



.SU ZONE CONTENT RESEARCH
APTLD-85

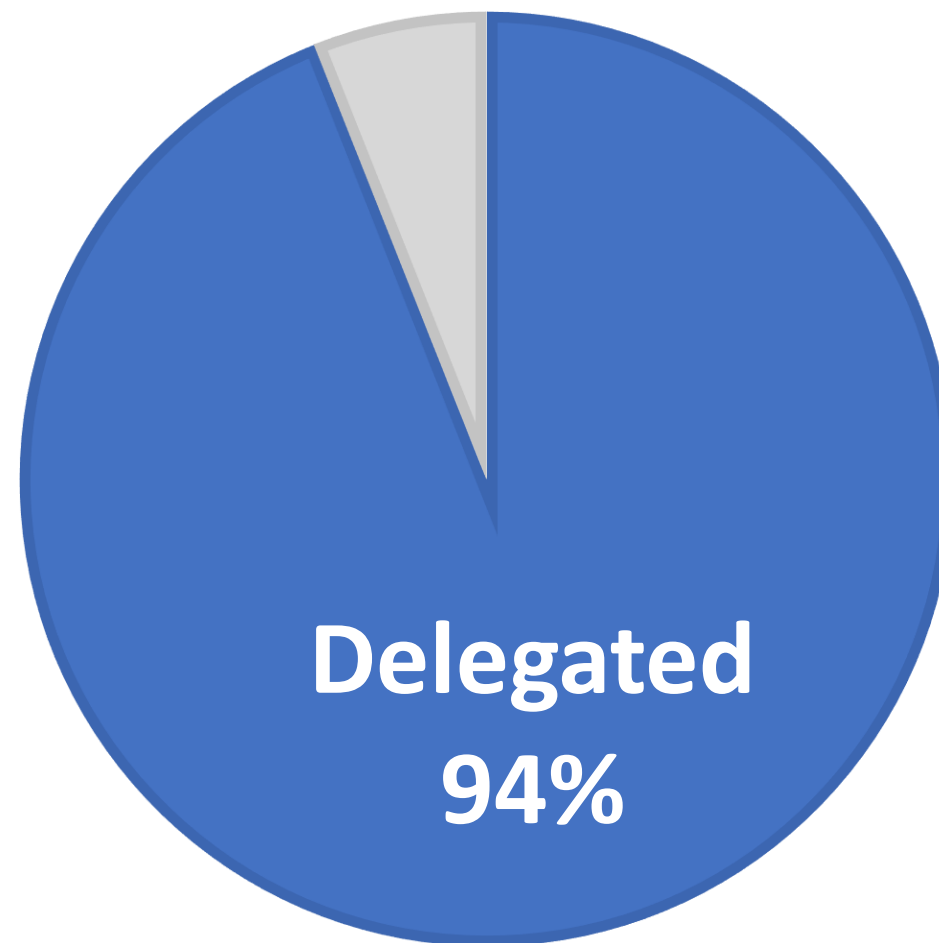


**WHAT IS YOUR
CUSTOMER ?**



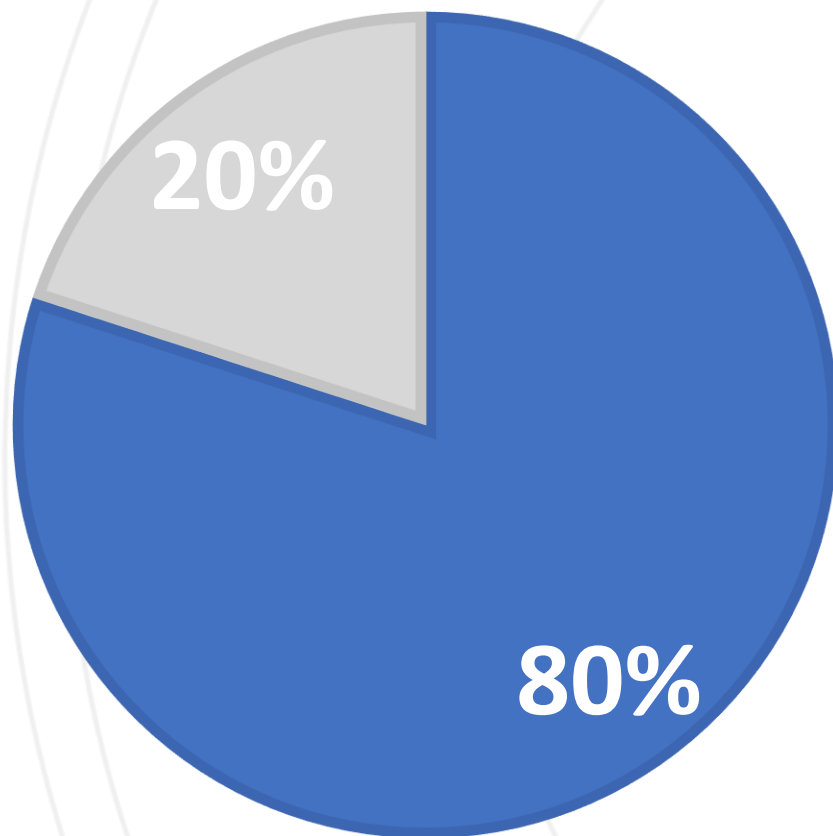
.SU

109 000+
active
registrations





Registrants



- Individuals
- Organizations



Content Research

- Get a list of all domains delegated
