ICANN
Transcription
Next-Gen RDS PDP Working group call
Wednesday, 21 December 2016 at 06:00 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. The audio is also available at: http://audio.icann.org/gnso/gnso-nextgen-rds-pdp-21dec16-en.mp3

Attendance is located on agenda wiki page: https://community.icann.org/x/i6TDAw

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

Nathalie Peregrine: Thank you very much, (Jamie). Good morning, good afternoon and good evening, everybody, and welcome to the GNSO Next Gen RDS PDP Working Group call on the 21st of December at 6 UTC.

In the interest of time today there will be no roll call taken. Attendance will be taken only via the Adobe Connect room so if you're on the audio bridge only today, as in connected via the telephone only, please let yourselves be known right now.

Hearing no one, I would like to remind you all to please remember to state your names before speaking for transcription purposes. And to keep your phones and microphones on mute when not speaking to avoid any background noise. Thank you ever so much and with this I'll hand it over to Chuck Gomes. Go ahead, Chuck.

Chuck Gomes: Thank you very much, Nathalie. And welcome, everyone, to our call this week, our last one before the New Year. Appreciate everyone on the call but

in particular those of you who are calling in the middle of the night. And as I look through the names on the list there are lots of you so much appreciated.

You can see the agenda up on the upper right. The first thing I'll do is just ask if there are any updates to statement of interests. Okay, not hearing or seeing anyone. Let's go on to our first agenda item, which is to review the poll results for the poll last week. If we can pull those up in Adobe that would be appreciated.

Those were sent out. Hopefully you had a chance to look at those. You should have controls over the size of the font to help you read what's on the screen. I'm enlarging mine so I can see them. Okay it looks good to me. And you have scrolling capabilities too so it'll be your responsibility to move along with us.

So let me slide that over a little bit. Okay. So I won't go through the introductory information in the poll. Hopefully all of you took the poll and you've read that already. The first question was, "During the Working Group call it was suggested that Sub Question 2.1," which is listed up above if you scroll up there, "is teased apart to allow for other possibilities. Which of the following alternative statements, if any, do you agree with?"

And as you can see, for Option A in that question, there were 18.2% that agreed with that one; 40.9% with Option B; 9.1% for Option C. And by the way, that - since there were 22 people that responded, every 4.6% or so is a person so that - for C, in other words, there were two people that chose C.

Nobody chose D, which was kind of an enlightening item there. Not enlightening in terms of surprise but the fact that nobody supports that one chances are we might be able to eliminate that as an option fairly quickly. 20% for Item E, which was a write-in, okay, suggested by Greg Shatan, so four people supported that.

Four point five percent did another write-in, was A and B, in other words, not just one of the above, nothing was chosen for A through D in Item 1 but A and B was mentioned as a possibility for - depending on the thin data element. And then someone else put B and C as a possibility. So again, as you can see, there are 22 responses.

And that's not a great sample, but as a leadership team we did look at the distribution of the groups that those responses came from and they were fairly - it was a pretty good distribution of different groups. Not that there was a lot for any one group but out of 22 that couldn't be the case, but at least we had at least one person for every group represented in most cases more than one.

Question 2 then gave the opportunity for you to elaborate either on your answer for Number 1 or to give a different answer. We'll try to always have one like that so that if you find the options in the main part of the poll to be unsatisfactory from your point of view, say that and suggest alternatives. That's what happened here.

And we got a lot of good information out of that exercise. And I think helped guide us in terms of how we're going to pursue the discussion today. I'm not going to go through the comments, but you can see, and again this was attributed so you can read the comments for yourself. There were - I can tell you the leadership team did read all of them. And we tried to take those into consideration as we were mapping out the agenda for today.

Let me pause there and see if there are any questions or comments about the poll, the poll results, or something else you'd like to talk to related to the poll that we took this past week. Yes, Greg, true but with regard to the fact that you suggested E, but the specific wording, if I - did - were you the one that suggested the specific wording that people chose? That's where probably I missed up. Anyway, thanks, Greg, for the input on that and your comments with regard to the poll. Thanks. My mistake. Apologize for that.

All right, so I'm not seeing any hands raised so let's go ahead then and go to the next agenda item, Agenda Item 3. And the way we're going to start that in our lead - on our leadership list after we got most of the poll results in, Lisa did a - what I thought was a very helpful analysis of the results. And so I'm going to ask her - and if we can bring that up now that would be good. But I'm going to ask Lisa to share her analysis of the results. And she understands it's just her analysis and so - but I thought it was helpful.

And so what we're going to do on each item of her analysis is to open it up for discussion on each item if people have comments about particular thoughts that she shares. The - Stephanie, the questions are - well they're on the screen right now. But we're not going to do the poll right now.

In order to have time for sharing the results and so forth, we need to have a cutoff, usually no later than the Sunday before so that staff has time to compile the results. At the same time, you're welcome in our discussion today to share your thoughts so we'll ask you to do it that way. And the results of the poll themselves, the numbers while they give us some guidance, they're not going to be used as a vote so you don't need to worry about losing out on that regard. There will be plenty of opportunities for people to chime in.

So for future reference, when we do a poll, and we'll probably be trying to do a short poll after each meeting, please try to respond within the timeframe and so that we can summarize the results like you've just seen. And like Lisa points out, a PDF of the poll was posted with - at the last meeting page so - and they're on the screen now. But now is it possible to bring up your analysis, Lisa? Or you want to just talk to that?

Lisa Phifer: This is Lisa Phifer for the record. I'd prefer to just speak to it. I didn't really actually present something to display on the screen.

Chuck Gomes: Okay and that's okay. In fact, that way the poll questions will stay up there and the results will stay up there so that's actually good, I think. So, Lisa, I'm going to turn it over to you.

