SUE SCHULER: Great. Thanks. Okay, Jim.

JIM GALVIN: Thanks, Sue. Welcome everyone. This is Jim Galvin, now claiming to be with Donuts. I'm sure everyone has heard about the lovely change of ownership of Afilias over the holidays, December 30th. So, good times. Good times. And my partner in crime, I don't know, Alan, are we criminals, Alan? Could be, I suppose.

Yeah. Thank you, Donna. We're not criminals yet, as Alan says. Very true. Anyway, thanks everyone. Happy New Year. Welcome to 2021 and our first meeting of our Registry DNS Abuse Working Group. Apologies from my co-chair, Brian. He's not able to be present today but I think that we will press on here nonetheless.

Samaneh is with us here. Well, we wanted to have three things but we ended up with just two things on our agenda here today. We're going to take this opportunity to talk to Samaneh and bring us all back up to date here with where we left off in our discussions with OCTO. Those who were part of the DAAR Working Group and participated in that, this is just a good opportunity to make sure that we don't miss anything.

These are the three topic areas that we had been carrying forward from DAAR when we were working with OCTO back in the early fall. And so we took this opportunity to invite John and Samaneh and she's joined us to come and say a little bit about these things.

I know that we have quite a number of more people in this group than we had in the DAAR group. So, we explicitly decided we'll spend a few minutes talking, just introducing these topics and then Samaneh will add what she wants to.

And then this is a little bit of a level set for people to ask questions. We're looking for opportunities for us to think about what it would mean to work together with OCTO to do some things in these topic areas or not. This is our opportunity to come up to speed here, understand from Samaneh's point of view, where she is on some of these issues.

And then we'll have some time afterwards for our own discussions if we want to say any more about them. And then very quickly, Brian, and myself, and Keith, and Graeme, just to expose it to people, we had created a draft invitation to use for other SOs and ACs.

As we talked about before, we want to engage in some outreach jointly with the registrars. Those meetings will happen on the Tuesday timeslot. We have a drafting invitation. We'll just show that to you. We'll focus on the questions that we are going to ask of those SOs and ACs as part of that meeting when we send it to them and make sure that people are on board with that, if you have any questions or comments about it. And otherwise we'll take—we'll just press that process along.

Reminder again that Keith Drazek had graciously volunteered to help facilitate and organize that so we'll just turn it over to him after that and then we'll start to organize some meetings with the

SOs and ACs. So, we'll show this off here in a bit when we get there.

So, I think with that, I'll just jump right in here to the first bullet item. Maybe I'll pause momentarily if anybody wants to add anything quickly. Not seeing any hands. Okay. Let's jump right into the discussions with Samaneh.

We have three bullet items there. The first item is this question of persistence of abuse at a domain name. And I know, Samaneh's already told me that she really doesn't have an update for us on this topic but I wanted to take a few minutes to just talk about the problem space again, just to bring on board other folks here who have not been tracking that discussion and are familiar with what we talked about before.

This is an interesting problem space. If you look at DAAR, what DAAR is basically giving you is a count or a percentage, in absolute or relative terms, of allegations of abuse in the TLD. And there's a couple of different slices across that presentation in the data.

So, from a face value point of view, it's certainly an accurate representation of what's there. The concern that we were having back then in the DAAR group and that we'll continue forward down this discussion is it's not really a complete picture of what's going on in the background.

And those numbers by themselves leave out information that would be helpful to people who really want to use it to assess DNS abuse. In particular, the characteristic that it doesn't show is

the fact that there's turnover, there's churn in the names that are actually counted.

And so, the question that we were exploring is, how do we show that and what are the ways in which we can show that? And is there data available that can be used to speak to the issue of the names that are alleged to have—show abuse one month, are not necessarily the same names the following month.

So, although the number of names that are alleging abuse might not change that much, okay, the actual names under me could change quite dramatically. And then in addition to that part of the problem space, the other piece of it is the fact that the DAAR really shows you a point in time. Maybe Samaneh, you can speak a little bit to this. We had talked about going to different kinds of averages in the values and so maybe you can update us here a bit about where that is, if you wouldn't mind when you jump in here.

But those are the two problem spaces about this persistence of abuse. Trying to add more information to what DAAR is showing for the community at large so that they can better assess what abuse is. I'm going to give Samaneh a chance to add to that if she wants in the explanation. And then just open the floor for questions from anyone if you don't fully understand what I said or for folks who were part of DAAR, if you want to add something to that, there were a number of people here in this group who were there during all these discussions, during the DAAR discussions.

And if you want to add some information to that to help explain it, that would be helpful too. I'm sure that everyone—this is a space

that we want to get into. OCTO has welcomed the opportunity to try to find a way to do something about this question.

