
ICANN Transcription

Registration Data Accuracy Scoping Team

Thursday, 10 February 2022 at 14:00 UTC

Note: Although the transcription is largely accurate, in some cases it is incomplete or inaccurate due to inaudible passages or transcription errors. It is posted as an aid to understanding the proceedings at the meeting, but should not be treated as an authoritative record. Attendance and recordings of the call are posted on agenda wiki page: <https://community.icann.org/x/YD3Cg>

The recordings and transcriptions of the calls are posted on the GNSO Master Calendar page
<http://gnso.icann.org/en/group-activities/calendar>

TERRI AGNEW:

Good morning, good afternoon, good evening and welcome to the Registration Data Accuracy Scoping Team taking place on Thursday the 10th of February 2022 at 14:00 UTC.

In the interest of time, there'll be no roll call, attendance will be taken by the Zoom room. If you're only on the telephone, could you please identify yourselves now?

Hearing no one, we do have listed apologies from Lori Schulman, Marika Konings and Roger Carney. Roger Carney has put forward the alternative Owen Smigelski. Statements of interest must be kept up to date. If anyone has any updates to share, please raise your hand or speak up now. Seeing or hearing no one, if you do need assistance, please email the GNSO Secretariat.

Note: The following is the output resulting from transcribing an audio file into a word/text document. Although the transcription is largely accurate, in some cases may be incomplete or inaccurate due to inaudible passages and grammatical corrections. It is posted as an aid to the original audio file, but should not be treated as an authoritative record.

All members will be promoted to panelists for today's call. Members, when using chat. Please select everyone in order for all to see your chat. Observers will have you only to the chat access. All documentation and information can be found on the Wiki space. Recordings will be posted on the public Wiki space shortly after the end of the call.

Please remember to state your name before speaking. As a reminder, those who take part in ICANN multistakeholder process are to comply with the expected standards of behavior. With this, I'll turn it back to our chair, Michael Palage. Please begin.

MICHAEL PALAGE:

Thank you, Terri. Good morning, good afternoon, good evening, everyone. As usual, we'll start off with a quick admin update. This is going to be rather short, there is not a lot to discuss from an administrative standpoint, we are just going to be moving forward with a continuation of our work from last week regarding the measurements of accuracy in the existing data sources. And I do appreciate that a number of stakeholder groups have stepped up and have done their homework. I greatly appreciate that. If we do manage to finish that portion of our work early today, we will discuss the topic of reaching out to UDRP providers. That's something there has been some I believe traffic on the list. So with that, if we could turn it over to—just a quick heads up. We do not have Marika with us, although we are in the capable hands of our other ICANN Org colleagues. And today, Berry will be driving the decks and presentation. And with that, Berry, who is up first? I believe last week, we did get through the registrars and did we complete the registries?

CAITLIN TUBERGEN: Michael, if I remember correctly, we did not complete the registries. Although I think we did complete registrars. Registries are next in the document so we could start with them unless any of the registrar reps had anything to add to their summary from last week.

MICHAEL PALAGE: Okay, and Sarah, it's always good when we have consensus. Okay, so according to Sarah in the chat, she believes that the registrars have completed their presentation of their assignment. So with that, I would ask—Sophie, yes.

CAITLIN TUBERGEN: Apologies for interrupting. I just wanted to flag what Berry has on the screen before we jump into the registries. And that was pursuant to a request from Sarah and others, where staff support was asked to call all of the references to data sources that folks had mentioned in previous calls, or that was mentioned in the briefing materials. We now have this circulated as a document from last week. So the action here would be for groups to review this and ensure that it's comprehensive. If they think anything is missing from this master list, please feel free to let us know. But I know that that was a request for the support team to come up with all of the studies and resources that are currently in existence that the team could reference. So that's what Berry's currently showing. But of course, if something's missing, feel free to flag it to us and we're happy to update that. Thank you.

MICHAEL PALAGE: Okay, thank you, Caitlin. So, for everyone there, and I guess we're going to step back into the administrative role, this, I guess, would be a new homework assignment. So after we complete our current tasks, this was something as Caitlin had referenced that was requested by Sarah. And as I believe Marika and others had noted, many of these, if not most of these resources, were actually in the original briefing documents that were provided to the group. But what ICANN Org has done at the request of Sarah was to compile an augment. So I know there have been some additional data sources that I think we have talked about that have been added to this.

So this compilation document, as Caitlin has said, has now been shared with the group. I would ask all of the members to review it, take a look, see if there is anything that is missing and that needs to be augmented. And again, thank you again to our ICANN Org colleagues for doing this homework and heavy lift. Well, thanks for doing this homework and making our job a lot easier as members. Is there anything else you would want to add on that, Caitlin, that I have missed?

CAITLIN TUBERGEN: We're good. Thanks, Michael.

MICHAEL PALAGE: Excellent. And with that, now, Berry, since you're driving, yes, thank you, if we could go back to the screen which is going to turn over to the registry. So Beth, Marc, or Sophie, who will be walking

us through the—I see Marc Anderson with his hand up. Marc, are you leading, or do you have a question? I'm sorry.

MARC ANDERSON:

I'm going to take a stab at leading. Unfortunately, Sophie is late joining today. So you're stuck with me. So apologies for that. Let's see. This is the registrar input there, if you could scroll down to registries. Thank you. Okay. So I don't want to just repeat what registrars told us last week, there's a little bit of overlap so I'll highlight differences. At the top, we noted that the assignment was a little bit of a challenge, when we were first thinking about it. I guess we're still struggling with exactly what standard to measure against. I don't think we have an agreed upon current definition. That's an outstanding topic, as I recall.

And while we've discussed what the current obligations are in the RAA, what currently exists, I don't think we really came to a concrete agreement on what standard to measure accuracy against. So it's hard to measure accuracy when you haven't agreed on what standard definition to measure it against. So noting that, I'll put that aside and get to our suggestions.

