EN

ICANN Transcription

IDNs EPDP

Friday, 25 February 2022 at 13:30 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/sIH3Cg

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

DEVAN REED: Good morning, good afternoon, and good evening. Welcome to the IDNs EPDP call taking place on Friday 25th, February 2022 at 13:30 UTC. In the interest of time, there will be no roll call. Attendance will be taken by the Zoom room. If you're only on the telephone, could you please let yourselves be known now?

We have apologies from Anil Kumar Jain and Maxim Alzoba. All members and participants will be promoted to panelists for today's call. Members and participants, when using chat please select everyone in order for everyone to see the chat. Observers will remain as attendee and will have view only chat access. Statements interest must be kept up to date. If anyone has any updates to share, please raise your hand or speak up down. If you need assistance updating your statements of interest, 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 documentation and information can be found on the IDNs EPDP 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 for the transcript. As a reminder, those who take part in the ICANN multi-stakeholder process are to comply with the expected standards of behavior. Thank you and over to our chair, Donna Austin. Please begin.

DONNA AUSTIN: Thanks Devan. Welcome, everybody, to our IDN EPDP call today. We're just having a little bit of a chat that attendance was light, possibly because we moved this call out by 24 hours. But it looks like people are coming in now, so that's great.

Just by way of updates I think we had a couple of requests to extend the date to provide comments on the language for A5 and A6. and I think we've granted that. So Ariel, I think we're now working to—I can't remember the date—about two weeks' time? Next Wednesday, okay. So thanks for asking the question. We'll always be as flexible as we can on these things. So once we get to next Wednesday and have a look at the comments, we'll make a decision on whether it's substantive or not and whether we need to come back and have some discussions around A5 and A6.

So for today, we're going to go back a little bit and have a look again at charter questions at A7 and A9 that we discussed a little while ago and staff are doing some work on those. So we will go back and have a look at where we are in those.

And then last week we were talking about D1B, I think we identified part two. And so we'll go back and continue discussions on those topics. So I think that's all I have by way of updates. I had a hand up from Jeff. Jeff, go ahead.

JEFF NEUMAN: Yeah, thanks, Donna. And can I make a request? And I know you guys want to keep the document as clean as possible, but can I request the ability to put comments directly into the document? The reason I haven't been able to get my comments in is because it just takes a lot more time to do a separate email … To draft an email, basically explaining which provision I'm talking about and the context for it. And then to put my comment in. It's just a lot more onerous to make the comments. And that's why I haven't gotten mine in at this point, because I know it's going to be a bunch more work when it would be much easier to put it into the draft. And again, I don't have to touch the language, that's not what I mean to do. But I would like to put comments in.

DONNA AUSTIN: I think access to comments is okay just so long as we don't get into a red line if that makes sense, Jeff.

JEFF NEUMAN: Yeah, totally. And I don't plan to do red lines. It's just, like I said, so much easier to put a comment in than to draft a long email saying, "All right, in this paragraph that starts with these words and does this, this is my comment."

DONNA AUSTIN: Yeah, I understand the comment. I guess the challenge that we have is that when you open up a document and you've got multiple people providing comment, then it gets difficult to understand what was in the text to start with. So it's a balance.

JEFF NEUMAN: Yeah, totally understand. And I promise to just put comments in. But, again, when you're when you're asking for comments on the language itself, it's much easier to do that.

DONNA AUSTIN: Yeah, understood. And I guess to be fair to everybody else, if they would—-I think Ariel and Steve are going to hit me over the head here, but we need to be fair across the board. So if folks would prefer to put comments on the document rather than an email, then I think we need to provide that access to all. Okay, so I guess Ariel will—or I'm not sure who has control, maybe it's Emily. But we'll allow access to the documents to make comments.

And also, Steve's made a good comment. To anyone who provides comments on the document, please provide an email to the list, so that folks are aware that that's been done.

Okay, with that, I'm going to hand over to Ariel, who's going to provide our recap and updates of the topics that we're going to talk about today. So, Ariel, I'll hand it over to you.

EN

ARIEL LIANG: Thanks, Donna. So we're going to talk about A7 first. As a reminder, A7 is the question regarding the single character TLDs. So the question has two parts. Part one is a mechanism for identifying the scripts and languages that can potentially allow single character TLDs. And part two is about the mechanism of criteria to identify a specific list of allowable characters within the script or language for a single character TLD as well as implementation guidance, if any.

So for recap, the EPDP team has reached agreement for part one of the question. So essentially, it affirms SubPro's recommendation that single character TLDs may be allowed for ideographic script and language combinations. And for now Han script is the only ideographic script included in the RZ-LGR. And Chinese, Japanese, and Korean languages are the only languages that incorporates the Han script. So with that background, the Chinese language and the Han script portion of Japanese and Korean language are appropriate for single character TLDs.

So these are the agreements from the working group who have heard and then we see, Jeff, that you have a suggestion in terms of how the recommendation should be drafted, so we'll take that into account and we'll develop the appropriate recommendation language for the EPDP team to review.

In terms of part two. This part we believe though EPDP team hasn't fully addressed it. There's some preliminary agreement to outsource the work for identifying allowable characters within Han script to the Chinese, Japanese, and Korean Generation Panels in consultation with the Integration Panel. But what we need is some

instructions so that these Generation Panels and the Integration Panel know exactly what work they are tasked to do. So that's why we like to discuss some proposal from staff and then the team can consider whether this proposal for instruction is appropriate. And if we're missing anything, we welcome your input and feedback for this proposal so that the request can be clearly conveyed to the generation panels and Integration Panel for their consideration.