- Check website
- Detect webhosting software
- Check mail service
- Check TLS certificate



Content Research

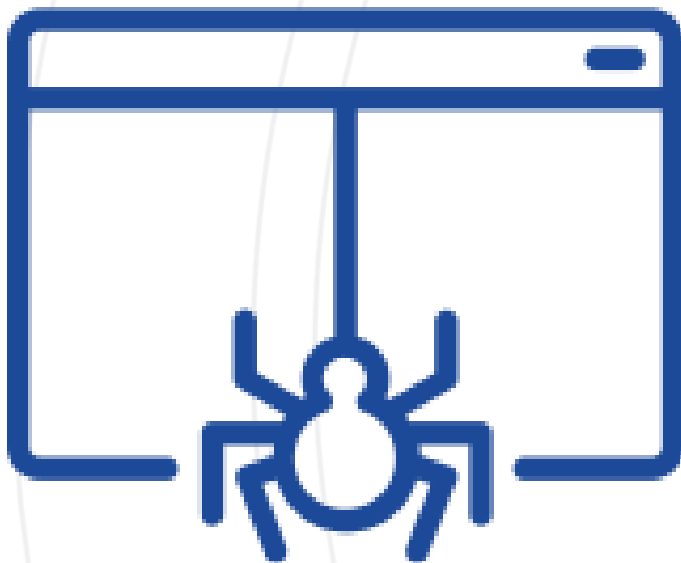
- Detect domains age
- Create classification by purposes of use
- Develop an algorithm for classification
- Assign domains to categories



