ICANN Transcription

IDNs EPDP

Thursday, 07 April 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/aQB1Cw

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 Thursday, 7th April 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?

All members and participants will be promoted to panelists for today call. Members and participants, when using the chat, please select "everyone" in order for everyone to see the chat. Observers will remain as view only status.

Statements of interest must be kept up to date. If anyone has any updates to share, please raise your hand or speak up now. If you need assistance updating your statements of interest, please e-mail 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 multistakeholder 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. And welcome, everybody, to today's call. We are moving into a new topic this week and I'm going to be extremely honest that this process caper could do my head in a little bit today so we'll see where we get to.

Just one thing about next week's call. We're coming up to … The time of the call next week will conflict with the GNSO Council call. So we need to move it because we know we'll miss some folks from staff, and Justine, and some other members of our team. So we're going to have to move the time of next week's call. Usually, we just move it 24 hours later but we can't do that because it's Good Friday and we know that in some places of the world, that is a holiday. And we want to be respectful of that. So we don't want to go down that path of moving it 24 hours later.

What I'd like folks to think about is do we think it would work if we do it at UTC 11:30 next week, which would be two hours earlier than this current time? If you are on the West Coast of the US, that's a pretty horrible time and it might be a little bit early on the East Coast as well. But if folks can think about that, then I'll come back to it at the end of the call and see if that would work for folks.

EN

And I'll get somebody to put that in the chat, that we think that might be the best option.

We could do it a couple of days earlier but I don't think that—in my mind, at least, doesn't make particular sense because we're compressed too much. So let's see if, at the end of the call, whether Thursday, the 14th of April from 11:30 UTC to 13:00 would work for folks. We'll come back to that.

Okay. So with that, I think we will get going unless I've missed something in the update. And Ariel will remind me if I have.

ARIEL LIANG: Donna, we're okay for now, although there are some developments for the ICANN74 schedule. But we can touch base on that and maybe let the team know after—at least let you guys know about the developments or not.

DONNA AUSTIN: Okay. All right. Terrific. All righty. So with that, Ariel, I will hand you over to Ariel to start us off on the charter questions for today. This is another area where we are a little bit disadvantaged because the SubPro IRT is not operating as we thought it would be. So we're going to have to make some assumptions about process. And I think, at least for me, this is going to get pretty confusing so we'll have to have our best conceptual minds working here today. So with that, Ariel, I will hand it over to you.

ARIEL LIANG:	Thanks, Donna. So today we're going to start a new topic. It's topic E. That's related to the detailed stages in the new gTLD application and evaluation process. We're starting on E3 because this is about string similarity. And we're following the 2012 round. Basically, that's one of the first stages in the initial evaluation of an applied-for string so we're starting from there.

Just to show you the question. Basically, we're proposing to tackle E3 and E1 at the same time because they both talk about string similarity review and they're interrelated. So it seems to be logical to do it together. E3, the question is the working group and SubPro IRT to coordinate to ensure consistency in the implementation of the string similarity review procedure for variant label applications of existing and future gTLDs.

E1, it has two parts but we're only tackling the part related to string similarity review. So the question is what role, if any, do TLD labels withheld for possible allocation, or withheld for the same entity, play vis-à-vis string similarity review process.

So before we dive into these two questions, we'd just like to clarify the scope of discussion today. We're only focusing on the future new gTLD aspects only. So we're not tackling the variant labels—the implication for existing gTLDs. So we're only focusing on future new gTLD application.

And another thing we want to confirm is that, for the SubPro, they have developed some recommendation and implementation guidance related to the string similarity review. They did not propose to change the criteria or standards for the string similarity review but owning some, I guess, incremental improvements and

some other aspects to enhance the process. So for our deliberation, we should not try to amend the structures or framework of process established by the SubPro already. We should only focus on the variant labels' role in the string similarity review process in future gTLD applications.

Here, I want to try to make the question clearer. What is being asked is, first, whether and how should the string similarity review be adjusted due to the implementation of variant labels. So that's the general or main question. And the sub-question is should the applied-for gTLD string be compared against withheld same entity variant labels. Basically, that's the role with how same entity labels play. Should they be used for comparison for the applied-for labels? So that's staff's interpretation of the charter question. That's what is being asked.

Then, before we go into the questions themselves, let's just do a quick refresher of what string similarity review is about. You probably have seen this chart before. That's the 2012 round—the process flow related to that. So the top row is about the administrative steps for application submission. Then when you go through the background screening, everything looks fine and then there's no early warning from the GAC. Then you go into the initial evaluation. So you see the box with the dotted red line around it is the string similarity review. That's one part of the initial evaluation process.

So what is the string similarity review about? This is a summary or recap in the 2012 round so I will just help everyone get up to speed to that. We tried to simplify the introduction into the four Ws and the H. So what, why, when, who, how?

EN

What is string similarity review? This process assesses whether a proposed gTLD string creates a probability of user confusion due to similarity with any reserved name, any existing TLD, any requested IDN ccTLD, or any new gTLD string applied for in the same application round. So that's basically the content of the review.

Why do we need the string similarity review? It's because this process will help prevent user confusion and loss of confidence in the DNS, resulting from delegation of many similar strings.

When does the review happen? It occurs during the initial evaluation. So that's what you saw in the previous slide, in the flowchart, where it happens.

Who will do the review, then? It's conducted by an independent string similarity panel. These are the basic background for string similarity review.