So now I will talk about some of the staff suggestions. First, we believe that it may be more feasible or reasonable to ask the generation panel to develop an exclusion list rather than an inclusion list. So that means consider developing mechanism or criteria to disqualify Han characters from being applied for single character gTLDs rather than an inclusion list that include all allowable characters. As you know the Han script has thousands of characters, and it probably will be a daunting task to review every one of the characters and then provide a list which one is allowed, which on is not allowed. But ask them to specify what is not allowed may be a more reasonable request. So that's the first general point.

And then, as a detail, the second point staff would like to suggest is perhaps the three Generation Panels in consultation with the Integration Panel can develop a hard no list of the disqualifying Han characters. So this can be due to security stability implications. So some ideas could be if a character has only one stroke or two strokes, that may cause confusion with some other symbols. For example, there's one Chinese letter is one stroke and it looks like a parallel line, and that means one. And then there's another one stroke Chinese character that looks like zero

EN

or O, so something like that to develop a hard no list. And then the Generation Panel and Integration Panel can tag these code points in the RZ-LGR. So with that the outputs can be reviewed by the EPDP team and to decide whether this hard no list looks appropriate. So that's the first kind of sub suggestion under that overall umbrella.

And then the second sub suggestion is the EPDP team to provide clear instruction to the Generation Panels and Integration Panel to develop additional criteria for evaluating single character gTLD applications. So those additional criterias can be principles and recommendations and such. So these are to be used by the DNS stability panel, the string similarity review panel in their consideration for the single character gTLD applications.

And again, that output is still subject to the EPDP teams review, in particular because the Generation Panels and Integration Panel, they are not exactly trained for this job. And it's really for the EPDP team to develop consensus recommendation based on what they produce in terms of these additional criteria.

So some of these additional criterias, we have some kind of rough idea what they could be. So first of all, character is an ideograph, ideogram. That's the SubPro recommendation. So one thing I like to kind of emphasize here is that we know that the Han script is an ideographic graphics script, but I'm not sure whether all Han characters are ideographic characters, and that's something the generation panels will be qualified to answer that.

And I'm a native speaker for Chinese but, for example, I know some Chinese characters by itself don't make much sense, they

don't have exact meanings. They have to combine with other characters to have meanings. So would those characters be considered ideographic characters are not? I don't have an answer to that, but the linguistic experts and the GPS, they will have answers to that. So that's one example.

And then other criteria for the GPs consideration is that the characters do not introduce confusion risks that rise above commonplace similarities. So this is also the SubPro recommendation, and that's some criteria that the EPDP team can relay to the GPS and IP and then they can further deliberate on that and consider whether that's appropriate criteria to recommend.

And then some other ideas, for example, the character needs to demonstrate association or meaningfulness, need to be composed via generally available input method editors, so the target user can type it using your keyboard in a local language and follow orthographic conventions. In addition, there are some existing SSAC advice—it's SSAC052—that's specifically related to single character TLDs. So perhaps the Generation Panels and Integration Panel can take a look at that advice again and see whether any of these criteria is worth recommending.

So these are some kind of instruction that the EPDP team needs to consider to provide the GPs so they have a clear idea of the problem their need to solve and then develop a solution to it.

And then lastly, as a default, the single character gTLD applications should not be allowed until the list and criteria and mechanism become EPDP consensus recommendations and

adopted and implemented by ICANN as consensus policy. So just chew on the side of cautious and keep things conservative. That's a recommended path in terms of what to do with single character TLD applications. So these are some staff proposals for consideration and now I will stop here. I see, Jeff, he has his hand up.

DONNA AUSTIN: Jeff, just before you go ahead, sorry to interrupt. Devan, did you want to just talk to Hamza—sorry, if I get that wrong—to see if he can move them across?

DEVAN REED: Sure thing, I've already sent them a message.

DONNA AUSTIN: But if Hamza is listening now, if you could just restate what you put in the chat?

DEVAN REED: Oh, absolutely. In order to participate in the discussion, Hamza, I'm going to send you a pop up in Zoom. If you can accept the promote to panelist button, that way you can participate in chat.

DONNA AUSTIN: Thanks, Devan. Sorry, for that, Jeff. Go ahead.

EN

JEFF NEUMAN:
Yeah, no problem. Let me say what I agree with. I agree with the fact that we should not be outsourcing this work at this point in time to come up with some sort of allowable list. I think that's absolutely right. I think it's just not worth the work. But other than that, I just don't see why we're being so prescriptive, right? There were technical reasons as to why we did not allow single character ASCII top level domains. And there were some security issues with that and that's why single character ASCII names are not allowed.

But those, those reasons don't exist for single character IDN names, right? So the reason why SubPro allowed single character IDN names is because they are all XN dash dash something. So they're really, in the ASCII, they're really at least five characters, if not a lot more. So with all that being said, I don't think we need any of this except for just saying that it's going to follow the same string similarity reviews that everything else follows in the new gTLD process, at least for new gTLDs. And I don't think we need anything. I don't think it matters whether it really mean something or not.

I mean, I think that the point for excluding single character TLDs, again, was a technical reason and a security reason within ASCII of having one character. But those don't exist when you have IDN single character TLDs. And therefore, why does it matter that it actually means something? Why does it matter that we make them go through all these hurdles? So I would say that the only thing it needs to be is in the third bullet, the first requirement where it says it's an ideograph or ideogram and that's it. I mean everything else just goes through the normal process.

Now, for existing TLDs I guess we can maybe list some of the criteria that we have for new gTLDs with string similarity and everything else. And then the other thing is, I'm not sure what SSAC52 actually adds to this. I mean, SSAC52 is mostly about ensuring that they meant something and that they were valid code points and things, which I think is the same for any kind of IDN TLD.

So again, I'd like us to be as flexible as possible without being so prescriptive as to what is … We don't need to develop a "cannot apply for" list. I think that's too much work and not necessary. Thanks.