Domains Research Tasks

Easy	Medium	Hard
Get a list Check Age	HTTP Mail TLS	Content Classification

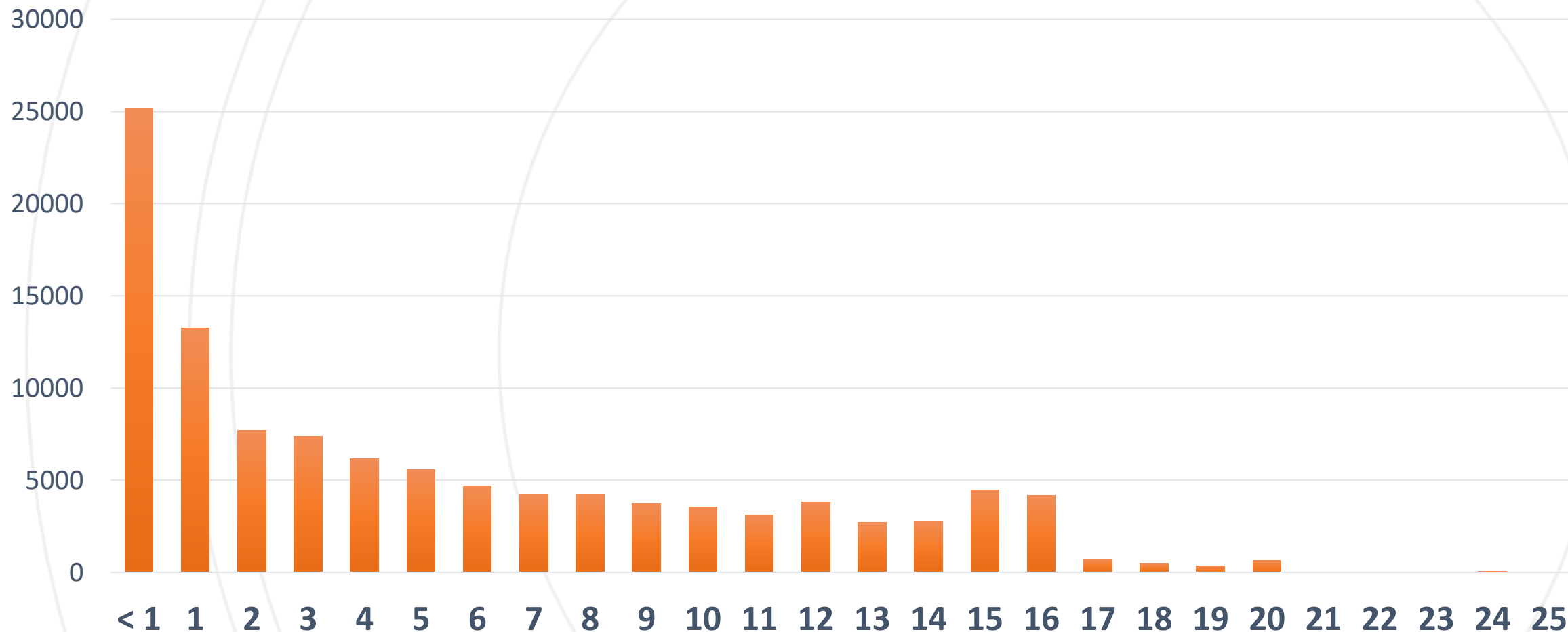
Crawler Bot



- Download website main page
- Get key words from headers, tags, meta etc.
- Analyze page text for frequently used words
- Build clusters of words
- Manual work: analyze clusters and define categories

