Meeting at the Intersection of Anti-Abuse and Infrastructure

Contributors:

Dave Crocker	Brandenburg InternetWorking and M^{3}AAWG Technical Advisor
Dennis Dayman	Chief Privacy and Security Officer at Return Path Messaging, Malware, and Mobile Anti-Abuse Working Group (M^{3}AAWG) Board
Tobias Knecht	Abusix CEO and RIPE Anti-Abuse Working Group Co-Chair
Jared Mauch	NTT
Tom Shaw	SURBL and M^{3}AAWG Hosting SIG
Foy Shiver	Deputy Secretary-General Anti-Phishing Working Group (APWG)

Moderator:

Jesse Sowell MIT Research Affiliate and M3AAWG Advisor
(jsowell@mit.edu)

Anti-Abuse and Attribution
The Blame Game

Most of the folks in here have encountered a blocking list (BL)

Unraveling precisely why a network landed on a list was not always easy

Mailing lists are full of “tense” discussions and complaints

Today’s objective:
Acknowledge the history, focus on the pragmatics of modern anti-abuse operations:

- ➔ What constitutes abuse
- ➔ How abuse indicators have evolved
- ➔ Economics of anti-abuse operations
- ➔ Q&A with panelists

Anti-Abuse and Attribution Prescriptive Ethos

What is Anti-Abuse? Let's start with a longstanding definition:

"all information exchange on the Internet *should be consensual*, and unless you choose to receive [traffic] from a third party, you should not *have to* accept it"[1]

Just because there is a *legitimate route* to a destination doesn't mean all traffic *using that route* is legitimate

Provides a ***prescriptive ethos***, but doesn't help with ***practical application***

1 Adapted from early definition from MAPS

Anti-Abuse and Attribution Evolution and Pragmatics

A more pragmatic definition:

"abuse is what customers complain about"[2]

What are the modern anti-abuse themes we are developing for the panel discussion?

1. Subjective → Objective indicators
2. Indicators are ***always*** error-prone
3. Focus has shifted from inbound to outbound (attribution)
4. Indicators development and use
5. Who bears the burden?
6. Economics of indicators and anti-abuse operations
7. Discuss!

2 Definition offered by Dave Crocker leading up to this panel discussion

State of Anti-Abuse
Evolution of Abuse Mitigation and Remediation

- ➔ Perceptions Influenced by a Few Bad Apples
- ➔ Early Indicator Development
- ➔ Bayes and False Positives
- ➔ Market Saturation
- ➔ Market Rationalization
- ➔ Effects of Modern Indicators

Perceptions of Anti-Abuse Attribution and Extortion

Signalling:

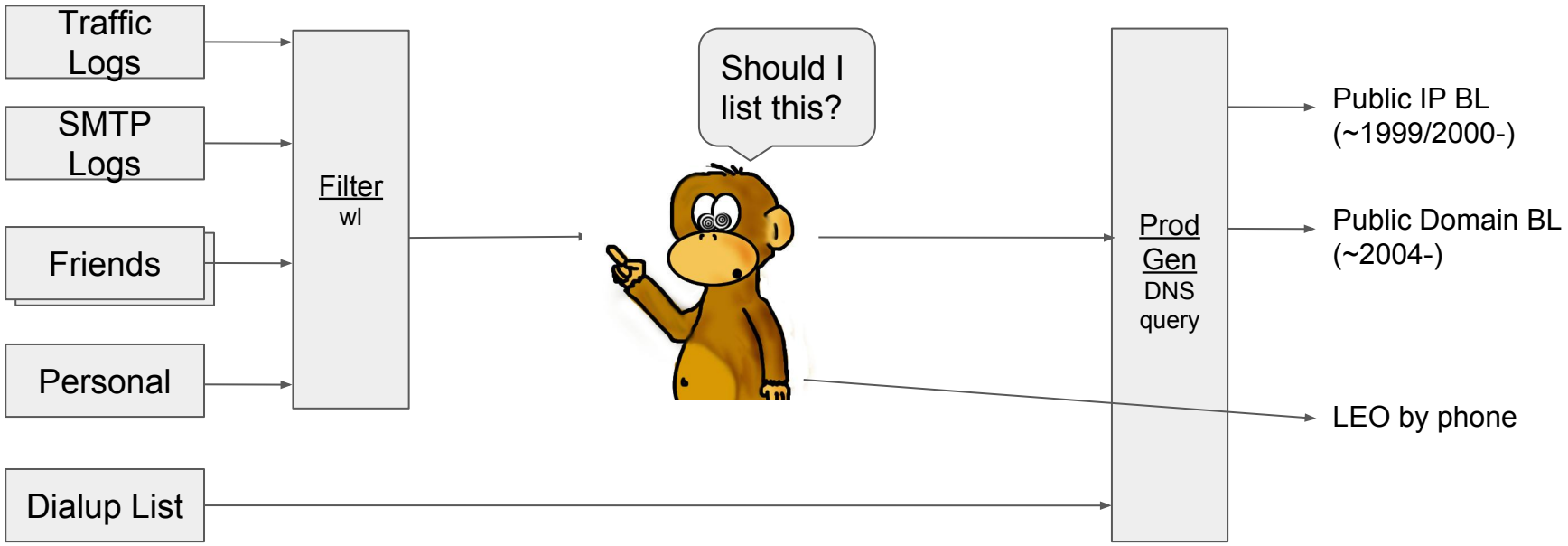
BL [listing]:
“someone using your network resources is generating unwanted traffic, you’ve been listed as abusive

Network:
“what do I need to do to correct this problem?”

BL:
For this fee, I’ll remove you from my list

This practice has happened, but has been long condemned as both unethical and counterproductive.

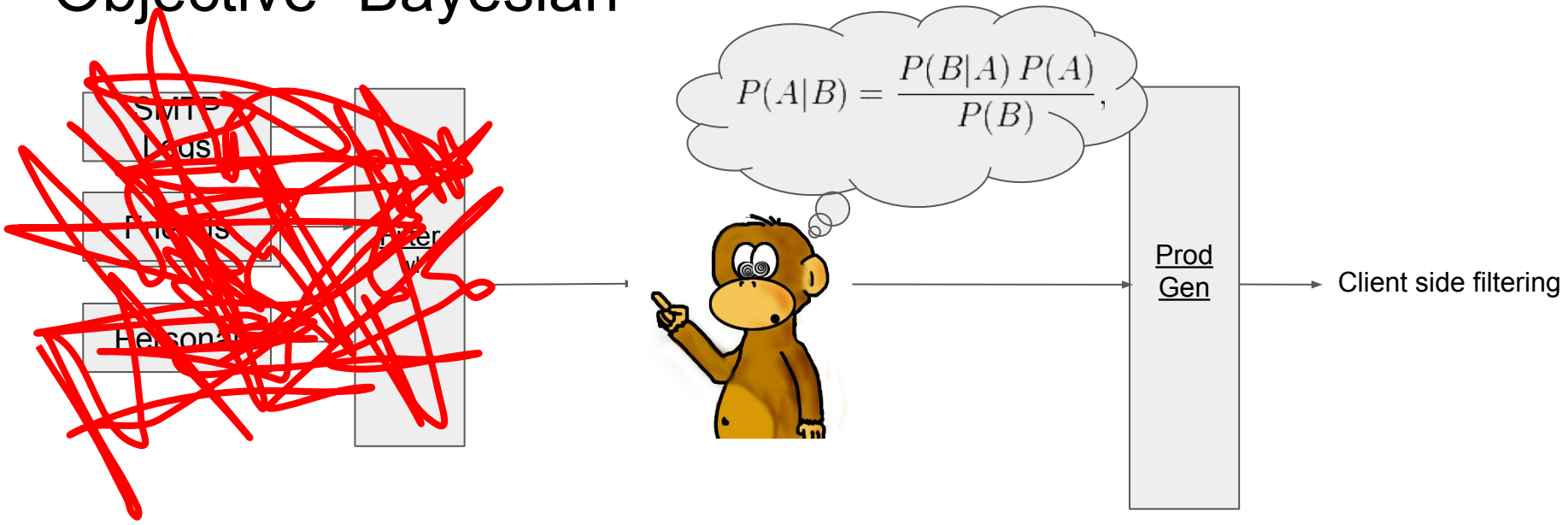
Early State of Anti-Abuse
Subjective Reputation and Signalling



Learning How to Signal Reputation

- ➔ **benevolent, but not necessarily warm and fuzzy, BL operators**
- ➔ **coping largely with inbound**
- ➔ **limited sources**