Now we're going to the how—how that's being done. In the 2012 round, they adopted a standard for string confusion. What the standard is, that string confusion exists when a string so nearly resembles another visually that it's likely to deceive or cause confusion. Basically, the likelihood of confusion, it must be probable but not merely possible that the confusion will arise in the mind of the average reasonable Internet user. Mere association in the sense that the string brings another string to mind is insufficient to find a likelihood of confusion. So that's the standard for string confusion in the 2012 round.

Just to continue on the how aspects—how the review is conducted. So applied-for string is compared against the following other types of strings. Existing TLDs—I believe that includes both gTLDs and ccTLDs. Reserved names. And then the third is other applied-for gTLD strings in the same application round. So in the later date, if there is string similarity found, then the contention set will be used in the later stage of evaluation and we have another question related to that in the charter. And also, strings requested as IDN ccTLDs. So that's the four main types being compared to.

If the applied-for gTLD string is a two-character gTLD string, it's also compared against any one character label in any script and any other two-character ASCII strings. It's to protect possible future ccTLD delegations. So there are some additional items for comparison if the applied-for string has only two characters.

In the 2012, I also want to mention that, as you may recall, some IDN applications indicate a self-identified variant. These were taken into account. Basically, those self-identified variants were subject to string similarity analyses and they're mainly to be reviewed to confirm they're indeed variants as defined in the relevant IDN table submitted by the applicant. Then they were treated, essentially, the same as reserved names.

I just wanted to quickly mention that variants did have a role in the string similarity review in 2012 round. But essentially, they are information only and they had no legal standing. So the applicant had no claim to rights over that. So that's a kind of recap of the string similarity review in the 2012 round.

EN

In the staff paper, there is a proposal in terms of what string similarity review should be modified or adjusted, taking into account the implementation of variant labels. They are proposing a maximally conservative approach.

Basically, what they're proposing is that the applied-for string needs to be compared against all variant labels, including withheld same entity and blocked variant labels of existing gTLDs and ccTLDs, other applied for gTLDs in the same round, strings requested as ccTLDs, and reserved names, including all two-character ASCII strings. So it's not just about the labels requested by an applicant but all variant labels, including the blocked ones and withheld same entity ones. So that's the staff proposal in the staff paper.

They provided some examples to explain the rationale for this proposal. For example, if string A is a withheld same entity and string B is visually similar to string A, then allocating string B would undermine the predictability of the outcome of variant processing from the RZLGR. Another example they provided is that, for example, if string C is blocked under RZLGR but the visually similar string D is allocatable, then the string D might become a workaround for the blocked string C. That's why the staff paper is proposing a maximally conservative approach to compare all variant labels against the applied-for string.

But they did realize there are some consequences for this proposal. It will expand considerably the number of strings that might need to be considered. If you recall, for certain scripts, like Arabic, they could generate tens, or hundreds, or even more variant labels—not just the blocked ones but also the allocatable

ones. So it will become a considerably large number of strings that might need to be considered. As a consequence, the entire string similarity review process will become very expensive to operate. So the staff did recognize the consequence of this suggested approach.

I see Jeff has his hand up. Maybe I can pause for a moment and see whether he has some input or questions for the presentation for now.

JEFF NEUMAN: Yeah. Thanks, Ariel. Great explanations. The way I've thought about this is almost like we have to do a matrix of the different inputs and the outcomes. If we're just talking about new applications going forward, right? Is that what we're discussing first as opposed to variants for existing ones, right? We're just talking going forward for new gTLDs?

ARIEL LIANG: Yes. That's correct.

JEFF NEUMAN: Okay. So then, what I think, at least in my mind, the way I would think about is that when an entity applies for a string—and we'll call that, I guess, the primary string, which … Wait. Did you call that string A or something like that? They should also indicate in their proposal whether they are interested, at that time, to also apply for any of the variants.

And then, when they do that, I think that the midway between what the staff proposal is, is to only evaluate the strings for with the TLD has indicated that it may want. This way, it won't necessarily be hundreds of variant labels that need to be checked.

Then, after that, in drawing a matrix out, at least in my mind, you would say, "Okay. Does the primary string …? Is it similar in the definition that's here to …?" You check existing gTLDs, other applied-for gTLDs, reserved names, and strings requested as ccTLDs.

If it's the first, third, or fourth bullet there, then it would be … Again, this is the primary string. Then the application would be rejected because that's what the impact is of being similar in the first, third, and fourth bullet. If the primary string is similar to other applied-for gTLDs, then it just goes into the string contention.

If the primary string has no issues—like there's no contention and it's not similar—then you go on to the next step, which is to evaluate the other applied-for variant labels. If the variant labels are similar to existing gTLDs, ccTLDs—the first, third, or fourth bullet—I believe, and this is just me, that the applicant should be allowed to, at that point, say, "Okay. Forget it. I don't want my whole application to be destroyed so I will commit to never being allowed to launch that particular variant," because then that variant would be similar to all those things.

Similarly, if the applicant wants to avoid string contention for all of its strings, it could do the same thing and say, "Okay. I understand that variant is similar to other applied-for gTLDs but my primary is not so I'm just going to waive my right to never apply for that

EN

variant." Or it could say, "No. I really want that variant, in which case it would be thrown in the same contention set with the other applied-for gTLDs.

