ITHI:
Identifier Technology Health Indicators

Alain Durand
ICANN, Office of the CTO

CA-IGF

2018

ICANN Strategic Plan 2016-2020

https://www.icann.org/en/system/files/files/strategic-plan-2016-2020-10oct14-en.pdf

2.1 Foster and coordinate a **healthy**, secure, **stable, and resilient** identifier ecosystem.

What is **ITHI**? Why Should You be Interested?

ITHI, or Identifier Technologies Health Indicators is an ICANN initiative to "**measure**" the "**health**" of the "**identifiers**" that "**ICANN helps coordinate**".

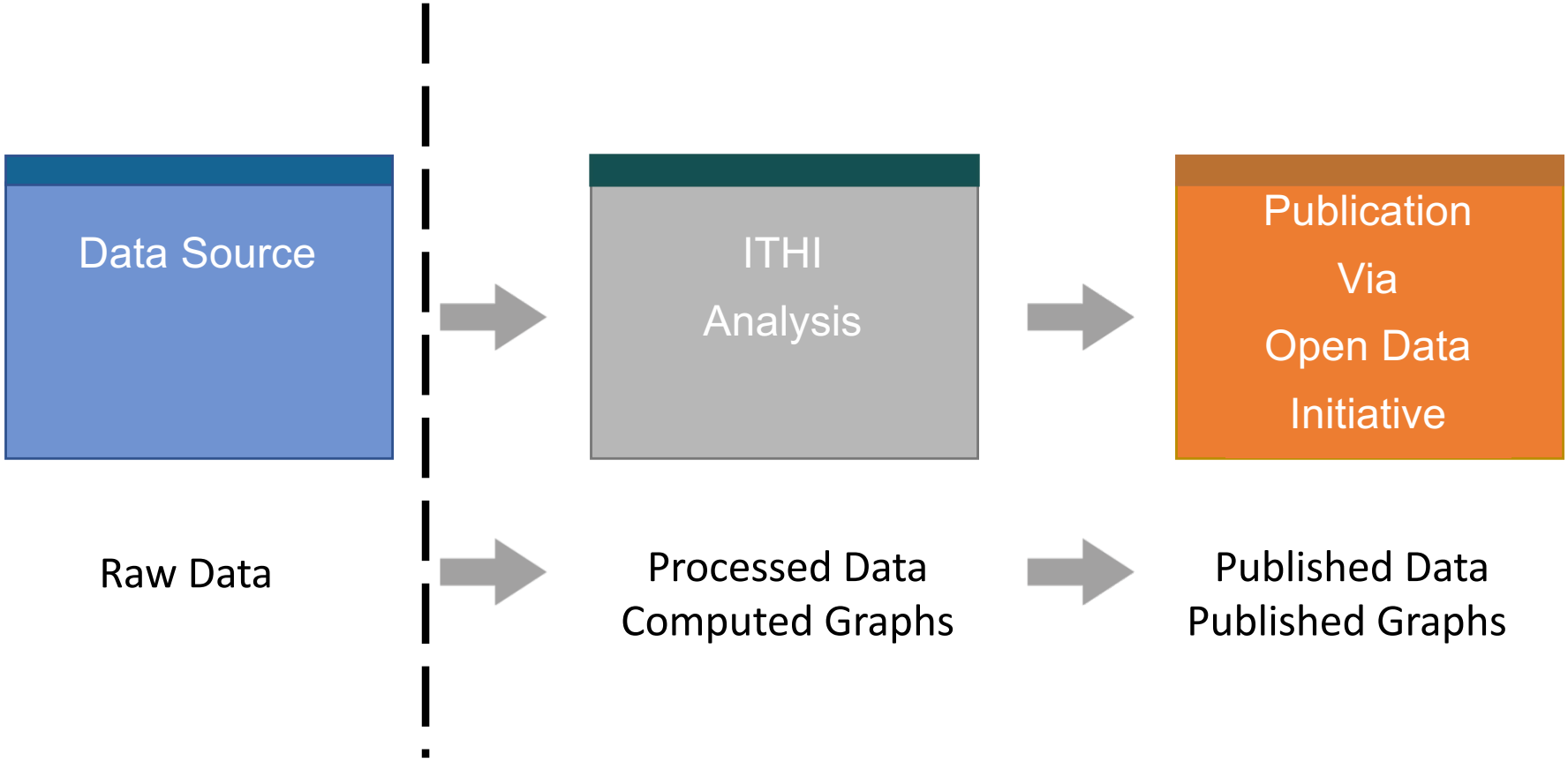
The goal is to produce a set of indicators that will be **measured and tracked over time** that will help determine if the set of identifiers is overall doing better or worse.