And we just haven't been able to figure out exactly what we can do, nor really committed to what is best thing to do. So, Samaneh, if you want, let me give you a chance to speak and then anyone else who wants to put their hand up, that'd be great. Go ahead, please.

SAMANEH TAJALIZADEHKHOOB: Thank you, Jim. First of all, I'm happy to be within this newly formed group, a bigger group. For those who do not know me, I'm Samaneh. It's been two years that I'm the project lead on the DAAR system. I work within OCTO, within the security and stability and resiliency team with John Crane.

I have background in academic research on abuse and on mathematical models to basically observe security incidents and predict them. Within the work that we did with the DAAR Working Group, as Jim already explained, we had open space to go through all the points that the group members raised as points that they saw incomplete or could be improved.

And just a slight update on that is that, if all goes well—and they published a recommendation document based on that. And we are planning to publish the first draft based on the recommendations that were made for this month, so for January 2021, and send that draft to the group to do checks and further improve it before doing the final publication.

Now, this was a bit general. Is there any questions so far before—because I seem to talk fast? Okay. So, Jim mentioned two things. One of them was regarding persistence metric. I'll leave that for second part. The second thing was the fact that DAAR at the moment shows basically what DAAR has as the indication of abuse is the number of times that a domain is listed based on the blacklist, reputation list that we use—that DAAR uses, actually.

At the moment, the reports are published based on the last day of the month. So, just one day on the last day of the month, the counts are based on that. Based on the recommendations we have received, we concluded that it's a fair point to have an average over months.

We did analysis and we concluded that median is better [inaudible] than mean so we changed that in the new report. Hopefully, you will see the draft soon. Now, there is another point where the metrics that explain abuse can be improved and that is not only to see how many times there are or how many domains within a space, let's say a TLD space, are listed as abuse security threat domains but also how long these domains stay listed.

Together with the group members, we spent one session. I worked on it before the session and we spent one session discussing how would such a metric look like, let's call it persistence for now. Part of the discussion was that, what is persistence?

Can we call persistence the amount of time that the domain is listed? The problem with that kind of metric is that then that timing is really dependent on the methodology with which the reputation

provider lists domains which is individual—which then would be depending on individual providers.

We looked at a bit how the space looked like. And the next step was to further see how we can improve such metrics. As you noted, before, I didn't do much work on that due to being—focusing on finishing the new draft of DAAR based off recommendations that were already made.

But we spent together with our group, we spent some time and energy on exploring optional ways where we can actually improve accuracy of the timing where a domain is up, let's call it uptime, for the sake of it. One of the ways—I'm not going to go through it but just to mention it and later, hopefully, we will have more time in future sessions, I can bring some material to discuss with you and then you will be able to see it more in action, but just as an idea—was that instead of calculating an uptime metric based on the amount of time or days where one domain is listed, one could see the amount of time that one domain is up in general or it has not been taken down.

We could use passive DNS data to achieve that and we have access to passive DNS data at the moment. So, we were busy last year to gain that access for this reason. Now, that also comes with its own up and downsides which I love to discuss with you and hear your ideas but let's leave that for the next session because I also know that timing is limited for me this time. That was about the first point. I'm happy to take questions if there are any.

JIM GALVIN: Thank you, Samaneh. And you reminded me as you were talking, I apologize, I said average amount of time that a domain is represented in the reputation list, and you're right. You had talked to us before about using median instead of average and I do remember that discussion now. I'd let that slip my mind here.

But I think, again, the comment that I would make to folks is, we know that this is a difficult and challenging space. And I just, I welcome the fact that Samaneh is willing to talk to us about this so that we can continue this discussion and see if we can find a way that we're all comfortable with—to move this forward. And Kurt, you have your hand up, please go ahead.

KURT PRITZ: Yeah. First, I'm going to start with an apology because I might have missed part of what was been said so this might be repetitive. But I want to reiterate a discussion we had earlier that it's not just a median or mean of persistence, the length of time an abusive domain stays registered but rather it's the shape of that population.

So, it's because our role as contracted parties is more about mitigating rather than preventing. It's important for us to identify how quickly abusive domains are taken down. And so, if the median or mean of an abusive domain is, say, eight days, some people might think, well, that's way too long because the abusive behavior occurs in the first hours or days.

But if we looked at that population of data and we saw that 60% of the domains are taken down in one day and the other 40% are still

up after 20 days, that would be a completely different problem. It would be maybe those outliers are really not abusive domains or it might be that the reputation providers choose not to delete the names off their list even though they're taken down because the reputation providers, as others on this list have pointed out, have different business goals than us. So, we might've already talked about it and I missed it, but the shape of that distribution is just as important as the mean number.

