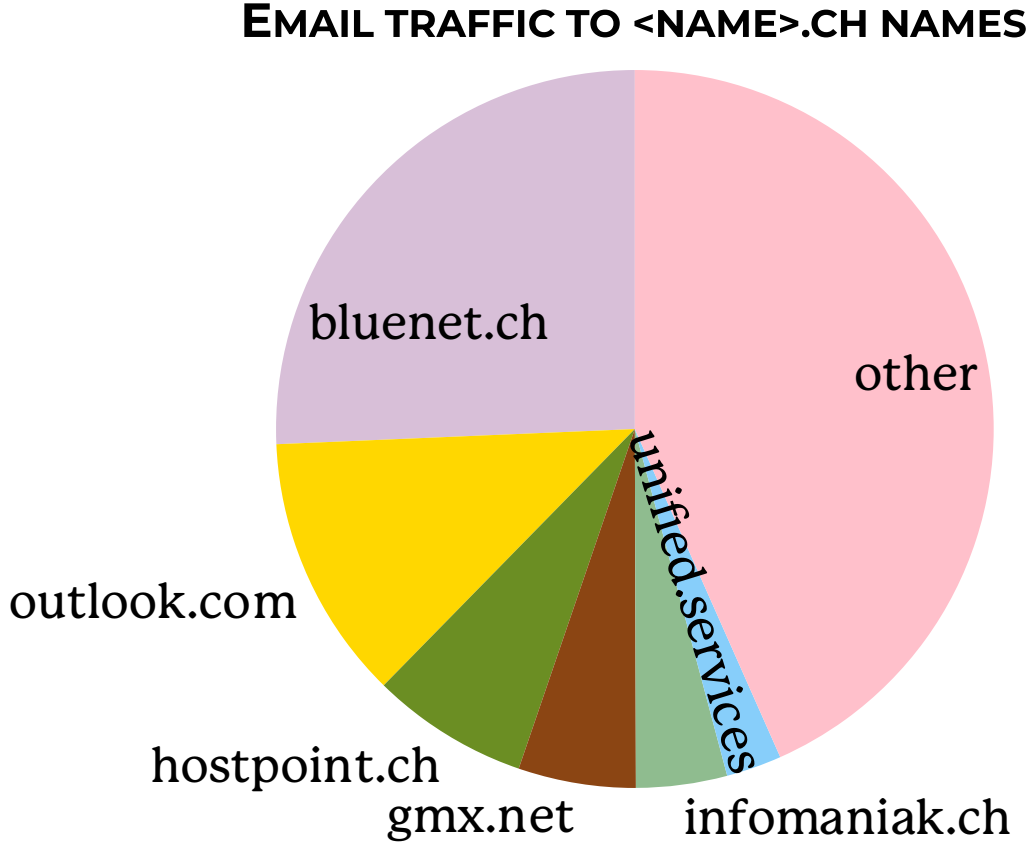


By TLD

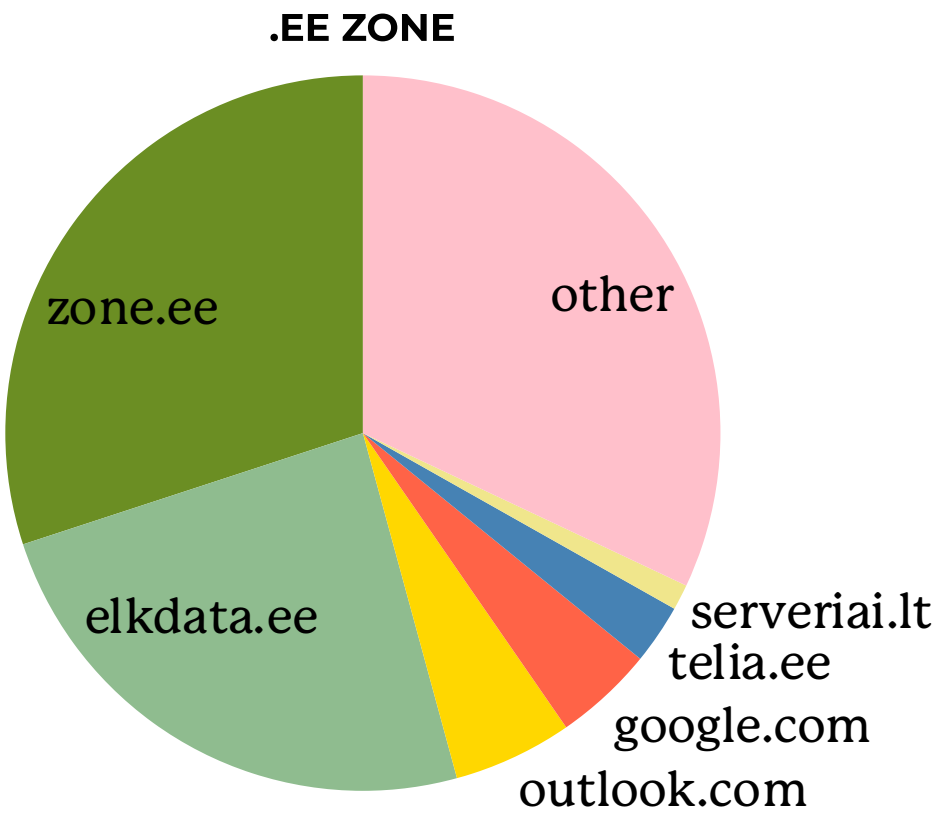


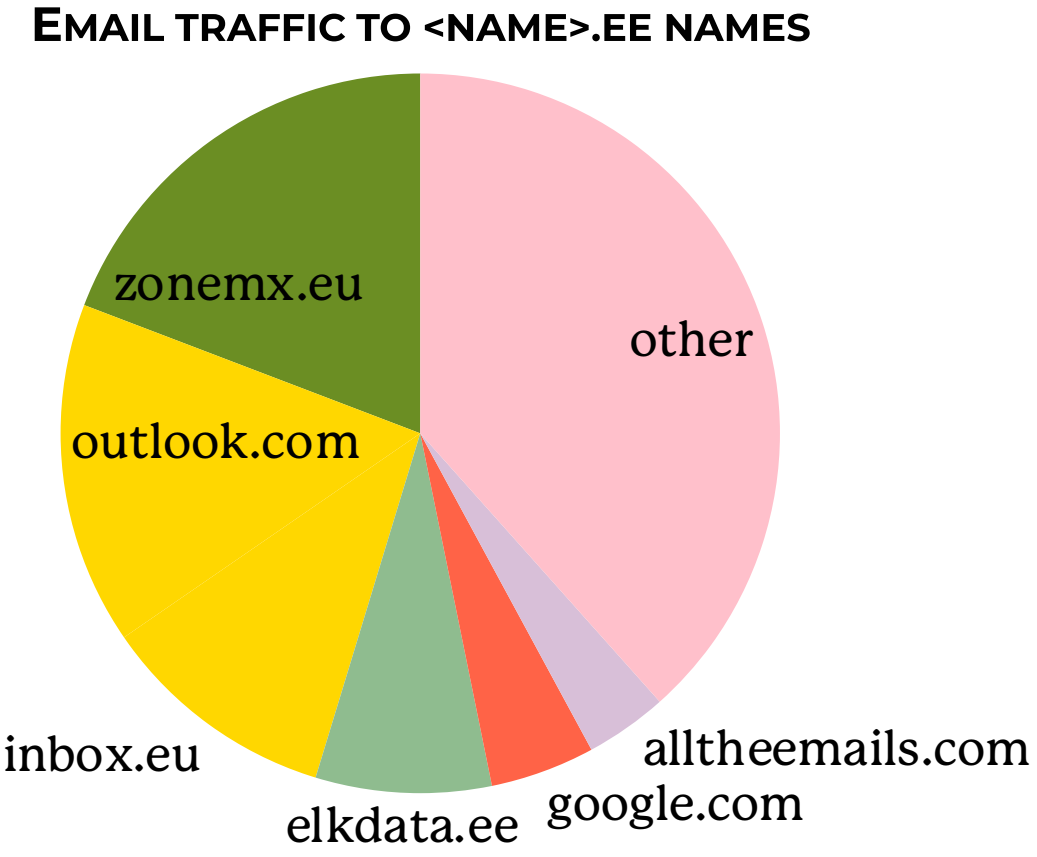
Comparing with Tobias' results





