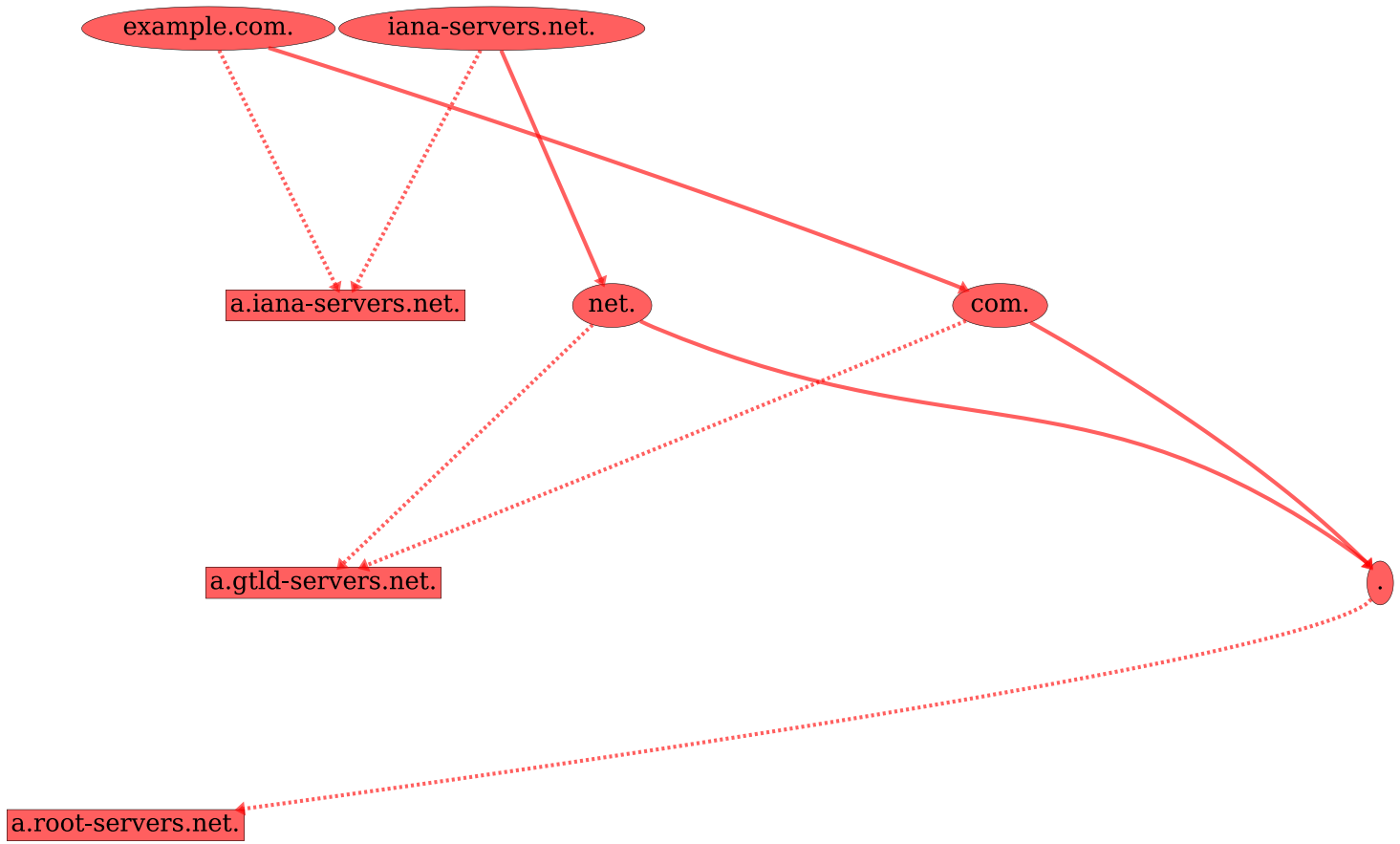
A Dataset of Comprehensive DNS Resolutions

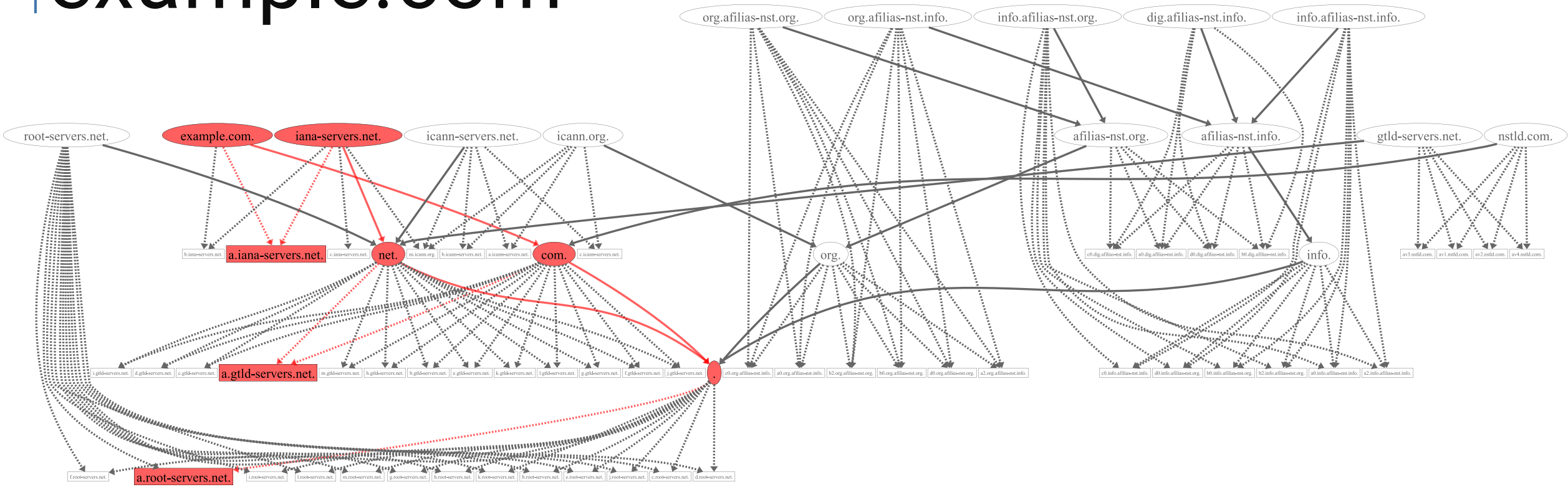
Florian Steurer
PhD Student
Max Planck Institute for Informatics