JIM GALVIN: No. Thank you, Kurt. Very good points. And because I know that you were really the principal who was pressing that when we were having the discussion last time. We had not mentioned that here, reminded folks that it's more—there are details.

I mean the shape of the population is important for all reasons that you just mentioned and we had talked through a lot of that before. And we're going to have to get back into that as we consider this question of, what is persistence? What is mean? How do we represent it? What are we trying to show the community?

And in some sense, we are trying to show that we were looking for a mechanism for showing that abuse is actually being handled. It is being mitigated. That's not to say that—I mean, we all know that abuse exists. There are still some problem areas, but you're not getting any kind of representation of the fact that there's a lot going on and a lot does happen.

And we're looking for a way to be able to show that. So hopefully this gives people a sense that there really is some interesting

discussion to be had here about what this looks like. And then as Kurt just reminded us too, a little bit it's about the quality of the data.

I mean, the reputation providers have their own methods for what they do and don't include and when they include it and when they remove it. And that affects the quality of what we can do here. So those are all issues too that have to be re-explored and dug into.

So, there aren't any other questions at the moment and Kurt, I'm going to assume that's an old hand. We'll wait. As Samaneh has said, she's certainly interested and willing to continue this discussion with us. They are interested in trying to do more, to do better if you will, for the community, so—and she has some stuff that she wants to put together for us.

So, we'll have another opportunity to have her come and talk in more detail about this particular topic and then we'll get more into some of these discussions and details. So, and I see Crystal's comment and first, I'll go back to Donna, the quality and purpose of the data. Yeah, the RBLs are collecting, yes, I was making that comment and you're right, Donna.

And Crystal, has anyone done a holistic review of the major sources? I mean, I'll give part of an answer to that. I don't really want to put words in Samaneh's mouth but it's my understanding even with DAAR, they did a pretty good examination themselves about where to get data from and what to use.

So, I don't know how far reaching all of that discussion was. I don't know if Samaneh wants to speak to that but they certainly did do

some investigation and they made some choices. And there is the document that describes their methodology and what they use and how they do it. I see that Samaneh came off mute so I'll give her a chance to respond a bit to that. Go ahead.

SAMANEH TAJALIZADEHKHOOB: Thank you, Jim. Yeah, I've mentioned some in the chat. First of all, I mean, this was done before I joined the group by the initial group that started with DAAR, so I'm talking in place of them. But the criteria was that the fields are well known in the academic—well referenced in the other academic research because there has been some research that reviewed fields and their differences and their level of accuracy and false positives, and that they are listed by industry players for white or blacklisting.

So, basically, they [evaluated] feeds based on these two criterias. This was the initial criteria. Later—and we are still not done with that but we are finishing it almost, that we are publishing the document in which we looked at different feeds, not only the ones that we're using for DAAR but several ones that are out there in terms of accuracy and several metrics that we defined how comprehensive they are.

Of course it's, yeah, from different perspectives and, yeah, you will have more information on that. But there has also been some examples by different, more neutral parties that I just pasted an example here. This one looked semi-reliable but of course their methodology is a bit vague. That's why we decided to publish our own document.

[CRYSTAL ONDO]: Great. Thanks, Samaneh. Do you have any timing on when ICANN will be publishing that information?

SAMANEH TAJALIZADEHKHOOB: Unfortunately, I don't. It was on my plate by the end of last year. I started it and it's sort of in the finishing stage but the work now, because I'm super overloaded, it's left to another colleague. John is still traveling so we have not yet started planning for this year. But hopefully in the future sessions, I can give you more precise timing on that.

[CRYSTAL ONDO]: Great. Sounds good. And I'm just curious because I know that VirusTotal is one of the things that's often cited by DAAR and it's just an aggregator of 80 different various sources. So, within each of the VirusTotal flags, the question is, of those 80 sources, which ones are more reliable. So, I guess that's where I'm coming from and it'd be interesting to see if a third-party had reviewed any of that.

SAMANEH TAJALIZADEHKHOOB: No. Definitely. I see your point. Yeah. I think the document will be interesting for you to see. At least I think, let's see how you think about it.

[CRYSTAL ONDO]: Thank you.

JIM GALVIN: Great. Thank you. Thank you, folks for jumping in. Good questions. And thanks, Samaneh, for being responsive. So, we'll have some new stuff to look at and think about when we come back and talk about this. So, moving on to the next sub item here, I know that DAAR has been doing some work with respect to ccTLDs.

