NCAP Discussion Group Teleconference | 24 February
Agenda:

1. Welcome and roll call
2. Update to SOI
3. Update on Study 2
4. Standing Agenda: **https://docs.google.com/presentation/d/1VjuBhPYQ7K59vppKxHSM-hpLnOIov_94-boWprFllMg/edit#slide=id.p**
5. AOB

Table of Contents

SOI: no updates

SLIDEDECK:

https://docs.google.com/presentation/d/1VjuBhPYQ7K59vppKxHSM-hpLnOIov_94-boWprFllMg/edit

This slide deck presents how NCAP Discussion group will do their work

Name Collision Analysis Standing Agenda

Slide 1: standing agenda

Standing Agenda

1. Maintain a list of Study 2 deliverables and their status
2. Actions towards the content of those deliverables
   a. Board questions
   b. Name Collision risk flowchart / algorithm
3. Data questions for other sources
4. Data questions still outstanding
5. CORP, HOME, and MAIL analysis case study and seperate Board question
6. Overarching principle of identifying 'harm'
   a. Security and privacy issues

SLIDE 2: Study 2 Deliverables and Their Status

Study 2 Deliverables and Their Status

Study Two goals include the following:

1. Understand the root cause of most name collisions
2. Understand the impact of name collisions

Study Two deliverables include the following:

1. Produce reports on the results of Study Two
   a. Report on root cause analysis, results of Study Two Task 1 (Root cause analysis)
      i. ***Status***: TBD
   b. Report on .CORP, .HOME, and .MAIL with the objective of being responsive to the Board specific resolution
      i. ***Status***: Initial case studies of strings measured against A and J data
   c. Report on impact analysis, with specific focus on the changes in the ecosystem
      i. ***Status***: TBD
   d. Final Report on the results of Study Two, which will include both prior reports and any additional results arising from their public consultations
      i. ***Status***: TBD
2. Undertake public consultations on the reports of Study Two
   a. ***Status***: TBD

We want to use structure on #2 slide. Consider details of each item on this slide.

SLIDE 3: Actions Towards the Content of Study 2 Deliverables

Actions Towards the Content of Study 2 Deliverables

Study Two actions include the following:

1. Conduct root cause analysis
   This is to be a review, study, and detailed analysis of all name collision reports that ICANN has received.

2. Conduct impact analysis

   a. Conduct case study of select names (including .corp, .home, and .mail) using at least root server data and global resolver data, if feasible.

   b. Conduct a data sensitivity analysis to determine what data can and should be used when evaluating name collisions.

   c. Answer research questions as proposed in Appendix 3, and allow for iterative analysis as needed according to intermediate results.

SLIDE 4: Actions Towards the Content of Study 2 Deliverables

Actions Towards the Content of Study 2 Deliverables

Risk Assessment Activities:

- NCAP DG has identified numerous activities to help answe Board questions.
- To execute those tasks, it may be beneficial for the NCAP DG to leverage existing processes for doing risk assessments.
- The purpose of NIST Special Publication 800-30 is to provide guidance for conducting risk assessments of federa information systems and organizations, amplifying the guidance in Special Publication 800-39.

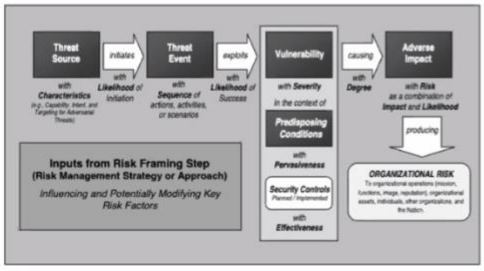
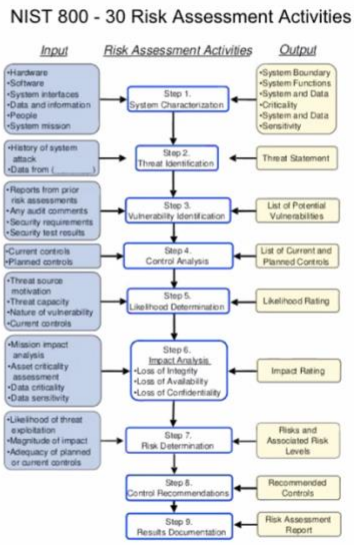