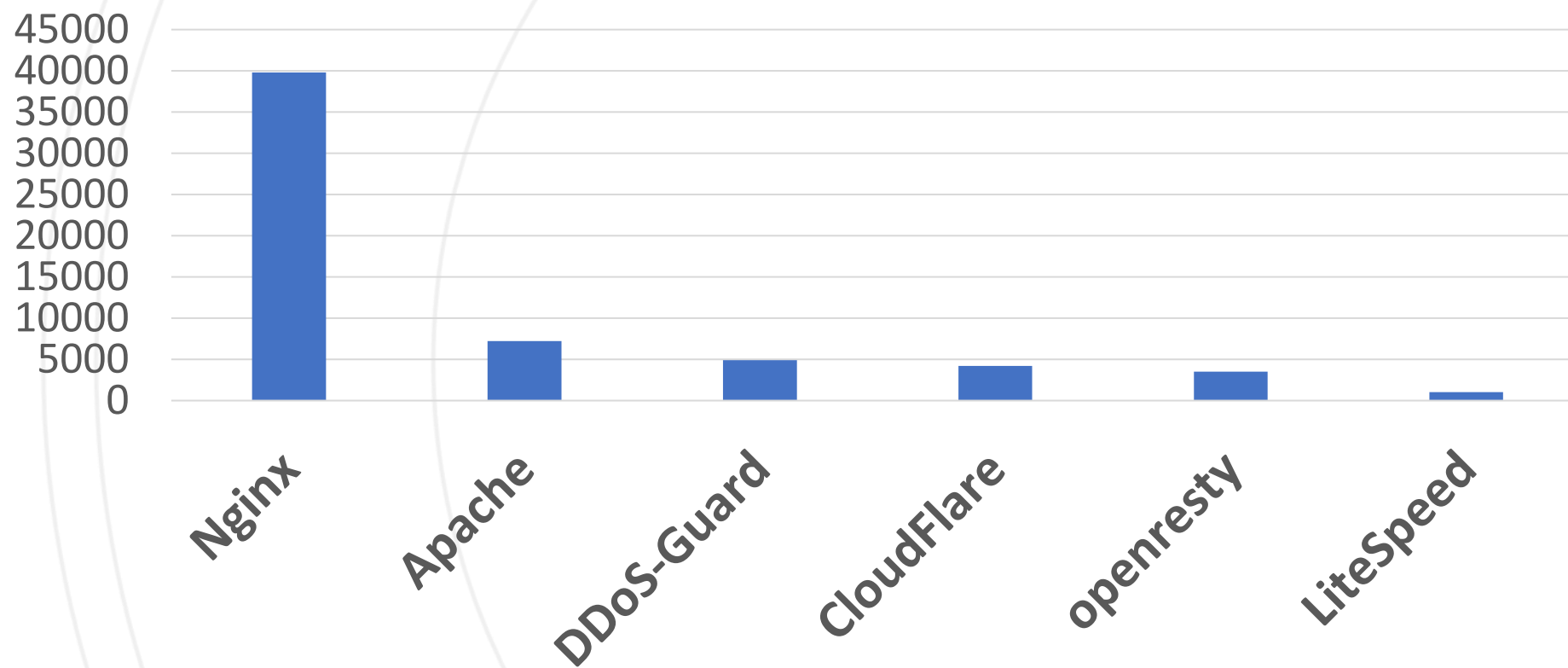


Domains Age



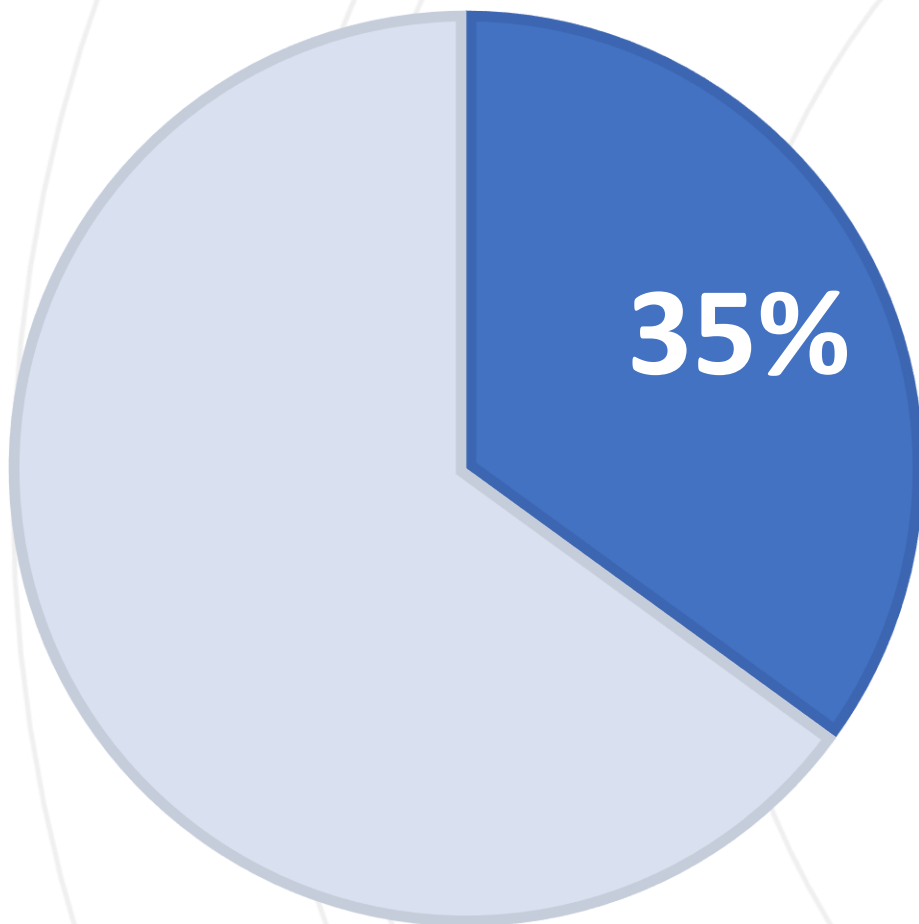


Most popular Web-servers

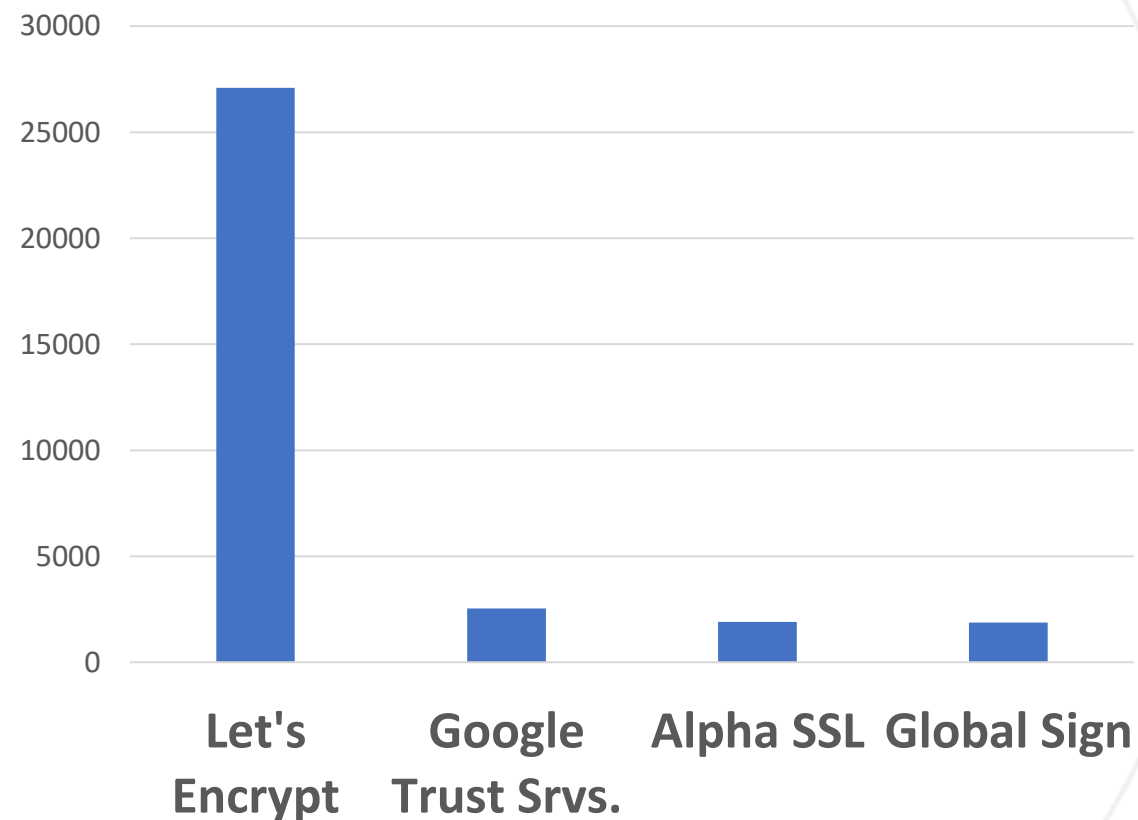