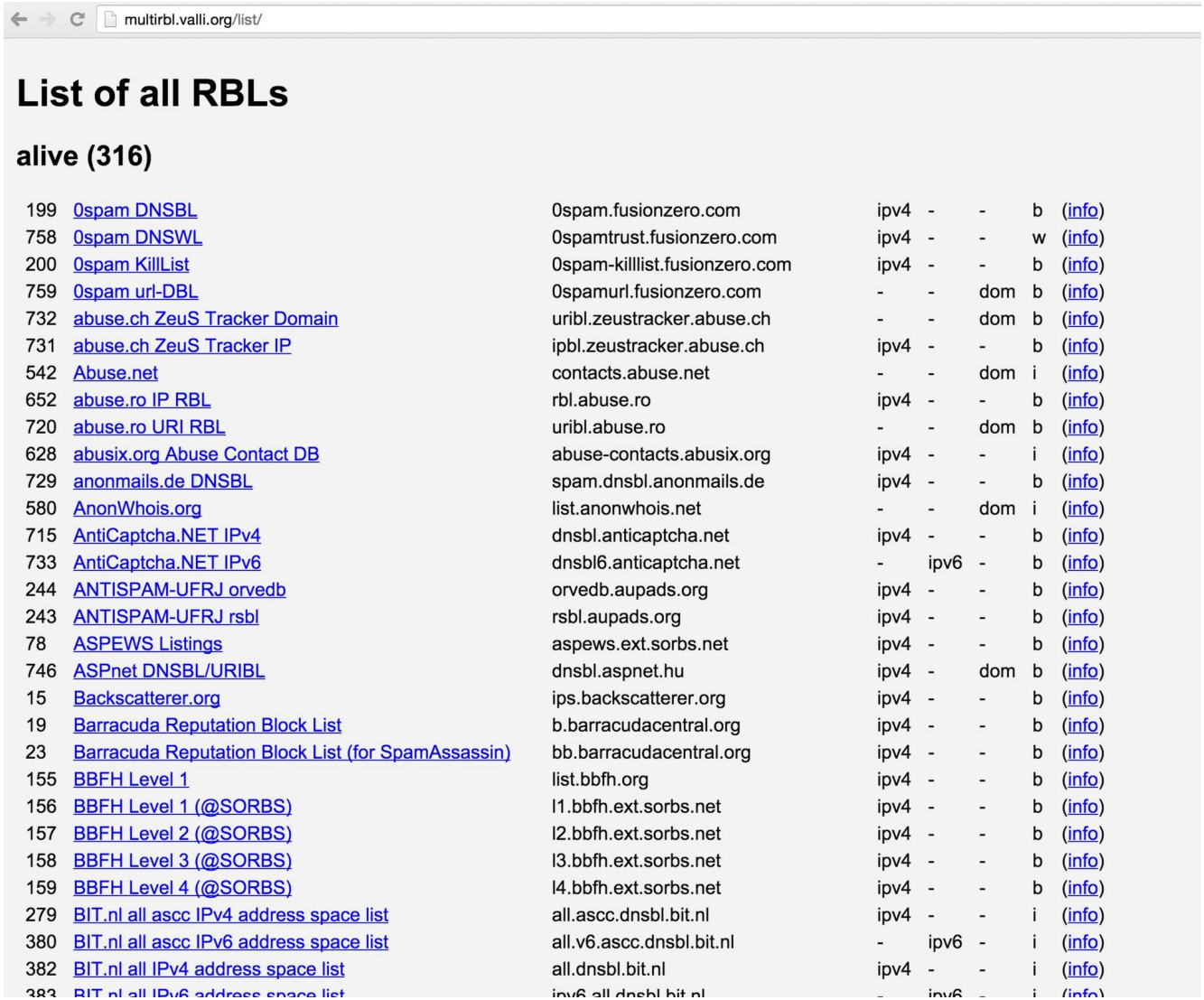
State of Anti-Abuse
"Objective" Bayesian



Problems with "automagic" client side filtering:

- ➔ burden of training and management shifted to users
- ➔ regardless of training, bayesian has high false positives
- ➔ band aid
  - ◆ purely inbound mitigation
  - ◆ no remediation of outbound sources
- ➔ **in response to Bayesian false positives, relative scoring of other sources**
- ➔ **so...back to "softer" reputation mechanisms**

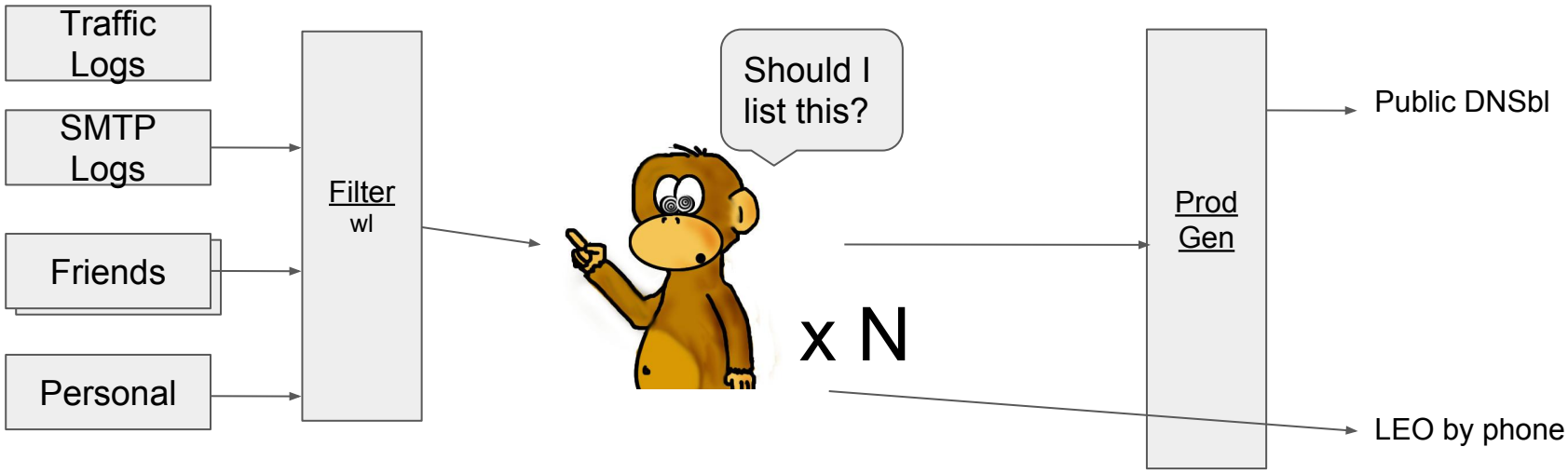
State of Anti-Abuse
Plethora of Blocking Lists



199	0spam DNSBL	0spam.fusionzero.com	ipv4	-	-	b	(info)
758	0spam DNSWL	0spamtrust.fusionzero.com	ipv4	-	-	w	(info)
200	0spam KillList	0spam-killlist.fusionzero.com	ipv4	-	-	b	(info)
759	0spam url-DBL	0spamurl.fusionzero.com	-	-	dom	b	(info)
732	abuse.ch ZeuS Tracker Domain	uribl.zeustracker.abuse.ch	-	-	dom	b	(info)
731	abuse.ch ZeuS Tracker IP	ipbl.zeustracker.abuse.ch	ipv4	-	-	b	(info)
542	Abuse.net	contacts.abuse.net	-	-	dom	i	(info)
652	abuse.ro IP RBL	rbl.abuse.ro	ipv4	-	-	b	(info)
720	abuse.ro URI RBL	uribl.abuse.ro	-	-	dom	b	(info)
628	abusix.org Abuse Contact DB	abuse-contacts.abusix.org	ipv4	-	-	i	(info)
729	anonmails.de DNSBL	spam.dnsbl.anonmails.de	ipv4	-	-	b	(info)
580	AnonWhois.org	list.anonwhois.net	-	-	dom	i	(info)
715	AntiCaptcha.NET IPv4	dnsbl.anticaptcha.net	ipv4	-	-	b	(info)
733	AntiCaptcha.NET IPv6	dnsbl6.anticaptcha.net	-	ipv6	-	b	(info)
244	ANTISPAM-UFRJ orvedb	orvedb.aupads.org	ipv4	-	-	b	(info)
243	ANTISPAM-UFRJ rsbl	rsbl.aupads.org	ipv4	-	-	b	(info)
78	ASPEWS Listings	aspews.ext.sorbs.net	ipv4	-	-	b	(info)
746	ASPnet DNSBL/URIBL	dnsbl.aspnet.hu	ipv4	-	dom	b	(info)
15	Backscatterer.org	ips.backscatterer.org	ipv4	-	-	b	(info)
19	Barracuda Reputation Block List	b.barracudacentral.org	ipv4	-	-	b	(info)
23	Barracuda Reputation Block List (for SpamAssassin)	bb.barracudacentral.org	ipv4	-	-	b	(info)
155	BBFH Level 1	list.bbfh.org	ipv4	-	-	b	(info)
156	BBFH Level 1 (@SORBS)	l1.bbfh.ext.sorbs.net	ipv4	-	-	b	(info)
157	BBFH Level 2 (@SORBS)	l2.bbfh.ext.sorbs.net	ipv4	-	-	b	(info)
158	BBFH Level 3 (@SORBS)	l3.bbfh.ext.sorbs.net	ipv4	-	-	b	(info)
159	BBFH Level 4 (@SORBS)	l4.bbfh.ext.sorbs.net	ipv4	-	-	b	(info)
279	BIT.nl all ascc IPv4 address space list	all.ascc.dnsbl.bit.nl	ipv4	-	-	i	(info)
380	BIT.nl all ascc IPv6 address space list	all.v6.ascc.dnsbl.bit.nl	-	ipv6	-	i	(info)
382	BIT.nl all IPv4 address space list	all.dnsbl.bit.nl	ipv4	-	-	i	(info)
383	BIT.nl all IPv6 address space list	ipv6.all.dnsbl.bit.nl	-	ipv6	-	i	(info)