Lisa Phifer: Thanks, Chuck. Again, Lisa Phifer for the record. In thinking about what we can learn from this poll, the first thought I had was that the goal of polling really at this very early stage is not just to find points of consensus, although certainly would be great if we could do that, but our goals were really to find a productive path forward for deliberation. So that's the way I approached reviewing the results that Chuck just went over.

We can try to find a path forward by trying to uncover differences in interpretation or assumptions, which is what we saw when Greg Aaron suggested a different alternative than one of the four that we had listed initially. We can also do this by trying to unravel interdependencies and strip down concepts to more basic building blocks that we hope might be able to be deliberated upon separately and then recombined.

And we can also do this by eliminating obvious nonstarters in hopes of allowing greater focus on concepts that while we don't have unanimous support or even near unanimous support for many of the options we can focus on the ones that at least had some support.

So keeping those thoughts in mind, when I reviewed these results, I saw a few possible ways that these results might help guide our next step deliberation. And we talked about this in the leadership call on Monday.

So first of all, although results were of course scattered over three of the four options for that first question, if you look at it, over 90% of the respondents, the 22 that did respond to the poll, thought that purpose should at least play some kind of role in policy for thin data. And to us, that seemed like a strong indication that this working group, as part of its next steps, should at least

examine both legitimate and illegitimate purposes for thin data. And that that might be a useful next step as part of our deliberation.

But in thinking about that, of course there were 9% who supported Option C and that is thought that policy should not take into purpose - take purpose into consideration. So in thinking about that it might be useful to hear from the people who supported Option C, why shouldn't policy consider purpose or look at purpose as a limiting factor?

Chuck Gomes: Thanks, Lisa. This is Chuck. So Shane, you've got your hand up. You get to go first.

Shane Kerr: Hi, good early morning, everyone. So, yes, I was one of the two people who said that - I believe I was one of the two people who said that we shouldn't take purpose into consideration. Part of that comes from previous background running a Whois registry, both at (Aaron) and (Ripence UC), and my thinking is looking at the actual types of data that we've defined in thin registries, it's not personally identifiable. It's all operationally useful.

So if we decide that it's only available for specific purposes, then that implies to me that there needs to be some sort of enforcement mechanism. Someone needs to be busy blocking access and punishing people who use it for reasons that are undefined. That enforcement mechanism is going to be expensive for everyone. It would be expensive for ICANN as a community because they'll have to be constantly reviewing and refining that. It'll be expensive for the organizations involved because they're going to have to try to find and - to find legal and other enforcement mechanisms, they're going to have to hire staff for this.

It'll be expensive for the people using the data because they're going to have to try to make sure that they're not violating any sort of terms of use and things like that. So to me, the potential gain by defining an explicit purpose is quite limited. And I realize someone on the list did mention that there's a

chance that this could be used to identify personal information and to track people down, and I recognize that is a potential risk.

I think it's quite a small risk compared to the actual overhead and cost of creating any kind of enforcement mechanism where it's basically anonymous data about services that are running on the Internet so that was my reason for saying that there shouldn't be any limits because I don't think there's any benefit for limiting and I think there's a huge cost.

((Crosstalk))

Chuck Gomes: Thanks, Shane. This is Chuck again. And I'm going to - after I say a couple things in response, in fact I'm going to come back and ask you a follow up question, Shane, warning. I don't think it's a hard one but I'll do that next. But then I'm going to - see if anybody else would like to support express why you would support C, whether you participated in the poll or not. And then I'm going to give opportunity for those who look at it in a different way so that we can kind of - all of us see different perspectives there.

So, Shane, the question back to you is this, the fact that we want to have - or let me start over. I'm not saying that very well at all. So if we take a position that thin data should have a purpose, that doesn't necessarily mean that we require people to share that purpose or that registrars have to manage that purpose.

So my question for you is, if - and we're not going to make this decision right now, we will get to it before very long, but if we assumed that even though we're not going to consider any elements, if we haven't identified some purpose for them, we're still not going to require a user who's seeing the data to report that purpose. Would that change your thinking at all?

Shane Kerr: Hi, this is Shane again. Honestly, I don't think so. That seems - I mean, it's a little bit early in the morning here to be having a philosophical discussion, but

it seems like a - a somewhat pointless activity. I mean, of course, anyone accessing the data is going - what they consider some purpose for it, right, they're not going to do it just as a, you know, a hobby. Well some people may. But then that's their purpose.

I mean, the very fact that you are trying to find this information out means that you want it for some reason. And I think, yes, and that's another cost which I hadn't mentioned which is that defining these purposes and limiting access based on them means prohibiting - possibly creative and possibly quite useful uses of the information.

You know, I can easily imagine some researcher wanting to do some data analysis for whatever reason and all the sudden not being able to do that because they have to try to delve into the acceptable uses policies at, you know, 50 different registries. So basically the long answer is no, I don't think that makes sense.

Chuck Gomes: But you would agree that the cost probably would be less in that scenario that I described?

Shane Kerr: Absolutely. Absolutely. If it was just sort of something where you said, hey, make sure that you have some reason for this because it, yes, then that would be lighter weight, yes.

Chuck Gomes: Yes, okay. Thanks a lot. Excellent sharing, much appreciated. Michele, you're next.

Shane Kerr: Sure.

Michele Neylon: Thanks. Michele for the record. I'm as equally kind of perplexed by this as I think Shane is. The amount of data that's available from - in a thin registry is negligible. I mean, you have the domain name, the create date, the expiry,

the name servers, that's about it. I mean, there's also the EPP status, I mean, but that's it. There's nothing in there.

So I don't understand the concern and I agree, I mean, in terms of starting to add in extra layers of access controls for something like that just doesn't make any sense to me whatsoever. Thanks.