I think that analysis would need to apply to every single variant that—again, that the applicant, at that point in time, thinks it's going to want. Let's say that's all done and the primary is delegated. In the future, if the applicant would like to use a variant that it didn't previously apply for, then I think that needs to undergo another string similarity analysis with the same thing— with existing gTLDs, ccTLDs, strings requested at ccTLDs, reserved names. And if there are still other applied-for gTLDs in the queue, because in theory we could have overlapping rounds, then the same thing would need … It could enter a contention set for that variant.

So to just summarize, I think it depends on which string is similar. If it's the primary, then it's no different than any other application in the new gTLD round. If it's not the primary one, then I believe the applicant should have the choice of whether to just say, "Okay. Forget that variant. I understand," or say, "I can't move forward if it's similar to an existing or string requested as cc-TLD." But if it's in that second category, I think the applicant can elect to have the entire set grouped in, in a contention set.

So I hope that makes sense. Again, it's probably easier if we drew a matrix on the different possibilities. But I don't think that every possible variant should be considered in the string similarity review if the applicant doesn't put every possible variant in its application. Hopefully that would reduce some of the expense. Thanks.

DONNA AUSTIN: Thanks, Jeff. Ariel, you don't have any more to run through on this?

ARIEL LIANG: Yeah. Thanks, Donna, and thanks, Jeff. Actually, there is one more slide about the contention sets. I stopped earlier because I thought Jeff may have some additional comments about the slides at hand. But he did mention contention sets so maybe I can also just quickly run through this one. It's the staff paper proposal that's an additional recommendation about contention-related issues.

Basically, the staff paper did mention there are several scenarios involving variant labels in the contention sets. The contention may occur in case two or more strings applied by different entities are visually similar to one another, variants of each other, or both visually similar and variants of each other. Basically, there is a new case. If two applicants, they are applying for two labels and they are variants to each other, that needs to go into the contention set.

The additional point mentioned in the staff paper is that when two or more applied-for variant strings are visually similar, they may only be allocated if they are associated with the same variant set and not being requested by the same entity. So it's basically emphasizing the same entity principle for that.

Another recommendation the staff paper proposed is that the entire set of variant labels needs to get processed as one contention set. If one of the labels in the set is already allocated,

then the contention is resolved in favor of the current operator. So the key point here is that, in the contention set, the entire set of variant labels need to go into that, not just the variant label in question itself.

So just to provide some additional points here, the visual similarity check needs to be performed for every requested-to-be-allocated variant against all the possible variants in every other set. If any labels are found to be visually similar, their entire variant label set must be placed into the contention set. So that's emphasizing or reiterating the previous recommendation here.

Then this graphic illustrates the example here. If applicant A requests allocation for only two labels here—t1v1, t1v2. These two labels need to be tested against all the variant labels in applicant B's, even applicant B only requested allocation for t2v1.

So that's the essence of this staff paper recommendation. And if my explanation is not clear, I will welcome Sarmad or Pitinan to provide their further clarification of that. So basically, this staff paper proposal is in line with the maximally conservative approach as well. So yeah. That's the last slide for this section of the meeting.

DONNA AUSTIN: Thanks, Ariel. Just a couple of things. I assume that the staff proposal was done by Sarmad's team. And I think it might be helpful for the group to understand the rationale of why the staff paper thinks this should be all-encompassing to give us an

EN

understanding of whether Jeff's proposal to constrain the number in some way makes sense and is possible.

So thinking about this from the perspective of what we're trying to do with this PDP in terms of having recommendations that are implementable and making it as easy for the implementation review team as possible. Then it would be good to be mindful of that. So I'd be interested to hear from Sarmad, if he's on the call, just the rationale and the thinking behind the staff proposal. Sarmad, before you do that I'll just go to Justine and Edmon. So Justine and then Edmon.

JUSTINE CHEW: Thank you, Donna. I just had a quick question. I wonder if someone like Sarmad or Pitinan could just remind us about the probability of a blocked label becoming unblocked. Thanks.

DONNA AUSTIN: Thanks, Justine. Sarmad, if you can answer Justine … Sorry, Sarmad. Sorry. I meant when I get to you. But I'll go to Edmon first and then we'll come to you, Sarmad. Thanks. Edmon?

EDMON CHUNG: Yeah. Speaking personally here, now. I guess, quickly, on Justine's point, I think it would require an update of the generation panel and the table to do that but I'll let Sarmad add to that.

Just thinking through what Jeff was talking about and also how we run through string similarity review or not, I will especially caution

that, for allocatable variants, we probably should run it through at least, as the concept for allocatable variants is that they are equivalent to the applied-for string. If it's not equivalent to the applied-for string, the IDN tables would not have identified it as a variant itself.

So not evaluating string similarity would potentially cause not only downstream issues, but in fact, potential issues for existing—I mean not existing—the users for that particular gTLD potentially going to another, confusing it with another, if there is a string similarity situation there. So that's point number one.

Point number two is we also need to think about third and fourth rounds. When there are existing applications, and then there are blocked variants, and then a new application coming in, whether the entire set of variants needs to be evaluated with all the existing sets of variants, that's something I think we should think a little bit more about.

I understand where Jeff is coming from and I probably don't disagree with some of the general approach. But I caution these two point. One is the allocatable ones which are considered equivalent to the applied-for string linguistically and technically— well, techo-policy-wise, I should say. And then future rounds where it then compares with then-existing blocked variants and so on. So those might be a couple points that we should think through before making this decision.