Here, I think we have sort of similar responses to registrars, although maybe a little less detail. Our first one is there are different ways in which we could ask registrars to report on their accuracy rates. We have an example there as of ask registrars to track the rate of email bounces for WHOIS data reminder policy notices, right, that would be a data point. Yeah, I think that would be something useful or interesting to have.

I think we don't have this in our submissions, but I will note registrars told us last week that the problem, or one of challenges here, I think, would be that we as an accuracy scoping team don't have a way compelling registrars to provide that data. And so at least through an accuracy scoping effort, that would be a voluntary ask near as we can tell, and sort since it would be self-reporting, I don't know if the data provided would be accepted by all members of the community. So I think those challenges exist, this is option.

The next one we know is third-party audit of existing data. Again, noting that data processing agreements would need to be in place. Volker noted this on the call week that while that is possible, there are challenges with cross jurisdictional transfers of data. It's not exactly clear how to work that out. Not saying it can't worked out, but this would be a challenge that would need to be addressed.

Again, also, there's an option of instead having a party do the audit, ask registrars to self-report on data there. Again, same challenges apply. Lastly, we have a note to basically compare against statistics provided by ICANN Compliance about accuracy complaints. This is something we've talked about on previous calls.

I guess I'll ad lib a little and add one last note. One of the registrars' submissions that they brought up last week, although I don't know that we got a chance to discuss it, was the possibility of having an audit dedicated to the topic of accuracy. I thought that was an interesting suggestion, and one that might be worth this group exploring further. But that's me ad libbing a little bit. So apologies for that. So that's it in a nutshell. Back to you, Michael.

MICHAEL PALAGE: Thank you, Marc. I'm wondering if I can ask you a question. One of the things that I was thinking about in doing my homework as chair, I always try to go through the same process, when you talk about the third party audit, and obviously the issue of the DPAs, which is something that is critically important, I was looking at instead of just doing a random sampling, would it perhaps be better to focus on perhaps those abusive domains that show up in the OCTO reports, particularly those involved with bots, botnets, malware, clearly illegal activity? Because that on its face would perhaps give a heightened legitimate interest test. The fact that one is going to try to do an audit or survey. I know, this is on the fly and you and Sophie and Beth are probably going to have to go back to the registry constituency. But that was one of the things I was thinking in my homework of instead of just casting out a net upon all hundreds of millions of registrations, do we perhaps look at casting the net to a smaller universe whereby most widely accepted definitions, there is illegal activity happening and perhaps looking at the accuracy of those domain names? Good idea, bad idea?

Owen says it's a bad idea. Steve is laughing. So Becky says "does not provide the size of the problem." So obviously, OCTO—I guess, okay, so there we go. So let's break this down. I'm glad. I thought things we're going to be quiet. So let's break this down accordingly. Let's go first. So Owen, you were first in the queue with it's a bad idea. Would you like to explain or would you like to defer to Becky with the size of the problem across the entire registration data set, her comment? You're up, Owen.

OWEN SMIGELSKI: Sure. Thanks. I completely agree with Becky, what you're doing there is just presupposing that we should only look at domains that have been identified as quote unquote bad as opposed to seeing whether it's representative across the entire gTLD registration space. I think focusing on just that then is a false assumption that accuracy equals bad or accuracy is something that's associated with bad domains, as opposed to taking a look at the whole ecosystem in there. And just focusing on that, yes, that's a place for attention. But if you start looking for bad stuff, you're going to find bad stuff. So I think we really should need to look at the whole thing holistically there, and not be too worried about it. And I do note that, Becky, you're posting your comments to hosts and panelists only, should be to everyone, just to make sure everyone sees these comments that I think we need to look at everything broadly. Thanks.

MICHAEL PALAGE: Okay, so what happens here is with that, Becky, would you like to speak to this? Go ahead.

BECKY BURR: Yeah, I think, Owen, I just think identifying, doing scoping based on things that have already been identified is bad and problematic, is not going to give you a real good look at what the size of the problem is. But I just have to say I honestly don't understand why you couldn't craft a carefully scoped DPA limiting the permitted processing to this analysis. And I also think that you could

probably solve the transporter issue but if not, you could do the analysis in Europe and you wouldn't have any cross-border issues.

MICHAEL PALAGE: So with regard to the narrowly scoped, Beth, would you mind if I put you on the spot on whether you agree or disagree with Becky's assessment? I know you've been leading a lot of the efforts on the DPA on behalf of the registries. Thoughts, if at all?

BETH BACON: I think I don't disagree with Becky. But I also think it's important because this was the registries' input, it's fairly high level. So it would very much depend who's doing the analysis, who has the data, if ICANN is going to do it, if we're going to do it, if a registry is going to do it, and then we combine it with something. The registries, registrars have a DPA. We're already covered. There's a DPA attached to the registry registrar accreditation agreement. It's just the challenge if we involve ICANN in that processing, because we do not have that DPA. I would love to craft a simple DPA. Just get that done, which is what we've been trying to do for a couple years now. So I don't disagree. I just think that maybe we need to think about it in more detail, to see if it's truly feasible. But I don't see why it shouldn't be in some way or combination doable.

MICHAEL PALAGE: Okay, so I guess the next person to put on the spot, Volker, I know you have generally raised a lot of questions or concerns is, do you see—Beth already cited the existence of a DPA between

registries and registrars. Do you think that there is any issue within that existing construct for either a registry or registrar to consider exploring this issue on a narrow basis?

VOLKER GREIMANN: Depends. The purposes have to be defined very narrowly, and the DPA will have to be crafted very carefully. And I cannot say whether this might be an issue before I have seen the DPA. But the trend being that any transfer of data across borders outside of Europe, especially to the US being more and more problematic, as seen with the recent Google Analytics decision by the Austrian GPAs, makes me wonder how this would be organized. So basically, it probably shouldn't be ICANN doing the analysis, but some third party that ICANN also authorizes. And then I am not excluding the possibility that it's possible. I think it's a good idea to explore that avenue, but ultimately, it's going to be not the easiest tasks that we have, or that ICANN has in organizing this.

