EN

ICANN Transcription

IDNs EPDP

Thursday, 27 April 2023 at 13:00 UTC

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

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 IDN's EPDP call taking place on Thursday, 27 April 2023 at 13:00 UTC. We do have apologies from Alan Barrett and Maxim Alzoba, and Edmon Chung will be joining us late.

All members and participants will be promoted to panelists for today's call. Members and participants, when using the chat, please select everyone in order for everyone to see the chat and so it is captured in the recording. Observers will remain as an attendee and have view-only chat access.

Statements of interest must be kept up to date. If anyone has any updates to share, please raise your hand or speak up now. All documentation and information can be found in the IDNs EPDP wiki space. Recordings will be posted shortly after the end of the call. Please remember to state your name before speaking for the

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.

transcript. As a reminder, those who take part in the ICANN monthly stakeholder process are to comply with the expected standards of behavior. Thank you, and over to our chair, Donna Austin, please begin.

DONNA AUSTIN:

Thank you, Devan, and welcome, everybody, to today's call. It looks like we have light attendance today. So I think what we'll do is push on anyway because I'd really like to get through the first part of the agenda, but if we don't, if the attendance doesn't pick up by the time we were going to get to the IDN table presentation, we may have a chat about whether to defer that to another time. But we'll see how we go.

So hopefully, folks have had a chance to have a look at what Ariel sent to the list, which is basically an analysis of where we think the phase two charter questions sit in terms of the board request and we'll go through that today. We're not going to go through it in a lot of detail, we're just going to talk about the high-level analysis that Ariel has largely done and see if we can agree or get some common understanding of what we think our project plan might look like and what we can reasonably tell the board at ICANN 77, or actually tell the GNSO council and then they can tell the board.

One of the things I did want to get some feedback on today is the meeting time for these calls. We're going to be pushing on for another 12 months, as seems to be the case to get through phase one by wrapping up our initial report and making that a final report and getting into phase two charter questions.

I was wondering whether folks would have any objection to starting these meetings an hour earlier. So, pretty much because it's late here for me and I find that I'm struggling a little bit to get through the second hour. So, what I don't know is what other commitments people have if we move it up an hour earlier and whether we're going to lose people. So, just interested whether folks are aware of conflicts that would happen if we did move it up an hour earlier. So, any feedback from people? Would that be okay if we did that or is that going to create us problems?

Okay, so it's okay with Satish. Yeah, Michael, two hours early would be great but I realize that if we go two hours earlier, we may lose some of our staff support on the West Coast, the US. So, if we move it one hour then it might be a little bit more palatable. Okay, so Jerry's okay. Nigel's okay. Lisa, Jennifer, Dennis, Anil, any thoughts? And Pitinan and Farell as well. Jennifer's okay. Dennis is good. Pitinan's okay.

Okay. So it looks like that starting an hour earlier will be okay with folks. So we'll still do two-hour meetings, but it just means that we'll be starting at UTC 12 instead of UTC 13. We'll confirm that on the list, because I appreciate there's quite a few people who haven't joined us on this call today. And we're not meeting next week so it would be from probably the 11th that we would start at UTC 12.

The other thing that Ariel has just reminded the leadership team of is part of PDP 3.0, which is a GNSO initiative to try to improve or enhance the PDP process, is that there is supposed to be a self-assessment at some point during the life of the PDP but after the initial report. So that's something that we may be kicking off in the

next couple of weeks. There's some work that has to be done to get the survey into a Survey Monkey, but that's something we will probably do in probably I would say the next four weeks or so.

Alrighty, with that, we will get started on the analysis of the phase two charter question. Okay. So you're good to drive, Ariel?

ARIEL LIANG: Yep. So, this is a quick reminder of the board resolution passed in ICANN 77. So the board asks us to develop a project plan that identifies all charter questions that will impact the next applicant guidebook, and also with consideration to ensure consistency with the ccPDP4's deliberation. And then as part of the project plan, provide a timeline by when this group will deliver the relevant recommendations to the GNSO Council.

We highlighted, identified all charter questions that will impact the next AGB, because this is the focus now for today's meeting. And I'm hoping we can get the group's input on analysis. This is also a recap of the analysis staff did. And then those bullet points drive our thinking. So, it's based on several assumptions.

The first assumption is for each charter question we, the group will develop corresponding recommendations that will result in a change to the status quo. So status quo means if there is a change to the 2012 AGB then there's a substantive change to that, or there be a change in the application question that didn't appear in the 2012 round but in the next round there will be a new kind of application corresponding to that particular charter question.

And then for a lot of charter questions, I guess all of them, if not every single one, they will have contractual obligation impacts. So, we're assuming that the group will develop recommendation that may change, for example, the registry agreement, registrar accreditation agreement, other consensus policies or temporary policies. So, basically we're assuming the worst case scenario from implementation standpoint that the group will develop recommendation that change the status quo. So that's the first major assumption we had in this analysis.

And the second assumption is that we don't know how exactly the next AGB will look like, but we chatted with the GDS team that's responsible for developing the next AGB. What we learned is that they will use the 2012 AGB as the basis for developing the next one, so we could expect to see similar structures and some content may be similar. So, that's another key assumption we made. And when we did analysis we checked against the 2012 AGB and tried to map the charter questions to the relevant section numbers. And then in that way we can explain why we think there's an impact or not.