DONNA AUSTIN:	Thanks, Edmon. Sarmad, I'll hand it over to you to just take us through the rationale and the thinking for the staff proposal. Thanks, Sarmad.
SARMAD HUSSAIN:	Thank you. Very quickly, first to respond to Justine's question. I think Edmon already addressed that. It is not going to be very frequent but if a generation panel re-forms itself and updates its proposal, it can potentially change a blocked label to an allocatable label. So as I said, that will be a bit infrequent.
	Coming back to this comparison, the comparison actually can be done at three levels. One—I guess the simplest form of that— would be that we only compare the applied-for strings with the other applied-for and delegated strings. When I say "applied-for," it means whatever is explicitly chosen by the applicant, which could be just the primary string or a primary string and a variant, depending on whatever the applicant wants to apply for. So that's a minimal case and that's not very different from what already exists today, where each string is compared with every other applied-for string and the existing delegated strings and reserved strings.
	The second level would be that the applied-for string gets compared with all the allocatable applied-for strings and the allocatable variants get compared with whatever is other applied-for strings and their allocatable variants, as well as the already-delegated strings and their allocatable variants.

Then thirdly, it would be that the whole set, which means applied-for strings, allocatable variants, and blocked variants. All of them get compared with all the other cases, of course.

So what, basically, the staff has proposed as far as the comparison is concerned, whenever a string is being applied for, it gets compared with all the other strings, and all the allocatable variants, and the blocked variants. The intention at that time was caution because we, at this time, don't really … Or especially at that time when the staff report was being developed, there was very little knowledge of how variants will operate.

Therefore, the approach we took was of maximal caution or maximal conservatism, where we say that we check against all the possible variants, including allocatable ones and the blocked ones, to make sure that there is—we go forward with the most caution.

The reason to compare with the blocked ones is that if a variant is blocked for whatever reason, the community does not want that string. For example, as a TLD, if there is a very similar string to it, but it becomes delegated, in a way it can become a workaround for that blocked string because you have a very similar-looking string to a block string which can eventually get delegated. That's at least the reasoning which has been documented in the report.

Again, to summarize, because we know less about variant labels, the staff approach was to proceed with maximum caution. Thank you.

DONNA AUSTIN:	Okay. Thanks, Sarmad. So back to Jeff. Jeff?

| JEFF NEUMAN: | Thanks. And thanks, Sarmad, for the explanation. I think, in trying to be consistent, if I go back a second, a bunch of weeks ago we had a discussion with the SSAC. We talked about being conservative in the amount of strings and we talked that we didn't think it was prudent to actually put a limit because that would be arbitrary. |

But I think we agreed that registries, when they wanted to have an allocatable string actually allocated, that the registry needed to demonstrate that it could handle, for lack of a better word—handle that. And by "handle that," it also means that they would have to be evaluated as to how they would educate users and registrars and all that kind of thing.

So if we apply the same logic, to me, it doesn't make sense to look at the entire set because the registry hasn't asked for it to be allocated. If, in the future, the registry—whether that's registry A or B in that second slide … When that registry wants, or if it wants, to get an allocatable string actually delegated, it, at that point in time, still has to prove to ICANN that it can handle the delegation of that allocatable string, and at that point in time, it could undergo another string similarity review.

So if there were any issues, it would come out at that point in time. And it's at the risk of the registry in later applying for it to be allocated because by that point in time, it could be after two more rounds and there could be a lot more to compare it against.

EN

So I just think, in our logic, we should probably try to be consistent. And if I'm right—and Donna, please correct me because I might be misremembering—but I thought that even thought something is "allocatable" doesn't mean that the registry automatically has the right to have it allocated when it wants. It still needs to prove that it can handle it.

So under all those scenarios, I don't see a reason why the entire set needs to be in the initial string similarity review if the registry hasn't indicated that it wants the entire set. Because if we try to do that, first of all, it's overly conservative and overly restrictive and I don't think we need to be that overly conservative and overly restrictive.

And second, as was pointed out in the con, I can't even imagine the expense that would be to have panels review every potential allocatable string, regardless of whether the registry has indicated a desire for it, especially, as Ariel said, in the Arabic scripts which could have hundreds of variants. So I think we're being consistent if we go with the approach that I've suggested. Thanks.

DONNA AUSTIN: Thanks, Jeff. So Anil and then Edmon.

ANIL JAIN: Thank you, Donna. I fully agree with Jeff, what he has indicated. In fact, I propose that the confusing similarity evaluation should be conducted at the time of delegation. Now, this will reduce the load, even allocatable variants for delegation, or even the blocked variants for allocation, because we don't know when a blocked

variant will become allocatable and then delegated. So my suggestion is that the evaluation with respect to the confusing similarity should be just before delegation. Thank you.

DONNA AUSTIN: Thanks, Anil. Edmon?

EDMON CHUNG: Yeah. Thank you, Donna. Here's a scenario. If we focus on reserved strings, then we will still have to compare the entire set with the reserved string because if a reserved string becomes un-reserved at some predictable point, it would have blocked that particular reserved string to be registered or applied for by another registry.

So at least for a reserved string, I would imagine that the entire set will probably have to be checked against in terms of string similarity or else then we would have disadvantaged anyone who might want a particular reserved string later on down the road.

DONNA AUSTIN: Thanks, Edmon. Dennis?