MICHAEL PALAGE: Okay. And I want to come back to the question, I think that Owen raised about the size. One of the things I think is interesting here we're going to be talking about size is, and maybe Steve could talk to this from an SSAC perspective, we obviously had been looking at data on the number of complaints that ICANN gets regarding accuracy. Obviously, there's some difference of opinion on why that number has dropped from pre-GDPR to post GDPR, we're looking at that. I know, Owen has been compiling some data. And we'll likely be sharing that with the broader group. And I think it's an interesting data point for us to look at.

But if we do look at some of the reports from OCTO, where we have tens of thousands of domain names, those numbers do seem to be substantially larger than what Compliance is reporting. And although that does not necessarily mean that those are compromised, they can be malicious and or compromised. I just think as we're looking for easy datasets to get to, and going to Volker's point on how we want to have a very tight, narrowly defined DPA, the fact that we are looking at if we are looking at illegal activity, that would give us a heightened legitimacy regarding having access to it.

So Owen, no, OCTO is not reporting on accuracy. What they are reporting on is domain names that are associated with illegal activity. And I think part of what would be helpful is, looking at that illegal activity, is that illegal activity associated with malicious domain names, and malicious domain names are probably not going to be putting the bad person's accurate data in there, one could assume, or it can be, or it could be synthetic identity, it could be stolen identity in connection with the malicious. In connection with the compromised, obviously, that would be something where having accurate data would be helpful so that people would be able to get in contact with the legitimate compromised domain name holder to inform them that their domain name has been used with illegal activity.

So with that, I want to go back, I think, I want to go back to Steve, and then Sarah, I see you have a number of comments in the chat. So Steve, you have the floor.

STEVE CROCKER:

Thank you. I have raised my hand before. Just very quickly, in response to what Marc was saying. A very specific point. Marc, you were raising the question of sort of what's the standard for measurement. And it wasn't clear to me whether you were referring to how you measure something or versus what the thresholds are that represent the goal of either how good something is or how bad something is. I'll just set that aside. That's why I'd raised my hand before.

With respect to the question that we're currently discussing, I just will echo in a certain way, the things that are being said here, there is basically two kinds of numbers that stick out where there's something large. One is the number of domains that are viewed as abusive, or being used for abusive purposes, versus the percentage of registrations per registrar, or per registry, I should say, that are being used for malicious purposes, or abusive purposes. And those two are different numbers, because for new gTLDs, the population, the set of registrations is relatively small, but the percentage, for some of them is noticeably high. Whereas .com, obviously, which is enormous and huge, has a very large number of abusive domains, but a very small percentage. And some of the dialog that I'm seeing in in multiple places, including inside SSAC are swinging back and forth between those two sort of measures, if you will, or statistics. And I think we need to be careful about where we focus our attention.

There's certainly a school, one faction that says new gTLD are a likely source of abusive domains and we should be ultra careful about dealing with that, versus a different point of view that says, Well, look, we got a lot of abusive domains no matter whether we

have new gTLDs or not, or whether we include the ones that were started before. So all in all, I'm in agreement with the several comments here that said we need to be very careful about our methodology.

A separate question is, how do you measure what's the relationship of accuracy to abusive domains? And that too, is a fairly complicated issue. I will say, however, that the drop in complaints being reported by ICANN Compliance does not translate, at least for me, directly into any positive signal that there's been improvements in the system. I think one has to look very carefully as to why there's been a drop. And one of the most obvious explanations is because people stop reporting or there's not enough feedback. So the whole question of reporting, and measuring the reporting and using that as an indicator of what the actual facts are, has to be examined very carefully.

MICHAEL PALAGE: And with that, thank you, Steve. And, Alan, my apologies for missing your hand being up, you have the floor.

ALAN GREENBERG: Thank you, it's been so long that I don't think I can cover half of the issues that I wanted to. I'll try to extract a few. It's my brain that I'm worried about, not the time. Number one, in terms of you're talking about DPS and stuff, my understanding, this discussion started way, way back when Marc was talking, and we were talking about voluntary contributions, since ICANN cannot require that registrars self-analyze their data and come up with anything.

So the concept of a DPA which is voluntarily going to be signed by specific registrars, I don't think has any merit. It's just not necessarily going to be done by bad actors, and perhaps not by anybody. So we've heard time and time again, if there's no benefit to a registrar, why will they spend money doing this? So it was an interesting discussion, but I'm not quite sure where it leads.

ICANN certainly could decide to do audits of various sorts. As Lori pointed out last time, many, many companies have set up data processing operations within jurisdictions to avoid cross border data flow. And I don't understand why that is problematic at all. Yes, it may be expensive for ICANN to do it but that avoids the moving data across borders.

In terms of the selective audit of just domains that have shown abuse, I think it's a fine idea, actually. Because the bottom line is accurate data is important as a concept, perhaps, but you have to look at when is it that we want that data. There's 200 million domains, I don't want to contact all 200 million people. I may want to contact those that have some relevance.

So if we study domains that have been shown to be abusive, or potentially abusive, it will not tell us anything about the overall picture. But it might help us understand something about the lack of accuracy and how it comes about. Or if there is indeed a lack of accuracy. It may well turn out that the data is all accurate, but not the registrant's data. That's an important thing to have learned, that we're not addressing an accuracy problem but some other problem that it's accurate as contact data, but not as contact data for that registrant. If we learned something like that, that would be useful.

So I really would like to, if we can come up with something that can be done, can be viably done, and give us some more information about what we're talking about, then I think that's moving the pointer forward. Thank you.

MICHAEL PALAGE:

Thank you. Thank you, Alan. And again, my apologies for missing your hand earlier. That's one of the things Marika is always good at when I'm missing virtual hands. To maybe follow up on that point. And I know Owen had raised this point in the chat, so I'd like to address it.