Chuck Gomes: Thank you, Michele. Greg, you're next. Are you on mute, Greg? We're not hearing anything. Okay, not sure okay, okay so he's - Greg's going to dial in. So let me open it up to those who have a different point of view on C. And then, Greg, just let me know when you're in and we'll get you back in the queue.

So would anybody like to explain why you disagree with any of the positions shared by those two that have talked in support of C? Stephanie, go ahead.

Stephanie Perrin: Thanks very much. Stephanie Perrin for the record. Can you hear me?

Chuck Gomes: Yes, we can. It all seems to be working, Stephanie.

Stephanie Perrin: Wonderful. Pardon my suspicion.

Chuck Gomes: I wonder why?

Stephanie Perrin: Yes, I just think it is probably a good idea as a general principle that we not give anonymous access. I mean, I can be persuaded that thin data is not personally identifiable. I still don't see even for research purposes, or general curiosity or hobbyist or whatever, why anybody should have access to this data anonymously.

So possibly we looked after that particular strand of enquiry separately on the poll. In other words, you still identify who you are, and authenticating, but you don't have to state your purpose because the general purpose is assumed

that it is for all of the above, general curiously, statistical, you work in the business, blah, blah. Thanks.

Chuck Gomes: Oh, my mistake. I was on mute all that time I was talking because I was - I got an echo when Stephanie was talking and I forgot to turn it off. My apologies. So, Greg, are you back in the queue? Are you on audio now? Okay, then let's go to Shane, and then we'll go to Lisa.

Shane Kerr: So I guess I'd like to respond to Stephanie's question or comment about...

Chuck Gomes: Sure.

Shane Kerr: ...why we would allow any anonymous access. In my mind, requiring authentication actually takes a step in one direction here because then you are creating a vast source of personally identifiable information that someone can mine or restrict. So what it means it's just to be some authenticating organization, there has to be some which - we know we're going to need in the long run with the vision that I understand for the RDS in that in order to access more sensitive information we want some sort of authentication.

But for - requiring it for all information, yes, it doesn't mean there's a huge burden on the organizations involved to track and somehow, I mean, you have to create privacy policies for all of this - all this access, right? So once you do that then there's privacy policy for the sort of very broad general access which I admit there's always some sort of identification possible, because you know the IP address, the person asking for the information is coming from. But that's quite easy to proxy away and obfuscate.

Whereas, if you require authentication, then basically you know who's looking at what and when and that's exactly the kind of information that it would prefer not be available to anyone at all, for example, I don't want my competition to know if I'm doing trademark searches. And I don't want the

government to know if I'm, you know, looking for likely-sounding names for, you know, privacy organizations or things like that.

So I think the idea of requiring authentication, while it seems like it protects people and protects the domain holders, is actually a mistake in that it creates a new problem area. That's it.

Chuck Gomes: Thanks, Shane. Lisa, your turn.

Lisa Phifer: Thanks, Chuck. Lisa Phifer for the record. I see that we're talking about access and access policies. And that was actually one of the findings, for me, of this poll was that when we posed this question we really were trying actually to separate the question of purpose and should purpose play a role in policy from the question of exactly the criteria would be for access, that is, you know, would access be allowed anonymous?

Would it require authentication? Would it require identification of the requestor or would it require a disclosure of purpose? All those are really good questions but they're actually questions that we anticipated would come up when we looked at the charter question on gated access, which poses those kinds of sub questions.

What we were trying to do initially with this poll was actually focus in on purpose and whether purpose should play a role and therefore whether we should look more at specific purposes. When Greg suggested his alternative, which is listed here as E, it was in part because he made an assumption that we were really looking to drill in on how purpose would play a role in accessing data.

And so there was a refinement basically of what you see here as Option A, a refinement of allowing access for any purposes except for illegitimate purposes, further refining that to state that users would be allowed to access

data anonymously and would not be required to declare their purpose and then, again, with illegitimate uses expressly prohibited.

So if you actually look at A and E as sort of variants of the same thing, of purpose being used to identify illegitimate uses and to somehow deter those illegitimate uses, that number actually is roughly on par with Option B, which is the opposite, looking at purpose as a way of allowing only those purposes that are expressly permitted by policy.

So we do have - the poll did sort of uncover that those diametrically opposing views but it did start going in the direction of access. And I guess the question that I would have for working group members is, is it really possible for us to deliberate on the role of purpose without delving into how that would play a role in access requirements?

And a couple of the comments that we got in response to Question 2 suggested that we really needed to be thinking about the purpose of collection before we could start thinking about the purpose for access and how access would be accomplished. So I'd love to hear working group members' thoughts on whether we could actually look at purpose without thinking about access yet even though we'll get there.

Chuck Gomes: Thanks, Lisa. This is Chuck again. Before I go back to Greg, if he has audio, it appears that Stephanie's problems with Adobe have been transferred to Greg so sorry about that, Greg. I'm glad it's now working for you, Stephanie. Anyway, before I do that though I want to come back to Michele. Michele, because I know you were on the Expert Working Group. And the Expert Working Group, as I understand it, suggested that there should be a purpose associated with data elements and the things we do with them.

But I understood you to be taking a different position with regard to thin data, and that's okay. But I thought maybe you could explain a little bit why - first of

all, did you agree with the Expert Working Group in the broader set of RDS data? And why do you make an exception for thin data?

Michele Neylon: Thanks, Chuck. Michele for the record. I agreed with the - I didn't descend from the overall findings of the EWG report. But I suppose the thing is in the - in some respects you're talking about kind of a slightly finessed view of things so that the thin data available in gTLD registries is useful, important, etcetera, etcetera, for a bunch of technical, operational reasons, practical, simple, straightforward reasons, which do not include divulging any personally identifiable information related to any registrant.

I mean, Lisa gave a list of the - well she gave a little snippet of text, which kind of - which covers pretty much everything that's available from thin. And I mean, sure, if you really, really, really, really want to, you can potentially find some kind of crazy nefarious usage for that data. But it's, you know, it does take a bit much.