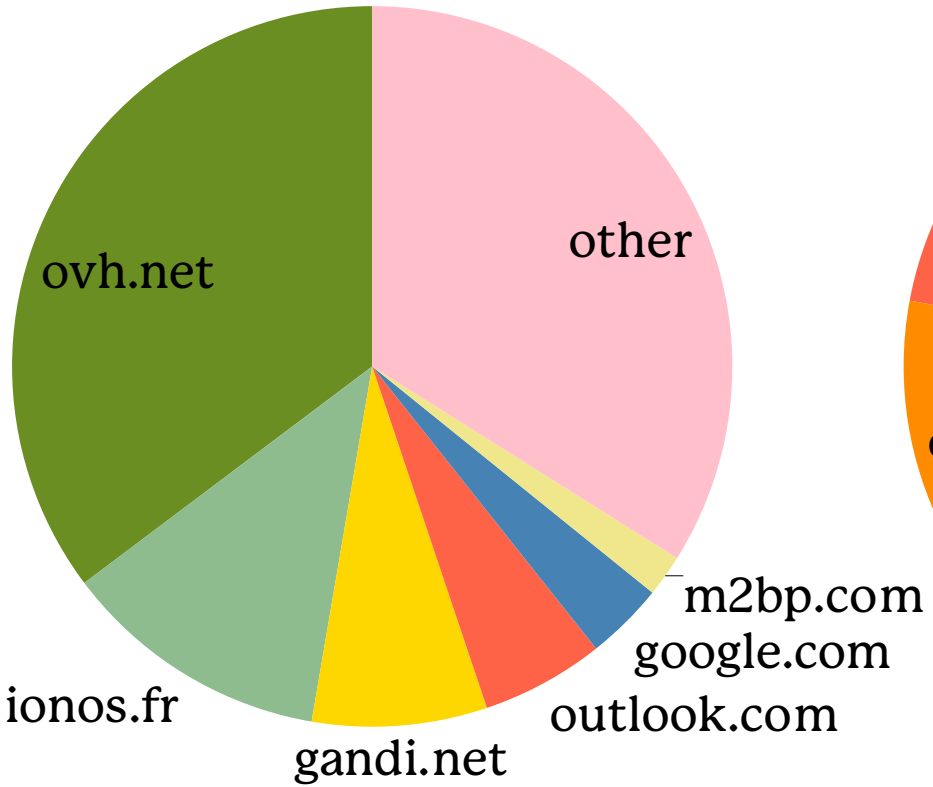
Comparing with Tobias' results



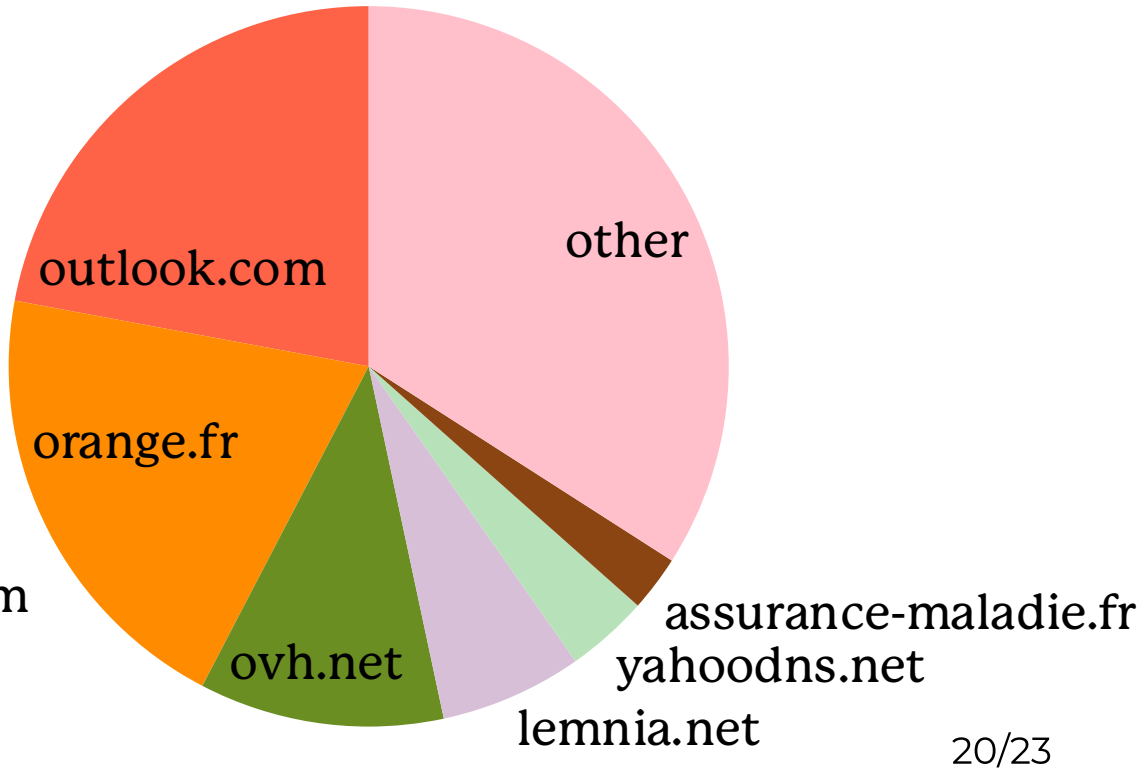


Comparing with Tobias' results

