ICANN75 | AGM – NCAP Discussion Working Group Meeting
Monday, September 19, 2022 – 13:15 to 14:30 KUL

JENNIFER BRYCE: Thank you. Good morning, good afternoon, good evening, everyone if you're joining us online. My name is Jennifer Bryce, ICANN Org project manager for the Office of the CTO. I'm here supporting the NCAP discussion group meeting on the 19th of September in person in Kuala Lumpur. I suggest that those of us here in the room go around and introduce ourselves and so the people that are in the Zoom room will capture your attendance and add that to our Wiki archive as well. I'm going to start over there with Julie, and then we can go around.

JULIE HAMMER: Julie Hammer, SSAC vice chair.

BARRY LEIBA: Barry Leiba, SSAC.

STEVE SHENG: Steve Sheng, ICANN staff in support of SSAC.

JIM GALVIN: Jim Galvin, NCAP co-chair.

MATT THOMAS: Matt Thomas, NCAP co-chair.

ROD RASMUSSEN: Rod Rasmussen, SSAC chair.

WARREN KUMARI: Warren Kumari, SSAC something-or-other.

JAAP AKKERHUIS: Jaap Akkerhuis, SSAC.

JENNIFER BRYCE: Thank you, everyone. And I will be monitoring the Zoom room as well. So, please do raise your hand in here or in chat if you have a comment, and we'll capture it here. So, with that, I will hand it over to Matt.

MATT THOMAS: Thank you, Jennifer, and welcome, everyone to the discussion group call. Rubens, welcome. Would you mind just quickly stating your name for the call so everyone's here?

RUBENS KUHL: Rubens Kuhl, Nic.br.

MATT THOMAS: Thank you, and welcome. Just a quick reminder, since this is an NCAP discussion group meeting at ICANN75, it is open. So, if you have questions, please raise your hand online, and we will take those. And with that, can we get the next slide, Jennifer, and we'll talk about our

agenda for today. So, it's been a little while since we've had a discussion group call. So, let's just go through the standard administrative items at the beginning. We've already gone through the welcome and roll call. I'll ask at this point if anyone has any updates to their SOI. I'm not seeing any hands. Okay. Thank you.

So, just going quickly over the agenda for the call today, we're going to get a status update from Jennifer around how the project is progressing, where we're at with the study two document, what the tentative timeline is going forward. And then, today's call is really focused on that study two report. We want to kind of give a high level overview with the current structure of the study two report, how it is currently structured in terms of sections and just discuss the overall mechanics of that report.

And then, we would like to open up the discussion to the group here to identify any kind of gaps or additional items for discussion, concerns. We'd like this to be an opportunity for people after looking at the document structure to see or state some things that they would like to make sure that gets included or if we need to have additional structures on that.

The other thing in between those two, we'll be presenting two new figures that have been developed. One of them is by ICANN staff, who has been very gracious and developed a very nice-looking timeline figure for the study two document. The other one is a first draft that ICANN staff is continuing to work and make that for us.

And then, like I said, we'll move into the gaps and subjects for additional discussions. I know during the writing calls one of the items that came up was coming up with a little bit more detail and substance for exactly what the technical review team or the TRT is supposed to do when it comes to assessing risk. So, I took an action item from that and kind of rough outlined a few potential heuristics for how such kind of analysis might go forward to help, hopefully, stimulate some conversation in here. And then, we'll close with any other business. So, with that, Jennifer, do you mind switching over and giving us a project status?

JENNIFER BRYCE: Of course. Thank you. So, I put together this very high level slide to give you an overview of where we are and what's been happening in the past couple of months. So, as Matt mentioned, the writing group has been meeting. They continue to meet twice a week. They've been spending that time building out the draft report text, together with Heather, the technical writer, based on discussion group discussions. And, hopefully, we will have the report ready for the discussion group to review as a whole sometime in September. Well, that was the plan, but I think it is delayed a little bit, given what Matt said. And there's some sections that still need a little bit of work. So September is probably more likely to be October. And

October, November, public comment on the draft report, which will include appendix documents. So, that's the case study of collision strings and the first [inaudible] study of the DNS queries for nonexistent TLDs. Those documents are finalized, and the discussion group reached

consensus on those in June or July, excuse me. But they will still be included in the public comment package there.

And then, the root cause analysis document, which is currently subject to the final call for comments, and at some point in the next couple of weeks, the discussion group will take a final consensus call on that document, as well.

And during the public comment period for the study two draft report, the discussion group will take some time to consider study three, our mitigation strategies, so that in the early calendar year of 2023, the discussion group will be able to provide a final study two report, which will include any adjustments based on the public comment periods and published together with that the consideration of study three.

So, that's where we are. Like I said, this is our target timeline. Already, we have the September dates that are likely to be moved, given that there's still some work to be done in the draft report. But once it's available, the discussion group will take the time to go through the document together and take a consensus call on that before it gets published for public comment. So, hopefully, that provides some clarity on where we are and what's been happening. And if anybody has any questions, I'm happy to answer them. Otherwise, Matt, back to you.

MATT THOMAS: Thank you, Jennifer, and thank you for keeping us on track and monitoring the status of the project as we progress. Like we said, we'll aim for Q4. But obviously, the situation is fluid, and we'll adjust accordingly. Jim, over to you.

JIM GALVIN: Yes. Thanks, Matt, and thanks, Jennifer. I wanted to make an appeal here in this open forum at ICANN75 and just remind people. Discussion group members, you do have access to the draft document as it's developing. And you are certainly welcome to take a peek at that as we go along. You'll see as Matt walks through the document that there's a fair amount of text at this point which is reasonably stable.

There's still a lot more to come. There is a lot of writing to be done. We have been saying that we are imminently done since before ICANN74. And the more writing we do, the more writing we discover we have to do. And so, we just keep slipping a little bit.