I mean, the - at a very high level, I mean, the idea that you would - you've got to be logging IP addresses. You can put in rate limits on the number of queries so that somebody isn't trying to hit the dotCom zone, for example, and try and - and try to pull back every single domains, data, over the space of an hour and a half. I mean, if we were to try and do that I'm sure VeriSign systems would block you. And that makes perfect sense.

So, I mean, you know, the amount of data, because there's so little in it, I mean, the amount of, you know, you can hold it, put it into a kind of a general purpose of I want to know something about this domain. I mean, I want to know - I want to know something generally about this domain without being that specific. Which, you know, if it fits within the entire kind of framework of the EWG's, you know, the purposes and all that kind of stuff. So because there's so little in that bit, in that data set that I find it difficult to try and come up with a deep meaningful set of reasons and purposes that I would need to provide to be able to give somebody access to that.

And conversely speaking, I can't think of any reason why I would deny anybody access to that information, because it's for practical purposes of, you know, having the Internet function. A domain name exists, which registry is it associated with? That information is not - it shouldn't be hidden. Which registrar is it associated with? Again, why should that be hidden? The name servers - you need to have them - if there are name servers set to the domain, in order for the damn thing to resolve on the Internet, you cannot hide that information.

So I just don't understand why anybody would, you know, want to try and make something simple complicated. I really, really don't. And apart from a kind of black helicopter type scenario, but even the black helicopters aren't going to help you because really the amount of information is so miniscule, you know, I just don't - I just can't see that, I'm sorry, Chuck, I just can't.

Chuck Gomes: No need to apologize. This is Chuck. But just one last thing, and we need to keep it short and get to the other people that are in the queue. But you really are - when you say you're supporting C, you're really talking about access. What I hear you saying, there are purposes, legitimate purposes for thin data. You're not suggesting that there aren't purposes. You're...

((Crosstalk))

Chuck Gomes: ...just saying that they should have access without declaring those.

Michele Neylon: Yes.

((Crosstalk))

Chuck Gomes: I just wanted to clarify that. Make sure that we all get that. Okay?

Michele Neylon: But that doesn't mean that - I think part of this is this thing about, you know, fully anonymous, blah, blah, blah. I mean, in general terms, you know, even if you're using proxies and everything else, I mean, you're still going to log - you're still going to have a log somewhere saying a particular IP address sent a certain number of queries to the VeriSign servers. So if those IP addresses are, you know, VPN endpoints, tor exit nodes or whatever, you'll still have something somewhere just in terms of rate limits and stuff like that.

But, no, exactly. I mean, sure, there are plenty of legitimate purposes. And I'm sure you can think of some nefarious purposes too. But the amount of data involved, I mean, as Shane put it, it's, no, sorry, I just can't.

Chuck Gomes: Okay.

Michele Neylon: Can't think of any reason to make it more complicated.

Chuck Gomes: Thanks, Michele. This is Chuck again. Greg, do you have audio now?

Greg Aaron: Chuck, can you hear me?

Chuck Gomes: Oh, I can. You sound great. Go ahead.

Greg Aaron: Okay, thank you. This is Greg Aaron for the record. I agree with what Michele said and I agree with what Shane said. And I think that the language used in the poll questions, I wonder if we're actually closer in some ways than is reflected in the poll results.

Let me explain. I don't think any of us are interested in having people use the data for illegal purposes. We're all against that, right? I didn't find A-D satisfactory, and I wrote E because E describes the current situation that we have to do as right now.

In that people can access the data anonymously, they don't have to declare their purpose. And there is a terms of service. And by that you say when you access this data you're not going to use it for certain prohibited purposes like spamming. That's one of - that's listed.

What that - that means I couldn't agree with - because that didn't...

Chuck Gomes: You couldn't agree with what? I missed something there. You couldn't agree with what?

Greg Aaron: So I could not agree with Option C.

Chuck Gomes: Oh okay thank you.

Greg Aaron: Shane was saying he liked C for the same reasons I like E. We're probably sympathetic with each other. The problem with C was that it didn't have anything about please don't use the data (unintelligible). But it otherwise I like C. I didn't agree with Lisa. I think I'm - E is closer to C than A. The problem with A is that it says the data - you can only see the data - you can only access it except for illegitimate purposes. And what that means is you have a gate. Somehow you're going to figure out which users are only going to use it for legitimate purposes. And doing that is basically impossible or very - for the reasons that Shane described.

The use of the word "accessible" in A, B and C does imply that there's a condition and that means it's sort of an access control. And that's why (unintelligible) write E. And I don't think that the four original choices even describe the current situation and so I don't think the group had a full range of things to choose from. And I'm wondering if people (unintelligible) if we also had under the current situation. Thank you.

Chuck Gomes: Greg, this is Chuck. And I think most of what you said was ascertainable but every - a word is dropping every so often so if we later make some

conclusion that maybe because of that, but I think most of it came through. Not sure why that's happening. But thanks for sharing.

And by the way, I'll let you know that in our leadership call on Monday we talked about the very thing you mentioned at the beginning of your comments and that is is that the wording caused some problems. And we'll get to that a little bit when we move on in the agenda today to try and deal with that a little bit, but we certainly think you're right on that.

Maxim, you're up.

Maxim Alzoba: Alzoba for the record. Actually we have live example (unintelligible) in which some technical data is stored, I mean, CZDS where lots of registrars, registries, and (unintelligible) parties have access. And guess what are the usual (unintelligible) information the user is supplied via this system, the majority is a personal project. Most probably they do something. And we cannot then justify is it legal or not. The second, I mean, in numbers is scientific research.