DENNIS TAN TANAKA: Thank you, Donna. I think—at least in my thought process, I'm thinking—I understand the conservative approach in terms of having all the labels, blocked, allocatable, and applied-for, put through the visual similarity review to find contention sets and whatnot.

But let's intersect that concept with what we have talked in the past about applications will have associated fees—cost recovery or whatnot. But I would assume that it's likely that an applicant is not going to be free-for-all. There's going to be costs associated with the application process. Therefore, it is reasonable to think that they will limit themselves to applying for a primary and maybe one variant, two variants, or what have you, but not everything allocatable based on the Root Zone LGR.

So bear with me and think about those two concepts, going back to what is the set that we want to put through the review similarity process. Is it everything? But what happens if an allocatable … And I put that in the chat. What happens if an allocatable or blocked variant but not applied-for is found in contention with another? What do we want to do with that label if it was not applied for—it just happens to be an allocatable one but the applicant really doesn't want it—are we going through that similarity process for nothing? Or what's the endgame, I think I'm looking for here.

I understand the conservative approach but I think we're looking at that in a vacuum. When we intersect that with what might be the anticipated application process, fees associated with it, I think we need to factor that in as well in our talk process. So thank you.

DONNA AUSTIN: Thanks, Dennis. Jeff?

EN

JEFF NEUMAN: Yeah. Thanks. To respond to Dennis the approach that I suggested, if and when the registry wants an allocatable variant to be actually delegated, at that point in time, it would have to apply to ICANN to do it. And if, at that point in time—like what you said in your example—if they had indicated that they wanted it initially, it would have been in a contention set. So now, by the time that they actually say, "You know what?

Yeah. I do want it delegated," it will have to depend on what's going on. If the contention set has already been resolved and there's already a delegated string, then at the time that the registry that now wants this allocatable string, it's going to get rejected because the contention set was resolved and now this variant that's allocatable, that the registry wants, can't be delegated because it's, at this point in time, visually similar to an existing gTLD because it's been delegated already.

So I think that scenario is actually taken care of easily in the proposal that I've made. But again, it requires all of us to agree that just because a string is "allocatable" doesn't mean that the registry can automatically get it. It means that when the registry wants that string, it needs to submit some sort of application or review at ICANN to actually have it delegated, which I think we did agree as the compromise to limiting the number of strings as what SSAC had initially indicated. So I think it's taken care of.

DONNA AUSTIN: Okay. Sarmad?

EN

SARMAD HUSSAIN: Thank you, Donna. I just wanted to make a distinction. We actually have a pair of strings here which come into this equation. We have an applied-for string and we are comparing it with another string, which is either existing or has been applied for by somebody else. So we have a string X and a string which is applied for and is being compared with another string, Y. X has its own variant set with allocatable variants and blocked variants. And Y has its own variants, allocatable and blocked.

And what the staff paper actually says is that we are not comparing all the variant sets of X to all the variant sets of Y. We are actually … At least what the staff paper suggests is that we just take the X which is applied for, not its variants, or one of its variants if it is applied for, and compare it with Y and its variant set. So we are actually only comparing only the applied-for strings against the other strings and their variant sets. I think that is the suggestion in the staff paper. It's slightly different than comparing all the variants of applied-for string with all the variants for other strings.

I guess I'm adding that explanation to respond to a comment which Dennis just made. In that case, if an applicant is applying for X, it will only get blocked if X itself is in contention with another string or its variants. But if a variant of X is in contention, it will not block X. Thank you.

DONNA AUSTIN: Thanks, Sarmad. Jeff?

EN

JEFF NEUMAN: Yeah. Thanks, Sarmad. That actually does clarify the fourth bullet. I think, then, that is similar to what I had suggested. But what I'm saying is I think the third bullet should be the same. I don't think we should just say that every … Just because a current operator has a TLD doesn't mean that it should have the right of first refusal on every single variant that it has. I think that would be an unfair advantage. If the existing or current registry operator wants that variant, it needs to apply for it. And at the time it applies for it, then that is when it would then fall into the fourth bullet, if someone else has applied for it.

So in other words, the current registry operator should not have carte blanche to every single variant simply because it has a current registry. A current registry still needs to apply for a variant and have it delegated. If it's able to do that before another applicant comes forward, then great. But if it's not, then like I said, it shouldn't have this right of first refusal, which I think is what the third bullet establishes. Again, that's just my view. Others may have a different view. So thanks.

DONNA AUSTIN: Okay. I knew this was going to challenge my head. I'm wondering, Sarmad. Can you scope this for us in terms of parameters? I think the problematic script here is Arabic. I'm just trying to understand the bigness, if I could put it that way, of the problem.

Most of the scripts have constraints within the label generation rules in terms of variants. When we were discussing a ceiling value, one of the things that we took into account was the fact that the scope of the problem wasn't as large as we thought it was. I'm

EN

just trying to wonder if there's some way that we can understand the largeness of the problem we're trying to deal with here. I think it's only probably seven scripts that we're dealing with and Arabic is the one that could provide the biggest challenge.

So Sarmad, can you answer that question? Then, Maxim, I'll come to you.

SARMAD HUSSAIN: Thank you. I think each variant set for any label, we can divide in three parts. We have the applied-for string itself, we have the allocatable variants of that string, and then we have the blocked variants of that string. So there are three, I guess, subsets we can look at for each variant set.