And we're focused on getting it right rather than meeting the deadline as compared to our commercial product offerings we often find in this world meeting the deadline. But I do want to encourage people. Please, on behalf of being a co-chair, help contribute or please do feel free to look at the work that we're doing. And now, you're going to get a good look at it because now we have a good chunk of stuff to show. And Matt will be going through all of that and remind everyone to do that.

And then, the next thing I'm going to do just, yes, I should get these people over here, I'm sorry just to ask but we're asking everyone in the room to state their name for the record. And so, Peter and Suzanne, if you wouldn't mind just doing that, and then back to Matt to continue the meeting.

PETER THOMASSEN: Yes, sorry for being late. I'm Peter Thomassen from deSEC.

SUZANNE WOOLF: I'm Suzanne Woolf, SSAC member.

MATT THOMAS: Thank you, and welcome, and thank you, Jim. And like Jim said, we'll continue to do our best working forward on the document. And I think there's the famous triangle good, cheap, and fast. And obviously, we want it to be one of our priorities. So, we'll keep pushing forward.

JIM GALVIN: [inaudible]

MATT THOMAS: Yes. Do you mind giving an introduction?

KIM DAVIES: Kim Davies, IANA.

MATT THOMAS: Thank you, and welcome. Jennifer, do you mind going onto the next slide? So, let me just at a very high level try another walk through the current structure of the study two report. There's six main sections as it stands right now. I'm just going to kind of provide a super high level overview of those six right here.

The executive summary is the standard intro text of it that is still to be done and completed. But the main portion of it starts off in the introduction. And here, we define the background. And the background is a pretty extensive 17, 20-page document or section looking at the historical context of name collisions within ICANN, the genesis of the work that has been done, when certain decisions were made, and how those influenced future works going into it.

And I think that's a very important portion of this document because it really kind of lays the groundwork, the context in terms of what happened back in the previous round and what are we proposing differently in this name collision discussion group and why is that substantive and better than what was back in 2012. So, I think it's a very important section of this document to keep the framing correct.

As part of that background, there is one of the two figures that I mentioned earlier. It's a historical timeline figure that clearly depicts all of the major publications and events that occurred since 2012 up to the present day. And I think that's a nice way for us to look at and better understand when certain advices were given to the community and what different methodologies were actually applied to name collision assessments, mitigation strategies, and so forth.

The next section in the introduction is really the methodology talking about how the discussion group functions, the roles that the technical contractor—[CASEY's] work in this—how the work was conducted, put out for public review, and the overall structure of the discussion group in terms of the substance and matter that were discussed and going onto it.

And then, the last section within that introduction is the terminology. And as all of you have known over the last several years, the terminology has been I think one of the more difficult components for us. We've changed terms a few times going through this between active and passive and honeypot and sink hole and whatnot. And it's a very important section of the document that we clearly define the names and what they reference in terms of name collisions going forward.

Within the document in the terminology section, there is an earmark for a separate section specifically talking about the terminology around harm and impact. A lot of our conversations, we tried to come up with what the connotation and the denotation of those two terms are, especially within the context of name collision.

And while we don't have any text exactly for that section right now, we think that having that explicitly stated at the beginning of this document helps provide some of the context and the guardrails for the readers of the report to better understand what is being said in the rest of this study two report.

Section two is really a summary of the NCAP's study two reports that Jennifer mentioned that we have a consensus on the first two and outstanding consensus call on [CASEY's.] This is just really pretty much an executive summary of the case study document and a perspective study and the root cause reports.

All three of those reports are going to be included as appendixes in the study two report. This section really is just kind of is more of a synopsis

of the executive summaries of all of those reports put into the main body of the study two report.

Then, that brings us into item three, which is a summary of the NCAP discussion group activities. As we were developing the writing of the study two report, initially, the report was primarily focused on just the case study, the perspective study, and root cause reports. And it kind of dawned on us that there has been an exorbitant amount of effort by the discussion group community over the years. I think we're up to a hundred meetings or so, of discussions, presentations, review of material.

And it would definitely behoove us to do a summary of that discussion material and put it into the study two report to reflect all of the effort that this community has placed on the name collision project over the years. And this includes the original NCAP gap analysis that's stimulated the changes or motivated the changes in the study two goals and objectives that we're currently executing on.

It also talks about the review of available data sets, some mentioning of how the issues around manipulation, and how that can influence name collision assessments and what data can be used and how so. Our conversations around control of interruption, the efficacy of it, the intent of it, as well as a review of the other additional proposals for a new measurements and assessment mechanism such as the ad-based measurement that provides a different perspective on name collision based off of an alternative to looking at just passive DNS data at the root server system.

Then, obviously, we also have the main work of the discussion group has focused on over the last year. And that is the work flow development in terms of revising what kind of advice that the discussion group believes the right sequence of events is for ICANN to go forward in the next round to be able to assess name collisions in a sustainable, repeatable fashion that facilitates the risk management framework that we've been discussing that this is really the core problem of the name collision project, that we need to provide the guidance in terms of how risk can be either quantitatively or qualitatively assessed for these applied strings and to allow ICANN to proceed with that type of work in our particular guardrails and context.

And finally, then, about a month or so ago, we ran through a series of tabletop exercises looking at nonspecific numbers but things like labels of high query volume, low query volume, and running them through the "corner cases" of the workflow to understand what scenarios might those types of situations entail when engaging with the technical review team as well as passive collision assessment and the active collision assessment. Any questions at this point, or should I continue onto the last three sections?

I'm not seeing any hands. Perfect. Next slide, please. So then, section four is really pulling out the collective findings from both the study two reports, which are the case study, the perspective study, and the root cause analysis, as well as section three's discussion group activities. What were the main findings that we have all identified over the last several years in doing that research? And how does that influence making the name collision problem a risk assessment problem?

This section's still kind of a work in progress. This is probably where most of the writing is occurring right now. The first three sections are fairly well I would say fleshed out. There is 60 to 70 pages of those sections right there. If you definitely want to get into the first three to start reviewing, that is an excellent place.