DONNA AUSTIN: Thanks, Jeff. And I'm interested to hear from others whether you share the same thoughts as Jeff on this. But the reason for the staff proposal is the last time we discussed this we agreed that we didn't have expertise within the working group to work out whether … I think one of the thing that's identified in the charter question is about a list of allowable characters. What we're suggesting here is maybe identify a list of disqualifies, so we go the opposite way. So some of this is because it's within the charter question. And also, because there was agreement that we didn't have the expertise within this working group and we should outsource this. And I think Sarmad might have said that this is something that the Generation Panels could possibly do.

So, interested to hear from others whether they feel the same as Jeff. Actually, it's really important that we hear from others at this

ARIEL LIANG:

Thanks, Donna. And I just wanted to quickly touch on SSAC052. This slide doesn't include every single recommendation or guidance from SSAC02, but highlights some points I believe may be helpful for the team to know what is being recommended. So, SSAC052 recommends a very conservative approach to the delegation of single character IDN TLDs. And basically, in terms of the confusing similarity point SSAC052 is a warning about single character TLDs have an even higher risk to cause user word confusion. And one specific point they mention is that those characters need to be carefully considered, not just within the script itself, but also across scripts.

So, if we want to put in the current context, if we do allow Han script for single character TLD, that seems to be that the characters need to be evaluated not just within Han script, but also other scripts that might be relevant. I am not a linguistic expert, I cannot say which ones will be relevant or impacted, but I do know like, for example, some single stroke Chinese characters may be confusable if you compare to some other, like katakana in Japanese or some other scripts, or even numbers, that type of thing.

So basically, 052 provides additional detail beyond what the standard should be. It's basically honing the fact that single character TLDs could cause a lot more confusability compared to other types. So they need to be examined very carefully. But I'll

EN

stop here now and probably someone can explain this better than me.

DONNA AUSTIN: Okay, so we've got a queue. So we have Jeff, Michael, Dennis, and Hadia.

JEFF NEUMAN: Yeah, thanks. Again, understand that some characters may be confusing with other characters in other scripts. But when we test confusability for new gTLDs, which this would be, we only test confusability against delegated strings or other applications. I don't know why we would ever do a confusability test against an undelegated string. So the fact that one character in the Han script looks like another character in a Han, or even non Han script, is irrelevant unless it's a delegated string or another application.

So that's part of the problem I had and SubPro had with this SSAC recommendation, was that very fact that it was way too conservative and way too prescriptive and treating IDNs differently than ASCII TLDs. We don't take an ASCII three character TLD, for example, and measure it against confusability with an undelegated string that may look like a three character in another script. We don't do that unless the three characters have been delegated already or it's another application. So again, I don't understand why we would do anything differently. I didn't understand it when the SSAC52 came out, and neither did the rest of SubPro for that matter. Thanks.

DONNA AUSTIN:	Thanks Jeff. Michael, Dennis, Hadia, and Abdulkarim. Go ahead, Michael.
MICHAEL BAULAND:	Thanks, Michael Bauland for the record. In general I'm not so sure why single character TLDs should be prohibited in the first place, even if it's just ASCII. But the rule is there so we have to live with it. Now, I don't really know much about the Han script, but I think the reason to have an exception here is because otherwise you wouldn't be able to register top level domains that are very common in ASCII like flower, or car, or house, or whatever because, please correct me if I'm wrong, all or many of these words are just a single character. So because we don't want to restrict those languages too strongly we should allow the single character TLDs.
	But at the same time, I think since it's restricted for all other scripts and languages, we should also have the restriction that these words should be real words and not just any Han character should be allowed as a TLD label. Thanks.
DENNIS TAN:	I guess I'm next. This is Dennis for the transcript. So first I just want to put in my own hat as the Chair of the charter drafting team. And looking at the questions here I just wanted to give a little bit of context as to how the charter drafting team getting into these questions. And in this case, right, and you need to understand that charter drafting team is tasked with coming up with the questions that the working group should discuss.

We didn't get into the substance of discussion that might as well help pony in or tweaking some of the questions. So we didn't have the time to get into substance discussion, we did discuss a little bit and trying to come up with the reasonable questions that the working group might want to discuss. But they're not to be prescriptive, they're meant to be directional. And in this case, for example, question about identifying specific characters does not seem reasonable based on the discussion, right? It's, I think, leaning towards a hard no type of criteria. Especially we're talking about Han script group with thousands of characters in it.

That said, I think I tend to agree with some of the conversation that's happening here. One thing, first of all, from a technical standpoint there should not be any issues, right? Single category IDNs on the wire, on the back end are at least five characters. Well, four if technically we just count the X N and hyphen, hyphen, right? Any single character after the [inaudible] transformation, they start with XN dash, dash and the [inaudible] output. So from a technical standpoint there is no issue.

But that remains the confusability issue, right? That's one of the items that SSAC052 discusses, and that is that single characters in general, single characters are most likely to introduce confusion than two or more characters in a string. And that's generally true because you lack the context. Fewer characters are in a label, you have less context. So I think that's reasonable to think that, for example, as Ariel put as an example, single stroke that can resemble a number or a letter in the Latin script, for example. That might introduce higher confusion or a likelihood to spoof a domain

name. In the case of numbers, that's one reason we don't have number in a TLD, because you could spoof an IP address.

Now, these have been said, right, so technical issues, none as far as I'm aware, again, because single character IDNs in the back end on the wire are not single characters, right? They at least have to start with the prefix, X and hyphen, hyphen. So it's more on the confusability aspect of it that might introduce security issues. And so how can we address that? I mean, today in the fast track process, because that's an ongoing and back in the day in the [inaudible] program 2012 and in the future, we envision a DNS stability planning review, which will introduce confusing similarity and such.