The allocatable variants are normally in … It's going to be intense. So it will be 7, 10, 15, 20, depending on the string. The blocked variants can be quite large. They can be in the hundreds or thousands, actually, or even larger. Blocked variants actually can be very large, not only just for Arabic Script. They can actually be very large, even for scripts like Latin, because Latin also has blocked variants.

So if you're looking at the problem size, we're talking about, of course, the allocatable variant. If you're just looking at allocatable variant sets, maybe tens of labels. If you're looking at blocked as well, then we're looking at thousands of variant labels. Thank you.

DONNA AUSTIN: Okay. All right. Thanks, Sarmad. Okay, Maxim.

EN

MAXIM ALZOBA: Just a formal note for the avoidance of confusion. If we're speaking about two sets of strings, X and Y, and one set of strings, X, contains strings which are variants to each other, and the second set of strings, Y, contains strings which are variants of each other in Y, then in situation where any string from X is a variant of any string of Y, effectively, is equal to a situation where X and Y, all strings are variants of each other, just for avoidance of doubt. Thanks.

DONNA AUSTIN: Thanks, Maxim. Edmon, go ahead.

EDMON CHUNG: Just quickly, yes. I agree with Maxim. But we are also talking about confusing these visually similar strings. In those cases, the sets could be a little bit different because a variant may or may not be visually similar. So just want to clarify that difference there.

DONNA AUSTIN: Thanks, Edmon. I'm wondering the best say forward here. I'm wondering whether this is a good place to stop this as a first conversation on this topic and give folks some time to digest what the staff paper proposal is and what Jeff has proposed.

I think, from a pragmatic and an operational perspective, if there's some way that whatever recommendation we come up with can be narrowed—and I think that's what Jeff was trying to do is take

EN

out some of the complexity and size of the problem—then let's see if there's a way to do that. But I think, for now, there's a lot of information to think about and I don't know that we'll get to far in continuing this conversation. I think we just need to let this set. It's the first time we've had a look at this.

And then, when we come back to it, I remember Ariel did a wonderful slide back in the beginning where she took us through a process. I'm wondering if there's some way we can try to capture this. Jeff talked about a matrix but I wonder if there's another way we can visualize this for folks to see if we can come up with a different way of doing this. But Sarmad, go ahead.

SARMAD HUSSAIN: Thank you. I just made a comment in the chat. Basically, I think, just to probably help explain the problem, we have X on the application side, and allocatable variants of X, and blocked variants of X. So those are three different sets. And they can be potentially compared with just Y, or allocatable variants of Y, or also blocked variants of Y.

So the group, of course, needs to discuss which sets to pick on the X side and which sets to pick on the Y side for the string similarity comparison. Is it going to be just X compared to Y, or X plus alloc-X with Y plus alloc-Y and so on? So there can be many possibilities and that's the answer we're trying to find. Thank you.

DONNA AUSTIN: Thank, Sarmad. Jeff?

EN

JEFF NEUMAN: Yeah. I have a question for Sarmad, I guess. When you were talking about the blocked variants, and you were saying that if there was an application for a string that happened to be—for another string—or a variant happened to be confusingly similar with a blocked string, you were making it sound like the registry was essentially getting around … Well, first you had said, "The community has made a determination that they don't want the blocked string or anything similar to the blocked string."

I'm testing that out to say if the community didn't want the blocked string, then wouldn't the community have also then made a statement that if anything was confusingly similar to it, it didn't want that blocked string?

I guess I'm not seeing what the issue is and how it's a workaround because if it's confusingly similar to a blocked string, then it would be blocked. It wouldn't be allowed to go forward. But if it was just a variant of a blocked string but not confusingly similar—which I don't even know if there's an example of that happening. But if the community really didn't want it, as you say, then why wouldn't the community have also blocked the variant? And if the community hasn't blocked the variant, then couldn't we say that the community has specifically chosen not to block the variant under your theory? Therefore, why do we care that it goes through?

You made the assumption that if the community blocked a string, it intended to block all variants. But I'm making the other assumption that if the community wanted to block all the variants,

EN

it would have done so. Thanks. Sorry. It might have been confusing, what I asked. But I'm testing the assumption, Sarmad.

DONNA AUSTIN: Yeah. Go ahead, Sarmad.

SARMAD HUSSAIN: Can I try to respond to that? I'm not sure I completely understood the question but I'll try to still maybe come up with an example. Again, this is off the top of my head so let's see if this works. I'm not sure whether it exactly fits this situation but think of a string like … Let's say there's a Latin string which is blocked. Is it, for example, possible for somebody to apply for a string which looks exactly the same or is similar to it in Cyrillic script, for example?

So, I guess, what would be the answer to that? If the Cyrillic string is similar to the Latin string which is blocked, should the Cyrillic string pass through? And if it shouldn't, then could that be extrapolated to a blocked variant example? Thank you.

DONNA AUSTIN: Okay. So very quickly, Jeff and Michael. I would like to try to wrap this up so we can move on.

JEFF NEUMAN: Sure. Thanks Donna. To respond to Sarmad, if it's confusingly similar, then absolutely, it should be blocked. If it just so happens to be a variant that's not confusingly similar, it should be allowed. I don't think we should be taking the strategy that anything other

than what's specifically blocked should be blocked. So others may disagree with me but the only thing that should be prevented are strings that are confusingly similar, not the fact that they're variants and not confusingly similar, if that makes sense. Thanks.