Sections four, five, six are clearly where the main work in progress is right now in terms of writing. But before we get into section five, we need to finish section four to understand... I'm sorry. Before we get into the recommendations, we need to finish section four and document and codify all of our findings that we've established so far in the work.

Section five right now is really kind of almost a stand-alone piece that tries to build upon the previous four sections. And that is directly answering the nine Board questions that were given to the SSAC in the name collision discussion group to answer. This was, obviously, one of the main charges for the discussion group and the obligations for us to fulfill here.

And so, we're trying to figure out best how to structure and represent the knowledge that we're documenting in the first four sections and relating it into these Board questions here. There's still some concerns as to the flow and structure of best positioning these Board questions in this document.

It might be that the study two report actually gets split into two different ones, one containing mostly the workflow suggestions, and a second containing a direct answer to the Board questions instead I don't think we want to move those into the annex at all, since that is a

direct objective of what we were supposed to achieve here in the NCAP discussion group.

Then, we would get into probably the meat of the document at the end in terms of actionable items. And that is the recommendations that we want to put forth. This is still, obviously, a work in progress. We want to make those recommendations solid based off of the findings that we have identified in section four. And this is where some additional work needs to be placed right now into the document after we finish up section four.

Section seven, that advice to the ICANN Board/conclusion. I'm not sure if we'll actually end up with advice to the ICANN Board or if that will just kind of naturally flow into the recommendations or the executive summary portion. It's mainly there as a placeholder right now. Regardless, we'll obviously have a conclusion. And that will need to be done at the end of the document as well.

So, that is probably the study two report from a 10,000-foot view. Like Jim mentioned before, the document is available to all of you within the discussion group shared folder or shared drive. So, please get in there, take a look at it, make comments, write some text. Any kind of edits, additions, they're all welcome and greatly appreciated. Does anyone have any questions or concerns about the study two document structure or flow at this point?

All right. I'm not seeing anything. So, here's one of the first two new figures then that I mentioned while going over the agenda. I had drafted the original timeline of this. And Jennifer was able to get this over to

ICANN staff who really made a much more attractive-looking figure than my PowerPoint stick figure that I had originally drafted.

But here, you can see the collision assessment timeline that we've been talking about before. It represents the last timeline that we had talked about in the previous discussion group where we start off with the applicant's static assessment. There's always the offramp where the applicant doesn't absolutely have to apply for the string. And it gives them that option to go not submit there.

And then, there's the application processing begins. And then, at some point, the technical review team will start its static assessment of the string, which will give an opportunity for offramp number two before the string then goes into the passive collision assessment, which we've highlighted that it runs there for the 90 days.

And then, again, with an offramp before the string proceeds into the active collision assessment for another 90 days, again, with another offramp before the final string goes into assessment package submitted to the Board with the Board decision and then any kind of host name collision assessment activities agreed there.

So, we're hoping that this figure better clarifies the workflow for folks and gives them something more visual to ingest and to understand then the verbiage and some of the bullet points that we've had on the slides going forward. Any feedback or suggestions on this is always welcome.

But this will be heavily used tomorrow in the NCAP general open session talking about the workflow [inaudible] to the public again tomorrow. I know this is the first time the discussion group ahs seen it.

But yes, this is hopefully the new graphic going forward. Next slide, please, Jennifer.

STEVE SHENG: Matt, this is Steve. For the audience, could you briefly recap the static assessment, passive collision, and the active collision, what those entail, just briefly? Thanks.

JIM GALVIN: I will take that question. So, remember, right, this is just the picture of our workflow overall. And this workflow has not changed since really early this year, and it may have been even earlier than that. We have certainly been focused on some elements of it. But, in essence, there's this notion of static assessment. And an applicant gets an opportunity at step one there to do their own static assessment as well as TRT doing a static assessment. And what that entails is looking at the publicly-available published data, which ICANN already does.

There's already indicators of what are the top NX domain queries into the DNS, for example, on its website. And so, this is really just an opportunity for applicants and the TRT to look at what that data may or may not have to say about the string that's being applied for and to have some thoughts about it. So, that's just a step one. It's just a quick look at whether or not your name obviously falls into a special case category. And you can do with that whatever you would like as an applicant. And then, of course, the TRT will have to do what it's going to do.

And a passive collision assessment is the model where the name is actually delegated but it's delegate with an empty zone. So, the intent there is to set up a situation where you pull out of the rest DNS infrastructure, you pull data up to the authoritative servers for the TLD so that you're getting more data than is generally available in the static assessment because there are, we know, at a minimum today, recursive resolvers, they essentially hide some data from the authoritative server so the TLD. So, again, this is just an opportunity to both see if there's any bad things that happen as a result of just the delegation all by itself. But, again, you get additional leading indicators of the volume of the queries. And the TRT is going to have to assess, and Matt will get to that when he gets later on, as to how they're going to assess that, how they're going to use that data and what they're going to do with it, at least some suggestions for how to look at that. Again, just a leading indicator as to whether or not you are going to fall into a special case category or not.

The active collision assessment then is where you, again, do a delegation. This time, you put some things in the zone. And, in particular, you put a real IP address, both that IPv4 and IPv6, which is one of the improvements over controlled interruption from the 2012 round, which only did wild carding at IPv4 in a single location.

But you actually give an address, a wildcard address, an actual address for the names. And you collect not just DNS queries but you also look to see if a few other kinds of queries are being used by the name. So, you try to capture data on whether, for example, it's web queries, is it email

queries. You want to get a sense of the volume and the distribution of those queries.

And that's what we're looking for here. Again, it's just a little bit more information. Well, maybe, it might be a lot more information, depending on the name. But you are both trying to see if the delegation itself causes harm or impact if that becomes visible. And then, you want to get a sense of whatever the harm or impact is behind the query by getting some knowledge about what the query actually is. And that's the purpose of the active collision assessment. And with that, the TRT will then, again, assess that data. That assessment will be made available, of course, to the applicant.