Screenshot from http://multirbl.valli.org/list/

State of Anti-Abuse
Plethora of Blocking Lists

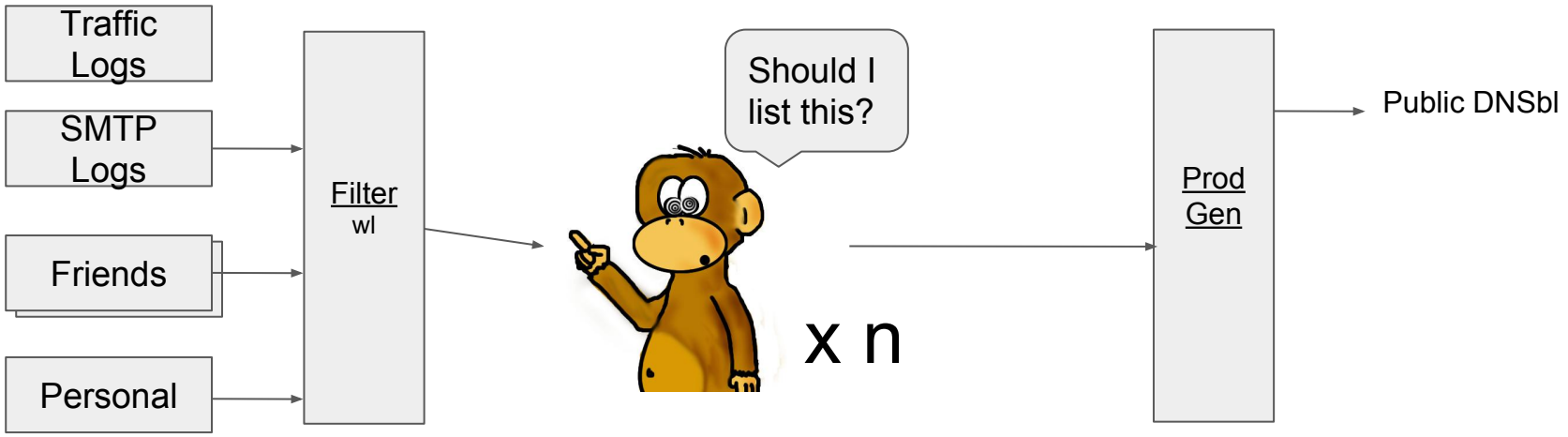


Market Quickly Saturated with RBLs of Dubious Quality:

- ➔ **some regionally localized BLs**
- ➔ **some little better than noisy Bayesian**
- ➔ **a number of extortionists**
- ➔ **variety of de-listing (remediation) policies**

- ➔ **within this chaos is a much smaller, effective, actionable set of sources and indicators**

State of Anti-Abuse
Markets + Institutions: Culling the Herd

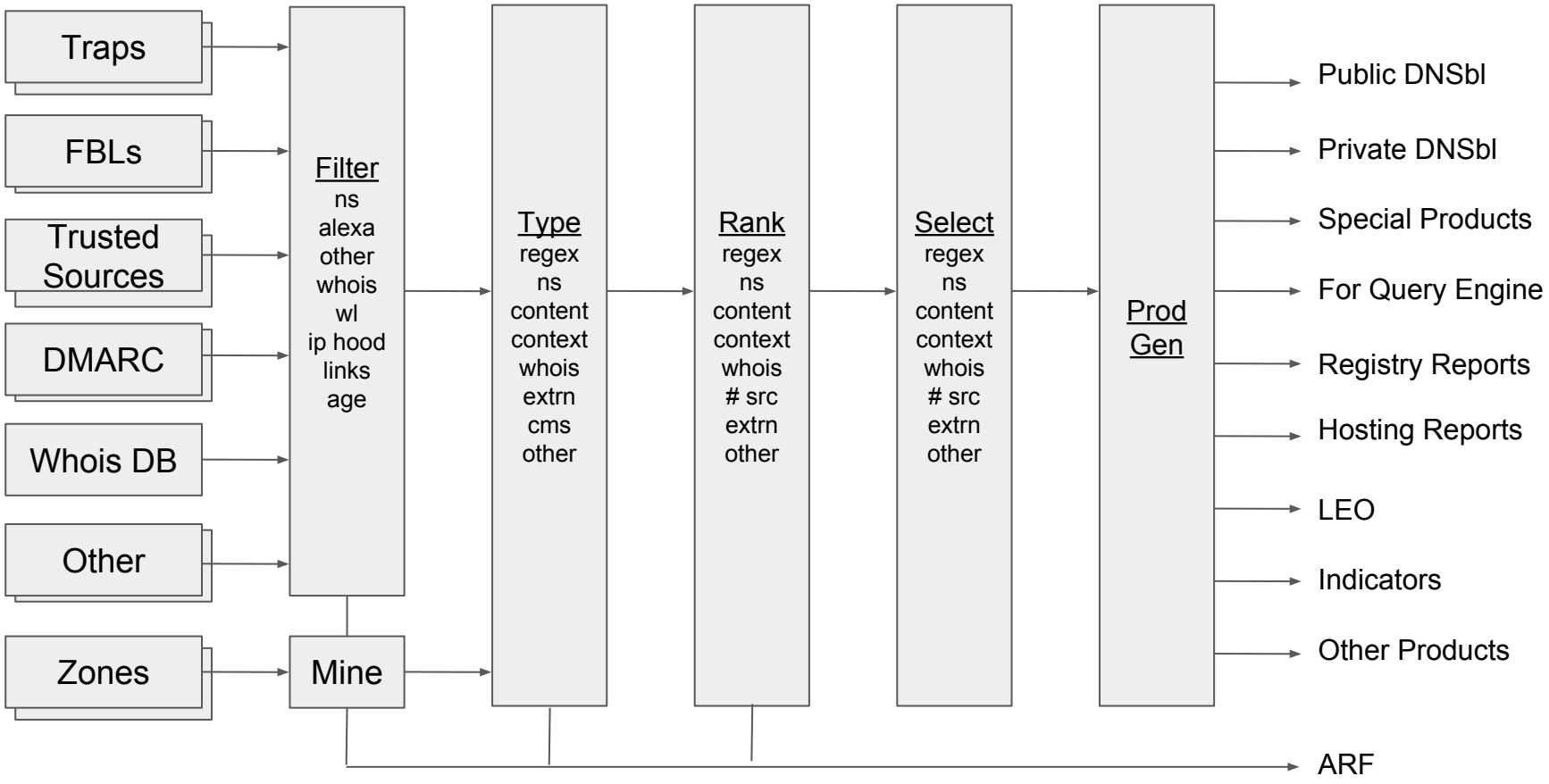


Market Rationalization of Reputation:*

- ➔ extortionists shunned
- ➔ sources of methods diversified
- ➔ lower false positives becomes market differentiator
- ➔ **responsiveness to credible remediators becomes a market differentiator**
- ➔ **anti-abuse operations transition from ideological position on consent to a response to market demand for credible reputation indicators**