DONNA AUSTIN: Okay. Michael?

MICHAEL KARAKASH: Yes. To Sarmad's question, if the label is the same or very much the same to the blocked variant, then it should have been considered a variant itself and would have been blocked itself, due to the fact that it should be a variant, like with ASCII and with Latin and Cyrillic. But if it's just confusingly similar to a blocked variant but not confusingly similar to the original label, then I do not see the need to prohibit the registration of it because it's obviously not similar enough to the original label to be considered confusingly similar to that label, just to the blocked label. So that's my view. Thanks.

DONNA AUSTIN: Thanks, Michael. And thanks, everybody, for the discussion. I think it's been a really good first discussion. And I think if we could try to—as Jeff suggested, matrix. If Ariel can do her magic in putting process in picture that makes it easier for us to understand, we will come back to this and see if we can make it a little bit easier to understand where the challenging pieces are and whether we can find a breakthrough here.

EN

I think we need to be mindful. The conservative approach could break the bank and be extremely difficult from a process and operational perspective. So how can we perhaps constrain that a little bit, which I think is what Jeff is trying to do—put some parameters around it to make the process a little bit easier to work through.

So Ariel and the team have said they're willing to take on that task, to try to break this down for us. I'm not sure whether we'll have that ready to come back to you next week but we'll try and come back to it soon so that we don't lose the good discussion we've had today.

All right. So with that, I'll hand it back to Ariel. I think we've got one other item to try to get through here today. So, Ariel.

ARIEL LIANG: Thanks, Donna. I'm not sure whether we could discuss the next one because the next question is regarding the potential consequences of the result of the string similarity review. But perhaps I can at least introduce this question and provide a context. Then that can be considered together with the previous two questions.

The question is E3a. After requested variant string is rejected as a result of string similarity review, should the other variant strings in the same variant set remain allocatable? In other words, should individual labels be allowed to have different outcomes/actions— e.g. some labels be blocked and some be allowed to continue with the application process? So that's the question. Again, we are

EN

focusing on the future new gTLD application aspect, not existing gTLDs, for this question.

In the staff paper, basically it laid out two possible outcomes for the string similarity review—the rejected string. One outcome is only the applied-for variant string is rejected and then the other variant labels in the set can continue to remain allocatable. Then the second potential outcome is that if an applied-for variant string is rejected, then the entire variant set is rejected. So basically, it laid out these two possible outcomes for consideration.

I don't know whether the team is at the place yet to talk about this question but at least we know what this question is about and the context of it. Then I will stop here.

DONNA AUSTIN: Thanks, Ariel. Maybe we can have a theoretical discussion around this and see if anyone has any thoughts on it, notwithstanding that we need to come back to our previous discussion. But maybe we can consider this from the perspective, if we go with the conservative staff approach, what would be the consequence here. In my mind, I don't actually … I haven't got in my mind what Jeff's proposal really was but I wonder if folks have any thoughts on this. Jeff, go ahead.

JEFF NEUMAN: Yeah. Thanks. So in my proposal, the only thing that is … I guess I shouldn't use "blocked." The only thing that's rejected is the variant that is actually found to be confusingly similar. All of the other variants, if they're not confusingly similar to whatever it was

that caused the first variant to be rejected, then those should remain allocatable. I think that that's the whole crux of my proposal—that the only thing that's prevented from being delegated is something that is actually deemed to be confusingly similar.

So that goes along with what I was saying to Sarmad, that if variant of X is a variant of Y, but the variant of X is not confusingly similar to the variant of Y, then both should be allowed to be delegated, even if to different owners. That was the whole crux of my proposal.

DONNA AUSTIN: Thanks, Jeff. Anyone else have any thoughts on this? Edmon and then Justine.

EDMON CHUNG: Thank you, Donna. I think, in general, I don't want to come across as trying to disagree with Jeff too much. I think there are a few things that we need to think through. One of them that came to mind is the concept of atomicity of the IDN variant set. If we treat an application, or a TLD, or a domain as one atomic whole, then we really should look at the entire set as one whole. That's where my reservations are.

Once we try to split it up and say different things might happen to different parts of it, then we run into a potential bigger issue in how we conceptualize these variants. So I think maybe, at the end of the day, it doesn't matter. But I think I'm just throwing up these

EN

questions to caution us as we think about it, especially with the atomicity principal as well.

DONNA AUSTIN: Thanks, Edmon. I just want to clarify. Maybe you need to put this into chat because I'm not sure I understood completely what you said. Atomicity?

EDMOND CHUNG: Yeah. The idea is that it's inseparable. The whole set is inseparable with the applied-for string. Atomicity, yes.

DONNA AUSTIN: Right. Okay. So it's a whole. It's not something you can separate and break down. It's the full set. Okay. Justine?

JUSTINE CHEW: Thanks, Donna. I'm just going to comment in my personal capacity, obviously. I understand where Edmon is coming from in terms of the concept of atomicity. But I'm just reminded that string similarity is a visual test. So I don't know how you can justify rejecting a string if it's not visually confusingly similar to another string. So regardless of whether it's in the set or not, if two strings are not visually similar, such that it raises a risk of confusing people, then why should we reject one or the other, depending on what you're applying for and what you're trying to reject. That's where I'm coming from. Thanks.