I acknowledge that the scope of this working group is not abusive domain names. We are focused on accuracy, no disagreement there, Owen. What I think I was trying to articulate earlier, though, was if we are going to have a narrowly scoped DPA, and if we're looking for the most likely scenario with the highest legitimate interest, so that would likely survive any type of scrutiny regarding the processing of that data, looking at domain names that are clearly involved in illegal activity under universally recognized definition, that would probably be the safest data sampling to do.

Now, I think that would be important. And to Alan's point, it would be interesting to look at the data or the results of that data analysis, particularly on the issue of compromised versus maliciously registered domain names. And to be clear, we would not sit there and say that represents scaled out across all domain names, obviously, we're looking at the subset, but I think that is the most legally viable subset that we could query under any DPA.

And again, that's just one of the reasons I kind of threw it out there for consideration. Volker, you have the floor.

VOLKER GREIMANN: Yes, I was a bit puzzled by Steve's comments, because I don't see how abuse statistics between TLDs really relate to accuracy. I think you said that abuse is not what we're looking at, we're looking at accuracy, regardless of what the domain is used for. And that's why I also agree with Alan to a point that looking at a specific subset of data that has been proven to be problematic from another perspective, will also give us very skewed data with regards to the accuracy question. So any analysis of registration data would need to be from a neutral pool of domain names, that has not been previously flagged for any reason. Otherwise, our results will probably look very, very complicated.

But one thing that Steve raised was that new gTLDs are in general, more regulated than legacy TLDs. And yet, those less regulated TLDs see less abuse, apparently, at least on a percentage base of new registrations. If, for example, new GTLD A has very high abuse, a very high number of abuse whereas .com has a rather moderate number of abusive domain names, then that leads to the conclusion that more regulation could actually—and that would include extra regulation that we will be coming up or the working group will be coming up with at a later stage—does not actually lead to less abuse. Or the other way around, the more you regulate, the more abusable you will get. I'm not sure if that's the correct conclusion. But it sounds like that, and I'm not sure that was the comment that Steve was trying to make. Thank you.

MICHAEL PALAGE: So if I can, Volker, just to set the record straight, and as someone that has worked with legacy as well as new gTLDs, I've actually posted it in the comment, when people make that broad sweeping statement about new gTLDs being the source, I think if you look at it, most of the new gTLDs combined account for 40%—Hold on, let me pull the exact statistics up here. The two most abused new gTLDs account for 41%. So when you look at upwards of 1000 new gTLDs, that those two are accounting for 41%. I really want to be careful here, and again, respectful to the people that are working on the SubPro to advance new gTLDs.

I really want to be careful when we potentially are casting broad brushstrokes. So again, I'm all about data and I think the EU DNS abuse report, while not perfect, I think actually has a lot of good, interesting data points. And I would encourage everyone to plow through all 173 pages of it. Going down, Beth, I believe you were next in the queue.

BETH BACON: Yep. Hello, friends. So I wanted to just really quickly bring it back to the registries' contribution that has started this giant snowball of discussion. I think that these are all very interesting points, and they're valid. But they're not helpful unless we actually decide if this is a good idea and what we want to study. So if it seems like people are like, yeah, this is some we do need to get some more data, this is maybe a way to do that. I want to suggest that maybe the registries and the registrars put our heads together, figure out what data we have, what we could share with whom, who could

maybe be the analyzer, aggregator of data, and maybe come back with something that is a light scope, so that instead of thinking about every single possible permutation of this process, or this effort, we actually are commenting on something concrete, because I think otherwise, we're just going to keep going around and around and about what to study and what not to study and talking about just looking at abusive domains or just looking at new gTLDs—which I didn't think was Steve's suggestion. So I think that maybe it's a practical way forward.

And I had other comments on the rest of this discussion, but I think I'll leave that aside and just ask if anyone thinks that that concrete way forward so that we don't just spin around for another 20 minutes would be acceptable and helpful.

MICHAEL PALAGE:

So while people are waiting to raise their hand or comment in the chat, Beth, I appreciate what the registries and registrars have done kind of being the tip of the spear on this particular assignment. I would not characterize us as going around and around, I would say that this has been hopefully a thoughtful discussion that will spur other members and stakeholder groups to perhaps make their contributions on where additional work can be done. So I'm going to take the glass half full that this dialogue has been constructive.

I'm just looking through. So Beth, I see support from Melina and Becky in the chat. So that's good. What I want to do now is basically try to perhaps wrap up. Yes, I see. Susan, I see both you and Scott have your hand up. So this will be wrapping up on the

registry. And then we'll be moving to either a ALAC or the IPC. I don't know who is next in the queue as far as their assignment for this week. But Scott, you have the floor right now.

SCOTT AUSTIN:

Thank you, Michael. I'm really glad you brought up the EU abuse study, I'm only going to touch on one thing, because we've been using the word "abuse." And the various varieties of that became very clear to me last week when I was going through it. INTA actually has a committee, and we're going through the various aspects of it.

But I sent it to a colleague and got an unsuspected response, which was basically, "Gee, it looks like only 3% or some extremely small fraction of the abuse issue deals with trademark infringement. So I guess, why is ICANN spending so much time on it?"

Which may be music to some people's ears, but I went and looked at what he was talking about. And it has to do with a chart that is on page 155 of that, that deals with a specific automated detection system. And that detection system basically does not include or it excludes, because—and Volker mentioned the word "skewing of data." It excludes trademark infringement as a basis for abuse. It basically focuses on malicious abuse. And if you go to the bar charts, it shows spam, phishing, other, which I'm not sure—I haven't gone into those details, a scam, pharma, counterfeit, which could be infringement. Impersonation.

And I only bring this up because when you get the trademark, it says zero. And I guess that there's a lot of different ways to define abuse, just like there is accuracy. And I just think we need to be mindful of that. And you've used the word "compromised," Michael. And I just wondered what the scope of that term was, in terms of forms of abuse, if that was beyond malicious.