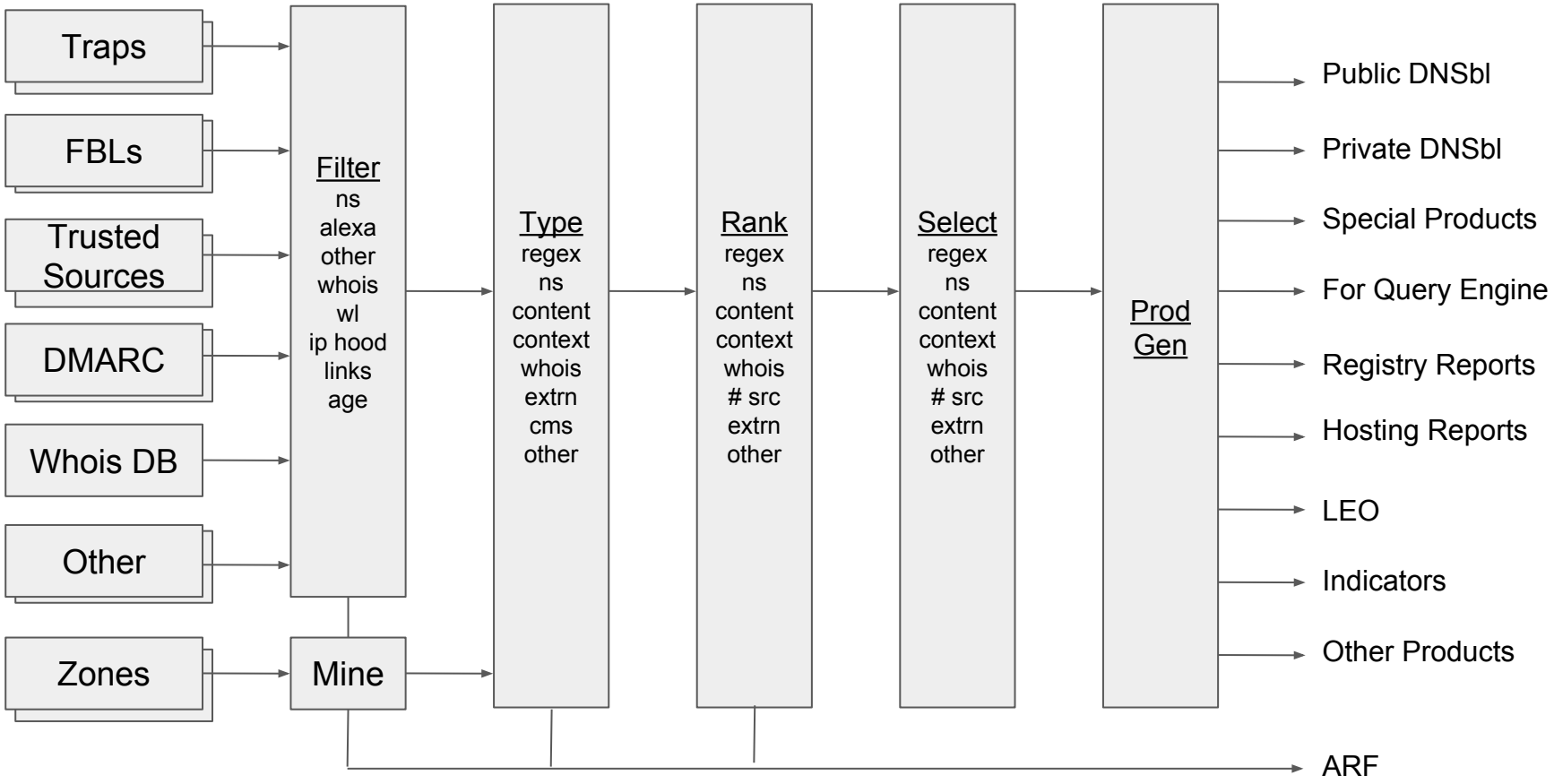
*Not the ideological magic of the “pure” market models of Econ 101, but a combination of private institutions fostering markets for credible reputation indicators, methods for analyzing these, and tools for automating those methods.

State of Anti-Abuse
Modern Indicators and Attribution

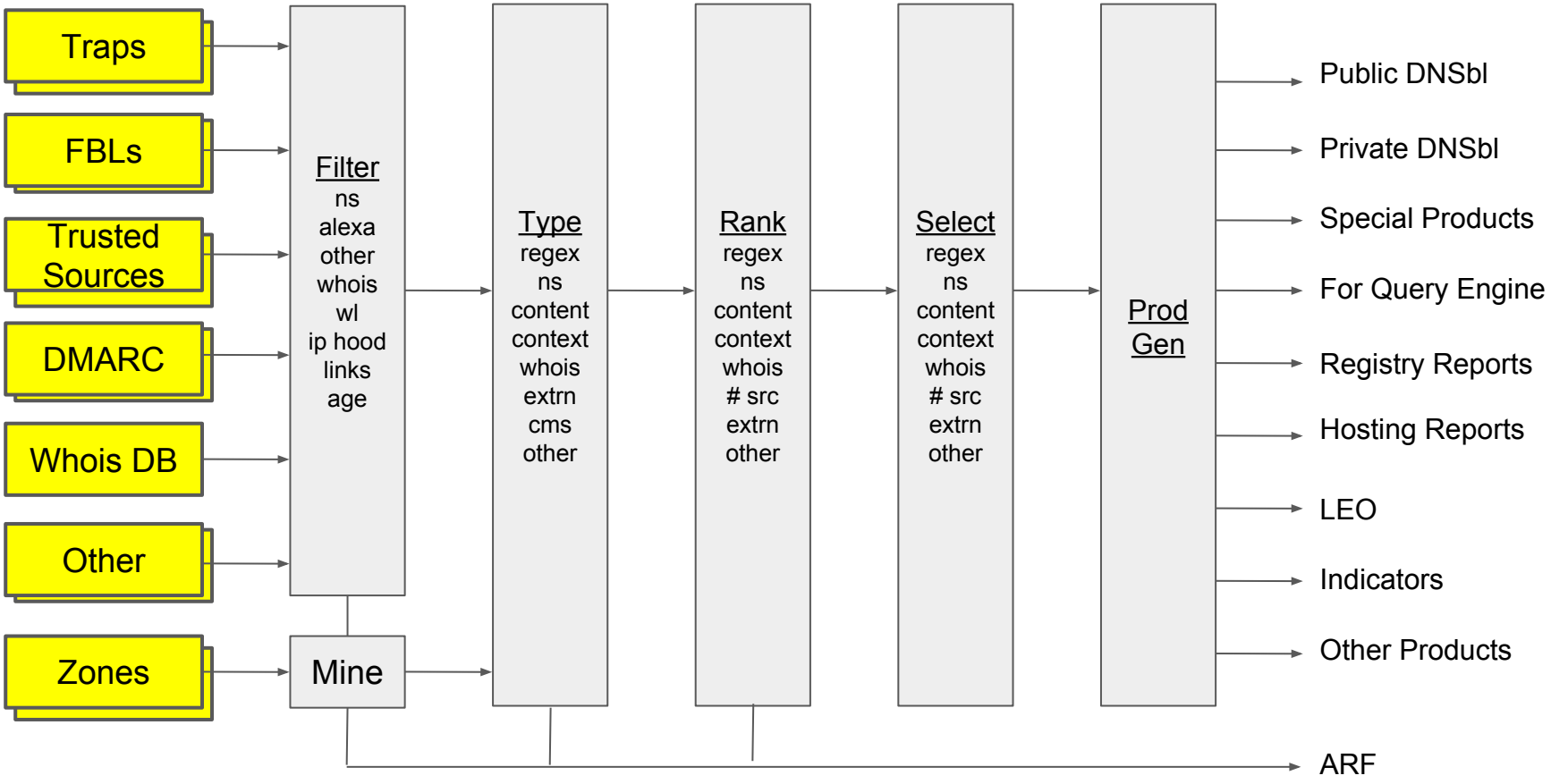


We’ve covered a good bit of history and background, any quick questions before we move on to indicators and operations?