So would that be the right place to perhaps an enhanced process to look at confusing and similarity issues for single characters? That they could just test and do the analysis instead for us to come up with criteria that we don't have the expertise? And I don't think we can rely on Generational Panels because they are not standing panels. They are formed for a task, it's completed and they are disbanded.

So, again, just to sum up, no technical issues. It's more on the confusing and similarity issues, but the processes and our envision for future rounds may be able to handle what we're talking about here. And yes, and one last point, yeah, I do agree that the meaning lists, I think I tend to agree with, I don't know who made that remark. We don't ask for meaning lists in ASCII TLDs, why would you test meaning lists in single characters? Either way I just want to put that out. Thank you.

EN

DONNA AUSTIN: Thank Dennis. So just before we go to Hadia, just to be clear the reason we had suggested previously that it would be helpful to go the Generation Panels to have a look at this issue is not really to provide prescriptive rules around the script, but, I guess, just to review the script from an ideogram or a single letter perspective to see whether there's anything that that could be disqualified.

So it's not that the Generation Panel, we're asking them to really develop a set of prescriptive rules for a future round, it's just to help us understand if there is anything that could be done that might help in the chartering question. Which was once the script's languages are identified, what mechanisms or criteria should be used to identify a specific list of allowable characters which can be used as a single character TLD within such scripts and languages. So that's why we were going to go back to the Generation Panels, to see if they could help us out with that.

And I think Sarmad had indicated that that was a possibility, it wouldn't be too hard to get the Generation Panels back together, because it's only three generation panels as I understand it, to have a look at this on our behalf. So we're not throwing this over to the Generation Panels to develop the rules, it's just to give us some insight into whether there is a list of allowable characters, or what we're suggesting here is go the other way, a list of characters that would disqualify the Han characters. So that was the intent behind this. What we're trying to get to here is what are the set of instructions that we would give to the Generation Panel to help us out with this question?

So I'm not sure whether we had confusion around that, whether we're handing this off to a Generation Panel to be involved in evaluating TLD strings in a process, we're certainly not doing that. All we're doing here is seeking the expertise of the Generation Panels to help us out with this specific question in the charter question. I'm not sure if that makes sense at all or not. Hadia and then Abdulkarim. Go ahead, Hadia:

HADIA ELMINIAWI: Okay, thank you. This is Hadia for the record. And first I would like to say that in Arabic there's no single character that exists that I know of. So I'm not speaking because of the Arabic script, right? So I'm really confused because I'm not sure exactly what problem are we trying to solve? And my understanding that the problem that we're trying to solve is actually the similarity and confusability that could actually exist with single characters.

However, I'm not sure that giving specific instructions to a Generation Panel, or actually the Generation Panel look into this and coming up with a list of restrictions is actually a good idea. And I think what Jeff was saying, or maybe at least this is how I look at it, that it is possible to allow candidates to apply for single characters. But then that does not mean that whoever applies for a single character will get it. Because after the candidate applies for a single character TLD, then would come the role of the confusability and similarity panel or group that will look into this and see if this actually causes risks to the ecosystem, if it causes kinds of confusability or similarity.

EN

So again, I think that this is the easiest way to go forward. And not only the easiest, I'm not sure if we do otherwise we will actually get better results. So yeah, thank you.

DONNA AUSTIN: Thanks Hadia. Jeff?

JEFF NEUMAN: Yeah, I think I agree with everything Hadia just said that confusability is only measured after there's an application for a specific TLD. And the confusion is tested against existing delegated strings or other applications in that round, and then they have different consequences if they are. I don't think there's anything we need to do differently other than SubPro's recommendation that the character that's being applied for is an actual ideograph or ideogram, and that could be one of the application requirements.

But to do any of this before applications is just too much work, unnecessary. And I'm afraid the Generation Panels are going to get into policy debates about not just what alphabet characters are allowed, which is sort of what they're doing now. But then they're going to get into things like well, I don't like this string because for any host of reasons unrelated to what the Generation Panels are supposed to be doing.

So just to sum it up, agree with Hadia. Confusability is only measured after the application and that's it. It's pretty simple. Well, simple from our perspective. Probably not simple for the evaluators, but simple for the policy.

EN

DONNA AUSITN: Well, speaking as the chair here I don't find any of this simple at the moment. Okay, so sorry, Abdulkarim, your hand keeps going up and down. So if you wanted to go now, please do

ABDULKARIM OLOYEDE: Sorry, I lost connection for a bit. I didn't hear Hadia's comment because I lost connection when she was making her comment. But my comment was, number one, I believe it's a good idea to have a Generation Panel because confusability, it's something from the ordinary users, it's something that might be a big problem. And especially if we don't understand the effect. Yes, I probably guess what Hadia said was it can be measured after, but I'm thinking if there is a possibility for a single character to be part of the application. If we leave it to probably after, then it might lead to another confusion. That's my comment on that.

And then Jeff mentioned the fact that if the character is delegated, then you cannot measure the confusability. I think that takes us back to the fact that we're saying there shouldn't be any restriction to the facts that would be delegated is going to be a variant or not. And if there is a possibility to restrict that, then I think it's a good thing that we need to analyze that now and we'll probably need an expert opinion on what is going to be the impact on single characters. Thank you.

DONNA AUSTIN: So I'm just trying to unpack this. So I think the general sense that I'm hearing is that were over engineering this by asking the

Generation Panels to consider a set of questions that we would develop as the EDP team to help us respond to this charter question. So I think that's genuinely what I'm hearing, although I think there were one or two that thought that there is value in going and asking the Generation Panel to help us out with this one.

So what I'm going to suggest is that the leadership team will take the time to go back and review the discussion that's been had and see what the path forward on this is. And I see that there's a little bit going on in chat that I haven't had time to digest. Hadia?