I mean, as we know, the DAAR was originally launched really just about gTLDs and they didn't have any other TLDs in there. So, Samaneh is going to tell us a little bit about what they've been doing and what their plans are with respect to ccTLDs and DAAR, and also about the DAAR TLD data being in MoSAPI so that registries can actually access their own data directly and you can look at what DAAR has about you. But I'll let her speak to what's going on there and if she has any questions for us, she can put those out there too. Samaneh?

SAMANEH TAJALIZADEHKHOOB: Yes. So, I think it was November that we—or maybe it was December. Anyways, either November or December, we started sending individualized report to each ccTLDs. These reports are similar to the ones that are published on the ICANN DAAR website which are anonymous for gTLDs.

But then with a difference that in each report that we send to each ccTLD, we only highlight them in the report, individuals and in the stats numbers so that they can see where they are in that space and that basically the idea was to make the data more useful. So

far, the feedback we've received was really good. I think most of the ccTLDs really like the report.

This was the initial draft, so this is not something finalized. We just had this idea, made it, send it to see what they think and ask them for feedback till the end of this month. So, we will see actually detailed feedback so we will see what are the areas to improve.

But the original idea was to make this individual drafts for the whole space, not only ccTLDs but also for gTLDs, individual gTLDs so that each TLD can only see themselves, not others but can at least see where they stand in comparison to others.

The problem we have at the moment is more technical and logistics with gTLDs. We have to figure out, first of all, one step in front of us is to publish the new DAAR report based on the feedback we have received from this working group. So, we want to have that first, have that confirmed and then go to the next step which is publishing individualized report.

But another logistic problem we have is that at the moment, the report is of course generated automatically per individual report, per individual ccTLDs but we still eyeball all the reports before sending them one by one. I do that. And when it comes to 1,200 gTLDs, it's not—yeah, we don't have the logistics in place to do it yet and that is not something I want to spend time on also.

So, within the group, we are figuring out how would the optimal way of doing it be when we have the content, when we finalize it with you and with them, then soon this would be also for the rest

of the TLDs. About MoSAPI, so at the moment—and ever since DAAR…

JIM GALVIN: Hi, Samaneh? I'm sorry. If I can just jump in because we have a couple of questions here. Well, I have one and there's one in the chatroom there from Donna on this particular issue. I was just going to point out to people. Donna is asking the question in the chatroom about, is the ccTLD information public or intended to be?

I only became aware myself personally of the ccTLD reports because as a service provider, we had some of our ccTLD customers come to us and say, "Hey, look at this great report here. What's going on?" So, we had it given to us in that way. I don't know that it's published in DAAR. I don't.

And I guess that's the question that I have that Donna has too. Will you be adding them to the DAAR list? And then in the tail end of what you were saying there, Samaneh, you were talking about being able to give any gTLD access to a similar kind of report for themselves and you were talking about the logistics of that. But yeah, would the ccTLD stuff be in the DAAR reports on the website, and what's that going to look like?

SAMANEH TAJALIZADEHKHOOB: Sorry to interrupt. That's a very good question, Jim. So, let me go back one step before being able to fully answer that. You remember that when ccTLD started to join, we had a discussion with you guys that we promised we won't publish any one-to-one

comparison of TLDs because it's not fair towards gTLDs and that we have to, together with you guys and the rest of the community, we have to—if we were going to publish such a report, we will have to figure out what would be the format of it, right? I don't know if you remember this discussion. Along with that discussion, one solution that we came up so far for not doing that unfair analysis and not publishing it was to just send individualized reports to each TLDs to avoid that public confusion, comparison, unfair thing.

But I think ultimately, we would—and I am not 100% sure about it because we focus mostly on finishing, sending personalized reports first, than to publish something about ccTLDs so far. But I think for now, we will leave the reports as they are for gTLDs and we'll focus to enhance this mechanism, this individualized report also for ccTLDs.

Now, if you and the rest of the community still wants also anonymous reports for ccTLDs, the same as they are at the moment for gTLDs, then obviously that's the easiest thing we can do. We can discuss it and see what would come out of it.

And I also want to still emphasize that the reports we are sending are not finalized. They are also only sent for feedback and comments, etc. And I would also love to hear from you guys, what do you think about receiving individualized reports versus public reports, etc.?

They are all open for discussion because we are going to still process that, review it, and then do whatever needs to be done. It's all about what you guys find more useful.

JIM GALVIN: So, let me just very clearly put those questions to the group here if anyone wants to comment it now, or we'll take these questions on board from you, Samaneh. The two questions—the first one is whether or not we would like to see ccTLDs included in the DAAR reporting, that's public information in the same way the gTLD is reported.

And the second question is, would gTLDs be interested in getting individualized reports, more detailed reports about their TLDs? So, would anyone like to speak to that or ask anything additional about that at this time? And I'm not seeing any hands but I've noted those two questions, Samaneh. Donna, please go ahead.