And there'll be some discussion about it. If there was some issues there to be addressed. But that's what this step four up there is. There's an offramp opportunity. The idea is to allow for the applicant to say, "Gee, this is starting to look too complicated. Maybe I don't want to do this," or it would be an opportunity for them to submit a mitigation or a mediation plan, which then has to be considered as part of the package that the Board goes to on the last step. So, that's kind of a broad brush look at those individual steps. And ideally, we're going to have a lot more to say about that in the document. And I'd be really interested in people's comments and reviews about actual steps. Back to you, Matt.

MATT THOMAS: Thanks, Jim. Warren?

WARREN KUMARI: I'm assuming you're not going to be shocked by what I'm going to ask. But for the static analysis stuff where we think the applicant's going to go and look at things and infer something about the string, are we going to give them a whole bunch of guidance? Because in the top 20 now, there is a string which did not show up last quarter. It's just started showing up in the last month or so.

It's in the top 20 in the list. And I don't know. To me, it seems that's clearer evidence that that string might have issues. But there is no way to tell if that string is a gaming query or anything else. So, I don't really know how an applicant is supposed to infer things or the TRT really is supposed to infer things from the static public data.

MATT THOMAS: So, in the specific, no, we're not going to give them any advice on how to evaluate the numbers. The only thing that we will comment on currently in the document is that if your name appears in these lists, whether it's as a result of static assessment or even the PCA or ACA assessments, if your name pops up in the list and it's there, then you now know that if volume itself is significant, then you're going to end up being put into a special category, become a special case. And the TRT will have to make some additional assessment itself. And that's where the expertise will lie. So, we're not telling the applicant what to do with the data or not. But it is important to understand and recall that from the subsequent procedures recommendations, there is actually a recommendation, there's part of a recommendation, that does require that the applicant gets some information in advance about the status of their name.

And this is the best that we can do, is give them some numbers and they get to do with that what they want. We're not telling them that it's a reason not to apply. We're simply saying that if you're got high numbers, you should know you're going to be subject to additional scrutiny, and then we'll see how it goes. And that's it.

WARREN KUMARI: Okay. To me, it still seems like that creates the perverse incentive for everyone to start querying everyone, like all of their competitor's names as fast as they can to cause them to show up high in the list and cause them to be special handling. And to me, it seems like the outcome of that for the Internet is really bad.

So, it seems like we're explicitly creating an incentive for people to pollute the name space and pollute the root service system, etc. by doing queries. So, to me, that feels like fairly irresponsible. But I have said that a number of times.

MATT THOMAS: So, two things. One is in an ordinary situation, you don't know what someone's going to apply for. So, it's hard to see the system unless you've got some kind of insider knowledge or yes, maybe you can see all of the global multinational brands. And you might have started querying for all of those to make that difficult for them.

And you're right. And there's a part of me that says, okay, so people might do that. You're right. They might try to game the system in that way, which leads me to my second point, which is it's the responsibility

of the technical review team to observe that that's what's going on and know to ignore it or dig into it and see it.

So, if a name suddenly appears and they are supposed to maintain longitudinal data about the status of names in this public data, then they will observe that a name suddenly appeared. And it's up to them to figure out whether this is something being gamed and is a serious concern that has to be investigated now, or whether they're going to let things go forward without any further consideration.

WARREN KUMARI: Okay. I guess I'll just make a clarification then. Maybe it wasn't clear. I don't really care about the TRT having to do it for something. What I am concerned about is the pollution from all of the gaming queries that will be created. Like the IOIO one which I just listed in the top 20, that's a little bit over 2 million queries per day. You don't have no way of knowing that's not gaming. It would be relatively trivial for people to game high query volumes on the order of hundreds of millions of queries per day.

And somebody has to answer those queries. And I am looking at Matt as one of those people who has to suck up all the queries which are going to hit the name space. And the incentive is clearly to stop gaming and stop gaming as soon as you can so you have the longitudinal data. So, at the end of the meeting, I'm going to stand up a company that sells off gaming queries. For only a million dollars, I'll get you to the top of the list. And just the incentive model here feels we're explicitly creating

a market for people to create gaming systems. And the pollution ends up being sort of carried by the Internet.

JIM GALVIN: So, just a quick response to it, and I see Rod's got his hand up here too, Matt. Pollution is a risk. And I think that we should at least acknowledge that and make a note of that. I would welcome any suggestions that you have if there's anything you think that we can or should do in this process overall to address this pollution concern. But right now, all of our discussions have just gotten us to this place, and yes.

WARREN KUMARI: I've made the suggestion a bunch of times. Don't use things like the magnitude list because they're trivial to game.

JIM GALVIN: That doesn't change if you move to PCA. You've got to get the same data. So, anyway, we shouldn't have this fight here now. Why don't I kind of get you...?

WARREN KUMARI: You get the data for [inaudible].

JIM GALVIN: Will you speak into a mic?

WARREN KUMARI: If you move to something like PCA, you only get the gaming queries while you're actually doing the passive collision assessment. So, yes, you might have some gaming for the 14 days or however long you run the PCA thing. That's much less harmful to the Internet than a couple hundred million queries starting from now until the round closes.

JIM GAVIN: Well, probably, the only other thing that's worth mentioning is one of the things we don't know is at what point in time relative to all of this will the names that have been applied for be publicly known. So, that's just part of this risk of pollution too. Anyway, we're not going to solve this here. So, I think we'll just leave that as a point and do Matt.

MATT THOMAS: Rod, over to you.

ROD RASMUSSEN: Right, and I don't want to belabor the point, but I would game theory this out because I don't actually think that we're going to have those names and root servers crashing because of pollution or any major problems because it will be fricking obvious if somebody sets up a company to pollute the data over time. So that's something we can actually be taking a look at. But I'm not a game theory specialist myself. But it's kind of like war games, mutually assured, destruction, the best game [inaudible]. While you're saying, what was it, two million queries or whatever to get to the top of the list.

JIM GALVIN: Can you repeat the question into the Zoom room, please?