FIGURE 3: GENERIC RISK MODEL WITH KEY RISK FACTORS



SLIDE 5: Actions Towards the Content of Study 2 Deliverables 2

Actions Towards the Content of Study 2 Deliverables

Board Questions	Study Two Tasks
(1) a proper definition for name collision and the underlying reasons why strings that manifest name collisions are so heavily used;	Completed during Study One but subject to revision according to analysis in Study Two
(2) the role that negative answers currently returned from queries to the root for these strings play in the experience of the end user, including in the operation of existing end systems;	Conduct impact analysis
(3) the harm to existing users that may occur if Collision Strings were to be delegated, including harm due to end systems no longer receiving a negative response and additional potential harm if the delegated registry accidentally or purposely exploited subsequent queries from these end systems, and any other types of harm;	Conduct root cause analysis Conduct impact analysis
(4) possible courses of action that might mitigate harm;	Conduct root cause analysis Conduct impact analysis Study Three Tasks to follow
(5) factors that affect potential success of the courses of actions to mitigate harm;	Study Three Tasks to follow
(6) potential residual risks of delegating Collision Strings even after taking actions to mitigate harm;	Conduct impact analysis Study Three Tasks to follow
(7) suggested criteria for determining whether an undelegated string should be considered a string that manifest name collisions, (i.e.) placed in the category of a Collision String;	Produce a report on the results of Study Two
(8) suggested criteria for determining whether a Collision String should not be delegated, and suggested criteria for determining how remove an undelegated string from the list of Collision Strings; and	Produce a report on the results of Study Two
(9) measures to protect against intentional or unintentional creation of situations, such as queries for undelegated strings, which might cause such strings to be placed in a Collision String category, and research into risk of possible negative effects, if any, of creation of such a collision string list. **(10)** to present data, analysis and points of view, and provide advice to the Board regarding the risks posed to users and end systems if .CORP, .HOME, .MAIL strings were to be delegated in the root, as well as possible courses of action that might mitigate the identified risks.	Produce a report on the results of Study Two Study Three Tasks to follow Produce a report on the results of Study Two

10 Board Questions we need to answer
Some have dependencies on study 3 to follow study 2, such as #4 & 6. Keep this in mind.

SLIDE 6: Data Questions for Other Sources

Data Questions for Other Sources

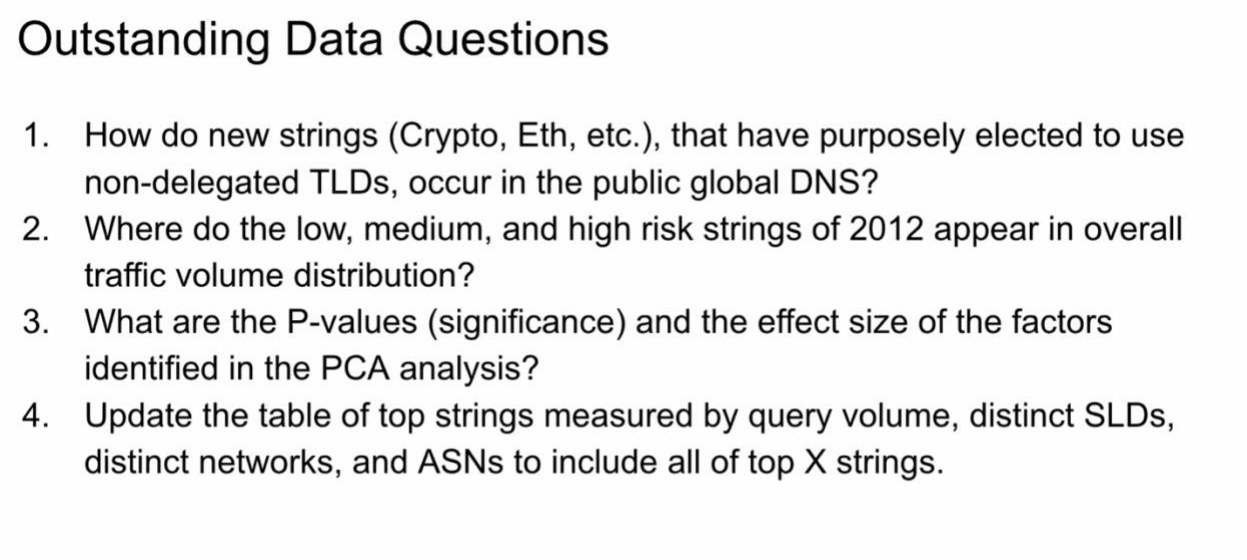
Questions	Desired Measurement Source(s)	Status
1.) What additional signal is gained by adding additional roots?	DITL	
2.) Given a set of strings and a set of DNS-based risk measurements, how does a Monte Carlo simulation of random data sources influence risk assessment?	DITL	
3.) Given a set of strings observed at the root from an open recursive resolver, how does the traffic volume and other characteristics appear at the recursive?	One or more root letters and an open recursive resolver.	
4.) Given a set of strings observed at the root that are mainly Qname Minimized, how does that traffic differ at the recursive?	One or more root letters and an open recursive resolver.	
5.) Given the set of ICANN collision reports, is there any signal pre/post-delegation of the string that in hindsight could have indicated name collision risk?	DITL	

Encouraging team to edit within slides (?)