DONNA AUSTIN: Thanks, Jim. Donna Austin. Thanks, Samaneh. So, I just want to clarify one part of the question about making the ccTLD information public within the DAAR reports. I think what you're saying is it would be aggregated within the gTLD records as well. Or would you separate the cc out from the gTLD information?

SAMANEH TAJALIZADEHKHOOB: Yeah. I mean, the only way we could think of was to separate it. And there's also another problem with doing it now because the ccTLD space is still very small. I think at the moment, we have 10 or so, maybe 11. And putting them next to each other in a comparison, at least scientifically and statistically, it doesn't make any sense. That's why we thought even if the community

wants this, let's wait to at least get some representable population so that we can meaningfully show something.

DONNA AUSTIN: Okay. Thanks, Samaneh.

SAMANEH TAJALIZADEHKHOOB: Thank you.

JIM GALVIN: Thanks, Donna. And to the comment in the chatroom there, I'll just jump out. Kurt had asked about suggesting about formalizing the question and discussing it at an RySG meeting. Yeah, at a minimum, Kurt, I think first, I do want to take the question on board for us and we'll talk about it more within this group before we would bring it to the larger Registry Stakeholder Group meeting.

Let's make sure that we have a consensus here in our group and an understanding of what we're proposing. That would be my response to that, Kurt. Martin, you have your hand up. Go ahead please.

MARTIN SUTTON: Thanks, Jim. And thanks, Samaneh. This has been really useful. I was just going to suggest that if there are considerations towards combining information with the ccTLDs and gTLDs and actually also the individual reports, it might be useful just to make it a little bit easier to visualize, to issue a sample report for the group to have a look at, anonymized so that even if it's an individualized

report, it's just trying to give a flavor of how to represent that information and perhaps then it will draw out how useful it would be for individual registries to ask for that information. Thanks.

SAMANEH TAJALIZADEHKHOOB: Sure, Martin. Definitely. If you guys want for future sessions, I can bring up an example of individualized report which is anonymous so at least you can see how it looks. Like it's not much different from the current report, except that we already applied some of the recommendations by the DAAR Working Group so it's monthly averages and different things. But yeah, why not? I can bring that.

JIM GALVIN: That would be great, Samaneh. Appreciate that. So, maybe when we get a chance to meet again, we'll continue our discussions about persistence in this. So, good question, Martin. Thank you. Be good to be able to see it. Actually, if you have a sample that you might be willing to share, you could send that along. We'll bring it, put it in front of the group here so that folks can look at it and then we'll be better able to say more to you and have a chat about it ourselves before the next time that we meet, when you get a chance. That would be very helpful.

SAMANEH TAJALIZADEHKHOOB: Yeah, I think so too. Sounds good. I will do that.

JIM GALVIN: Thank you. Okay. So, I apologize. I interrupted you before you were about to start talking about TLD data in MoSAPI. So, given that we seem to have gotten past all the hands that were up, let me just turn it back over to you to pick up from that point.

SAMANEH TAJALIZADEHKHOOB: Okay. Thank you, Jim. Yeah. So, basically, almost from when DAAR started publishing reports, monthly PDFs, at the same time in theory, gTLD should have been able to see their daily stats, so their daily scores per security threat in MoSAPI.

Now, I used to have some stats of how many TLDs on average access that information per day from our internal resources. But in theory, they all should be able to access it because it's the same system that they can access other things.

Now, we created the same mechanism for ccTLDs also. I want to emphasize that this is separate from the monthly reports so it's not PDF. They just see basically a line which shows how many phishing they had, how many spam and what is the overall score in comparison to their size.

And each TLD only sees their own data regardless of being cc or g. Just that gTLDs, originally, all should have access to this system because ICANN use it for other purposes. ccTLDs didn't have or some have but most don't.

So, as part of onboarding process, when they volunteer to provide their zones, they also set up an access to the system. Still, they only see their own data but it's the same data as what gTLDs see.

I'm not sure—did I answer any questions that was regarding MoSAPI? I wasn't sure what was the original concern about this.

JIM GALVIN: I'm not aware that everyone knows that they can go get at their data there. And so, I just wanted to make sure to call that out to people. And yeah, and so just knowing that it's there, folks might want to check with your own internal teams on that point. Is there a documentation, Samaneh, about how to get it, what's there, and what it looks like?

SAMANEH TAJALIZADEHKHOOB: Yeah. There are several documentations about that but it's actually great that I know that not everybody knows, even though it's been announced several times but, yeah, we can do it more. Why not? We can publish another announcement with the information, how they can access it and what does it look like? [inaudible].

JIM GALVIN: Okay. Thank you. Any questions from anyone about that? Any concerns? Donna, you have your hand, go ahead, please.