DONNA AUSTIN: Thanks, Justine. The confusingly similar does have a visual test, which some might say is a subjective test. But to Edmon's point, if it's … I like the way that Satish put this is the indivisibility of the set. Does that hold in the confusingly similar discussion? Okay, Justine. Is that a new hand?

JUSTINE CHEW: Yes, it is.

DONNA AUSTIN: Okay. Go ahead.

JUSTINE CHEW: Sorry. I forgot to include a qualification for the first bullet. Say I were to agree with only where the applied-for string is rejected. The other allocatable variants continue to remain allocatable unless those other allocatable variants are also found confusingly similar. So there's a proviso behind that as well. Thanks.

DONNA AUSTIN: Thanks, Justine. Jeff?

JEFF NEUMAN: Yeah. I really plus-one what Justine said. I think we're not protecting anything other than confusingly similar. If we did so, we're almost getting—not almost—we're getting into content and we're judging that things that may have equivalent meanings should be blocked or whatever.

But we don't do that in ASCII. If someone wanted to have shoes and another person wanted sneakers, we would allow that because that's not confusingly similar. Just because they could mean the same thing … Well, they don't. I used a bad example. But just because they may mean the same thing doesn't mean we block those two strings that mean the same thing. So applying that to variants, we should allow variants unless, as Justine said, it's confusingly similar.

So yeah. Edmon put an example in. But Edmon, in the variant situation you described there, shoes, shoes (lowercase), and all that kind of stuff, that would absolutely be confusingly similar. I don't think a panel would find those not to be confusingly similar. I do believe that they are visually similar enough in the standard that we've adopted.

If we want to change what's confusingly similar and that standard with respect to variants, that's one thing. But I do believe that confusingly similar is the standard we've adopted for ASCII script and that should be the same for all other scripts. If it's not confusingly similar in the meaning in which ICANN has established, then we let it go.

And Edmon, remember. There's not just … I don't think we're talking about it now. And Ariel, you stop me. I think it's another question, anyway. But remember, there's a right to challenge under a string confusion objection. So I think I'll stop there because I think we will get to that. But just remember that we have that as a failsafe as well. Thanks.

DONNA AUSTIN: Okay. This has been interesting. I think this has been a good discussion around this charter question as well. Maybe, when we put the two proposals side by side … To Edmon's point about the—and I'm not going to say it correctly—atomicity of the label and the variants, and considering them as a whole, is that an important principle? What happens if we break that? I think that's probably an important question for us as well, to understand the consequence of doing that. Maybe there is no consequence and maybe it will be okay. But perhaps there is that bigger question that we should think about as well.

Okay. So good discussion. Ariel's going to do some magic and when we come back to this discussion again, we'll have a little bit more texture, I suppose, to consider.

We've got six more minutes left. I just want to go back to the beginning of the call. I suggested that because this time next week is going to be in conflict with the GNSO Council session, there's a proposal that we move this call two hours earlier. So we start at UTC 11:30 and go for the 90 minutes.

I just wanted to … I think there was some agreement to that at the start of the call but I wonder if folks could use their checkboxes—the tick or the cross—to see whether that would work. We can't vote here. There you go. Thanks, Michael. That's why I can't see it. So can folks just put in chat whether they agree to 11:30 on the 14th? Maxim, go ahead.

MAXIM ALZOBA: I fail to find this but I agree. Thanks.

EN

DONNA AUSTIN: Thank you, Maxim. All right.

JUSTINE CHEW: Maybe it's better to ask if anybody has any issues with it?

DONNA AUSTIN: Yeah. I was just thinking that myself, Justine. We'll put this to the list because I know I only sprang it on you 90 minutes ago and people need to consider their calendars. So we will put this to the list. If we don't have a lot of objection, then I think we'll go forward with this time.

Also, I'm mindful that this is a representative group. So if we have one or two that can attend from each of the community groups, then I think we can go forward with that as well. So we'll put it to the list for objection. But I think, based on what I'm seeing here, that we are good to go for next week at 11:30 UTC. So thanks, everybody. Good discussion today. Ariel, go ahead.

ARIEL LIANG: Thanks, Donna. I guess we are going to the AOB section. I just want to remind folks about two items. Is that okay?

DONNA AUSTIN: Go ahead. Yeah. Absolutely.

EN

ARIEL LIANG: Okay. Thank you. One reminder is about the A5 and A6 redline added. So there is a second reading going on and we encourage the members to review the proposed new redline on the Google Doc and to provide your input and comment. The deadline for your review is this Friday. So hopefully we can receive your input by tomorrow.

The second reminder is that staff also circulated a proposal to reach out to the Chinese, Japanese, Korean generation panels regarding the single-character TLD question. The proposal is on the list and we'd like to hear your feedback as well, by Friday. And with your input, we will try to finalize that proposal and present again to the team before we do that official outreach. So there are two items being circulated on the list.

DONNA AUSTIN: Thanks, Ariel. Justine, not to put you on the spot but could you— because my brain's gone. Can you remind me what we intend to do with A5 and A6 after tomorrow's deadline?

JUSTINE CHEW: If my memory serves me correctly, the leadership will just have a look at what other comments have come in after the closing time tomorrow. And then the leadership will decide what to do with those comments. But we don't think we will bring that back to a call for discussion.

DONNA AUSTIN: Okay. Thanks, Justine. So we've got a minute to spare. Thank you, everybody. We will see you a couple of hours earlier next week.

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 rest your day.

[END OF TRANSCRIPTION]