This is a long term project, expected to run for many years.

ITHI Principles of Operation

- Technical focus
- Problem areas → Metrics → Measurement
- Current value and trend over time
  - Automated process to collect & analyze data
- Measurement, not interpretation
- Extraction of statistics to avoid data privacy issues
- Open source tools & results

ITHI: Process



7 Metrics and Data Sources

Metric	Name	Data Source
M1:	Inaccuracy of Whois Data	ICANN compliance dept.
M2:	Domain Name Abuse	ICANN's DAAR Project https://www.icann.org/octo-ssr/daar
M3:	DNS Root Traffic Analysis	Samples of DNS root traffic
M4:	DNS Recursive Server Analysis	Summaries of recursive resolvers traffic
M5:	DNS Resolver Behavior	APNIC
M6:	IANA registries for DNS parameters	Scan of recursive resolvers traffic
M7:	DNSSEC Deployment	Snapshots of DNS root zone

ITHI Metrics Related to Domain Abuse

M2:
Data from the Domain Abuse Activity Reporting system

What is the Domain Abuse Activity Reporting system?

- A system for reporting on domain name registration and abuse data across TLD registries and registrars

How does DAAR differ from other reporting systems?

- Studies all gTLD registries and registrars for which we can collect zone and registration data
- Employs a large set of reputation feeds (e.g., blocklists)
- Accommodates historical studies
- Studies multiple threats: phishing, botnet, malware, spam
- Takes a scientific approach: transparent, reproducible

DAAR Project Goals

- DAAR data can be used to
  - Report on threat activity at TLD or registrar level
  - Study histories of security threats or domain registration activity
  - Help operators understand or consider how to manage their reputations, their anti-abuse programs, or terms of service
  - Study malicious registration behaviors
  - Assist operational security communities

The purpose of DAAR is to provide data to support community, academic, or sponsored research and analysis for informed policy consideration

Current Reputation Data Sets

- SURBL lists (domains only)
- Spamhaus Domain Block List
- Anti-Phishing Working Group
- Malware Patrol (Composite list)
- Phishtank
- Ransomware Tracker
- Feodotracker

SpamAssassin: malware URLs list
Carbon Black Malicious Domains
Postfix MTA
Squid Web proxy blocklist
Symantec Email Security for SMTP
Symantec Web Security
Firekeeper
DansGuardian
ClamAV Virus blocklist
Mozilla Firefox Adblock
Smoothwall
MailWasher

M2.*: Number of Abused Domain per 10,000 Registrations

Data from 01/31/2018

M2 metric name	Global Average
M2.1 = number of Phishing Domains per 10000 registered domain names	4.28
M2.2 = number of Malware Domains per 10,000 registered domain names	3.28
M2.3 = number of Botnet C&C Domains per 10,000 registered domain names	2.89
M2.4 = number of Spam Domains per 10,000 registered domain names	86.73

Total number of gTLDs: 1143, Total number of registrars: 1952

M2.*: Concentration of Abuse

Abuse	gTLD50	Registrar50	gTLD90	Registrar90
Phishing	1	7	11	45
Malware	1	2	7	9
Botnet	2	3	5	28
Spam	4	3	18	18