ROD RASMUSSEN: Yes.

WARREN KUMARI: I was unclear how it would be obvious because if you look at the queries, as an example IOIO, which I do not think is being gamed, but it's in the top 20.

ROD RASMUSSEN: Right. So, in my game theory model, if you have a company selling or if lots of people are interested in this, the only way you get in the top 20 is to keep increasing the volume again and again and again and again. It becomes pretty obvious that gaming is going on at that point because the only way you stay on top and get your names in the red zone, so to speak, is to continually...

WARREN KUMARI: Only if there's multiple people selling the service.

ROD RASMUSSEN: Okay. Well, then, what do we care at that point? The damage doesn't get done if it's just one name or two names, right? So, if the risk is crashing [inaudible].

WARREN KUMARI: No, I don't really think the root server system will crash. But there will be an additional load, an additional cost, and additional service that everybody has to provide for and a degradation of service quality.

ROD RASMUSSEN: Well, I'd actually ask Verisign whether that's a problem. But I'm looking at it like if the real risk is that lots of people do this, then it becomes obvious. If it's just a couple people doing this, then that's up to the design team or the TRT to figure it out. I don't know. That's just my thoughts. I'm belaboring the point.

MATT THOMAS: Okay. Suzanne?

SUZANNE WOOLF: Sure. Thanks. I do want to sort of point out that there's more to the question of impacts of polluting the name space than the quantitative things that are occurring first to us. And we've expressed concerns over time about expanding the name space. And in a way, this is an expansion of the name space.

It's an expansion of the name space that we're not likely to find useful. And I think that's also something to pay attention to. It just raises the noise in the system in a way that is maybe hard to capture quantitatively. And so, it might need to be looked at in addition to the quantitative measures you're suggesting because I agree with you about what the numbers will show. But there's sort of a qualitative concern here too.

MATT THOMAS: Thanks, Suzanne. Barry?

BARRY LEIBA: Yes, Suzanne said a little bit of what I was going to say. The other part of what I was going to say is another way of spinning this is that we're giving people numbers that... We have to give them numbers. So, we're giving them numbers. And is it really valuable to give them numbers that are of questionable use, that it's anybody's guess what they mean.

WARREN KUMARI: Yes. And what I keep hearing is SubPro said the public must have some data. Therefore, we must give them some data. This is data. Let's give them that. If SubPro's recommendation was not correct because there is no useful data, it seems much better to say, "There is no useful data that can be provided," instead of, "Here is a thing. It's meaningless," right? It seems like that's disingenuous, at best.

BARRY LEIBA: It's valid to say, "Here's what we've come up with. Here's why we don't think that's as useful as it looks." It's better not to do that than to mislead them.

MATT THOMAS: Yes. And let me just confirm and emphasize that point. You're right, Warren. If we can adequately produce some text that justifies that this is not a good thing and not useful and they should just move on from

that, and we shouldn't have this static assessment, I would really welcome someone else to write some of that text.

This is kind of the model that we're in at the moment and the path we're going down. But I'm very interested in other people picking up some writing assignments and proposing a different path here. And this seems like a nice, isolated thing that if you want to make that point, please do. And I'm happy to take that on Board and do this a little differently. So, thanks.

JIM GALVIN: Rubens.

RUBENS KUHL: Just a point from the subsequent procedures report, it's not said that ICANN must give something. What is written is to the extent possible, ICANN should supply a do not apply list or something. So, it's basically should, so not must. So, if we think it's worse to get this, then don't. There's nothing that requires anyone to provide bad data or things that would just make life worse.

MATT THOMAS: So, thank you for that clarity. I really very much appreciate that. And that just says that we really are in a place where thank you, we can go either way. Let's just get the text going with whatever direction we want to go in.

And I would welcome being able to see a well thought out, rational set of reasoning about what we're going to do or not do in this particular case. And then, sure, that makes sense to me. We can certainly go in that path instead. That will be fine.

BARRY LEIBA: Okay. I'll look at proposing some text.

MATT THOMAS: Thank you, everyone. Jennifer, can we go onto the next slide? Hopefully, this figure is a little less contentious or up for debate. This is history, and so, it's well-defined. This is the first draft of the historical timeline of all the events surrounding name collisions going back to 2012 up to the present date.

This really highlights the beginning of the application period, when the application period closed, all the way through various SSAC documents, the Interisle reports, the JAS reports, the OCTO advice, when alternate path to delegation was decided that that was okay, when it ended, and when controlled interruption actually went, and then where the NCAP study one and two works started to fall into this.

I think this is... At least, for me, when coupling this with the introduction and background text, it is very helpful in terms of framing what exactly happened back then. Here we are in 2022, a decade later. It's kind of sometimes hard to recall all of the events and all of the discussions and all of the research that has been going on since then and what's gone on when.

So, I think this is, hopefully, really helpful for the community. When they look at the beginning of this document, it gives them proper framing, the context, and understanding what happened when. I'll just note I had a conversation on the side with Suzanne earlier yesterday about really framing this study two report around making sure that we highlight the differences in what we are suggesting here compared to what happened historically and why the suggestions based off of the findings and ultimately the recommendations are not a zero sum game but they are a positive over what happened before.

And so, hopefully, this gives that information and insights to everyone to really kind of understand name collisions back then was really past the application period closing. It was in the throes of the TLDs actually wanting to be delegated. The applicants wanted their TLDs. ICANN wanted to get those out the door.

And here we are in 2022, and we have this green field opportunity to kind of step back, level set, and decide is there a more optimal sequence of events that creates less entropy or chaos in this workflow that happened to create that sustainable, repeatable model? So, hopefully, when you go into the document, and I think the text in the introduction is really good in terms of, hopefully, framing that going forward.

But yes, this is the second figure that ICANN will...that Jennifer is working with right now to get this look a little bit better than my stick figure right here. I'm sure they'll be doing a great job. But this will be one of the additional figures that goes into the study two report. Yes, Jim.