And the third is cybercrime investigation. It's majority. So after the first few months of the work of the system, we can pull zone files for new TLDs it became empty exercise. You know, people know what to print into these (frames) in the form to gain access for whatever purpose they want. So they just - the work was shared and anybody can have access and if we try to prevent you from doing so should we enforce access through the compliance, for example.

And, I mean, if just the example of current system which demands the purpose of information use to be identified, it doesn't work. Thanks.

Chuck Gomes: Thank you, Maxim. Let's go to Stephanie.

Stephanie Perrin: Thank you. Stephanie Perrin for the record. I understand this basic thing that everyone is saying that nobody can really do much harm here and when they do start doing harm the registry like then will block them. What does it actually cost to know who is doing this? Is it prohibitive at this stage? Obviously if VeriSign can block them they have access to data about where the guys are coming from. Is there a way you can game that?

I'd just like to understand a little bit more to make up for my lack of technical know-how as to how difficult it is to know who they are, because as far as I'm concerned what we're really talking about here is - and I don't want to ruin the angle at which you're trying to tease apart the data, I understand what you're trying to do here, but we're talking about a system of mutual accountability which is what transparency versus privacy is. And in this instance, we don't see enough risk in making the data available.

But if there is misuse, what would be the cost of knowing who was coming at that data? Thanks. I hope I've expressed that in some sort of logical frame.

Chuck Gomes: I think you did, Stephanie. Thanks. This is Chuck. Let's go to Rod.

Rod Rasmussen: Sorry, it took a while to unmute there. So let's see, why did I raise my hand? So a couple of things. First is - and this is Rod Rasmussen, by the way. The first is that a couple of the elements in the thin data are just parts of the DNS so if the domain is actually active in the DNS you have anonymous access to millions of name servers from around the world all the time, recursive name serves, they're looking things up.

So the domain name itself and the name servers, that part of thin data is just you can't regulate it, right? You just can't. It's part of - if you had to require a purpose you'd have to basically change the way DNS works in order to limit access, right? So I don't see that happening so it's basically a fact of life that two of those data elements are basically going to be anonymously accessed regardless of whatever you want to intend to do with thin data.

So given that, there are already two elements that are completely accessible via anonymous distributed around the world access. You're really talking about, okay, where are we going to set policy with respect to other data elements? So keep that in mind.

The second thing I wanted to bring up and just talk about or refer to is the way we got to where we're at with this particular thing and coming out of the EWG. Remember the EWG report was a complete - everything's kind of holistically put together at the end so the thought around requiring purpose was based on the fact that we come up with a system where you would be stating purpose as part of the access anyway.

So there is this kind of circular logic to it, right, where you've got a gate kind of assumed so purpose is built in. Now that said, one of the reasons for requiring purpose was, in fact, to try to make things work under data protection laws and the like that we were briefed on as being requirements going forward. So regardless of, you know, whether or not, you know, a particular purpose would be allowed or not you still have to declare your purpose as part of accessing any sort of data, at least that's what I remember - recollect from being taught by Stephanie and others in the group and experts that were brought in.

It was this is just part of having to deal with the reality of the future, which is if you're asking for data that is collected in some process you have to state your purpose for asking for the data.

You know, so I'll leave it to the data protection people to tell us whether or not for certain elements like this whether or not that's really truly something we have to worry about or not but it's, you know, from a technical perspective there's no need for it from a - it's more of a policy legal perspective that's sitting there. So that's my - those are my - I think what I raised my hand about

about half an hour ago so but it's good that we have a lot of people chiming in. So it's good to see. Thanks.

Chuck Gomes: Thank you, Rod. This is Chuck. And I'm going to go to Shane and then I'm going to let Lisa continue with any analysis that she still wants to cover before we go to the next - to move in a fairly specific direction and get some feedback from the whole working group. So, Shane, did you - oh there you are. Okay, yes, I thought maybe - okay go ahead, Shane.

Shane Kerr: Sorry.

Chuck Gomes: Go ahead.

Shane Kerr: Well actually, I don't know, is it worthwhile for me to try to address Stephanie's question about...

Chuck Gomes: Sure, go ahead.

Shane Kerr: ...complexities in - okay. So basically the answer to the question is how hard is it to identify someone to authenticate them is it's really, really, really hard, there really is no way to do it if you're looking at the Internet as a whole. If you want to try to - with a system that is open for everyone on the earth to use, and at the same time can identify a human being or organization or something like that that actually is using the system then there really is no way to do it right now.

And as it actually terrifies me that there might be a way to do it at some point, but certainly for now it's not. Just because how do you establish a relationship with someone? Certainly if they're a customer of yours, that's relatively straightforward because you have some sort of contract or other business arrangement, right? You have someone's credit card number or you have someone's payment details at the very least.

But in the generic case, that information is just not available. You can come up with proxies for that. I think someone mentioned pseudo anonymous sort of identification. You can ask for email addresses where people have to have access to an email account and things like that. But that's really quite a weak form of authentication. And in the end doesn't buy you anything since these things can be obtained anonymously requiring that you authenticate to some anonymous authentication doesn't really - it kind of pushes the problem out but doesn't really address it.

So unless you want to limit access to people who have a contractual or other business relationship, I don't know that there is any way to get identification or authentication on the Internet today. That's it.

Chuck Gomes: Thanks, Shane. Chuck again. And let me turn it back to Lisa to see if she wants to cover anything else on her analysis.

Lisa Phifer: Thanks, Chuck. I think some of the points that we discussed in our analysis have already come up during conversation, but one of the observations the I had that we haven't really delved into is trying to think about - in the set of answers that we got kind of showed some differences of opinion. We've heard a fair bit now from people who supported Option C or E or maybe A, but we haven't heard a lot from the 41% or so that supported Option B which felt that legitimate purposes should be explicitly permitted.