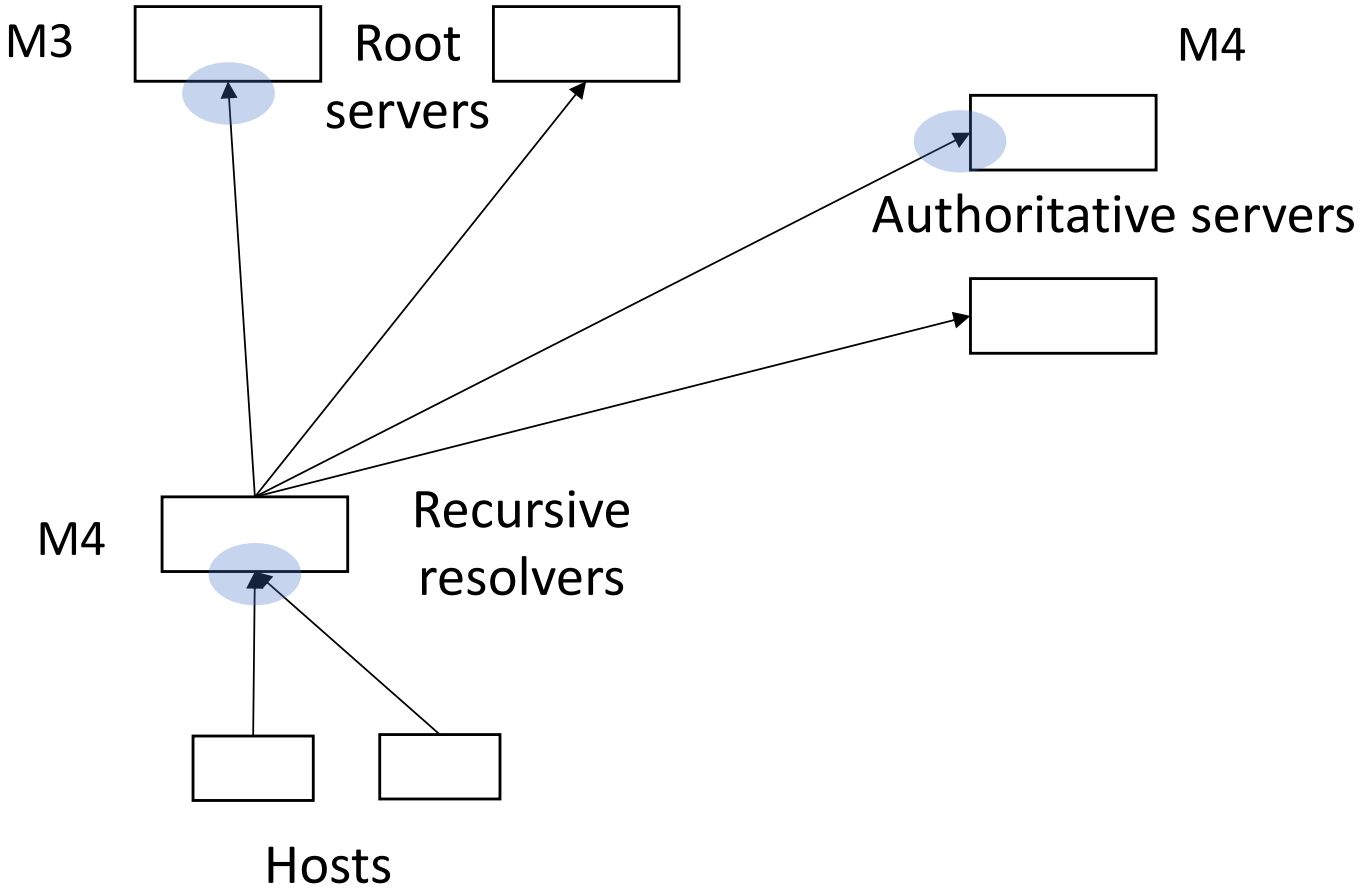
Table shows the number of TLDs/Registrars to account for > 50%/90% of all abuse of the specified type.