Reputation Mechanics for Hosts
Overall Architecture

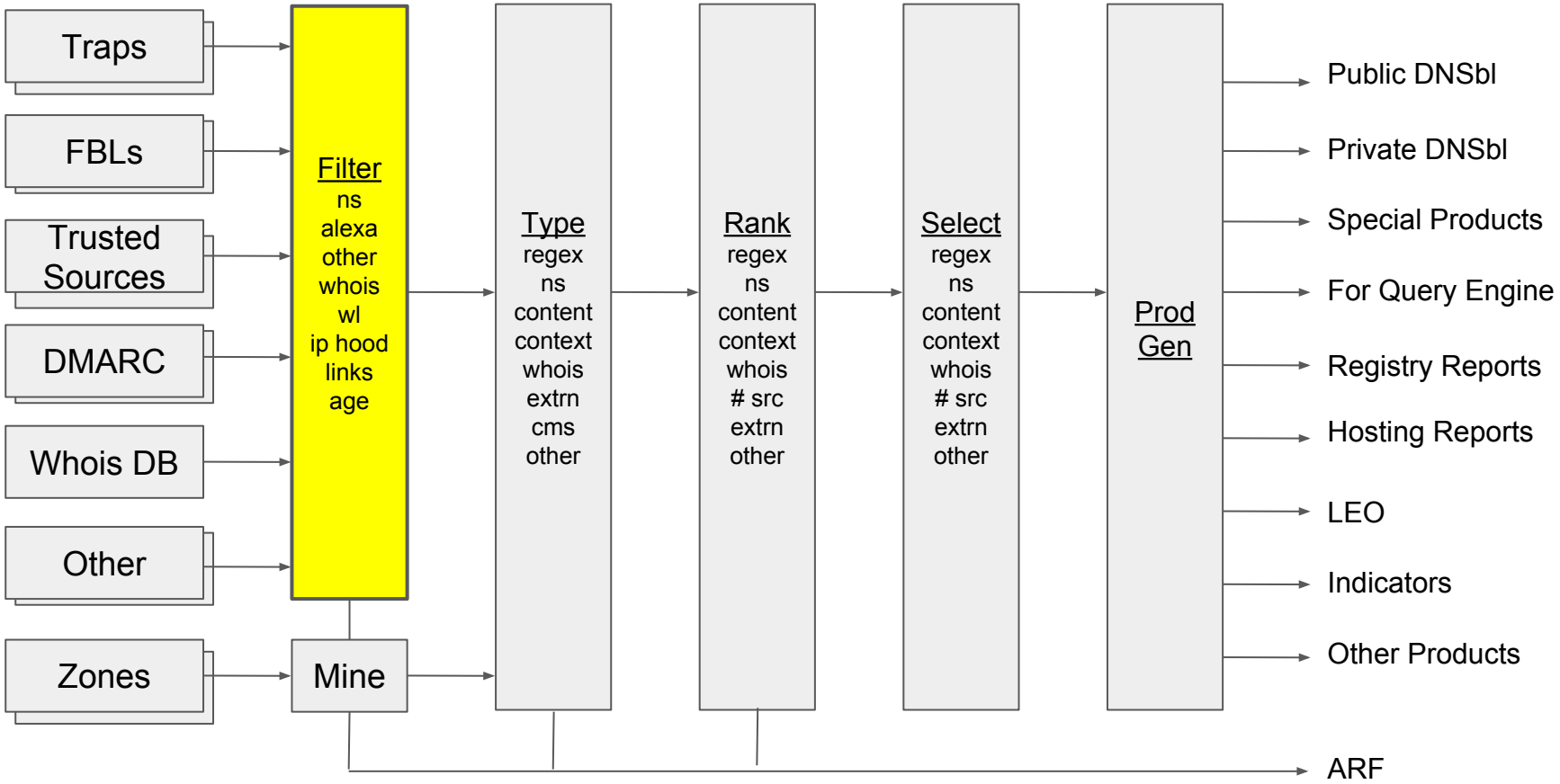


Reputation Mechanics for Hosts
Diversity of Sources



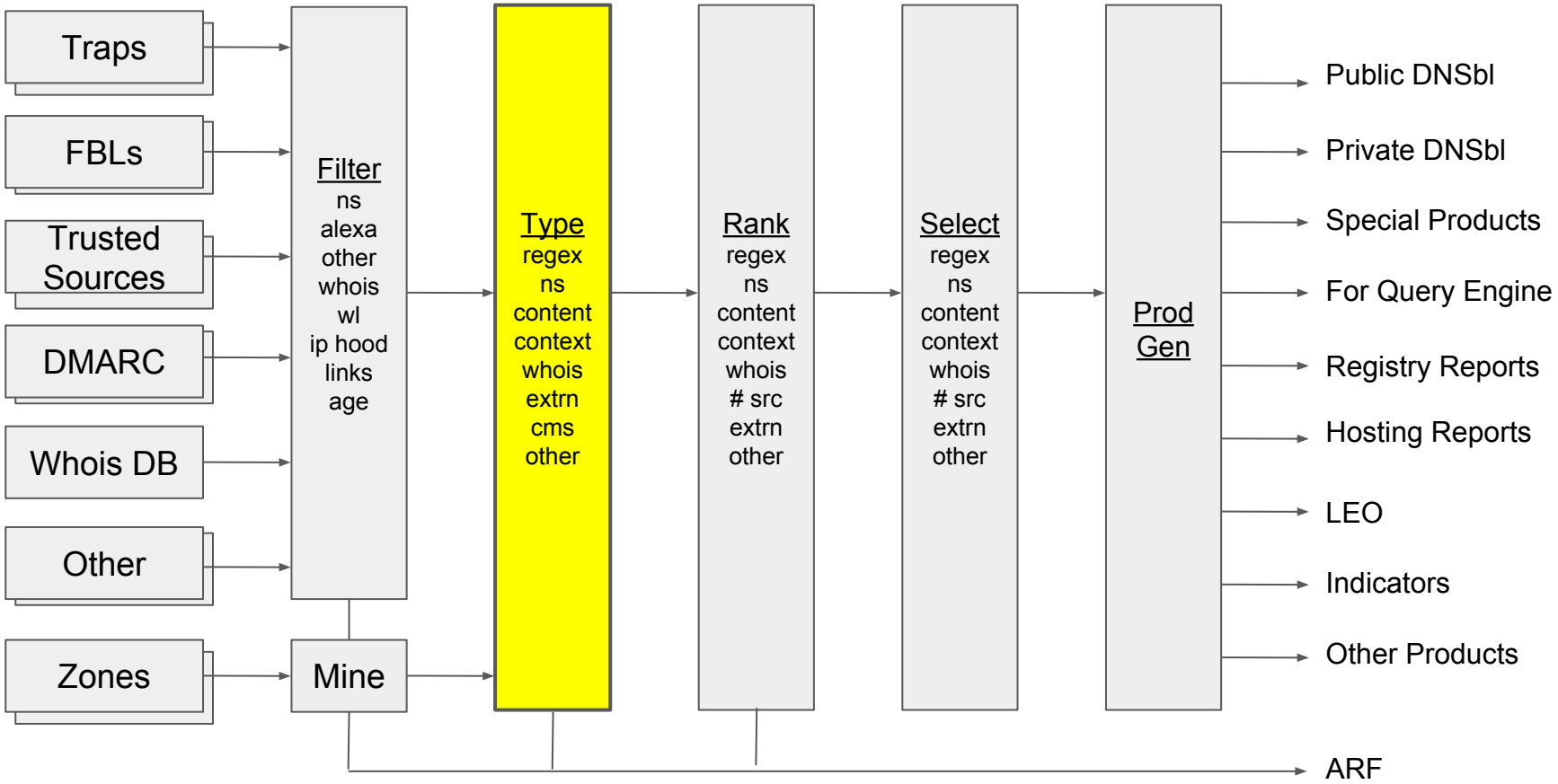
- ➔ Mail flow, raw e-mail data, traps and DMARC
- ➔ Trusted sources: ISACs, researchers, security organizations
- ➔ Mine zones, WHOIS