MICHAEL PALAGE:

Yeah. So what happens is I will give a plug out to the contracted parties, which I believe will be having a session at the upcoming ICANN 73 meeting that will be looking at maliciously registered. And let me rephrase that. There is going to be a session that will look at abuse, and distinguishing between whether that abuse is associated with maliciously registered domain names, or compromised domain names.

So this would be someone who registered a domain name in good faith, but somehow there was a security vulnerability that allowed a nefarious person to get access to their account or their website or their blog, and then somehow engage in the problematic activity. So again, if one of my colleagues from the registry or registrar constituency can perhaps pop that into the chat, I think that would be good. So that's one of the things I would encourage you—not only would I encourage you, I would actually encourage all members of this working group to attend that session, I think it will be highly informative. Susan, you have the floor.

SUSAN KAWAGUCHI: Thanks, Michael. I just wanted to respond to Beth's idea. I appreciate the fact that this group of registry and registrars that are participating here would go back and think about what they could do and provide as an analysis of inaccuracy or accuracy. But I think it's the age-old problem that—and I feel like I say that this on every group I've been on—our problems are usually not with the registrars and the registries that participate in ICANN and especially the registrars that are not represented here. Whatever study we could think of, if we are not receiving—if all registrars do not participate, then I think the data would be skewed.

I fully believe that the registrars here when they say they're compliant with the RAA are doing that, but it's always those fringe registrars and some of those have been identified in the EC report for abuse that may not equate to inaccurate, or not following the RAA, but I bet you would find a higher percentage of inaccurate domains in some of those registrars that have a higher level of abuse.

And I realize we're not studying abuse, but the two do go hand in hand for the most part. So my question would be, okay, we come up with a study, but how do we compel all registrars to provide the data? And that's the only way if we don't see the whole set of data then or an analysis of the whole set of data. I'm not sure how valuable it would be.

MICHAEL PALAGE: Thank you, Susan. Owen, go ahead.

OWEN SMIGELSKI: Yep, thanks. I really don't like how we keep going down this path, assuming that just because something is abusive means that there's got to be inaccuracy there. If we look at people who get speeding tickets, we're going to find some one set of data as opposed to whether we look at all drivers who drive cars on the road. I think it's very prejudicial to look there. And also, just because something isn't, then that makes the assumption that inaccurate data means it's abusive domain name there.

And then I think it's a really bad path for us to go down. I think we'd need to be looking at this in a more broad, general, holistic approach to see about what is the scope of the problem because again, this is an accuracy scoping teams, not an abuse scoping team. And we need to avoid that because if you start focusing on a very small subset of domains, we could probably find out that they're registered by left-handed people or a whole bunch of other things in there that don't apply to what we're trying to do here. Thank you.

MICHAEL PALAGE: As I said, Owen, I responded back, I hear your question, I hear your concern. And looking at some of the OCTO reports, there's 20,000, 50,000. I don't think those are incredibly small. And again, just from a legal perspective, the ability to process that subset of data would probably give us the highest legitimate interest, I think. But again, this is what I think we're going to put together and perhaps Council and others will decide where to go with this.

So with that, I want to go back to our chart. Berry, who is up next on our chart, which stakeholder group? IPC, thank you very much.

Scott, I believe you will be walking us through the presentation here today.

SCOTT AUSTIN:

Yes, thank you, Michael. I did put up my hand real quickly just to note that in that same report we were just talking about, there's a section about .DK and the ccTLD related to it. And one of the statements in that report, page 159, I think, says—and this is, again, their analysis, they being the folks related to that ccTLD. It sort of stood out to me that says the reasoning behind their algorithms is that online criminals tend to use inaccurate registrant data. And what definition of inaccurate they're using ... ? But again, I think to respond to the comment just a few minutes ago, that in some circles anyway, they look at inaccurate registrant data as related to abusive activity, and that there's some relationship there.

Anyway, the IPC submission for the homework. Let me bring my screen up to the side so that I can see everything.

MICHAEL PALAGE:

While you're doing that, real quick—I'm sorry, I'm going to give you an extra 30 seconds to find your location. Sarah, I believe it is on page 158 of 173 on the EU study where .DK is specifically addressed. And they talk about verification. And then there also is a reference. The registry shall investigate the accuracy of WHOIS information. And just as another data point, I believe the word accuracy was used 52 times. I don't know if that will help. So page

158 to 159 of the EU DNS abuse report is the specific reference that Scott was citing. Scott, hopefully that gave you enough time.

SCOTT AUSTIN:

Oh, thank you for filling. It did me give me a chance not only me a chance to get the full screen, but it also got a chance for I guess Berry to move the section on the screen, the desktop so I can see it. And by the way, as I've mentioned in this response, the .DK was referenced in a position paper that the IPC did back in August of 2021 as something that we felt was headed in the right direction or had dealt with accuracy in a way that was meaningful from our position.

Anyway, first, the IPC shares things that were submitted by others. Alan, in ALAC, had expressed a concern about the amount of time available for us to determine whether goals were met. We agree with that. We also believe that the reference he made to cross-field requirements, which we have seen referenced in some other discussions with some of the folks in the service providers, WIPO and so forth, that that is something that would be very helpful. So we agree with that.

In the GAC section on this, they use the definition for domain name registration data for the ARS, and as many of us have said, that's been suspended. But by the same token, its definition did include not only syntax and operability, but also identity. And so we felt that it was important to at least get on the table the fact that among the widening, as ALAC had referenced that for verification, I think that that is one of the concerns that we think that the identity validation, registrants and registering organization should

be considered, especially with recent rises in fraud. And we can all debate about that. But the amount of fraudulent registration data is something that should be considered and that also is given great analysis in .DK in terms of the way that they dealt with that. And significant reduction in fraudulent registration data also led to the reduction of what they call fraudulent web shops, online, that were dealing very unfairly with consumers, confusing them, and resulting in counterfeit goods being sold sufficiently that their [own country] was besieged with seizure orders to take down the counterfeit sites. So I think that that was something that we brought to the attention before and I think it shows up in the recent EU report from 31st, from the end of January.