So we definitely have an opposition of views that might be hard to resolve. But it seems like one of the things that we could take away from the poll is that if people do disagree when thinking about thin data as a sort of big clump and all purposes as a big nebulous concept, that if we tease apart those thin data elements and look at individual thin data elements we might find that there's more agreement on specific thin data elements. Rod mentioned a couple of them are already available anonymously regardless.

We might also find more agreement on a few specific and basic purposes. And so if we go down the path of deliberating on individual thin data elements and perhaps on the specific purposes for those individual thin data elements, we might actually find that we have some agreement looking at that in more granularity.

The other observation that I wanted to share is that I think one of the things we stumble over sometimes is making assumptions about possible implementations when we try to visualize what's the implication of a policy requirement. For example, we've been talking about authentication or anonymous access as though they were mutually exclusive.

But I think most of us have gone to a captive portal to gain access to a network or to the Internet and sometimes those captive portals have both an option to click through, view terms of service and get anonymous access to the Internet, as well as an option to fill in your user ID, authenticate yourself and then get access to additional resources.

So it is possible to have an implementation which allows more than one methodology. And we just need to keep in mind that when we're thinking about things and assuming implementations, that really policy should drive the implementation even though of course it's important to think about whether the implementation is practical.

Chuck Gomes: Thank you, Lisa. Any comments on that? Okay so I'm going to try and narrow - now we're still essentially focusing on Question 2.1 and you can scroll up if you want to relook at that, but we're just looking at a portion of Question 2.1 because as we found in this survey, or this poll, you - it delves off into several critical points where we're going to have to deliberate eventually. But if we try to do it all right now we're not going to get very far at this stage.

So am I correct in concluding - let me phrase it differently. Is there anyone on the call who thinks that we shouldn't be concerned with purposes as far as

thin data is concerned at all? And I'm not saying that we should require people to identify purposes to access data, all I'm saying is, you know, the EWG report basically suggested I think, putting it in my own words, probably not very well, but that there should be a purpose for any data in the RDS.

Whether we allow access to it or who we allow access to it or whether we require authentication or anything else, is there anybody that thinks that that purpose isn't an issue to consider when we're considering what datas would go into an RDS? Would you raise your hand if you think that we shouldn't consider purpose at all?

Okay, so I can conclude, at least for those on the call, that we do have consensus, the 22 of us, that purpose is a legitimate issue for us to consider. So then with that in mind, then what is the purpose - and Lisa led into this in her last comments - let's look at the thin data elements and what some purposes are for collecting them.

Now right now, and you can tell - you can see right through me I'm sure, I'm trying to keep it very narrow so that we keep our focus and can make some progress that way. So the question I'm asking for each of the thin data elements, and we'll look at those individually, is what is the purpose of collecting each thin data element? I'm not talking about the purpose of accessing, okay, I'm talking about the purpose of collecting each thin data element.

And we can look at the - so the first question I'm going to ask, and just - and you can respond in the chat or you can raise your hand, is what is the purpose of collecting the domain name sponsoring registrar? Michele.

Michele Neylon: Thanks, Chuck. Just before I answer that specific question, apart from the name servers, which was something that the registrant would choose to actually set, I mean, a lot of the data that goes into thin isn't something that's - collected so much as created. I mean, if you register a domain name, you

know, by function of the fact that you register it you create it, which means you have the create date and various other elements.

But when it comes to - but specifically to your thing about the registry, I mean, obviously VeriSign, and I keep on picking on VeriSign, doesn't want to field questions for every single domain name that's registered, that's why registrars exist. So providing the registrar of record means that if somebody has an issue, be that - whoever that may be they know who all of the ports of call would be.

Chuck Gomes: Thank you. This is Chuck again. Anybody else want to share a purpose for collecting the domain name sponsoring registrar? So okay, Greg.

Greg Aaron: It tells the registrant who's responsible for their domain name. And it's required by ICANN policy to help facilitate registrar to registrar transfers.

Chuck Gomes: Now, Greg, I'm going to discard your second reason. That doesn't sound very kind does it? But because keep in mind, we can change policy. We can come up with recommendations that change the policy. So you're right, in the current system is fulfills a contractual requirement. But we could change that contractual requirement because if we recommend policies that are ultimately adopted by the Board, then the policies will change in Registry and Registrar Agreements. But your point is well taken, okay so that's good.

                 Michele.

Michele Neylon: Just to disagree with Greg, since that's always fun, especially when he's in a weakened state and so am I. The assumption that there is a direct relationship between a registrant and a registrar is a dangerous assumption to make. A very large number of domain names are registered via what are called in the ICANN system, resellers, but, I mean, for practical purposes when you're talking to small businesses it's their web designer, web developer, IT services company or the quote unquote, web guy.

A lot of our clients wouldn't have a clue who we are and wouldn't realize that there - that we are the registrar. They know that they've been - that they paid Company X for a Website, for web design or to have email set up in their office or something like that. Thanks.

Chuck Gomes: But, Michele, would you agree with me - this is Chuck - would you agree with me that even though the registrant may not know who you are from - an ICANN's perspective, from a contractual perspective, they essentially do have a relationship with you at least if for no other reason from a contractual point of view.

Michele Neylon: I refer to your previous comments about policy. I mean, the fact that the policies and the contracts at the moment only speak of - only speak of the registrar registrant relationship, is actually kind of against reality. The reality is that yes, sure, officially contractually, whatever, there is a relationship, direct or indirect, and all that. But for practical purposes, saying that you need to go - that you as the registrant of Domain X need to go and talk to registrar whoever, the registrar often won't even know who you are because well, you know, they only know - they know of the reseller or it could even be a case that the Whois information doesn't even contain their details.

Chuck Gomes: Point taken. Okay. Good. Let me go to Stephanie. This is Chuck again.