Reputation Mechanics for Hosts
Cleaning Raw Data



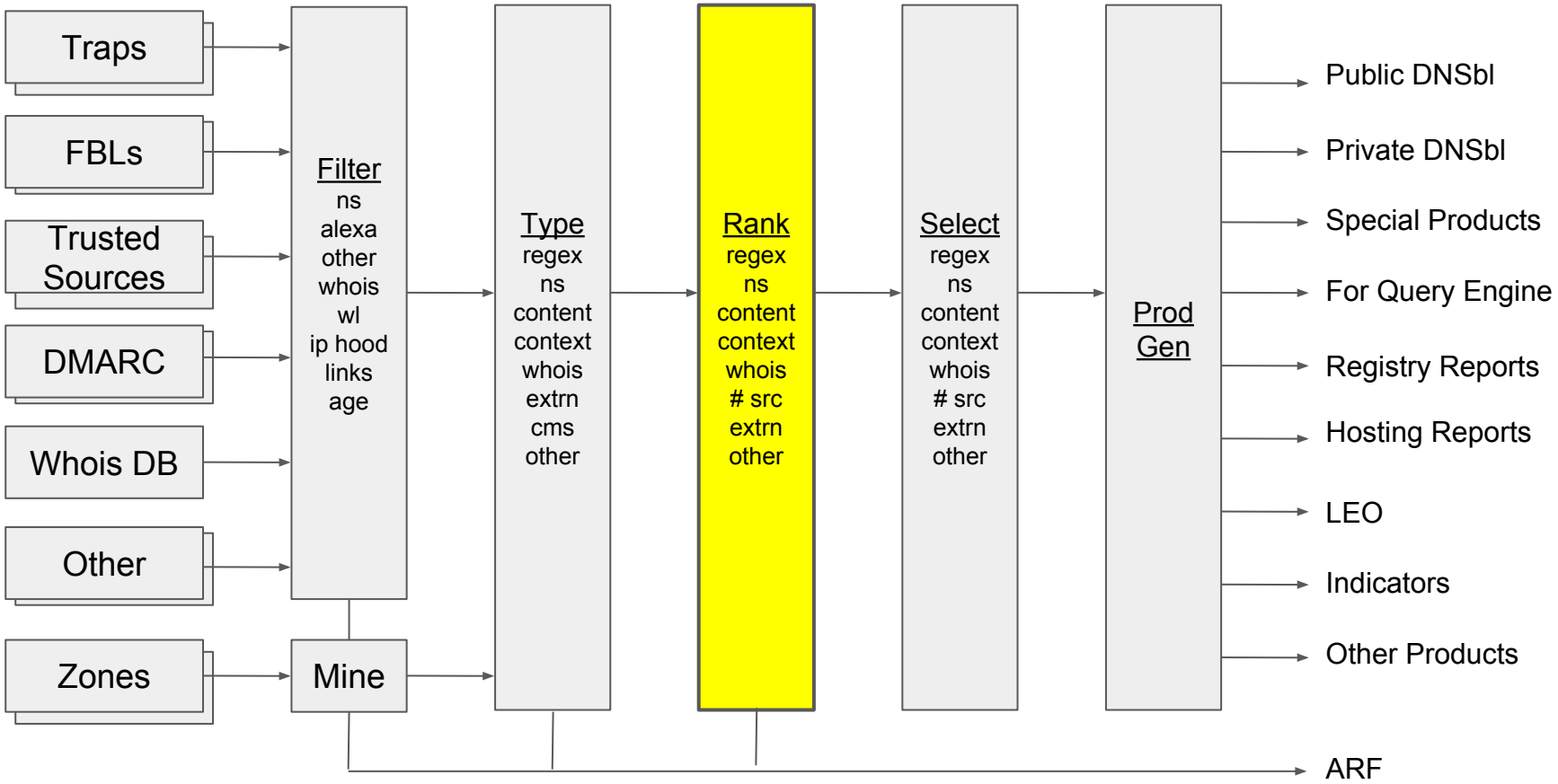
- ➔ In general, filtering is data source specific
- ➔ Reducing noise: false positives and endemically “dirty” data
- ➔ Flag the “800 lb gorillas” that need special treatment because of potential collateral damage

Reputation Mechanics for Hosts Categorizing Listings



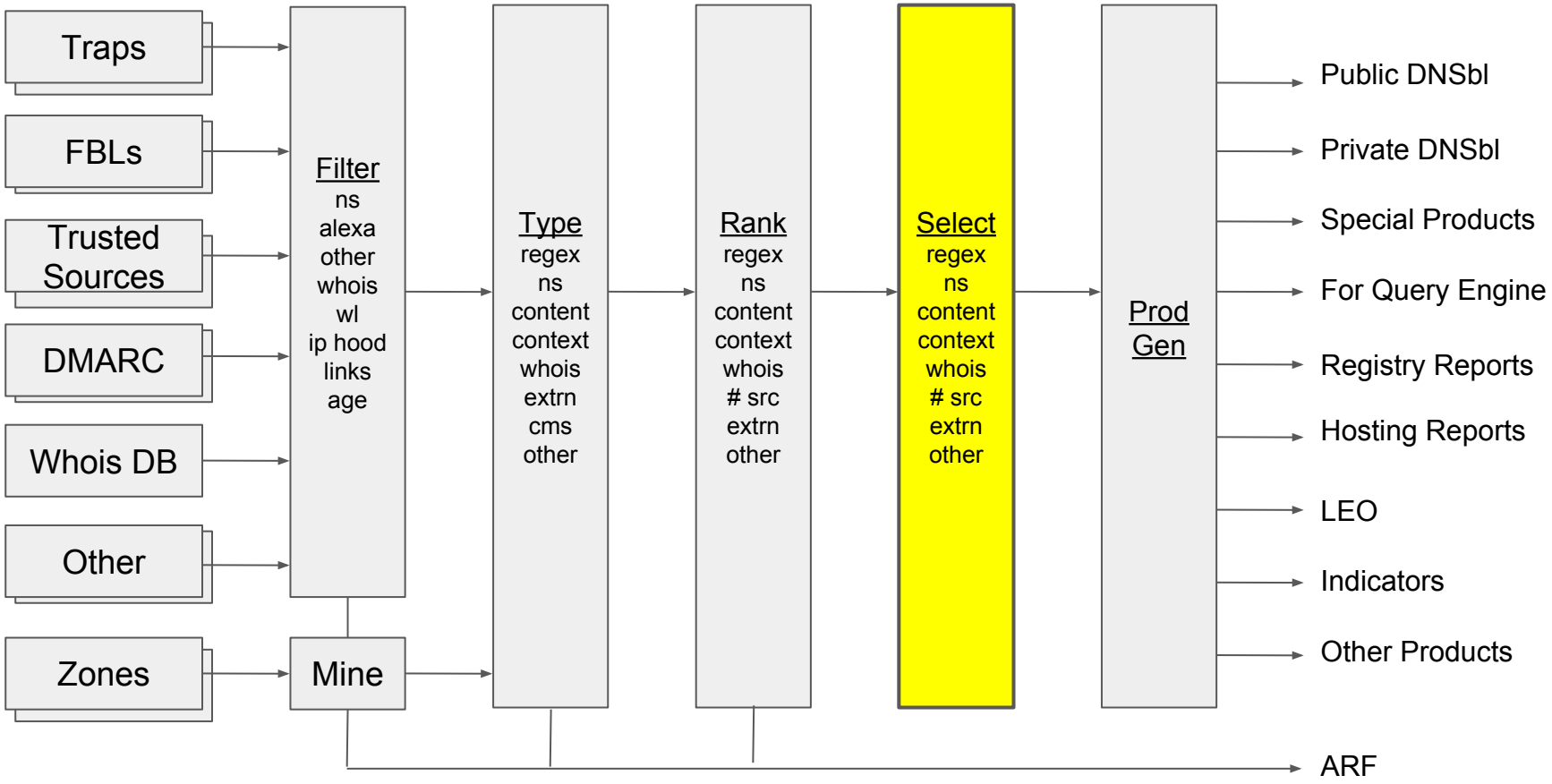
- ➔ Typing indicators
  - ◆ Patterns in URLS, content
  - ◆ Identifying listings from known bad actors
  - ◆ Correlations between IP space and nameserver