HADIA ELMINIAWI: Yes, thank you. This is Hadia for the record. So I just wanted to note that that actually, if we decide not to go to the Generation Panel or put those restrictions up front, that does not mean that we are going against the SSAC actually recommendation. Because the purpose of that recommendation is to avoid confusability, similarities, issues like that. And all of that will be looked at after the applicant applies and a panel starts looking into the applied for TLD and any possible risks to the community from that TLD. So I don't see that going against SSAC recommendation. Thank you.

DONNA AUSTIN: Thanks Hadia. And I don't think the concern here is that we're going against the SSAC, we're actually going against what this a EPDP team had previously discussed and agreed. So it seems that now that staff have put something in front of folks, we're

EN

changing our mind about the path forward on answering this question. So I think that's why I'm struggling with it. Sarmad and then Ariel.

SARMAD HUSSAIN: Thank you. So one benefit of doing some analysis before the application process is to actually help the applicants because once the applicant applies and then the DNS stability panel for some reason, or the string similarity review panel for some reason, which may actually be reasonably obvious, reject the application, that's obviously not a desirable conclusion. Whereas if there are some obvious things, which will potentially be rejected by such panels, it may actually be a good idea to identify them upfront so that applicants don't apply for them. So I think that's certainly one of the advantages which may actually come out of this exercise. Thank you.

DONNA AUSTIN: Thanks Sarmad. Ariel?

ARIEL LIANG: Thanks, Donna. And I'd like to supplement what Sarmad mentioned. And in fact, SubPro, in their recommendation for single character TLD suggested or welcomed the identification of characters that could potentially be single character gTLDs and that will substantially increase the predictability of what would likely still remain a case by case manual process by the DNS stability panel or string similarity review panels. So SubPro actually welcomed the identification of these characters to help

make the process more predictable. I just want to mention that and then it seems, this group is having some different opinion and not to do this, but I want to provide the context from SubPro's report.

DONNA AUSTIN: Thanks, Ariel. Jeff?

JEFF NEUMAN: Can you do display the SubPro recommendation if that's possible before I make my comment?

DONNA AUSTIN: Do we have that readily available, Ariel?

ARIEL LIANG: Yes, I do. I just need to figure out how to get it from my third screen to the second screen. So one second.

DONNA AUSTIN: While Ariel's doing that, can folks think about – And I understand it's in chat that we don't need a requirement about that the actual meaning, but given what Sarmad has said about the potential benefits of having the Generation or Integration Panels have a look at these, if we strip this back considerably about what we were asking the Generation and Integration Panels to help us out with, would there be some support for doing that? Okay, so Jeff, you had the recommendation? Go ahead.

EN

JEFF NEUMAN: Yeah, sorry. my screen is not showing it. Just give me one sec. Sorry. Apologies. Sorry. Apologies. Okay, right, so that's the rationale, but what's the recommendation? Yeah, I mean, I don't see what we're recommending as being against that. I mean, I understand that there's a rationale and discussion, but the only thing that was adopted, this is a weird thing about the GNSO, but the only thing that was adopted was the recommendation and not the rationale. So I don't see what we're recommending as being against that, number one.

But number two, is there's a lot of things that could be helpful to do for applicants. Why wouldn't we just have a string similarity review panel set up to look at things for applicants that have questions beforehand? There's a lot of things we can do. But we don't because it's expensive, time consuming, resource intensive. I understand there's benefits that we could have for having a string panel up in advance, but those benefits need to be weighed against the length of time it would take to do this, the complexity of what we're asking, the resources that would have to be expended, and whether that would delay the overall program.

I think the first three sub bullets, which were in that previous slide, were the relevant ones. And everything else, I don't think needs to be in there. Thanks.

DONNA AUSTIN: Thanks, Jeff. Hadia?

EN

HADIA ELMINIAWI::   Thank you. This is Hadia for the record. I was wondering if the workforce the Similarity Panel does always have to be manual and human based. And I was thinking that maybe actually having predictable outcomes is a good thing. But I was wondering if we could have a tool for that in order to give sort of a predictable outcomes so that applicants could use a certain tool to predict an outcome. And then after actually the application, then comes the human based review. Again, I'm just thinking out loud. Maybe a tool is a better way forward for our desire to have a predictable system. Thank you.

DONNA AUSTIN: Thanks Hadia. Jeff?

JEFF NEUMAN: Yeah, Hadia, a tool will be fantastic. ICANN tried that in the last roun, iIt built a tool called called Sword. And what Sword did was it assigned a percentage number to how similar one string was to another. And one string to another meant you could just type in two strings, it didn't have to be to an existing string. And it gave you this number score. ICANN ended up throwing that Sword tool out because it was just useless. And at the end of the day it hurt predictability because users thought that, oh, this has a 96% similarity rating to another string. Yet, it was allowed, because ICANN didn't really look at the numbers. And there's probably other reasons I just don't know about as to why else it didn't work, but maybe technology has advanced in the last decade to make it better, but SubPro did recommend against having an automated tool for confusability because of the disaster that Sword was.

DONNA AUSTIN: Sarmad, could I ask you just a question about timing if we strip this request back to seek some guidance from the GPS and IPs? How much time do you think that would take? And in the meantime I'm going to go to Justine, she's got her hand up. Justine?

JUSTINE CHEW: Yes, thanks, Donna. This is Justine. I just want to make a comment, I guess it needs to be clarified, and this is to what Jeff said last. We did in SubPro if I recall correctly recommend not to use Sword. But I don't actually remember SubPro recommending not to use an automated tool. We just recognize that Sword was a disaster, but we didn't actually say that an automated solution isn't possibly the right way to go. So I just want to make that clarification. Okay, thanks.