DONNA AUSTIN: Yeah. Thanks, Jim. Not a concern but just a question. Is the data in MoSAPI, would that be different from an individual requesting the information directly from OCTO? Not sure I'm understanding.

SAMANEH TAJALIZADEHKHOOB: Very good question, Donna. It will be slightly different, not in a sense that the numbers will be different. You would still—the number, for instance, gTLD A sees in the MoSAPI is still the same number that the gTLD sees in the monthly report except it's daily and that's monthly.

But the benefit of the individualized report in addition to the data in MoSAPI is that, there the TLD can also see herself in comparison to the rest. To the rest, that is anonymous. But still, it could add some visualization of where do I stand in terms of A and B.

And to just let you know, in the future, our ideal scenario is that we have a platform, we have a dynamic platform that actually TLDs can just go and select metrics and can see themselves in comparison to the rest. But before having that developed, this is the best we could do.

DONNA AUSTIN: Okay. Thanks, Samaneh. That's a helpful clarification.

SAMANEH TAJALIZADEHKHOOB: Thank you.

JIM GALVIN: Okay. Thank you for the questions. Thank you, Samaneh. Any other questions on this particular point? Okay. So, let's move on to the third point [in the dossier] about including registrars in DAAR. It's always been a background thing, something hanging out there

in the cloud so to speak about expanding DAAR to report on registrars in the same way that it's reporting on TLDs.

And Samaneh, there is a particular technical issue which makes that problematic. But let me let Samaneh speak to this for a bit and then we can come back around and add to it.

SAMANEH TAJALIZADEHKHOOB: Yeah. So, I don't know, maybe some of you already know that it's been a while that we are trying to tackle that problem either within the DAAR system which is a separate system, but we are seriously looking into this. Now we have a working group on this internal working group to just point out some of the struggles we have that I also used to point out at the end of my DAAR presentations, is that at the moment, DAAR provides daily stats on abuse, so daily concentration metrics for abuse per TLD, which is not dependent on WHOIS. So, we get data from zone files and data from reputation lists. Once we want to do that for registrars, we need the associated registrar ID which WHOIS is one source of that.

Of course, we are all aware of the problems of the WHOIS. If we want to do it in a way that is replicable by anybody in the community, we would need to go through the public way so basically calling WHOIS. And then, we either would overload the WHOIS server or we would get rate limited, which then the consequence of which is that we cannot have daily reports similar to the ones that DAAR has now.

Of course, we could go further. We could develop monthly numbers or like even quarterly numbers. We've considered the trade-offs but the ideal scenario would be that we could have a similar granularity that we present for DAAR.

Now, I know—and some of you might know—that there are reports by some other firms or industry players, that they did publish snapshot analysis of the registrar space which is something we could do as well. But the step that we are foreseeing is more grand than that so we are working on it.

At the moment, we are discussing internally whether we should, for the moment, leave the replicability issue aside and use our own internal data to create metrics for registrars. We still don't have legal access to such thing which is another project we need to start, if even internally it's approved.

Other things that were discussed was to have a mechanism that is beyond us as ICANN, just as a community to have the mechanism in which registrar IDs no longer only dependent on WHOIS or any other products that are upcoming that is related to that but can be found elsewhere.

For instance, maybe Jim has already aware of that discussion that whether the registrar ID could be in the zone file or some other alternative way that one without accessing PII can get this information without going through all the limitations of WHOIS and [RDAP,] etc.

To make a long story that I already explained short, is that, we are intensely working on this problem. If the community wants from

us, we can publish a snapshot analysis like the ones that are already out there. That's no problem.

But we are trying to find a more longer-term solution for this problem and we are getting closer than before, given that we have to find all kinds of sources from all directions. And one of the feedbacks that I would really love to hear from you is that, if you have alternative ways in which these metrics can be created on daily basis but not going through the same difficulties that I just explained to you, etc.

I would say that that needs another session but since the discussion is open now, if there are thoughts or—I don't know how Jim would approach this. I'll leave it up to you guys.

JIM GALVIN: Yeah. So, thank you, Samaneh for the summary and the overview. Just to focus the discussion a bit, yeah, the specific technical problem that I know that the folks over at OCTO have been doing this is, in order to do it something closer to near real time or daily, they need to be able to correlate the registrar of record, the registrar ID with the domain name, so where is that registration, in order to create all the data and really do that. And that's the problem they're trying to solve.

Samaneh made reference to one potential solution and I want to call it out here. I actually see it as two possibilities but we can split it into two different ideas. And this is a place where registries could participate if we were interested.