Reputation Mechanics for Hosts
Scoring Listings



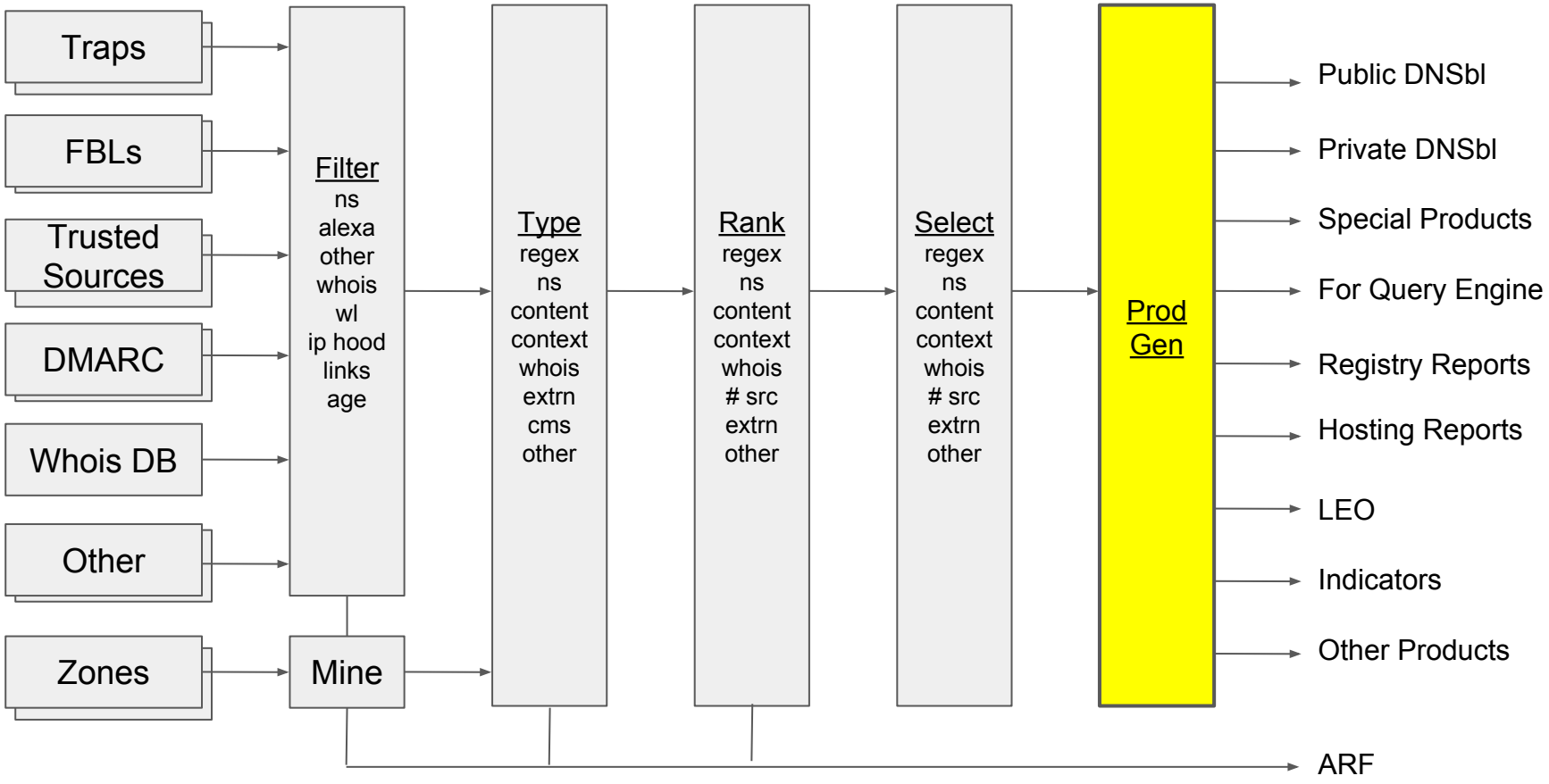
- ➔ Types condition mode of content analysis
- ➔ History of IP and domain space
- ➔ Correlation *across* diverse data sources
- ➔ Pattern identification

Reputation Mechanics for Hosts
Selecting Valuable Indicators



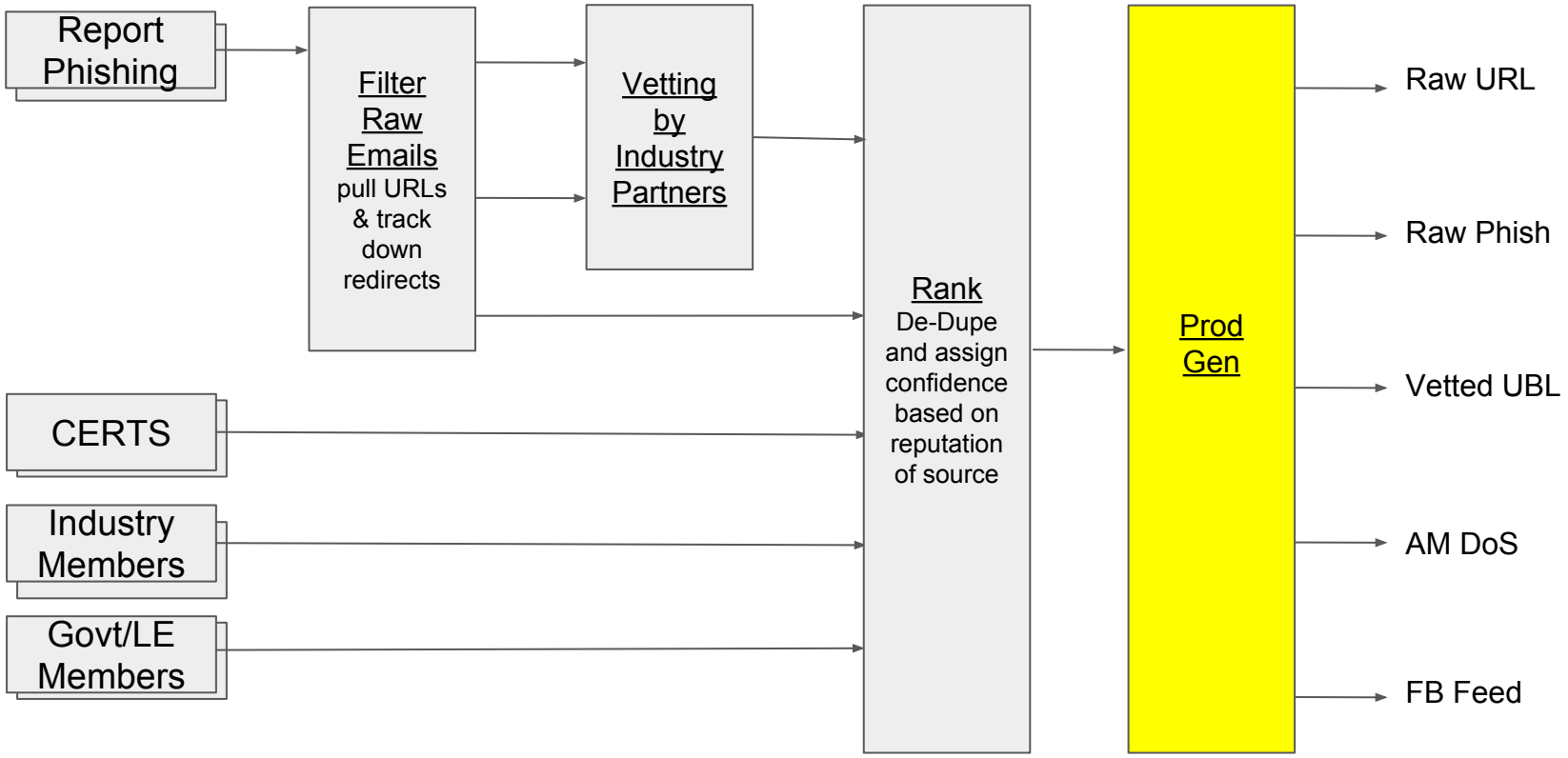
- ➔ Customer specific preference set
  - subset of salient types
  - where and how used in client value proposition
  - confidence score

Reputation Mechanics for Hosts
Selecting Valuable Indicators



- ➔ Formatting the data for particular customers
- ➔ Making the data *consumable*

Reputation Mechanics in phishing



Data sources vary and determine rankings

Background tasks to verify existence and expire entries

Final processing takes reputation from sources to determine confidence ranking

Anti-abuse Operations

Making sense of diverse indicators

Abuse Handling History

Outbound Abuse didn't matter!

- no automation
- working with mutt, yes mutt ;)
- 3 people staring at 100k reports per day (7 per second peaks)
- ~40% of the ip space was blacklisted

It was all about Inbound Security.

- very expensive Spam Filters and Firewalls
- 10+ people teams to maintain them
- topped with an even more expensive update fee
- and make sure you don't forget to pay the consulting fees

→ Arms race!

A few thoughts

- An Arms Race never solves problems.
- Filtering malicious traffic is not the same as fixing the source of malicious traffic.
- Your Inbound is your neighbor's Outbound and vice versa.

What's the plan?