35% have a valid TLS Certificate

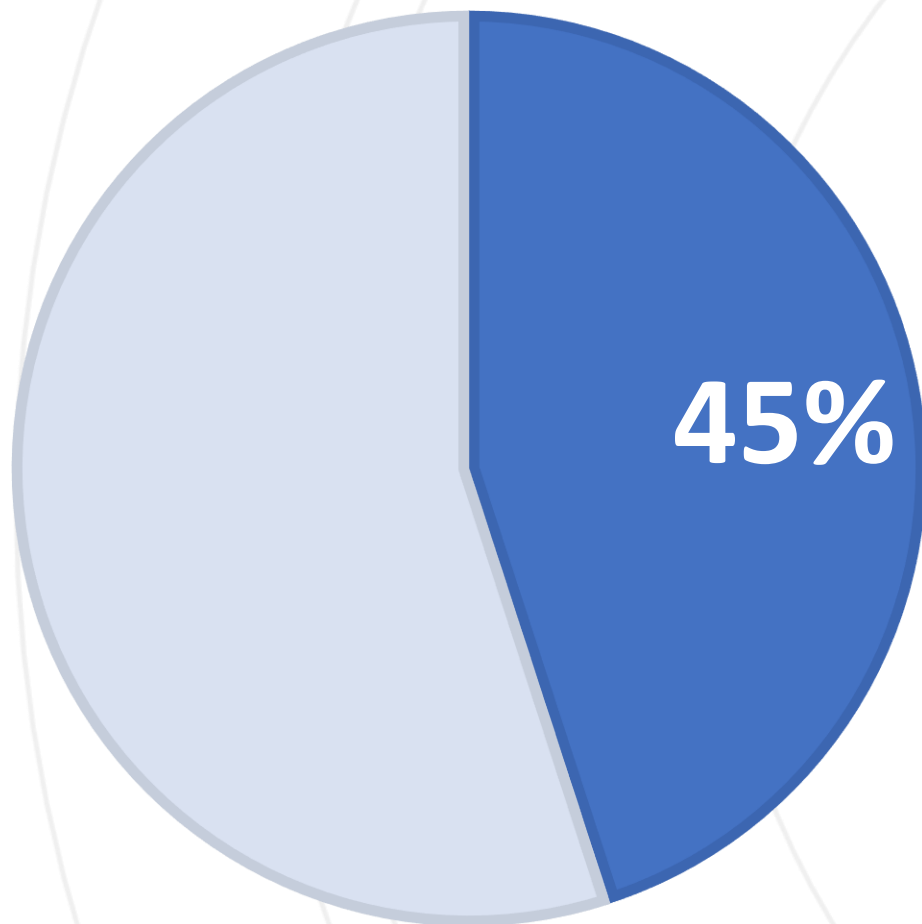


TOP CA

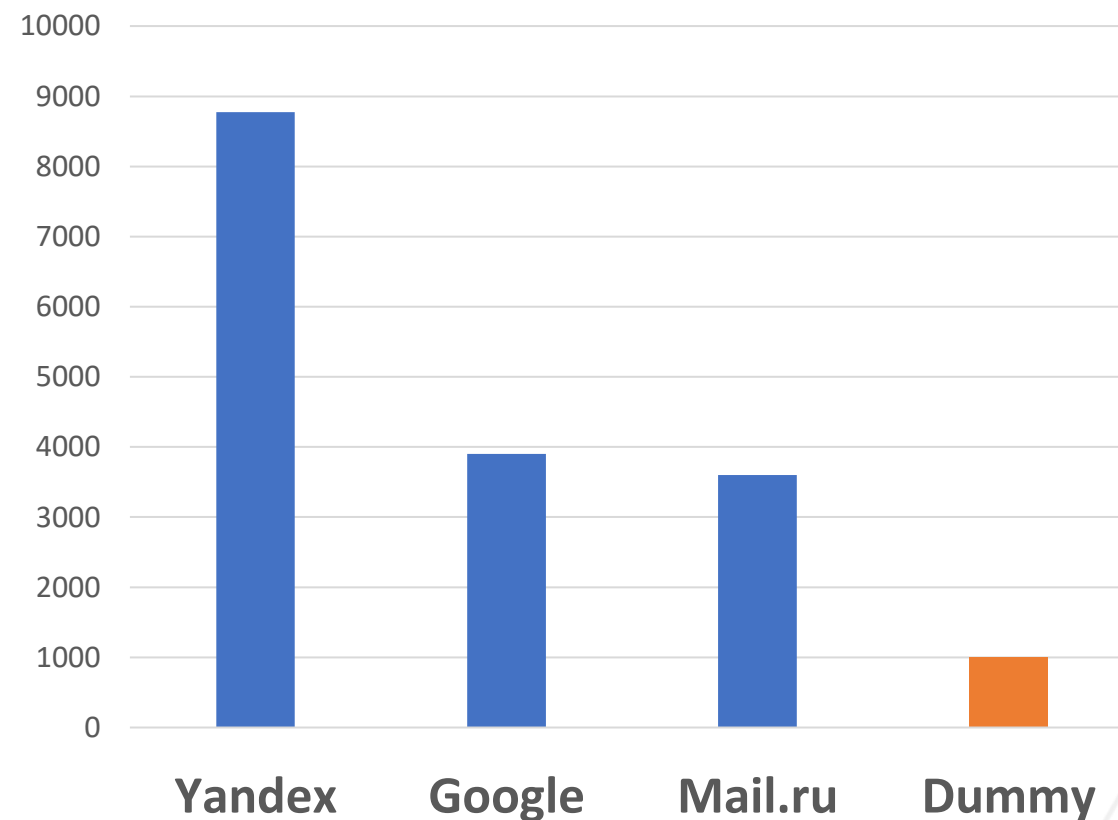




45% have MX-record

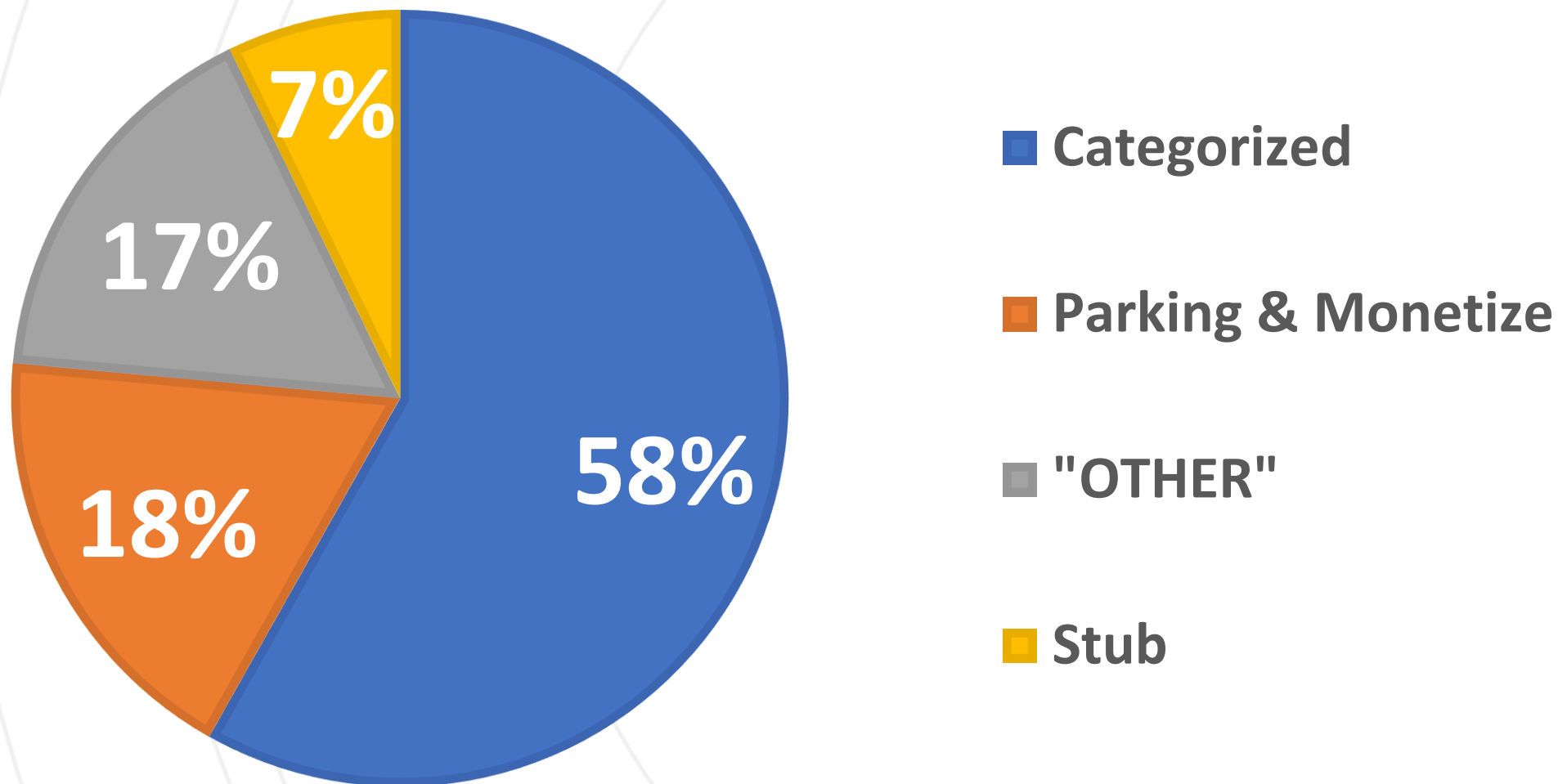