Stephanie Perrin: Thanks very much. Stephanie Perrin for the record. And I think actually I'm saying in my more clumsy way what Michele basically just said. In terms of what you were describing as the reasons to have the registrar name there, we seem to be straying into availability or display. If the name is generated by whoever puts their hands on the actual registration and assigns the domain, i.e. is the sponsoring registrar, you don't have to put Blacknight out there, you could put a number.

And only certain parties could have access to the key to those numbers which would not of course be an idiot-proof code but it would - it might actually tell you more, as Michele has just explained because you could have first and second and third level resellers and contact points. So you would know whether you're dealing with Blacknight who actually did it all the way through or Blacknight who was acting as the initial registrar and who handed it off to resellers.

I hope you follow my points. The point is that the purposes that were stated as for why that data element should be there, are - we're talking about display, not collection. Thanks.

Chuck Gomes: That's correct. And right now we're talking about collecting. So that's why we need to - please keep our focus narrowed onto collecting. We will get to, probably not in this meeting, we will get to display and access - we have to get to those but for right now we're talking about collecting. And let's go to - what is the purpose of collecting the domain name registration statuses?

In fact, let me see if I can speed this up a little bit because maybe it's not necessary to go through one by one. If it is, we will. And this is Chuck speaking. Is there anybody that thinks there's not a purpose for collecting any of the thin data elements? Shane?

Shane Kerr: Hi, this is Shane...

Chuck Gomes: I realize you raised your hand before I asked that question. So if you were going to say something else, go ahead.

Shane Kerr: No, I think it's related. I think looking at this data we can kind of - we can kind of put them into two or three different categories instead of going through each by one. In the case of the URL and the sponsoring registrar, those are both kind of referrals to more authoritative information, right? So this is how you get in touch with the people who may know more.

And then the other type of data is the one which is strictly required operationally which is the name servers. And you can argue that it's not required operationally because it's in the DNS, but it's kind of a check for that, right?

All of the other types of data I would argue are not strictly necessary but - or sort of helpful hints. Certainly the date information can provide, as meta data, certain kinds of information. You know, you can see if a domain is quite new you might want to question it and if it's quite old you may give it some more - may respect it more and trust it a little more and these kind of things.

So I think those three categories are probably how we can view all of it. And I think just by describing what they are we can say that they all have a purpose, at least one. So I think hopefully that answers your question.

Chuck Gomes: It does. Thank you, Shane. This is Chuck. Mark, you're next.

Marc Anderson: Thank you, Chuck. This is Marc. You know, on this one I wanted to jump in and, you know, agree with something Michele said initially is that with the exception of the name server information, none of this information is really collected. It really is created. You know, when the registrar connects to a registry and creates a domain, you know, adds a domain to a registry, you know, the registrar Whois server, referral URL, that information is automatically populated. You know, that information, you know, exists by virtue of the registrar that created the domain name.

Likewise, the updated creation expiration date, you know, that is all, you know, essentially system-generated data; it's never actually collected. You know, it really is...

((Crosstalk))

Chuck Gomes: So, Marc, let me...

Marc Anderson: ...that's being collected in any way.

Chuck Gomes: This is Chuck, Marc. Let me look at it from a different perspective. You're right, from the registry or registrar or maybe even the registrant's perspective that some of these things are created. But from the perspective of an RDS system, we're talking about collecting that data for that RDS system. So I don't think we're disagreeing, it's just the use of the word "collect" is from the perspective of an RDS system. Does that make sense?

Marc Anderson: You know, it does. But I think at that point you're crossing over from collecting to, you know, the purpose of displaying the data or providing the data.

Chuck Gomes: Well, we're not getting to display yet so...

Marc Anderson: Right.

Chuck Gomes: So bear with us, we'll get there. But your point is well taken. But I think collecting is okay, I'm willing to be corrected on that if you're looking at it from the perspective of an RDS. So Greg, go ahead.

Greg Aaron: Thank you, Chuck. This is Greg. The Expert Working Group's report has a really handy table. It's a table of data elements, the data fields at you see in Whois, for example. And then there are matching purposes, the purposes for which they're collected and the purposes that people use them for. And...

Chuck Gomes: It's interesting that you mention that, Greg. We lost our ability to scroll down. Can we scroll down? I think this is a subset of the table you're talking about.

Greg Aaron: Yes, there we go. That's great. I would suggest we use this list often because it's a nice piece of work. And might save us a little bit of time. And it's also useful for us to see it in front of us.

Chuck Gomes: And as you can tell, we were planning on using this because we also think it's a nice list. It doesn't mean that this is exact, that it - it covers everything, but it is - I think it was a good piece of work. And it's there. And so to try and get back to the last question I asked then, if we scroll back up to the - and well, yes, since I don't have scroll capability, scroll back up to the questions that we were just looking at and then we'll let people - if we can give people the scroll capability themselves. Good, all right, so I can do it myself.

But if you scroll back up and look at those questions, what is the purpose of collecting domain name, sponsoring registrar, statuses, creation date, expiration date, name servers, last updated date, and the Whois URL, okay? I'm getting the sense, and I want to test this right now, and we'll also test it in a poll especially for those who aren't on this call, but everyone will be welcome to participate.

Are we in agreement that there is at least one purpose, some purposes may be more critical than others like couple of people have pointed out, but there's at least one purpose for every one of these thin data elements. Is there any disagreement with that?

Okay, I see some agreements. I don't see any red Xs. Appreciate the green checkmarks. But so I think we have another point of agreement, at least for 22 of us, okay, and if we add Lisa into that, and I like to include Lisa as well. We could probably say 23. So good. All right so we made that much progress.

Now, again, referring to the table from the EWG report on the purposes, let's just spend a few moments looking at those. So if you would just scroll through them yourselves, is there anyone that disagrees with any of those purposes for any of the thin elements? We're not saying that that is necessarily - those are the only purposes. But for right now, is there any disagreement with the

work of the EWG on those purposes or any comments you want to make on those?