DONNA AUSTIN: Thanks, Justine. Sarmad?

SARMAD HUSSAIN: Thank you, Donna, visit Sarmad. Just to respond to your question from earlier, so we have actually, in some cases, gone back to Generation Panels to help on some of the work beyond the root zone as they've already done. And we actually have found them quite helpful and supportive. So there is a good chance that if we go to them with some extra request we would find them helpful.

As far as the time is concerned, actually, it really, of course, depends on the task. And if maybe the working group could help frame even at the higher level, what we could do is go to Generation Panel and maybe request some estimated time from them and come back to the working group. Without, I guess, requesting them for estimates, it's harder to, of course, suggest our time frame. Thank you.

DONNA AUSTIN: Thanks Sarmad, and I appreciate that. And just for the rest of the group, one of the things that I found really helpful when we're working through these questions is having good data or good information available to us about questions that we struggle with, particularly because we don't necessarily have the expertise within this group to consider some of these questions. So this is one of the ones where we thought previously that it would be valuable to seek assistance from the Generation Integration Panel.

So what I think I'd like to do is if the team can give leadership and staff some time to review the conversation that's happened today and see if there's a way that we can strip this back to just a minimal request from the Generation Integration Panels to see if there is something of value that we could ask them to take a look at that would help us in responding to this question. So we take onboard the comments that perhaps we've gone a little bit overboard with some of the specifics and that we might be asking the Generation Panels Integration Panels to get into policy. That's not the intent.

Our intent is just for the Generation Panel or Integration Panels that have specific expertise to help us out in responding to these question so we can strip this back considerably to what the ask is, and maybe just see if there is a hard no list or disqualified Han characters. And maybe the answer is that it doesn't doesn't matter. And even that is valuable information to us.

So we're going to be working on this PDP for another at least six or eight or nine months. So I think there is probably time for the Generation Panels to have a look at this if we strip it back. So I just want to ask the group, is that a reasonable ask and something that can be signed on to? We will come back, we may just take this to the manning list to see if what we come up with is suitable. But is that a reasonable approach to try to get past this question here? So Jeff?

JEFF NEUMAN: Yeah, unfortunately, I don't think that that's a reasonable request at this point until we have the actual question down as to what we would ask. And here's the reason, and I keep coming back to this, confusability as we deal with it with gTLDs is confusion to an existing delegation or other applications. So for the first part, if we're asking the Label Generation Panel to answer the question as to whether there's any single character that could possibly be in those Han scripts that would be confusingly similar to an existing delegation, okay, that's fine. But we can't ask them anything else because we don't know what applications they're going to be. And so you can't ask a Label Generation panel to test confusability against applications that haven't yet been submitted or you're asking them to do speculative work.

Now, if there are characters, like Dennis just pointed out, that looks like a hyphen, well then we can set a rule that says, just like hyphens aren't allowed as the third and fourth characters or the first or last. Perhaps you could set a rule like that, I guess. But that's not confusability, that's because of a technical reason why we don't allow hyphens. So before I could even say yes to the approach of a Label Generation Panel, we really need to see the question that we're going to ask them and how narrow it is.

DONNA AUSTIN: So what I'm asking is, in principle if we can strip this back and find something that's acceptable to the group, is that something that we're okay to work on? Or do we just draw a line under this now and say we don't think there's any value in going to the Generation Panels? And if not, then have we adequately answered this question? And do we have enough to develop a recommendation around it? And can we just move on? I don't want to spend time on things unnecessarily. And if the sense of the group here is that we don't need to do anything more, that we can answer this question and make a recommendation and move on, I don't think we're in that position. But if that's the sense of the group, we can do that. Satish?

SATISH BABU: Thanks Donna. Satish for the record. Well I agree quite a bit with Jeff, I still think in my personal opinion we should consider the option of going back to the GPs and the IPs. The reason is we are doing for the first time and they are experts and they are linguists, we are not. I mean it is good to have an opportunity on a

precautionary basis to see if there are some general principles or guidelines that might apply, and not for specific comparisons or usability, but general principles that might apply. Assuming of course that the cost and time aspects are not significant. So my position would be to support that we try to redefine this and then see what would be the time and cost implications. Thank you

DONNA AUSTIN: Thanks Satish. Jeff, new hand?

JEFF NEUMAN: Yeah, sorry, you're hearing too much from me, I guess. Yeah, I guess my last point on this is the conversation that's happening in the chat is one of the things I'm a little bit worried about with Label Generation Panels, right? So the fact that something looks like an equals sign, who cares, right? I know in ASCII we're not allowed to have equal signs because that's not in the ASCII characters. There's no technical rule against having an equals sign. So long story short, if we send this type of thing to a Label Generation Panel, it should only look at the existing rules and technical rules for new gTLDs. And should not comment on well, this looks like a plus sign or this looks like whatever else simply because it's not allowed in ASCII.

I think the reason characters aren't allowed is important. And the hyphen, for example, that was a technical rule because we didn't want confusion against whether something's an IDN name or not, because tools can be confused as to thinking it might be an IDN. Numbers aren't allowed because there's confusion as to whether

EN

it's an IP address. Those kind of rules are okay and can be applied by maybe Label Generation Panel. But any new rules of things like well, this looks like that, we really need to be careful with the Generation Panels.

DONNA AUSTIN: Okay, thanks, Jeff. So I think what we're going to do is the leadership team and Steve, and Emily, and Ariel, we'll go back and review this discussion. We'll see if there is a stripped back version that we could seek assistance from the Generation Integration Panels on. As Satish said they may have expertise that will identify something that we haven't thought of, and there's value in doing that. But I also take the point that we don't want to do something that is going to be significant costs or added burden, I suppose, on the Generation and IP. So, we will take that on board, and we will see what we can come up with and see if what we do come up with is acceptable to the group. And if it's not, then we'll just draw a line under it and we will move on. Okay. Thanks, everybody.