It would be something that we would have to do or we would have to make sure happens. She talked about putting the registrar ID in the zone file and for me, that really means one of two things actually. It could mean actually putting it in the DNS, right?

I mean, as registries, obviously we all have a DNS infrastructure, we could automatically add an additional record for a particular domain name in our zone files that indicates the registrar ID. Of course, that really only works for names that are actually delegated. But then again, if the name is not delegated, then there probably isn't much abuse going on anyway so maybe that's okay.

The other possibility rather than putting it directly in the zone file which of course would affect all of our traffic in the DNS infrastructure and that may or may not matter to some people, is putting the registrar ID in the CZDS file.

We're all, as gTLDs, obligated to produce the CZDS files. One possibility is to add that information as a comment directly into that file and then ICANN could simply subscribe to all of those files and get those. And of course, anyone in the community could get them too and they would have that data.

That, of course, would still be a change for us because you now have to produce a slightly different file, but it would also be a way of doing that which would be an offline mechanism for providing that information and making it available.

So, and Samaneh talked about a number of other solutions they're looking at. The zone file solution is something which really would fall to us if we wanted to participate in that and it's something

worth talking about and like I said, there's those two possibilities. Anyone have any questions or comments about that?

Any additional comments for Samaneh or questions? Otherwise, I think we'll take on board the question for ourselves here about what we think about this particular topic and whether or not and how we might proceed in the space. I'm not seeing any hands go up at the moment about that. So, I guess with that, Samaneh, any other closing comments or questions for us on these topics?

SAMANEH TAJALIZADEHKHOOB: No. Just that the point you made was also a good point that we haven't discussed. Yeah, let's just discuss it further after you pass it on board and we can [convene another session about it.] Other than that, I'm happy again to be to be involved in this discussion with you guys. Looking forward for more feedbacks, and with that, I would like to leave the session so that you can continue with the rest of your discussion, if that's fine with you.

JIM GALVIN: Yeah. Thank you, Samaneh. Appreciate your time and we'll look to invite you again. Let me know when you're ready of course to come join us again and we've got multiple activities going on here, we'll find an opportunity to slot you back in to join us again. Very much appreciated. Thank you.

SAMANEH TAJALIZADEHKHOOB: Perfect. Thank you, Jim. I'll be in touch. And thank you, everybody. Have a nice day. Bye-bye.

JIM GALVIN: Okay. So, being a little conscious of time here, we've only got about nine minutes left, but let me ask Sue if she wouldn't mind moving to that document and putting it up on the screen. And here, I'm going to put the link into the chatroom here for us. Folks can also go to it directly.

There shouldn't be anything too exciting here in this document, but I did want to call out the four questions that you see there in the middle. What we had done is put together those as the starting point to help guide the discussion with the SOs and ACs that we would invite.

And we'll turn this over to Keith to take this message and draft this up for all the SOs and ACs and go through the process of bringing all that on board and putting it together and go forward. So, I'm going to give people a chance. Maybe I'll just read very quickly, those four questions.

What are your pain points regarding DNS abuse, right? Are you seeing practices from registrars or registries that you find helpful? Trying to stay more on a positive note here, then asking them what they don't like. Are there practices not broadly implemented that you would like to see?

So, this sort of a follow-up to the pain point question. And then of course asking them quite directly, how is it that you're assessing DNS abuse? We do hear a lot of discussion in the community at large about the sky is falling. DNS abuse is everywhere. Urgent, it's got to be dealt with.

This felt like an interesting way to get at the issue of, how is it that you're deciding that abuse levels are as terrible as seemed to be reported? Since of course we have plenty of evidence to suggest the opposite and that being true so we just opened that door.

I wanted to give folks a chance just to look at that and think about it while you're continuing to read that or look at the document online. Keith, I think you're still with us there. Do you want to say anything? Now, please, go ahead.

KEITH DRAZEK: Yeah. Thanks, Jim. Hi everybody. I think you covered it real well and I think the little bit of work that we put into this ahead of time coordinating also with Graeme. I think the questions are pretty straightforward and I think the key here is that we want to have a meaningful, constructive engagement.

And I think that last point—the last question is really important as we identify the various sources of data that people are using. And the better we understand where their data is coming from, I think the better off we'll be.

So, I look forward to—once we get this thing finalized, I look forward to starting the outreach and starting to schedule opportunities for us to have these constructive discussions with the various SOs and ACs and presumably some of the other GNSO, SGs and Cs as well. So, happy to take any questions or guidance but I think you've covered it. Thanks.

JIM GALVIN: Thanks, Keith. Donna, you have your hand up. Go ahead, please.

DONNA AUSTIN: Thanks, Jim and thanks, Keith. So, just an initial reaction is that I'm concerned about questions two and three. We've expressed concerns before about being held up to the best practices being conducted by ccTLDs and I think that's where this will [hint].