JIM GALVIN: So, I just want to add my own personal perspective to this. I love this timeline when Matt put this together. I was the one who spent most of the hours trying to put the background section together, which took hours because it was really hard to find everything. That was really what most of the time was spent trying to find my way around the ICANN website to pull all this stuff out and find all the links to everything.

But you don't really appreciate it when you're reading the text. When you're seeing all the text and the events and you don't realize the bulk of things that was going on back in those early days. There really was a lot of effort put into it between the Interisle reports and the JAS reports and SSAC and even the choices that ICANN made along the way in order to keep the program moving along.

And I just like the visibility of all that and seeing that. And as Matt said, one of the conclusions that we came to along the way here, putting all that together is making sure that the message is not that we're replacing anything that's been done before. It's actually very clear to me at this point, once having gone through very carefully through the background and seeing all the details of some of those background events and being reminded of all of them, that we really are just improving what's been done all along.

It really is just a natural evolution, a natural step forward in what's been done. And it's important that we give due credit for all of the analysis that was done before. Nothing has happened to fundamentally change

what we learned from before. But it is important to evolve with the changing dynamics, not the least of which is the change in the DNS infrastructure.

Access to data at all is problematic at this point. And also, what do you do? And so, there's opportunity to improve. And I like to think that that's what we've gotten to. That's what we've created is just an improvement, incremental movement forward, and setting up a framework for something that can continue to evolve as it needs to going forward. And I think that's really what the Board was asking us for most.

So, just my own personal perspective on this. As you look at this material and look through the document, keep that picture in mind. And hopefully, that will resonate for you too. And if not, we definitely want to know that because maybe we should do something a little different in the document. So, thanks.

MATT THOMAS: Thanks, Jim. Jennifer, do you mind going onto the next slide? Thank you. So, this is where I kind of want to open it up to the group. Over the multitude of weeks and the last month or so, during the writing teams, there have been a couple of times where certain subjects or elements of the document kind of highlighted we talked about this a little bit in the discussion group or some people felt like we needed to have a little bit more discussion.

Not to put Warren on the point here, but I think one of his comments was around maybe some more concrete guidance to the TRT for name

collision risk assessment. And what are some real heuristics that we're going to provide them without being overly specific to them? But what can we help formulate and provide in real context and advice to make that role and their responsibility achievable and effective?

So, one of the things that we have on the next slide is a rough outline of what I kind of drafted for some of those. But I would like to just kind of open it up the floor here if anyone else has any other topics that they'd like to bring up at this point. Sure. Okay.

JIM GALVIN: So, I want to emphasize that point that Matt just asked. And this is really to the discussion group. And maybe it requires, Matt, that you and I should really press a question to the mailing list to make sure that everybody gets it.

This is kind of a last call for what is your pet question that has not been addressed to your satisfaction. That doesn't mean that we haven't already talked about it and had a fulsome discussion about it, and we're going to continue down the path that we're going. But it's important for us to make sure at this point that everybody has had their opportunity to speak up about any issues or concerns that they have so that we can make sure that it has been covered along the way here.

So, that's the emphasis here. And we really do appreciate... There are a couple of people who have had a lot to say along the way. Warren happens to be one of them, which is a good thing. I'm grateful for the fact that he clearly knows what he wants to say, and he's making his point. And we're being careful to make sure we cover it.

Casey another one who has a couple of issues on his mind all the time. He comes to our writing team calls and reminds us that we're not really quite getting his point. And that's okay, and it really is a last call to the entire discussion group. We just really are begging for people to make sure that we're either on track or we're not. What are your questions? What are you concerns and comments? This is the time. So, I can't say that enough. Thanks. Sorry.

MATT THOMAS: Thanks, Jim. Steve, please go ahead.

STEVE SHENG: I don't know how much in the report discuss about gaming and its impact not only in the static analysis but also in the PCA and ACA. So, there needs to be some discussion about how observable gaming is, right? Because if it's not observable in some sense, it's difficult to observe. And that impacts PCA and ACA.

That's going to make the TRT's job much harder, right, because I think, realistically, gaming is going to happen. And people are willing to set up infrastructure to continue to [inaudible] queries. And there needs to be guidance to the TRT on how to handle that. And if it's easy to identify, then that guidance, they can easily discard that. But if it's not as discernable as we might think, then I think it's going to pose problems. So, some discussion on that. Thanks.

MATT THOMAS: Thanks, Steve. Over to Warren.

WARREN KUMARI: Yes, thank you. I don't know where I put the rest of my list, but I have sort of a list of the what else bits. Somewhat responding to Steve's thing, I worked out how much it would cost to stand up a gaming service that would be, I think, fairly pretty much completely unstoppable, undetectable.

And I believe that it would cost about $24 per month because I need two servers, and they're $12 each, and probably about three or four hours of writing some bash script, and that's about it. So, $24 a month to build a gaming service, which is a primary server, a backup server. And I think there is no realistic way that anybody could detect it because [inaudible] could do some very good mapping of existing behavior of strings which appeared like .console, IO, etc. as examples.

Then, my sort of concerns is yes, I don't think we're really discussing gaming very much in the document at all. I think there is very little clarity on what all the TRT should do. It's very high level. There should be a TRT, and they will do the analysis, but there's very little concrete, "This is exactly the sort of things they should look at. This is kind of how they'll do it."

There's also, I don't think, nearly enough discussion on the expertise that we require from the TRT. I think there's a very small number of people who have the necessary technical background to understand all of the data. And also, for many of these, and Matt and I have done and Wes Hardaker have done some mitigations of colliding strings, a lot of it is not really technical stuff that you can find in a book.

And also, a huge amount of it is relationship-building. The console thing, the only reason Matt managed to fix it was he happened to know me, and I work and Google, and I went and poked someone. But there was no other relationship that could [inaudible]. For the set of Mac queries, the only way it was sort of hunted down is somebody knew somebody who knew somebody who worked for CnNIC who knew [inaudible] who... Anyway, whatever.