Okay, Shane.

Shane Kerr: Just a quick question about registrar and registrar URL. It seems to have, in that table there fewer purposes than the registration status and client status and server status. Where actually I think the registrar URL are probably the - one of the most broadly useful pieces of information because certainly those can also be used for criminal investigation and all these other things that are in the part about it. Or am I misreading the table?

Chuck Gomes: So you think - are you suggesting, Shane, that they maybe could have emphasized registrar and registrar URL more than shows in this table. Is that what you're saying?

Shane Kerr: Yes, but again maybe I'm just misreading the table, I don't know.

Chuck Gomes: That's okay, I don't know that the - I can't speak for the EWG because I wasn't on it, whether they - the number of lines for each one is significant in terms of importance, I suspect they didn't intend that. But if any of them want to comment on it I can. But it's a good point. Okay, Marc. Are you on mute, Marc?

Marc Anderson: Yes, sorry, I was on mute there. It took me a second to get off. Marc Anderson for the record. I guess I'm wondering where you're going with this one, Chuck. You seem to have a destination in mind but, you know, I think, you know, we can all agree that you can come up with a purpose for almost any set so I don't think on principle I can't, you know, I can't disagree with a statement that, you know, we agree that everything here has a purpose, you know, I think it's certainly possible that all of these things have a purpose.

On the other hand, the Whois server, right, in the thick registry, the Whois server, you know, that really serves a purpose for a thin registry or is intending to serve a purpose for a thin registry. You know, I think in a thick registry, you know, that's largely a worthless piece of information.

Chuck Gomes: You don't think that there would need to be a Whois server for the registry?

Marc Anderson: You know, if you're getting that data you're there.

Chuck Gomes: Okay, maybe, yes, all right.

Marc Anderson: Yes, so, you know, I think, you know, I guess I'm wondering where you're going with this. I feel like you're headed towards a...

Chuck Gomes: Well let me answer that question, okay? Keep in mind that we're looking at Question 2.1, if you scroll up to the top. We're trying to answer Question 2.1, "Should gTLD registry data," and forget accessible for now, "be in the RDS for any purpose or for only for specific purposes?"

All we're doing right now is if we think that there should be purposes for anything in the RDS, then we - and then we may need to identify those purposes. In fact later questions that we're going to deliberate on will assume that we know those purposes. So that's where we're going. So bear with us if you don't see it now, I think it'll become clearer later on.

Lisa, you want to add anything on that? Or any of the cochairs? Okay, so where do we go from here? So I think that we - we've certainly come to a few simple very narrow conclusions that we agree on today, so those will be stepping stones for the next steps. We obviously have to talk about access and display as well. And we will.

The - so the next steps then, okay, so we've covered - we've really just covered part of - we've only focused on thin data right now. Okay so we've

narrowed that. And we've only focused on collection. So the - we seem to be, I think, in pretty good agreement on those very narrow focuses, thin data, and collection only for the RDS. So now we need to expand that, and I'm not - I don't think we have enough time to really get into that.

We've kind of got into a little bit it the last meeting and this meeting when people talked about access and people talk about display and people talk about anonymous and so forth, so we were obviously getting into those areas and those are critical areas.

So let me ask Lisa if we have enough to do a short survey. You're going to have more time on this survey, everybody, since we don't have a meeting for two weeks, or maybe it's more than two weeks. Anyway, the - but, Lisa, do you have enough to create a short survey? And a big part of this survey this time, because I think we're pretty clear, although you'll have a chance to change your mind on the poll, I shouldn't call it a survey, it's a poll.

And unless you just kept quiet and didn't share your opinion, I think we're all pretty much on the same page. Now for the sake of getting valid data, for those on the call and not on the call, I'd encourage all of you to respond to the survey. It should be really quick for you since you've already been through the exercise. But, Lisa, do you have enough to do a short survey? I think there's at least two conclusions we came to in this meeting that we need to put out to the whole working group in a poll. Are you okay with what we have?

Lisa Phifer: Yes, Chuck. Lisa Phifer for the record.

Chuck Gomes: I'm sorry, I'm looking now at the chat.

Lisa Phifer: Right. I think, just as you said, there were a couple of points of agreement, apparent agreement, during this call but what we could do with the poll is confirm that we were in agreement. What we saw on the last poll was

sometimes wording makes a big difference. So let's see if we can get it really clear. And also give a chance for those not on today's call to weigh in.

Chuck Gomes: Thank you. Okay just want to make sure that was the case. And we'll try and turn that survey around by end of the day tomorrow at the latest. Well, I guess tomorrow is different for you than it is for me. I'm still on Tuesday. Not very much longer but I'm still on Tuesday. So we will do that. Look for that survey or poll, sorry. It's late for me and I'm definitely not a night person. So and take - probably just take a minute or two for those of us on this call. It may take a little bit longer for those that weren't on the call.

But take a few minutes, respond. You'll have a lot longer to respond on this one since we don't have another meeting for - until the 10th of January. Please be aware of that, okay, the 10th of January, so the second Tuesday in January. Is there anything else we need to cover in this meeting, any questions or final comments?

Okay, that being said I want to wish you all a great couple weeks of a break from the RDS, other than the poll. Please do the poll. If you get the poll done this week, it's all done you don't have to worry about it, okay? So but have a great holiday. I hope everybody gets some time off. And I wish you a very happy New Year and look forward to our continuing to work together. And our progress may be slow so be patient.

But even though it may have been small steps tonight, we reached a couple points of simple consensus. We're going to get into tougher areas when we get into access and display, I understand that. But if we'll all continue to cooperate like we're doing now I think we'll be on the right track. That said, I will adjourn the meeting and the recording can stop.

END