The practical considerations, because I think that's really what this question was getting at, we looked at ICANN should be the one that audits the RAA compliance. But we also noted that back in some of the early resources that Marika had provided, there was a reference to the NORC group that helped with the study that originally was the prototype for ARS and included validation. And it said they used private and public sector validation specialists. So perhaps there is those same specialists or later ones have developed since then that would be available to assist ICANN in their needs for either that audit and for the things that were necessary for this development of accuracy that we're trying to deal with here.

And we also felt it was consistent. And then I use that as the jump over to talking about the fact that third-party verification solutions in the Denmark approach and the reference to obviously this DK Hostmaster does have some governmental aspects and

governmental ties to it. So I provided some of the information there. So that would be for credibility purposes for to show the references where I got the information. I think that's pretty much it.

MICHAEL PALAGE: Thank you, Scott. Any questions, comments or concerns? Marc Anderson, you have your hand raised.

MARC ANDERSON: Thanks, Michael. Scott, thank you for the overview. So you have a lot in here about, I guess, additional accuracy measures that you would like to see put in place. But when it comes to how to measure accuracy, what I think I heard you talking about was a suggestion to have ICANN Compliance do an audit dthat the accuracy measures are being followed. I just want to make sure I'm—I guess I want to make sure I understand the IPC submission and the points you were making when it comes to how to measure accuracy. So I want to just make sure I heard that right. Your suggestion on measuring accuracy is an ICANN Compliance accuracy-focused audit. Did I have that right, did I miss anything? Thank you.

SCOTT AUSTIN: Should I respond, Michael?

MICHAEL PALAGE: Please do.

SCOTT AUSTIN: Yeah. I just didn't want speak out of turn if there's other others in line. But what I was trying to do, Marc, was really address the task I've been given, which was—and I did not fill out the column to the left of the column that I submitted my information in. And that seemed to me those were the primary points that were being made. So in order to try and remain consistent with that, that was the focus.

I also, like I said, had seen some excellent suggestions, I thought, by ALAC and some items in what I thought was the GAC, although Sarah is questioning where I saw that. That was further above. But it was really just for the definition, or for the reference to the ARS that included identity validation, or within that summary. But to explain, Marc, why I focused on those items, I was trying to do what Marika and I think Michael had basically told us to do, was that we were responding to what had been submitted in the left-hand column.

MICHAEL PALAGE: Marc, you have the floor.

MARC ANDERSON: Thanks, Scott, that's helpful. You mentioned the ARS. Can you maybe expand on that? Are you suggesting like restarting ARS or making changes? Especially given what we heard from ICANN about ARS.

SCOTT AUSTIN:

Yeah, I think that's extremely problematic. And no, I don't want to go there with specifically ARS as it stands, and then the situation where it appears to be suspended and it's going to be that way for a while. All I was trying to get at, again, was to respond to the very narrow question I was confronted with, and doing that using some of the tools that were available, which included the combination of operability and syntax and identity that was in that original creation and development of ARS.

And again, harkening back to the 2010 study that both talked about that study being the prototype, or the predecessor to the ARS. And in that particular case, there was clearly the participation of what were known as validation specialists or validation expert from both the private and public sector. And that was something that to me was new. And I thought it might be useful to assist ICANN in this exercise, whether it's in reviewing what kind of compliance is occurring, or in reviewing how much of the determination of accuracy should include something like the validation of identity.

MICHAEL PALAGE:

All right, Scott, I see no further hands up here. Berry, I know there's some interesting discussion going on the chat. But I want to be mindful of trying to get through the homework assignments. Would you mind remind me who is up next? Do we have ALAC? Yep. ALAC is next, I believe. So Alan, you have the floor.

ALAN GREENBERG: Thank you. A number of points here. First of all, the question asked that we're supposed to be filling out in this column was very different from the original one. The original one was, how do we determine if there are accuracy problems and who should do it? This one said, how does the scoping team do it? And I'm a little bit confused on that, because I didn't think the scoping team is going to be running studies. And if it is, our timeframe is going to expand radically. So I'd like a little bit of clarity about what we're really trying to achieve here. If we are going to commission a study of some sort, whether it's ICANN doing it or an outside group, that's going to blow the whole timeframe we're looking at completely. So I think we need to make sure we're all talking the same thing.

MICHAEL PALAGE: Excuse me, would you like me to address that first, or do you want to get through everything?

ALAN GREENBERG: Sure, go ahead.

MICHAEL PALAGE: Yeah. So what happened to this was, I fully agree and concur with you. And if you notice, early on when I was talking about trying to give a heads up to ICANN about budgetary considerations for potential studies, so I guess the best way to summarize the question on your first point is yes, if there are recommendations about what we need to do where we make a recommendation on a study and somehow that gets approved, that will inherently

provide our timeline. Yes. So hopefully that answers your question.

ALAN GREENBERG:

Okay, thank you. The second thing is, we seem to be conflating whether people are adhering to the RAA specifications and with the overall issue of accuracy. The RAA specifications are very, very targeted, they just look at specific fields for some kinds of verification. And moreover, they are only looking at domains that are newly registered or changed. The original accuracy studies looked at sampling of the overall domain name system, the gTLD system. And that's a very, very different group of domains.

And I'm not going to say whether we end up determining that the RAA specs, the RAA, very narrow target of only new domains is sufficient or not, but let's not conflate them. They're two very different areas and two very different scopes. And I think it's important that we know what we're talking about.

What else is there? Yeah, I think the rest of the things, I've said before, and I hope we said them clearly. We don't believe this is a self-reporting thing. If we're trying to assess the accuracy, we have to assume that we're doing it for a reason and simply self-reporting, having people say, "Yes, I was a good boy and girl and I followed the rules" is not going to be sufficient. And moreover, we are not in a position to compel registrars to do any sort of full audit of their own data or contract with someone else to do it.