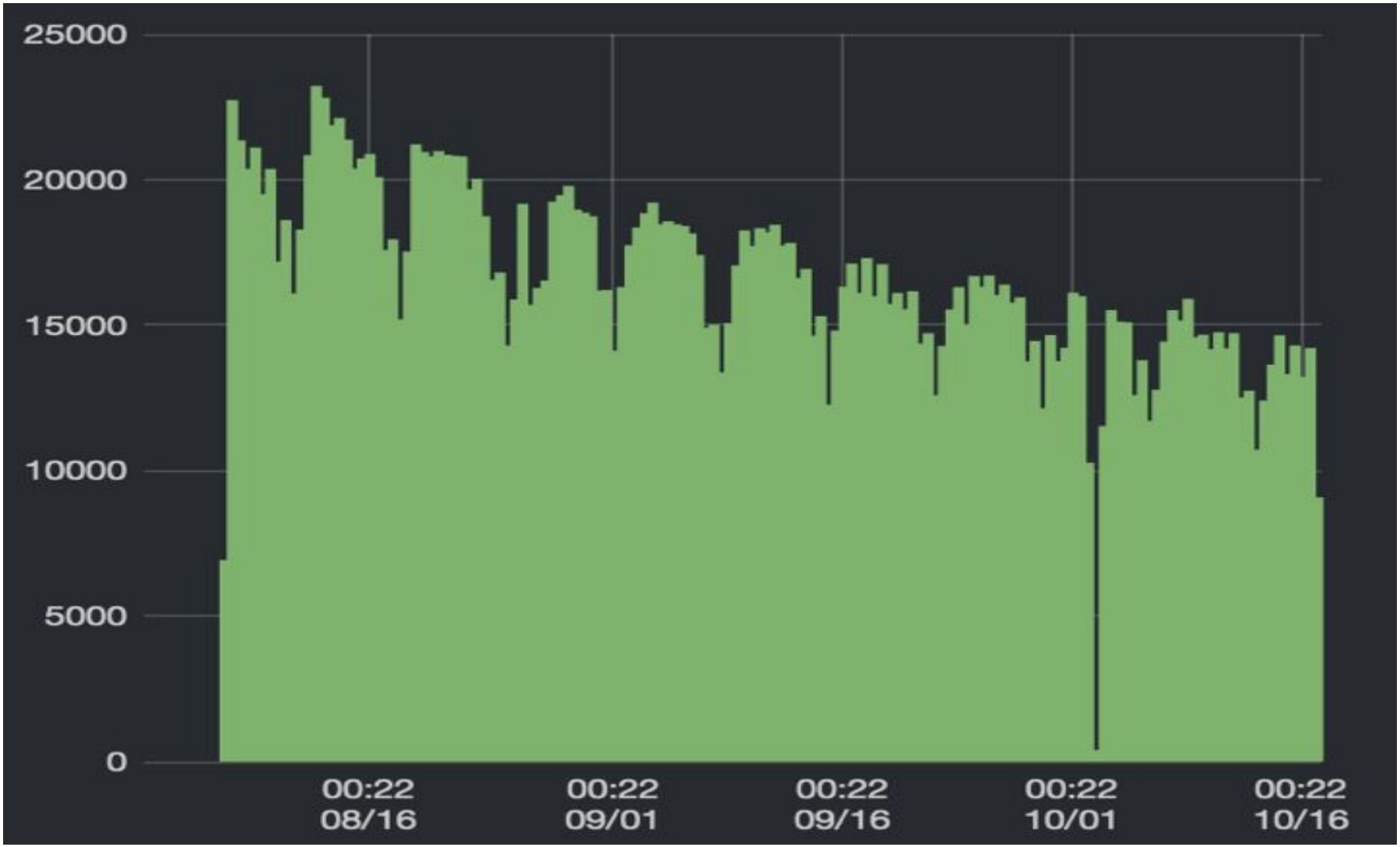
Keep it clean! Your network!

Be Fast! Automate your abuse handling.

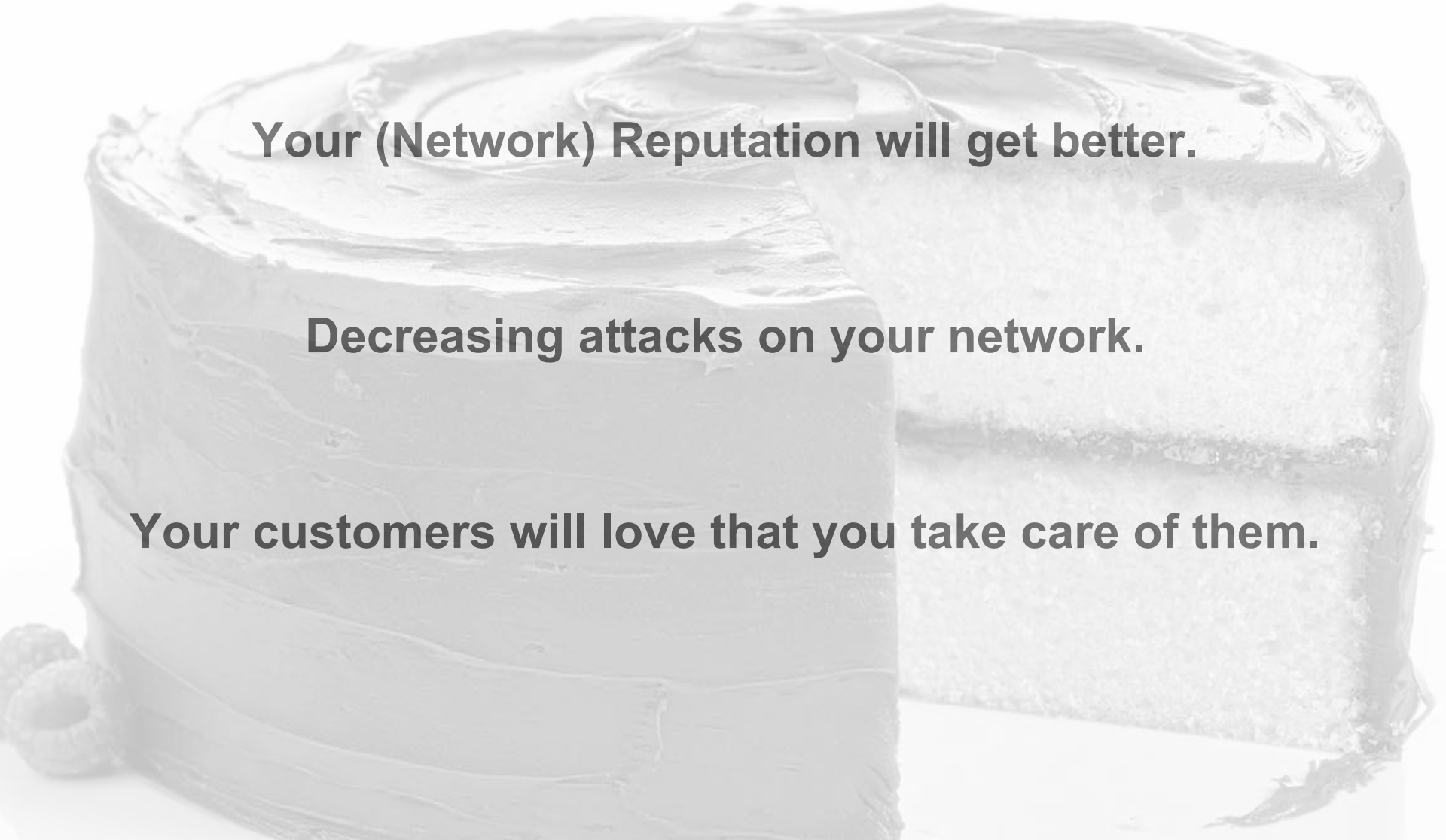
Be greedy!
Get your hands on as much information as possible.
(external, internal, 3rd Party, …)

Be generous!
Share as much information with others as possible.

Grumbot Example



What's the benefit?



Your (Network) Reputation will get better.

Decreasing attacks on your network.

Your customers will love that you take care of them.

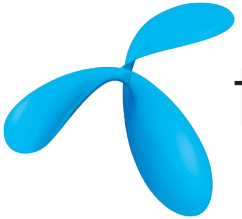
Good Examples

unitymedia

Compromise rate of less than 0.4%

TeliaSonera

One of the cleanest networks in the world.



telenor

99% of customers that have been contacted by the abuse department loved the fact that telenor took care about them.

Reputation Panel

Questions and Discussion