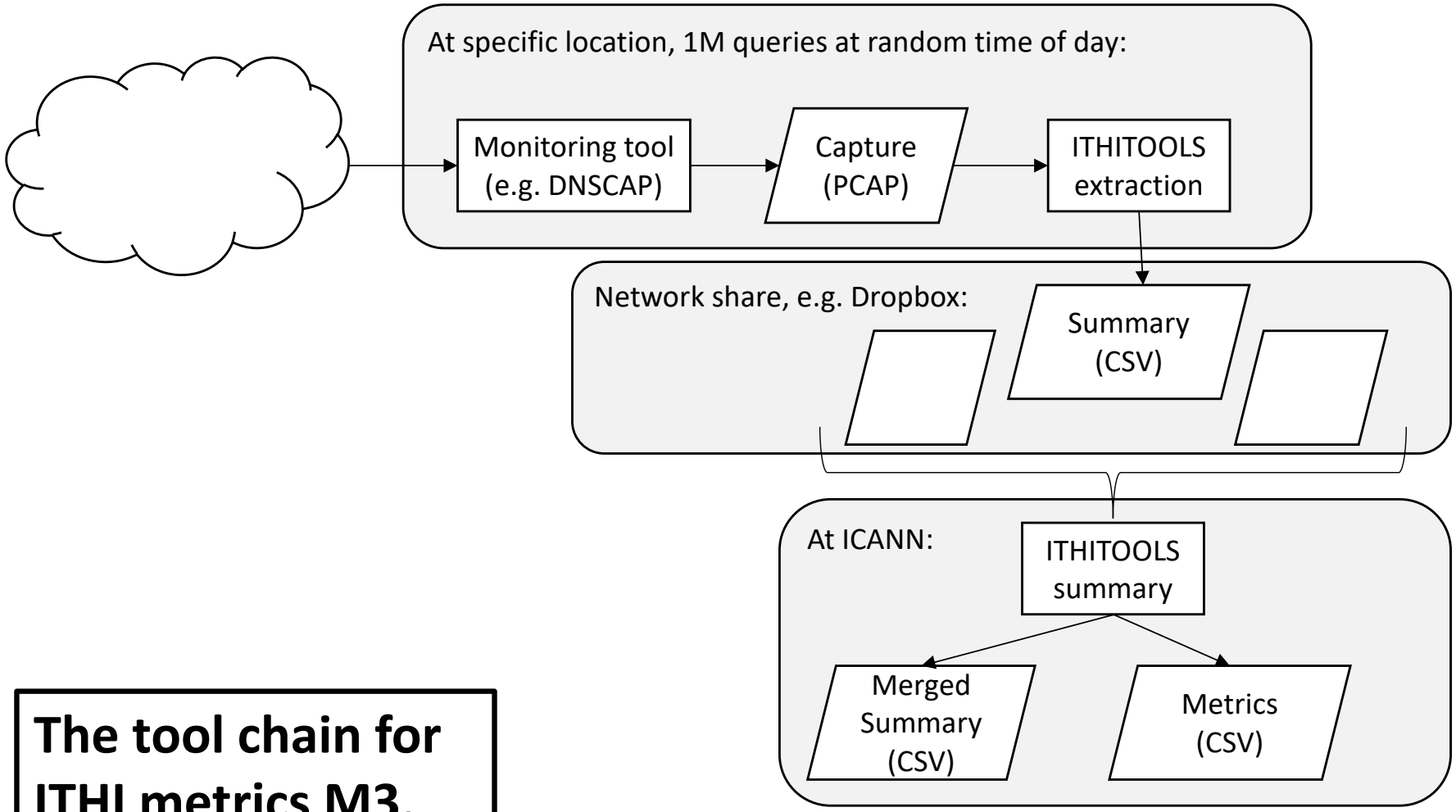
Total number of gTLDs: 1143, Total number of registrars: 1952*