And lastly, if there is a belief that this is out of scope in terms of ICANN does not have the right to do it, either without the right

data processing agreements or the right joint controller agreements or at all, I think we need to get clarity on that. My understanding is ICANN is at least some sort of controller of this data. And since many of us [sit in PDPs] setting the rules, it's hard to imagine ICANN is not a controller at all, that ICANN should have the ability to audit the data and make sure it's correct. There is a trans-border data issue that has to be addressed. But I believe that's addressable.

So going back to the discussion we had last week and Becky's suggestion of asking the European Data Protection Board, I think if there is any doubt whatsoever that ICANN has the right to do this and what the right mechanisms are, we need clarity on that. Thank you.

MICHAEL PALAGE: Excellent. I see a hand up. Volker.

VOLKER GREIMANN: Yes, thank you. And I found the comment from Alan a bit confusing when he says that the requirements of the 2013 RAA only applied to a certain subset of the domains because, Alan, if you recall, we had that same discussion as part of the WHOIS review team. And we basically concluded at that stage that from the data that we were presented with, actually a majority of domain names were at that time already under the 2013 RAA, not part of the grandfathered domain names simply because they had been transferred, many domain names have been deleted and re-registered, and all the processes in place in the 2013 RAA

guaranteed that that number would become smaller year by year and only a very small subset of domains remained under the grandfathered rules.

And as you also recall, we had the discussion at the time that those domain names that have been registered for such a long time and not been touched and not been transferred usually do not form a problem with regard to accuracy of data because those were long-term owned domain names that the owner had a very high interest in. And therefore, the accuracy issue basically did not arise at that point.

The second point I wanted to raise with that regard was something that I will come back to later because there are others in the queue and it doesn't fit here.

MICHAEL PALAGE: Marc Anderson, you have the floor.

MARC ANDERSON: Thanks, Michael. I guess I want to ask for a little bit of time to talk about Alan's concern about time for a study. Michael, you sort of addressed it, but I don't want to gloss over this one quickly, I think this is an important point that deserves a little bit of time. That's also something I feel kind of strongly about.

Based on my experience with previous working groups, which is different than what we're doing here, but if we identify a study or a survey or come up with a mechanism for measuring accuracy, and it'll take a certain amount of time, I think it's important that we have

the time to get the results of that, to be able to get to the other parts of our assignment as the accuracy scoping team.

We need good data, we need to have solid data to be able to provide firm and actionable recommendations. And so I think this is an important point and I don't want to just gloss past it. So could I maybe ask you to speak to that a little bit more and sort of make sure we all have a common understanding of what would be the process if we identified—or maybe I'll be glass half full, once we identify the process for measuring accuracy.

MICHAEL PALAGE:

And since I'm putting people on the spot here, Berry, I know, you and I have discussed, I know early on in the process, we had talked about this with timing, budgetary. Could you perhaps give your insight historically, on what that would look like from an overall process standpoint?

BERRY COBB:

Well, I guess, process procedurally is, once this group has identified kind of the guardrails or scope of any type of study, we would need to document it in a way and present the contents of a request to the GNSO Council first. And that would need to be basically approved by the GNSO Council whereby then it would be submitted to the Board or Org for consideration. And this presumes that there's a sizable resource requirement for a vendor and/or staff time to conduct such a study. And I'm also presuming that this is somewhat different than the studies that have been done in the past for ARS. So I'm trying to separate that out.

But ultimately, that request would be evaluated by the Board and Org to determine feasibility and funding and ultimately get to some point of an approval process by which something like that could get kicked off. As it stands right now and where we're at in the scoping team, this is kind of an unplanned expense. So their planning and finance team would have to consider how and where any funds to support this would come from. So in a very general nature, that's the process. I'm not saying that that's the exact process. I'm really kind of replicating how funds were requested from our phase one and phase two of the EPDPs. And hopefully, that provides a little bit of insight.

MICHAEL PALAGE:

It does. And again, I apologize for putting you on the spot there. And I guess my comment to Volker is Volker, now you understand why we may need that light drive engine we're talking about to transverse our galaxy. Unfortunately, to Marc's point, and I think Sarah's point, if we are going to do this, let's do it right. I think reading the first ODA, I think there was a lot of missing gaps that were not done previously. And hopefully, as part of this scoping group, we can perhaps learn from the ODA on how we need to do data gathering sooner in the process as opposed to later.

And even though that may involve more time, I think it's better to make that time investment earlier in the process, as opposed to later in the process. My personal thoughts. And yes, I have seen your comments many times, Sarah, and we will get to that on the accuracy definition. Okay, where are we at now? Marc, go ahead, you have the floor.

MARC ANDERSON: Thank you, Michael. And I appreciate that comment. And Berry, I appreciate the explanation about the path we would go through. I'm dropping in our third charge, third bullet point in the charge of the scoping team which tells us on the basis, the assessment under charge one and the data from charge number two, which is the measurement of accuracy, will undertake an analysis and really, that analysis is dependent on the data we get from step number two, the measurement of accuracy.

To Michael's point, I think it's critical that we have this information, we have good data that we can take concrete and actionable steps against. And if that means we have to wait a little bit to get those results, I think we should do that. I don't know what others think on that. And I know this is a little bit of a tangent to what Alan was talking about in his submission, but he brought up the timing, and I thought this was a good time to raise it. So as I said, if we're going to do this, we should do it right, we should have good actionable data to come up with concrete suggestions off of.

MICHAEL PALAGE: Thank you, Marc. Alan, you have the floor.