OARC 43

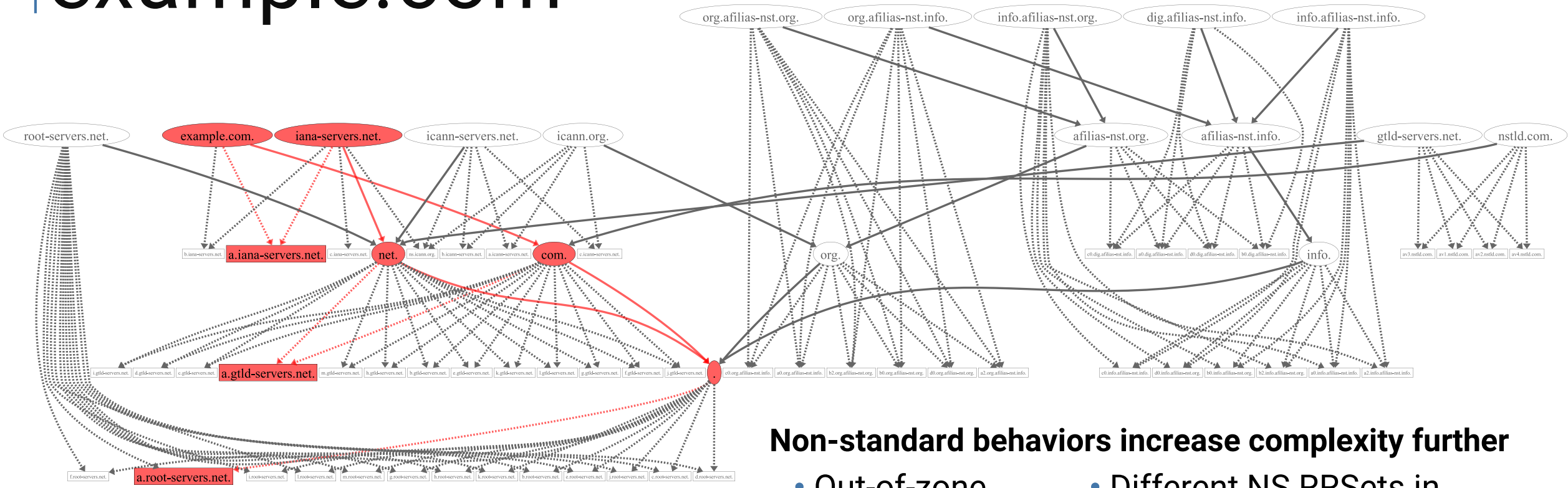
One Possible Resolution of `example.com`



All Possible Resolutions of example.com



All Possible Resolutions of example.com



Non-standard behaviors increase complexity further

- Out-of-zone GLUE
- NS is CNAME
- CNAME at apex
- Different NS RRSets in parent and apex
- Multiple CNAMES
- …

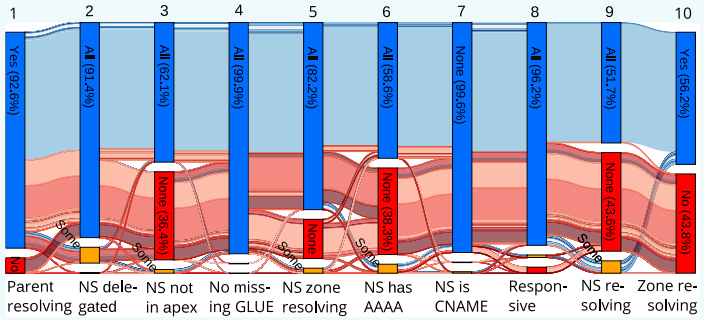
Data Collection

- 812,000,000 domains
- 89,000,000,000 queries
- 39 TB
- A, AAAA, TXT, DS, DNSKEY, CDS, CDNSKEY, CAA, TXT, MX, SOA
- December 2023 – Jan 2024
- Full dual-stack
- Capturing many incompliances and non-standard behaviors

What can we use this for?

Open-Sourcing
- Data
- Toolchain

IPv6-only resolution problems



Inconsistencies amplify effects deeper in the DNS tree

	Sampl.	Count	Unrespon.	P = C	P != C
com	Best	156M	9.24%	83.40%	7.36%
com	Worst	156M	9.47%	81.64%	8.89%
top	Best	3M	14.28%	75.90%	9.81%
top	Worst	3M	13.74%	73.56%	12.71%
> 2nd Lvl.	Best	37M	3.21%	96.49%	0.30%
> 2nd Lvl	Worst	37M	8.01%	89.53%	2.46%

What next?
- Impact of inconsistencies on DNSSEC?
- EDNS0 buffer sizes?
- IP-address level redundancy?