(*) We removed two parking registrars from those statistics

ITHI Metrics Related to DNS Operation

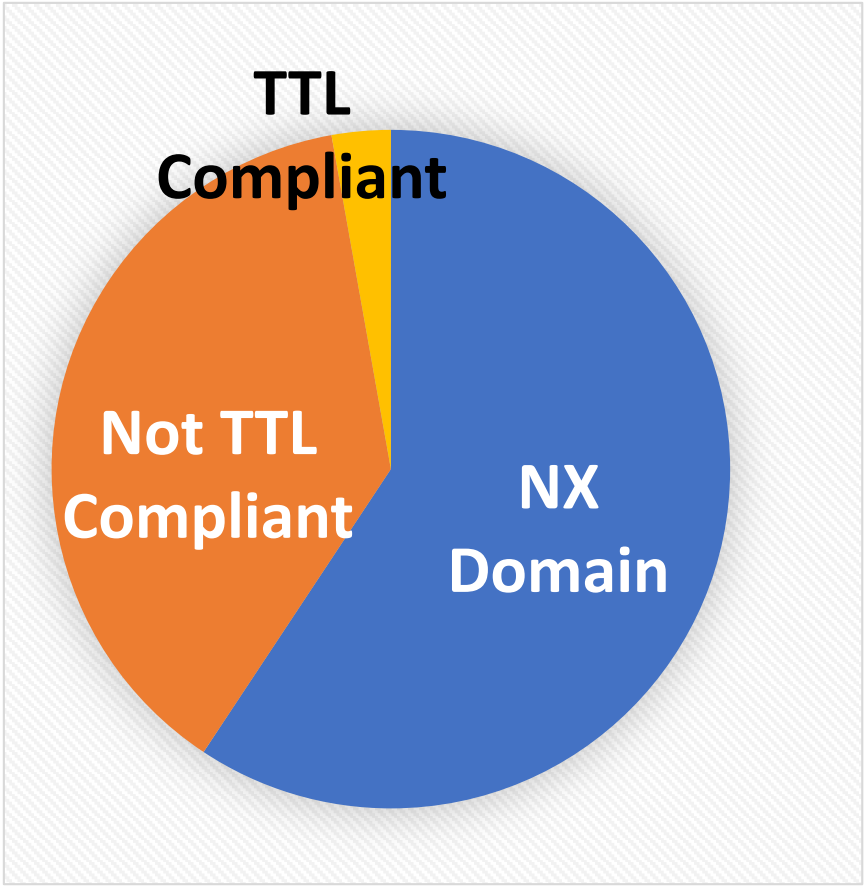
DNS Data Collection Points





The tool chain for ITHI metrics M3, M4 and M6

ITHI M3: Overhead in Root Traffic



Example of results, from the analysis of some B-Root traces

- Overhead at root needs tracking
  - Many “NX Domain” responses
  - Many queries not needed if resolver caches were TTL compliant
- Three metrics:
  - M3.1: NX Domains/Total Queries
  - M3.2: % not TTL compliant queries
  - M3.3: NX Domain per classes of TLD

DNS Operators: We Need Your Help!

- ITHI metrics help the whole community
  - M3: health of the DNS root
  - M4: analysis of TLD usage and leakage of strings
  - M6: health of IANA parameter registries for DNS
- Capture methodology is safe
  - Minimal load, no privacy issues
- Please contact us if you are interested!

Engage with ICANN

Thank You and Questions

Visit us at **icann.org**
Email: email


@icann


facebook.com/icannorg


youtube.com/icannnews


flickr.com/icann


linkedin/company/icann


slideshare/icannpresentations


soundcloud/icann