ALAN GREENBERG: Thank you. Two points. On Marc's comment right now, I don't disagree that it would be nice to do it. But I just wanted to acknowledge that it's going to change our timeline all together. The ARS, if in some magical world, we could restart it today, they still had a six-month cycle, that it took over six months from the

time they started collecting some data until the time they had the results. So we're talking a very significant time. If we commissioned a study, like the NORC study, again, assuming it's all possible and viable to do that—and that's far from clear—then, again, we are talking about a very significant amount of time to carry out that kind of study. I'm talking about six months, a year, a year and a half is what we've seen in past cases.

So I'm just noting that if we are going to do that, and I think it would be a dandy thing to do, it's a different world than what we were talking about before. In terms of Volker's comments on what the RDS review team found, I'm afraid my memory is a little bit different than his. My understanding is that there are still significant domains that are not treated as new and therefore not subject to the accuracy terms within the 2013 RAA. But so be it. Thank you.

MICHAEL PALAGE: Marc, you'd like to respond?

MARC ANDERSON: I would, and thanks, Alan, for that. I think your point's well made on the timing. My personal opinion is that if we have to pause and wait for the results of a study of some sort in order to get good actionable data, I think it would be worth it. I wonder if others agree. I'd like to hear from other people, but I realize that that pushes the timeline out potentially considerably and means that we would potentially have to pause our work for a period of time while we wait for the results. So I am curious what others'

thoughts are. Michael, where you are on that. But I think it's worth us considering and maybe important for us to get on the same page there, what everybody's expectations and thoughts are on that.

MICHAEL PALAGE:

Oh, that was, if you recall, one of the very first questions I asked everyone, was, what was your expectations of what you wanted out of this group? So with regard to the expectations out of the study, I think this is to Sarah's point, I think we're going to need to come back to that definition/explanation because I think that definition of accuracy is largely going to be driven by some people's view of assignment three and four of where they need to get to, right? So if we adopt the definition put forth by the registrars and supported by the registries, the fact that you have syntactical validation, and then you have one of the two operational contacts, either email or phone work, that is accurate per the 2013 registrar accreditation agreement. That will define the study one way.

I believe, though, there are others that have talked about a more expansive definition. And that is part of why we did the gap analysis not of where we're at, but where we need to be. And part of that, Marc, I think, will define what the study is and how it is crafted. I don't know. That's something that the group as a whole will have to come to. And my job is just to facilitate that journey, and hopefully reaching consensus. So that's my personal perspective. Did that answer your question?

MARC ANDERSON: Sorry, was that directed back at me?

MICHAEL PALAGE: Well, I think you asked me. I thought I heard a question directed at me towards the end of your statement. So I wanted to give you my insight as a chair of what I think I've been trying to achieve since the start of this journey and where I think we are at this inflection point, if in fact, we were to commission a study and what that would mean, and how that would progress. So I'm sorry if I miscommunicated.

MARC ANDERSON: No worries. I guess my question was more like where are you and where are others in the scoping team on the possibility of pausing our work while we wait for the results of a study. Is that acceptable? Is that not acceptable?

MICHAEL PALAGE: So I guess my response, and then I'll put it out to the floor, is if we do not have any other work to do, I think pausing should be something that we should consider. I don't think having weekly calls to do nothing would be productive for anyone. That being said, I do think there is some additional work that can be done. You know, just looking at the EU report, some of the work that's being done in the ccTLD community on accuracy and verification.

While ICANN is commissioning their study, I think this group could perhaps begin looking at those data points. Owen has compiled some of the statistics from ICANN Compliance. So to me, I think

there is a lot of additional work that we as a group could look into as part of assignments three and four, while that ICANN study is being scoped and conducted. But that's just my personal opinion. Melina.

MELINA STROUNGI:

Yes, thank you, Michael. And thanks, everyone for this interesting exchange. I've heard several times the argument that we need data in order to provide concrete and helpful recommendations. And I fully understand this. Maybe another idea of course, one option would be to stall or pause any work. But at the same time, a concrete proposal will be to address other issues such as, for instance, the fact that there is currently lack of ability to check the magnitude of the problem.

So I think I've heard from all groups, if I'm not mistaken, that we can't know, we can't assess—if there's no data, we can't assess if there is accuracy problem, or if there is, how big that is. So just the fact that there is inability to check at the first place how big this is, regardless of whether the argument is that ICANN has no access to this data or XYZ. The fact that there is currently inability to check whether there is an accuracy issue is an issue on its own, is a self-standing problem. And I think, as a group, we can see what solutions we can find to address the fact that there is inability to check the accuracy in the first place. And for that, I don't think we need any data, we just need to think of ways to tackle this issue. Thanks.

MICHAEL PALAGE: Thank you, Melina. And just a quick time check. We have about seven minutes left. I want to just see, are there any other hands? There are no other hands. I believe the next up is NCSG. Would you like to try to squeeze that in in seven minutes or would you prefer to wait till next week and have a full discussion? I don't want to put anyone on the spot there.

STEPHANIE PERRIN: We both have to leave sharp on, so we would prefer to leave it till next week if that's okay.

MICHAEL PALAGE: Totally. Put it this way, you did the assignment. I will give you the full time and any additional time you need next week. So you will be first up. Just a quick reminder. I believe we are still missing the GAC, SSAC, and the is ISPCP assignments. So hopefully, if after the presentation next week of the NCSG, those other groups will be in a position to present their statements on this particular assignment.

One other thing that I would like to do, just mindful of the many times that Sarah has repeated the concern about the definition, I would like us to perhaps work asynchronously via the mailing list to perhaps nail that down. I think we are close. And I would like to just kind of just now that down so that we have agreement on that to provide clarity on next steps. So with that, I am proposing to call this meeting to a close five minutes early unless there are any individuals with final thoughts, comments, questions or concerns.

Seeing none, thank you, everybody. I thought it was a good, constructive exchange today. And with that, enjoy the rest of your day. You can stop the recording.

TERRI AGNEW: Thank you, everyone. Once again, I will stop the recordings and disconnect all remaining lines. The meeting has been adjourned. Stay well.

[END OF TRANSCRIPTION]