Then, there is also—I don't think we really discussed the conflict of interest concerns on the TRT. Depending on how much money you think they're gonna be sloshing around in the next round, I don't think that the people who are serving on the TRT should be allowed to be working for a current applicant or very recently have worked for a current applicant or in the very near future work for a current applicant because if you look at the amount that some of these new gTLDs end up costing, people are willing to invest millions. I would be happy to allow a .web to go forward if somebody wants to hire me next month, for example.

I think, also, the PCA and ACA, we sort of describe what they should do but not with any detail because if you want to implement PCA or ACA, this is exactly what it does. This is how it should work. It kind of looks like this.

And then, there's also some discussion about people might want to go mitigate these things. But there's, once again, no detail on how mitigations can be done. I know that Matt did send me some text, which I haven't fully reviewed yet, where he wrote up sort of some background and at least one of the mitigations he'd done. So, I think that's sort of

the what else thing. And as I say, I don't know where the rest of my list is but...

MATT THOMAS: Thanks, Warren, and Rod, over to you.

ROD RASMUSSEN: Yes. So, this gets to the question I was going to bring up earlier. So, I may as well address it now is this is a fundamental question about the documents we deliver, both the discussion group report and an eventual SSAC recommendations. A good example is that last point that Warren brought up.

The SSAC's not going to touch that, right? That is outside of our remit, conflicts of interest. We could bring it up as a potential concern. But certainly not any recommendations around how to handle that.

And, fundamentally, it goes to how much solutioneering we want to provide from the study group and the SSAC in these versus providing good information. But at the end of the day, the Board's going to have to say, "Okay. We're going to go and implement this, this, and this." And we're going to have Org in some form or fashion put this program together and probably have a public comment on how the program's put together.

So, how much of that work do we do up front as inputs? And I am looking at it as we want to provide inputs and here's some ideas around how to do this versus, "Here's how to do this," right? And we've created the process where we're saying, "This is the process that we think we

should follow." But the details of that process, I think, we don't want to get too far down the path of prescribing that because A, it's going to take a lot of time, and B, it's all going to get re-reviewed anyway.

So, if we concentrate on what are the good, some practices and points that we want to make around things to look at and consider in these topics, and basically, I think Warren did a good job of enumerating those, I think that would be where we would want to be. That's my opinion. I don't know what others think, but I think that's where we want to end up.

MATT THOMAS: Thank you all for the excellent conversation here. And Steve, going back to your original question around the gaming, there is a small section right now in the document. We definitely should expand upon that just based off of this conversation.

But at a 10,000-foot view, it basically says gaming is a hard problem and, to date, that there is no known solution or clear deterministic algorithm that could easily identify some kind of behavior to that. But we should definitely find some additional text to give a little bit more color to that topic and concern.

ROD RASMUSSEN: [inaudible]

MATT THOMAS: Yes, yes, exactly. So, well, that was an excellent list. Thank you both, Warren, Steve, and Rod, for all that commentary. I think that's a great thing that Jim and I have captured here in terms of things that we need to make sure that we've highlighted it in the document going forward. Anyone else in the group or the...? I don't have the chat open here. Yes, Jim, go ahead.

JIM GALVIN: So, I just want to be really careful. So, I'm going to call you out for the way that you just said something, things that we're going to highlight in the text. No, these are topics which we're going to consider whether or not they've already had a fulsome discussion and then see where our consensus landed. And if it hasn't, then we will figure out whether or not we should have that discussion.

So, for example, I'll give you a response to the conflict of interest thing. We're not touching that in this document. It has no bearing, no relevance at all in my opinion. That is not a technical issue. This document and work product is only addressing technical concerns. Nothing to say about conflict of interest because, actually, even how much conflict of interest applies depends on how you implement the TRT because there are a couple of ways of doing that, one of which has less COI issues, and the other one has more. Not our problem. So, anyway...

BARRY LEIBA: Am I correct that there is nothing in the document that says how you propose the TRT be formed, right? Yes. Okay.

JIM GALVIN: It's all functional requirements. The path that we're headed down is this is the functional requirements that we're looking for in a TRT and a neutral service provider. We are defining the roles and not even indicating whether they have to end up being two roles. it could be one role, one organization. It could be insourced, outsourced, whatever. All those options are on the table. We're sticking to the technical expertise.

ROD RASMUSSEN: I do have a question to that point which Warren brought up earlier. Would it be useful to specify some skill sets for the TRT?

JIM GALVIN: And that is there.

MATT THOMAS: Any other comments or questions on this particular topic for gaps or subjects for additional discussion? I'm not seeing any hands in the room or the chat room. I'm not sure how much time we really have left to get into the next slide. But why don't we bring it up anyways? Wow, that's really small font from here, and I don't think I can even read it.

I can't read it on my screen. Oh, the joys of getting older. So, this was talking to Warren's point of what are some of the actual more concrete heuristics and guidance that we might want to suggest to the technical review team, especially when we're talking about their responsibility or their function of assessing risk.

Now, I'm going to speak personally here with my NCAP chair hat off and just talk about a lot of the work that I've done looking at name collision strings based off of telemetry at A and J Root Server and doing outreach to remediate those. This is roughly kind of the heuristic set that I would work through when looking at the data at that point of view in terms of figuring out how to potentially go out and conduct some kind of remediation or outreach to remediate that string.

Clearly, my analysis based off of that if we were to map that back into the workflow would be the equivalent of looking at more of a PCA type of data but within a limited context of only two of the root server identifies. But first, one of the things that I always looked at when I was looking at a particular collision string was source diversity, right?

And I was coming from a diverse set of networks or a diverse set of ASNs. And if it was coming from a specific one, the direct outreach to that network operator was very effective most of the times. Warren's brought this up multiple times, .console, seeing that traffic at A&J and knowing that it came out of AS15169, knowing it's Google, let me directly poke him and be able to identify that and conduct some kind of outreach and remediation.