So, Ariel, I think we're going to move on to A9.

ARIEL LIANG: Yes, thanks, Donna. This is Ariel. So, A9, just to remind folks what the question was. It's about the label states and the definitions. And then in the staff paper there are five that's delegated, withheld same entity, blocked, allocated, rejected. And there was an action item for staff to overlay the label states against the application

EN

process and see where the overlapping areas would be. So we look at that and this is just to present you what we found.

So based on the 2012 round there were already existing application status related to different labels. So the first point in this green box is the ones we saw on the application page. So initial evaluation, and extended evaluation, evaluation complete, so on so forth. So this is not a comprehensive list because it doesn't include the status in the contention stats, for example. But we found that there are potentially several application status that may overlap with the label state as proposed by the staff paper, and we're listing in the table below.

So there's one label state is delegated for [inaudible] label and then the application status is also delegated, so they're basically the same. And then the label state as rejected is actually encompassing not approved class, will not proceed application status. And that was already pointed out in the staff paper. And then the other three labels states, allocated, blocked, withheld same entity. These ones do not have the equivalent application status. But upon further review we found that allocated, it could be a precursor to in contracting for gTLD applications. But it's a different thing for ccTLDs because contracts do not really apply to them.

And then for blocked and withheld same entity they do not have applicable application status because these cannot be applied for and they need to be set aside. And there these are based on RZ-LGR calculations and there are new definitely for now, so understandably they do not have equivalent application status.

So, this is a what we did with the action item, but what we're trying to propose right now is to basically get this question unstuck because previously the group was talking about several ideas such as streamlining the label states into three, basically withheld same entity can include withheld same entity/rejected, withheld same entity/allocated. Or create something different so allocated can be better differentiated from delegated and those types of discussions. But upon some further consideration and discussion with the EPDP leadership team, staff liked to propose that the EPDP team accept the staff paper proposed label states as the preliminary recommendation for now, and then revisit them if new discoveries are made.

So we just want to give you some rationale why we're proposing this suggestion. So the first point is that, as you saw in the previous table, the label states, they do not depend on the application status. They're two separate things. They have some overlapping areas, but the term used, they don't have dependencies with each other. And also the team is not asked to further expand the application status to accommodate the varying levels. That is not what the charter question is asking the team to do.

And second, the label states, we want to understand how they will be used in the future. So, their main usage is to track the status of variant labels, and then based on the status of the variant labels may be applied or used in the different stages of the gTLD application process, and even other processes such as the IDN CCTLD processes. So you can probably envision that those states will be aggregated into some kind of list or stats, and they will be

EN

incorporated into some kind of algorithm so that in different stages of the application process for gTLDs they will be used to consider for whether they're applicable for no objection process, for example, similarity review, and those type of things.

So, so this is their usage. So when a variant label have a different label state, they may play different roles in the various steps of the gTLD application and evaluation process. So the main focus point is not really how they are called, is more about their effect and the role they played in the application and evaluation process. And another point I'd like to mention is that the staff proposed these labels states because they want to make sure it's TLD neutral, means they can be applied consistently across both gTLDs and CCTLDs to the extent possible due to their potential use cases.

And currently, based on Dallas report the CCTPP force variance subgroup, they already proposed to adopt the staff paper recommended to label states as a preliminary agreement and include that in a glossary. So they're basically not deviating from what the staff paper proposed. And due to this reason, we think the real focus is not really how to call the label states. It's more about how the variants will play in the similarity review, objection process, contention sets, and all these matters that will be tackled in the later part of the charter.

So what we're proposing is that consider agreeing with the staff paper proposal for now, and then revisit later so that we can get unstuck from the topic of trying to figure out the terminology. Because that's probably not something we need to focus not at the point. So I'll stop here.

DONNA AUSTIN: Thanks, Ariel. And I think one of the other reasons to consider this as a preliminary recommendation and see if we find other discoveries along the way is that this is a question that the working group is supposed to coordinate with the SubPro IRT to develop a consistent definition of variant level status and that ideal set. And because there is no SubPro IRT we can't have that discussion. So I'm interested to hear from folks whether they can live with this as a preliminary recommendation and something that we can come back to at a later point if we need to. Thanks, Michael. And Jeff has asked if we can put the question on the list because he can't remember the staff recommendation. And like I say that we have Michael and Dennis that support this as a preliminary recommendation. And so Ariel, I meant to.

ARIEL LIANG: Thanks, Donna. And I just quickly went to the slide where the staff definition or proposal are for these five level states as a refresher, but I will put the SAP paper link in the chat and on the list so folks can review that in their spare time later.

DONNA AUSTIN: Thanks, Ariel. Jeff?

JEFF NEUMAN: Yeah, so thanks. And I can't remember because it's been a while, are any of these definitions inconsistent with what the definitions

EN

are in the new gTLD process? And I think the only one probably would be delegated, I guess. That's the only same one.

DONNA AUSTIN: I think in the previous discussion we had allocated is a different nuances depending on whether you're in a G or a C space. Ariel, go ahead.

ARIEL LIANG: Actually, I do need to double check. So the application status for delegated, that's different from what is displayed on the slide. And what displayed on the slide is actually a definition extracted from the integrated issues report. That's the first report that all these labels states were defined. And in the staff paper and need to confirm that whether these definitions were iterated or not. But if so then there's some difference. But as mentioned earlier, I don't think there's necessary dependencies between these two. This is basically to tag the variant labels. And there may be some overlap with application status, but we think the terminology is not the most important thing to discuss right now.