Top Public Mail Servers



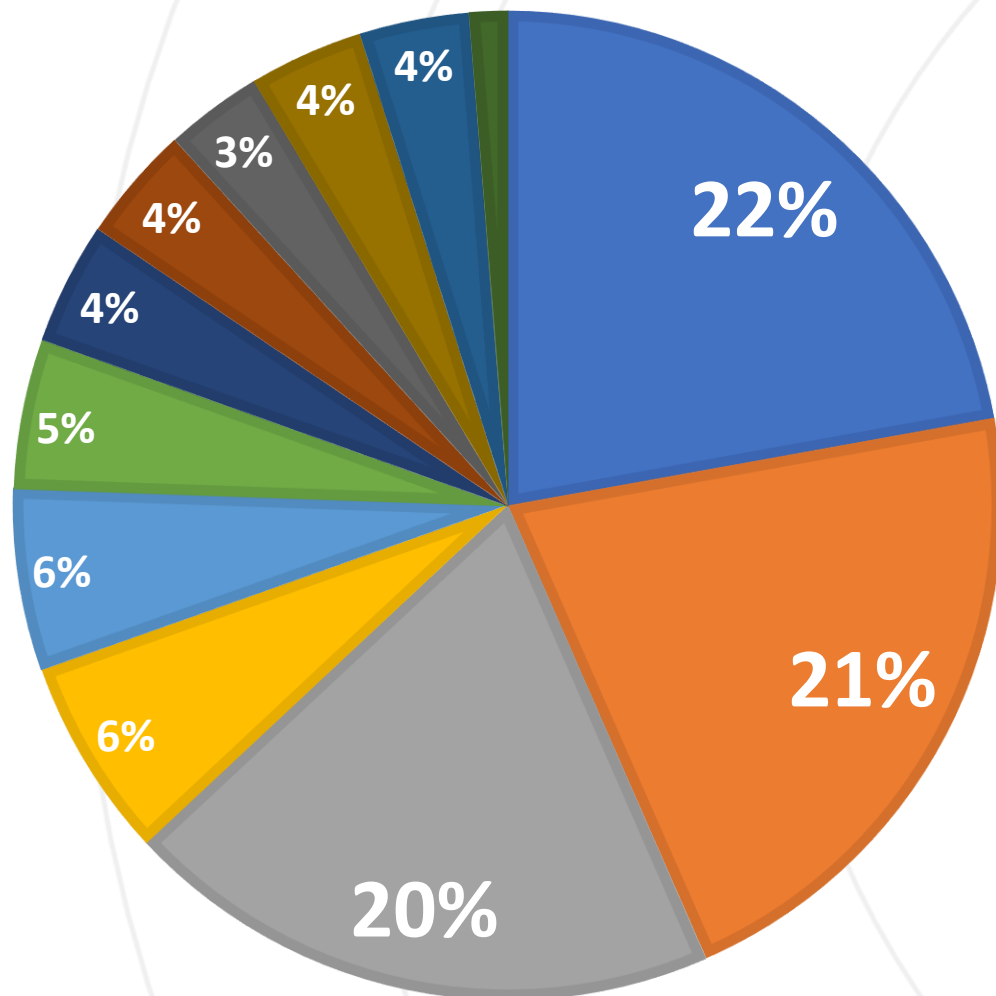


% from all delegated domains





Shares from 58%



- Business/Portfolio
- eCommerce
- Education
- Auto/Tech
- Legal services
- Real Estate
- Casino/Games
- Financial
- Furniture/Design
- Medicine
- Logistics/Transport
- Sex



TODO in (near) future

- Implement neural networks algorithm
- Improve categorization accuracy
- Add multilanguage content analysis



Contact for cooperation

Evgeny Kuskevich



evgekus@ripn.su