So, I am concerned about asking those open questions and then starting an expectation that whatever the groups decide are helpful that all of the registries and registrars will in some way be obligated to follow those. So, I'm not sure we want to lead with those open questions but I think questions one and four are really good. So, that's just my initial reaction. Thanks.

JIM GALVIN: Thanks, Donna. Excellent point. Let me try to channel Brian a little bit who tends to jump on that kind of topic when we talk about it. I think that he would probably say and emphasize the fact that we are talking about DNS abuse, and that's why the first paragraph opens with this, in the categories that we have defined.

And so, we want to try to stay in that context. You're right, Donna, that they're likely to go off and try to start comparing us to ccTLDs especially and some of the things that they do. And in the general case, I think that we can easily say that that doesn't apply because it doesn't really fall into these categories here.

On the other hand, the observation I make is, even on the gTLD side, we have some gTLDs with some pretty strict registration

policies. And as a result of that, they do employ in those particular cases a lot of the stuff that's done by ccTLDs.

So, it seems hard to avoid having these questions there. The topics are likely to get there anyway and we figured that if we at least had the questions here, we could at least directly drive the questions, guide the discussion. We can steel ourselves for whatever's likely to come in advance and know that it's going to be there and be prepared to talk about it.

And, yeah, Donna says, “So yeah, I'd prefer to get there anyway, rather than ask them upfront. Let us get there instead of upfront.” That's an interesting—let me just reflect on that for a minute. Kurt, you have your hand up. Go ahead please.

KURT PRITZ: So, it's in line with what Donna was suggesting. Maybe targeting on the information they use or the data they use or some information compilation first but include in the letter the idea that our goals are to establish practices for registries and registrars that would be helpful.

So, let's target the consultation on information and data gathering that's meaningful for measuring DNS abuse so we can determine which practices would be helpful rather than being confronted with practices that the other SOs and ACs would want to see us implement right away.

So, we could identify these end goals in the letter but say, this is our step-by-step approach for doing that. We want to establish this long-term—not long-term but we want to establish this working

relationship with SOs and ACs that build information and then arrive at conclusions. Thanks.

JIM GALVIN: Thanks, Kurt. Being a little conscious of time here, let me just jump to Sam. Go ahead, please.

SAM DEMETRIOU: Thanks guys. I don't mean to throw cold water on this discussion but I actually think that creating too many expectations about what the outcomes of these conversations are going to be in this initial outreach e-mail is what we do not want to be doing, right?

I think I like the way this is drafted in the sense that it just opens the door for discussions without jumping ahead to what the final outcome of this is going to be. I just don't want us to be in a position where we're overpromising things right up front.

So, I guess I would maybe push back a little bit on what Kurt just suggested. In terms of the questions that are presented in here, I think we can push up the question of what information do you use and such. But then again, I don't see as much of a problem with asking about other practices as long as we're very clear when we have these conversations.

And I think it's pretty easy for us to be clear in this way that we're just getting the information and we're not trying to replicate it for every single gTLD. Basically, what I'm counseling right now is expectation management for these conversations. But otherwise, I

think the outreach draft here as I've just skimmed, it looks really good.

JIM GALVIN: Thanks, Sam. And let me call out to the record Keith's comment in the chatroom which I'm actually thinking at the moment I kind of like. As you were saying, Sam, we really do want to be careful about expectation management.

As others have said here too, we don't want to promise anything. This really is information gathering and an open discussion on that information gathering. And Keith's suggestion in the chatroom is maybe we keep question two because again, that's just gathering information.

We already know what we're doing, let's see if they know what we're doing and give them a chance to tell us a little bit about it and what's useful to them and then drop question three. So, rather than looking for new things that they might somehow get the implication that we're promising to do those kinds of things, let's just stick to what we know is there and see what they know about it and not ask for open-ended non things.

And I think I'm liking that particular proposal at the moment. Hope my rationale made sense to people. And Donna is saying in the chat that "That might work, Keith, need more time to consider." So, we'll take these questions here, send them out to our mailing list here.

I definitely want to sync up with Brian and Graeme on these questions before we send this out, so we'll do that once and then

we'll say more on the mailing list here about what we're going to do and how to go forward.

So, with that, let me do a real quick Any Other Business because we're now after the hour here. I want to be respectful. And Keith, pretty sure that's an old hand. So, thanks very much everyone. We will meet again next week and pick up. And please watch on the mailing list for another proposal for what to do with this invitation message. And with that, we're adjourned.

SUE SCHULER: Thanks, Jim. Julie, we can have the recording.

[END OF TRANSCRIPTION]