JEFF NEUMAN: Well, it needs to be consistent, right? So I'm fine with having one set of definitions, but what's going to be critical, at least for our purposes, and not necessarily the CCTLD purpose, but for our purposes, we don't want to have delegated mean one thing for IDNs and one thing for everything else. Sorry, one thing for variance and then another thing for everything else. So I mean, I'm okay with the proposal to use staff definitions, but I think you

DONNA AUSTIN: Thanks, Jeff. Okay, I think this is a preliminary recommendation that folks can live with for now, but it may be something that we have to come back to in the future. Okay, so let's – I'm just reading Jeff's question in –

JEFF NEUMAN: Sorry, not a question set in the SubPro ODP, and Steve can correct me if I'm wrong, I believe, unless I'm remembering a different conversation, I believe there was a question about terminology. And the question asked of us was whether we were very strict on the terminology or whether it was conceptually from the SubPro report. Whether the staff could enhance it with additional terminology and other stuff. I think the answer was, look, conceptually, that's a detail for staff, but it needed to just be consistent. Steve, thanks, help me out here.

STEVE CHAN: Thanks, Jeff. Right, I think they did ask that question. This is Steve, sorry. I partially raised my hand to validate what you just said. But also to note, and I definitely welcome Sarmad and Ariel and others to correct me if I get this a little bit wrong. But the set of label states here is intended to be used for essentially tagging the variant set for a given gTLD. And it doesn't necessarily have to

match up exactly with the application statuses as they're used in the new gTLD application system, or the status system, I suppose.

In my mind what the application system would serve to do is end up being sort of a trigger to make sure that you have the accurate states for your set of variant statuses, or states I suppose. I don't know if that distinction makes sense. Maybe you need a visual representation to understand that distinction. But as I understand it, they don't need to match up one to one, you just need to know when to properly trigger the state change for your variant set. I hope that made some sense, and hopefully I actually got that accurate. And I'm sure Sarmad will tell me if I didn't. Thanks.

DONNA AUSTIN: Thanks, Dave. So to Jeff's point in chat, is that we should not have one word mean two different things. I think this is where we got caught up last time. And we tried to understand what is the purpose of the definitions that we have here. And what Steve has explained is that it seems that there are two different purposes here so we might have different meanings. And I understand that will create confusion. So maybe that confusion is still floating around in folks heads. So Edmon and then Sarmad.

EDMON CHUNG: Thank you, this is Edmon speaking personally, not as Board Liaison, I guess, just as a disclaimer. Yeah, so based on actually what Steve is saying, and also what Jeff is saying, then perhaps we should really think about these designations different from the

application status situation. And if we do have an overlap which is, for example, delegated, then we really should not use the same terminology if we're trying to describe different things. Or at least preface it with a prefix or a variant. Delegated variant, allocated to just to make sure that we don't try to use the same word for different things.

DONNA AUSTIN: Thanks, Edmon. Sarmad?

SARMAD HUSSAIN: Thank you. So I think one of the ways at least I look at it is that you have this cariant set in which you may actually have a hundred let's say variant strings for a particular TLD or TLD string. Those 100 strings can be in one of these five states. So there could be some variants which are blocked, there could be some variants which are withheld same entity, or what we used to call allocatable. And these are sort of non overlapping sets, right? So if there is a string out of those 100 variants, if there is a one string which is in blocked, then obviously it will not be in any other states. So it's not overlapping sets. But collectively, they formulate the complete set of 100 strings. Applications, an application process or a contracting process, or some of these other processes, could take one label from one state to another state. So at least that's, for example, one way or how I would look at it.

DONNA AUSTIN: Thanks Sarmad. Jeff?

JEFF NEUMAN: Yeah, thanks. I know that we can kind of differentiate and have one term mean different things, one for looking at variants, and then one for applications. But I think we need to design this system for people that don't do this every day and applicants, because they could be faced with this. So at the end of the day, I really don't care what the words are that we use, just that it's consistent and it means one thing regardless of the purpose. So can we just recommend not saying that we accept the staff paper definitions as is, but rather asking staff to coordinate with the ODP to come up with terms and definitions that are consistent regardless of the purpose for which they're being used? Thanks.

DONNA AUSTIN: Thanks Jeff. And I think we can use that as a recommendation. As you said, we had hoped we'd have a SubPro IRT that we could work this out with as well. So we'll change the recommendation a little bit to reflect that nuance that to the extent that we can ensure that there's not confusion around the definitions of certain terms. I can understand that having rejected mean three or four different things depending on the purpose is going to be very confusing. So let's try to streamline that in some way. And to the extent that we have control over maintaining those consistent definitions. So we will take that on board as an amendment to our original recommendation. Okay, I think we're good on A9, Ariel. So we can go back to, I guess, D1B.

ARIEL LIANG: Sorry, Donna, this is Ariel. We're one minute over time now.

DONNA AUSTIN: Oh, gosh. I thought we still had 30 minutes to go. I'm so sorry, everyone. 90 minutes goes really fast when you're having fun. All right, apologies everybody. It's Saturday here for me so have a great weekend everybody. And thanks again for your contributions and we will talk to you in a week. Thanks, everyone.

ARIEL LIANG: Sorry, Donna, it's actually not a week. It's a Wednesday next week for the ICANN73 session.

DONNA AUSTIN: Is next week ICANN week or is it the week after?

ARIEL LIANG: My apologies. I should check.

JUSTINE CHEW: We will confirm things on the mail list, okay?

ARIEL LIANG: Sorry for causing the confusion.

DONNA AUSTIN: Okay. Bye, everybody.

DEVAN REED: Thank you all for joining. Once again, this meeting is adjourned. I'll end the recording and disconnect all remaining lines. Have a great weekend.

[END OF TRANSCRIPTION]