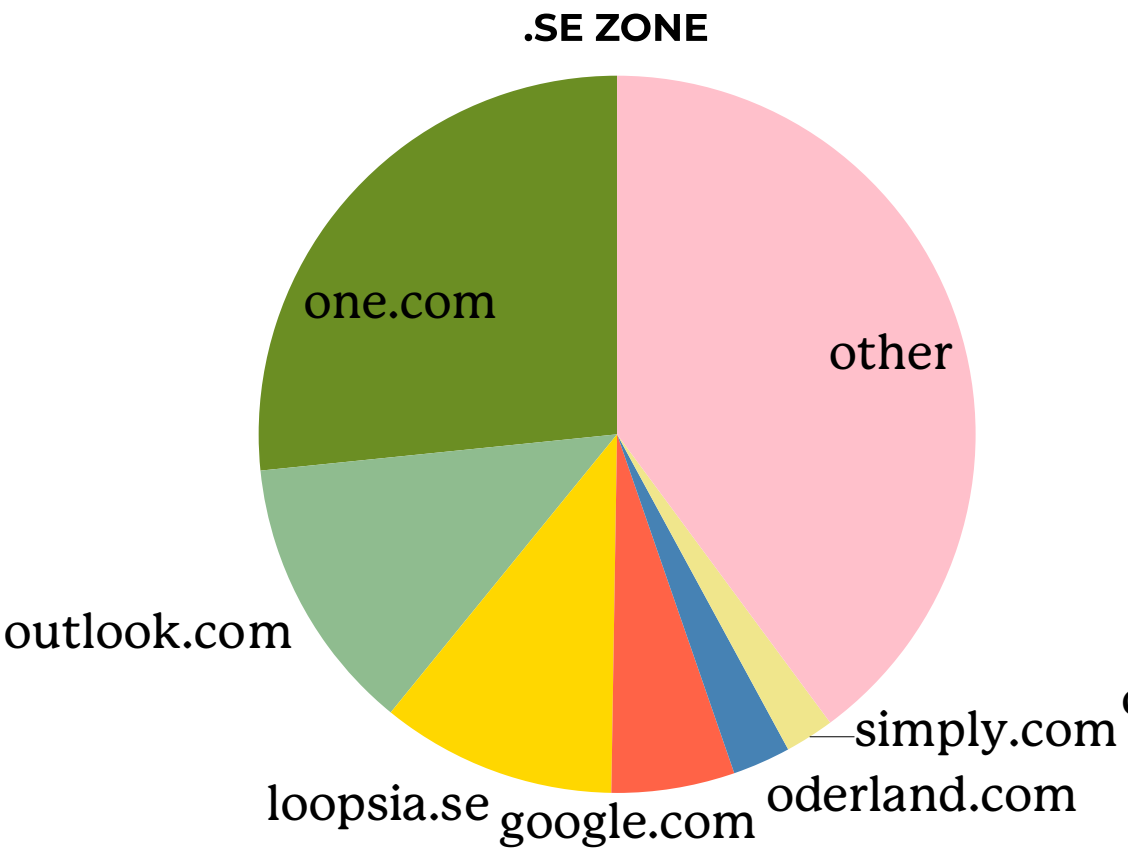
.FR ZONE

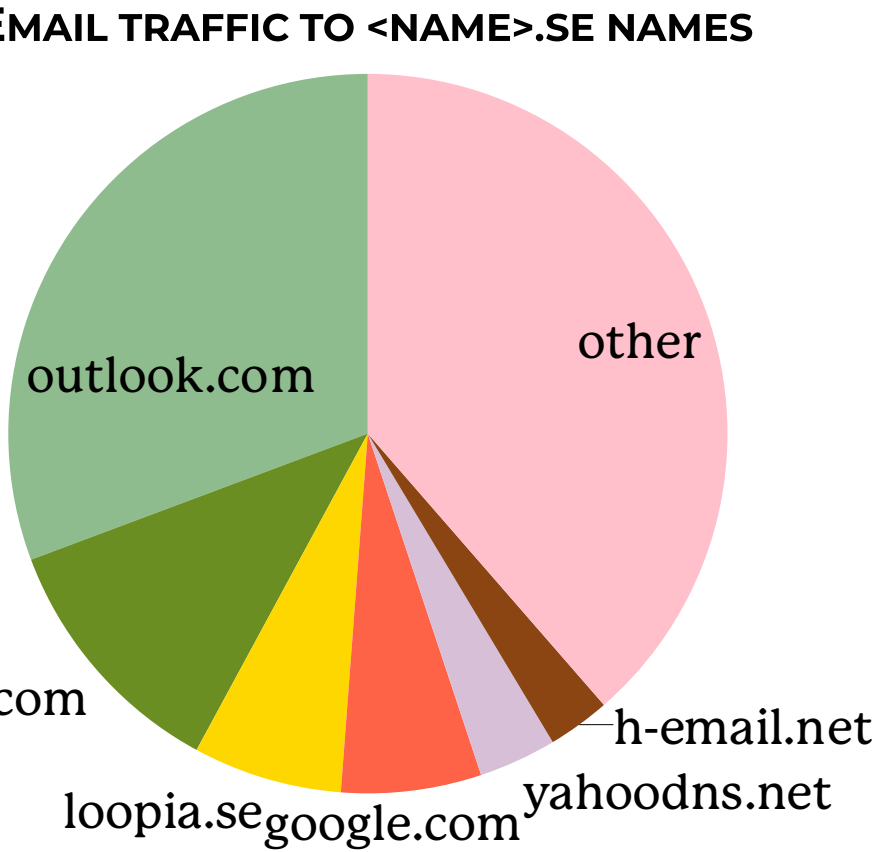


EMAIL TRAFFIC TO <NAME>.FR NAMES

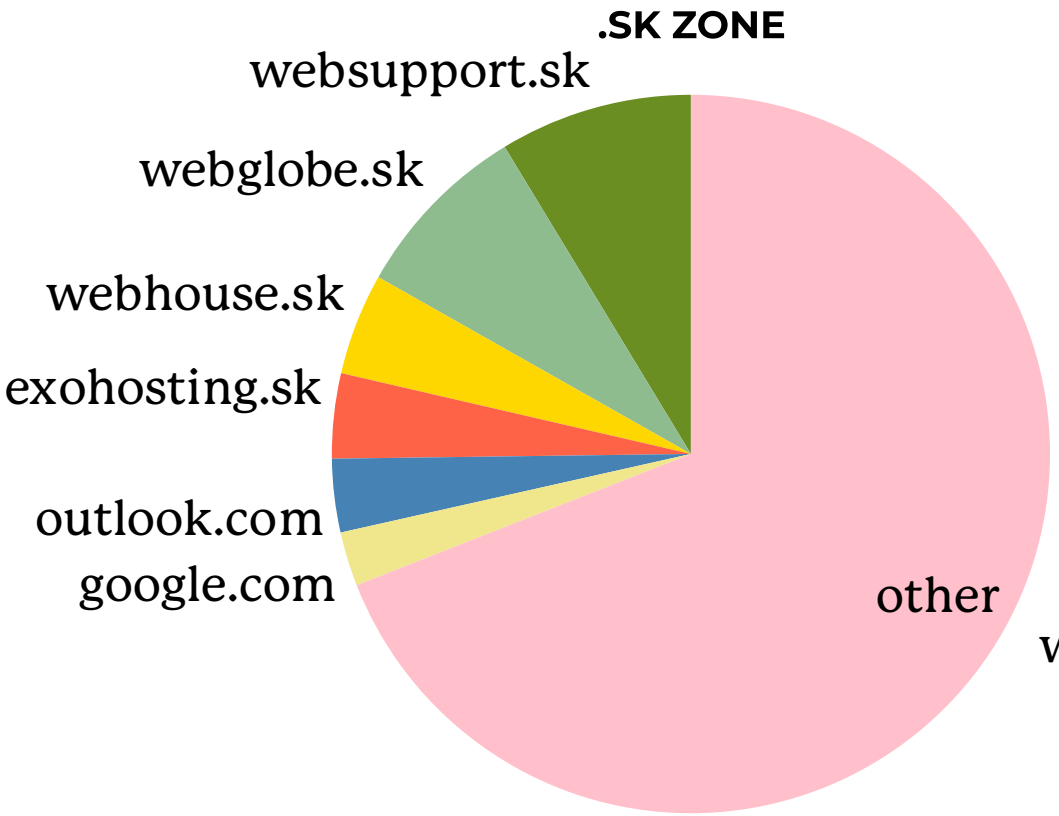