But when it starts to come from a broad set of networks and ASNs, it makes that next step of analysis a little bit more difficult. You need to start looking at the next types of properties of the collision strings. And specifically, I would typically look at the second-level domains and other types of labels in the queue names for commonalities.

Now, of course, that is more difficult now as time has progressed since we have things like Q&A minimization at the root where that telemetry is more obscured based off of that. But the next step in step two would be looking for those common SLDs. And do you see some kind of commonality?

One of the strings that I would say is when you start to see very common second-level domains for the entire thing, it sometimes often lets you very quickly identify the underlying source. So, in the instance of .tcs, it was all coming under Microsoft Windows Defender domain, which quickly allows you to identify the source, reach out to them, say, "Hey, you're clearly using something that's not intended. Could you please fix this," right?

But then, there are also other more qualitative elements to the strings that you would also clearly identify what's going on. This is in the case of other things that we've seen like with D-Link, Belkin, [BB Router,] or FRITZ!Box where those types of strings and the second-level domains in them clearly start to give you an indication that this is coming from some kind of consumer end device or some kind of a small office, home office, or router or networking device.

And so, that information starts to allow you to assess what are the potential impacts and the harms of those and what would happen if that TLD was to be allowed to be delegated, and it also gives you some kind of information in terms of how potential mitigations and outreach to fixing those can be done. I've worked with numerous ISPs who have leaked strings in their home router systems. And they've all

acknowledged that it is an issue, and it's usually embedded inside the device via some kind of service like DNS Mask.

But, unfortunately, addressing these types of problems, it's very different than something like a .console or a Microsoft Defender where a quick software or a configuration can be changed. Those types of devices require firmware updates in which then you're kind of playing into the long tail of it taking years for those devices to be updated or if they're ever updated at all, right?

And then, if you're not starting to see any kind of commonalities between either networks or a concentration in networks or some kind of commonality between the second-level domains or label itself, it starts to kind of go into more of a bespoke investigation to understand what the root cause of those are. I really don't know of a great way that I've been able to say heuristically when it's a very diverse set like, "99% of the times that this is always going to happen."

It relies on more of a little qualitative and open source searching, googling, investigating in the GitHub repos and identifying certain things that like .rancher that we saw with Kubernetes and some of that leaking out, that it's attributed to these various different things. But it's not directly understood based solely off of traffic data to be able to go into that.

And then, finally, some of the other things that we have identified when looking at some of these things, especially when you tie it to the home office routers and stuff like that, are there other types of labels that we know that there are clearly heightened levels of security risk, things like

queries for ISATAP or WPAD that are known. Is there any reason to believe that PCA would be impactful or harmful if it was deployed? And if there is any reason to believe that ACA would not succeed in disruption and notification.

This was my attempt at trying to put some real at least initial drafts down for what the TRT could look at in terms of doing an assessment. I don't think this is an exhaustive list by any means. But I wanted to use it as a starting point here for the discussion group to see if there's any additional thoughts or we think that there could be any other kind of expansion or areas that we'd like to expand on in here. Barry.

BARRY LEIBA: Yes, the bespoke part, I guess bullet four, is what makes me think. My initial thought was the TRT was something like designated experts in the IETF where we get IANA registration requests and designated experts take a look at them. And that works because we don't get a flood of those requests. We get a few at a time, maybe a dozen at a time.

For something like this, we may be talking about tens of thousands of these coming in. And that kind of analysis doesn't scale if the TRT is expected to do in-depth analysis of all of these requests or even a large percentage of them.

MATT THOMAS: So, I completely agree with that. And if I'm going to channel my inner Jeff Schmidt for a minute here on this, I think it's the framing that we're trying to identify those black swans. So, maybe the bespoke analysis

isn't done on every single one. It is there was significance in items one, two, and oh, there is no three on there, is there? Items one and two.

BARRY LEIBA: Five is right out.

MATT THOMAS: Yes, that would have motivated a reason to actually look into it. Maybe it is not required every time. Maybe this is...

BARRY LEIBA: No, sure. It's not going to be a hundred percent. The question is, is it going to be a tenth of a percent, or is it going to be twenty percent? And if it's twenty percent, that's probably not feasible. If it's a tenth of a percent, it might work.

MATT THOMAS: Warren.

WARREN KUMARI: Yes. I think sort of another difference between this and things like designated experts is from having done something like analysis of these and trying to mitigate some and things, the amount of tooling and background and access to data, etc. that you need is fairly significant. And then, the amount of work that actually goes in is also quite large.

So, for example, during the root key rollover stuff, Was Hardaker did some analysis. And it was... Actually, I don't know if he ever published,

but it was a significant number of hours. While bored, I tried hunting down one specific string, which just the string tickled my fancy. And I probably spent 60 or 70 hours trying to figure out what the source of that was. And that was with access to a set of large data that most people don't have access to.

There is also the problem of sort of the black swan analogy thing is you don't really know that there's a black swan till the first time you've see them because they're so weird and unusual events. But that means that you first need some sort of way to classify things into white swans and black swans, probably okay and not.

And I think a lot of the stuff we've been discussing is we don't know that a string is dangerous until we do the analysis. So, there is no easy way to do a bucket these into probably okay and bucket these into these may be black swans. If we had that, then we would just run that algorithm on all things, and we'd be like, "Done." The whole problem is. We don't know if a string is dangerous. If we could, we wouldn't need the TRT or anything else.

MATT THOMAS: On that last point, Warren, I completely agree with you. And I think that comes back to one of the tenets of why the workflow is designed in its way. And that was a risk management step function in which you're increasing the risk of exposed to ICANN Board in a substantial manner while still getting new additional data. So, it's a balancing act between the two, right?

So, just being cognizant of time, it's at 2:30 right now. I think we're at the end of the meeting. I'll do a quick call for AOB. No hands, nothing. Well, it was a great discussion group. It was great to see everyone back here. And tomorrow, we have the open plenary on where we're at with study two. If you have time, please stop by and attend. Thanks.

[END OF TRANSCRIPTION]