#1. Data sensitivity analysis. Does signal growth continue to expand or subside

#2. If you randomly picked L root or A root, how different would those assessments come out?
#3. Look at A & J data, find a big open recurser like Google, identify A & J going to them and open with that recursive resolver to see what other traffic properties they are seeing on those strings. How diverse?
#4. Root server operator data and a recursive resolver data. Look at non-Q named
#5. DITL is probably only data source. May have some A& J but is diminished pre-2016.

SLIDE 7: Outstanding Data Questions



Questions the team has brought up in last 2 months.

JEFF: revised study 2 proposal says ICANN has 40+ name collision reports they used on Study 1. Will we look at those. Pre-2012 data and determine trend.
Matt: technical investigator will have access to those reports, but there are admin issues for all team to get access

#2: look at distribution of how much query volume each string receives.
#3: look at p value

SLIDE 8: corp., home., and .mail Case Study/Board Question

CORP, HOME, and MAIL Case Study / Board Question

1. We've completed an initial case study of these strings against A and J data
   a. **CORP**: https://docs.google.com/presentation/d/1mcOpf-4bugrc_aqVQCn5LaC5CwF4QHdOT1MClxSzcY4/edit
   b. **HOME**: https://docs.google.com/presentation/d/1A8u1acNf85PMCKiAEC_inzfOIm3cQUkGzpSsKyI33FQ/edit
   c. **MAIL**: https://docs.google.com/presentation/d/1msCT0aZJ6fBuB7Xq5N7HUzQhLFGMCSlHnHchmvtQxFo/edit

2. Conduct case study of select names (including .corp, .home, and .mail) using at least root server data and global resolver data, if feasible.

3. Next steps?
   a. Re-examine these after establish a set of risk factors and flowchart?

Do we feel comfortable with the analysis already done?
Matt: suggests #3 s

SLIDE 9: Overarching Principle of Identifying 'Harm'

Overarching Principle of Identifying 'Harm'

- What is "harm"? Does it imply physical? Cyber? Reputational? Or is it compromised credentials, systems, or data? The connotation of "harm" may include numerous things making it difficult to appropriately apply scale and context to this otherwise broad term within the scope of name collisions.
- We propose the following broad categories based on our analysis of the literature and data reported:
  - **Interception and Manipulation**: Private queries leaking into the public DNS that were previously answered by the root servers can be subsequently received and answered by various parties, either purposefully or unknowingly, after the delegation of a TLD string. In such a scenario, an attacker's exploitation of name collisions will allow them to intercept and manipulate DNS queries. Through these name collision events, attackers may capitalize on a variety of passive and active attack vectors including reconnaissance/enumeration, MitM attacks, internal or personal document leakage, malicious code injection, and credential theft.Some of these attack vectors and corresponding risks stem from DNS-SD or zero-configuration protocols that utilize the DNS as a bootstrapping mechanism. Coupling those protocols with either intentional rooting of a namespace in an undelegated TLD or through unintended consequences of suffix search lists, these types of queries are often the most exploitable attack vector in a name collision scenario.
  - **Signaling Interruption**: This is likely a spillover of Board question #2 that discusses the role played by negative answers currently returned from queries to the root. Some things that come to my mind would be breakage of applications that utilize the DNS as a signaling tool rather than as a directory (e.g. Chrome startup, Mozilla DoH, etc.). These situations again are likely due to search list processing. Do we want to talk about the impacts of signal changes when controlled interruption is deployed or the TLD is delegated (with registrations)? For example, how a browser would change its user displayed error message from something like "Domain not found: NXDOMAIN" to something around "Cannot connect to…." Another scenario is one in which conditional logic of the returned DNS answer is baked into the application and can be handled in many different ways….making it difficult/impossible to assess/track/remediate/etc. (e.g. Mozilla encoding of 127.0.53.53 into their DoH logic within the application).

The slide text above was in Study 1's appendix.

Review what we drafted on “harm” and see if we want to refine and how this plays into our overall assessment.

Jim: Definition of HARM is critical for Study 2. How do you evaluate harm?

Jeff: provide guidance to Board on how to judge the magnitude of such harm. Scale of harm, measurements, likelihood to occur. What is ICANN’s role in harm?

Greg: severity of the harm (the depth) as well as magnitude (breadth).

Matt: how to start determining severity/magnitude based on data in study 1 report as well as case studies presented. Study 1 – academic papers looked at name collision vulnerabilities: what do we think of these known issues and how does that vulnerability then become weaponized in a name collision situation to determine the magnitude. How do you contextualize it?

Jeff S.: mitre attack framework (cyber-security framework) and categorizes what bad guys do.

https://attack.mitre.org/

Most big problems are where there is NO higher level cryptographic authentication. SMTP is vulnerable to collision types of things because people don’t use the cryptographic authentication.

Jeff N.: string has no context on it’s own prior to delegation per SSAC’s SAC103 & SAC114. Until SSAC changes their advice we can’t look at the context