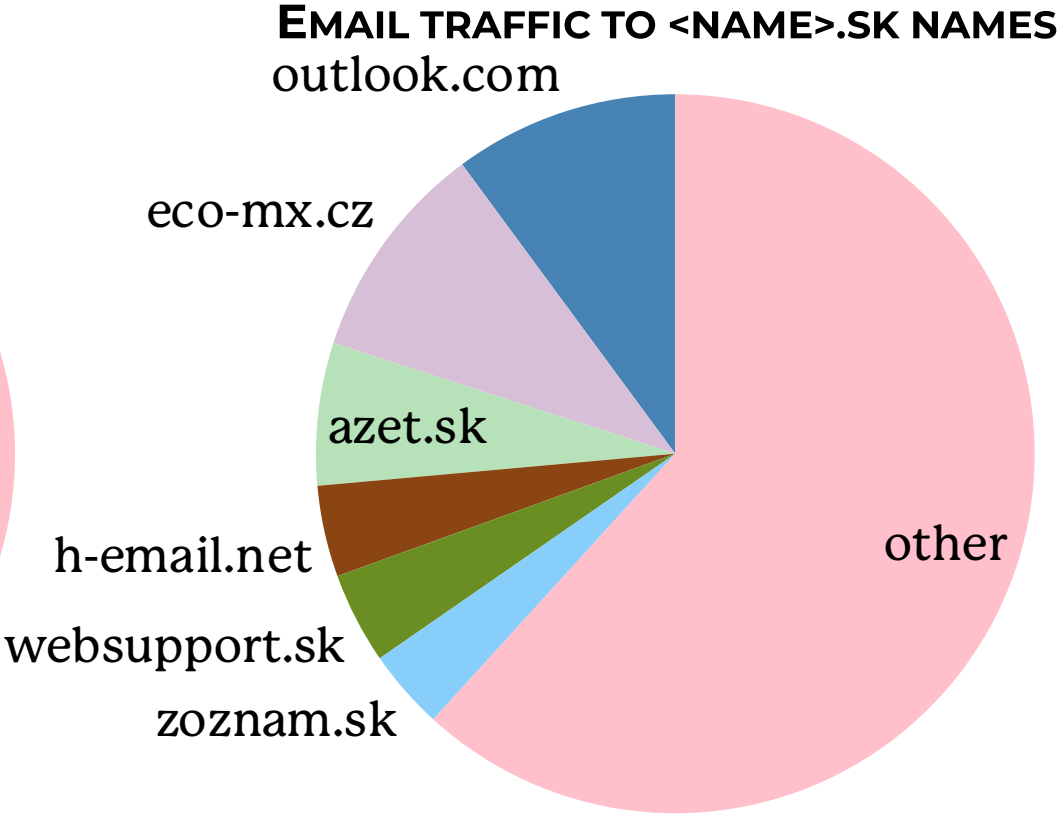
Comparing with Tobias' results





Comparing with Tobias' results





Who is actually really running email?

- An email service from a locality may be registered in any TLD
- MX registrations in a TLD are not necessarily from that TLD's locality
- A WHOIS or RDAP of the MX query name might be better … but cumbersome

NLnet *Labs*

Who is *actually really* running email?

Willem Toorop @ future RIPE meeting

? or !