And then certainly about the contractual obligation. So, as I noticed earlier, it's not just limited to the registry agreement. We also considered other agreements, policies, and procedures that may have a legal effect. So, it's in a generous sense we call it contractual obligation. We didn't say RA impact specifically.

And then the fourth bullet point is a new one that I didn't include in this slide, circulated in the list because we just got that input from Karen Lentz's team, is that if we identify any question that may have an impact on the registry agreement, in fact, it will have an

impact on the AGB as well because due to the ICANN procedure, the next AGB and the updated registry agreement will be published together for public comment. So, basically, they will go through the same public comment process. And so, they're basically a package deal. So, if we identify any substantive updates to the registry agreement, it will basically have an impact on the AGB. And that's something we take into consideration in the analysis. And then later you will see that's also reflected in our thinking for the project plan development.

And lastly, in this table, you see the highlighted question number here. These are the foundational questions that we identified in our previous analysis. And those questions may take four meetings for completing the deliberation. So, that's a quick overview.

And then in the table here, you will see in one big picture how the impact may look like. So, top is our preliminary thinking regarding to impact to the next AGB. In fact, the majority of the questions probably do not have a direct impact on the next AGB. But I think the majority of the foundational questions actually do. So, we put yes.

And then for application question, that's just what we think in terms of whether it's kind of appropriate for the applicant to answer any specific question that may be related to the charter question topic. And we believe the majority is either yes or maybe. It may impact how they answer the application question based on whatever recommendation this group developed.

And then in terms of contractual obligation, again, our thinking is almost all of the questions do have a contractual obligation impact, only with the exception of one, which is C4A, because we believe this already being covered by SubPro. And it actually won't result in any change to contractual agreements.

So, that's a quick overview here. And I'd like to just quickly pause here and see whether there's any questions or comments or concerns about this analysis at the high level. Donna

DONNA AUSTIN: Just to let folks know that the idea behind this analysis is so that we can maybe—like we did with the chunking before, which resulted in Phase 1 and Phase 2, that we may need to reorder some of the sequence of the charter questions to help us get through those that are related to the AGB or the registry agreement. But I think you can continue, Ariel.

ARIEL LIANG: Okay, sounds good. Thanks, Donna. And I want to provide a quick update on the analysis we did, because after checking with Karen's team, we made some change to our analysis to three charter questions, which are D5, G1, and G1A. So D5 is about the fees related to variant domain transaction. It's the $0.25 paid by the registry operator and the $0.18 paid by the registrar, I believe, so for each transaction.

So initially, we put no for the next AGB impact because we believe that this is only a reference to Section 5.4.1 regarding the expectation for a registry operator, that one of the expectations is

that they must pay reoccurring fees to ICANN, so it's only a reference. But the main, I guess, requirement is included in the registry agreement and/or registrar accreditation agreement. Actually, I couldn't find the reference to the $0.18 in the RAA, but if anybody is familiar where that $0.18 should be mentioned, I will check the relevant documents.

So in summary, our previous assessment is it should be a no to the next AGB impact, but then as what we chatted with Karen's team, if there's any substantive change to the registry agreement would occur, then it will actually be an impact to the next AGB as well. And then we did note that Section 6.1 in the registry agreement is regarding the registry level fees, and that concerns the $0.25 that's covered under this topic.

So if this group recommends any substantive change, not in terms of how much it will be changed, but based on what's the principle, like whether each transaction occurs, that fee will apply or it's applied to the value of label set, if we make that kind of recommendation, that could fundamentally change how the registry level fee is being charged, then there will be an impact to the registry agreement, of course, and then that will have a tripling impact to the next AGB. So that's why we updated the impact assessment of next AGB to maybe for D5. So that's a quick kind of update for D5.

And then for G1, that's regarding the IDN implementation guidelines. So basically, it's a foundational question for the group to discuss because there are some challenges regarding how it is being updated, especially for the recent version 4.0 and 4.1. And then initially, we thought there shouldn't be an impact to the next

EN

AGB because similar to D5, AGB may only reference the IDN implementation guideline in its section 1.3.2, that's regarding IDN tables. And then that's the part that mentions compliance with IDN implementation guidelines.

But if the group does recommend any substantive or foundational change to how the IDN implementation guidelines may be updated in the future, and also its enforceability, for example, then it will have a change to the registry agreement, specifically Specification 6. That's where the IDN implementation guidelines is mentioned. And then in the same vein, it will have a trickling effect on the next AGB as well. So that's why we updated G1 to maybe for next AGB impact.

And then similarly for G1A, that's a sub-question under G1 regarding whether a separate legal mechanism should be created for IDN related contractual obligation. So if EPDP does recommend any creation of a separate legal mechanism, then we foresee there will be a substantive update to the registry agreement, and then it will impact the AGB as well. So we updated next AGB to maybe as well for G1A.

So these are the kind of new updates we did between the time we circulated the presentation to today. So I just want to pause for a moment and see whether there's any comments or input from this group on this analysis.

Edmon has a comment. The base agreement is not in the AGB, and this seems to impact the base agreement but not the AGB. So, yeah, that's correct, Edmon. And we learned about this from Karen's team, is that the base agreement and AGB will be

EN

published together through public comment. And so if we identify any substantive change to the registry agreement, it basically needs to kind of—it has a trickling effect on the AGB. So basically what the board asks us to do is identify the questions that may impact the AGB. But in fact, we also need to identify the questions that may impact the registry agreement in a substantive manner because they will be published together before the next round. So that's why it's safer to include the questions that will also have a substantive impact on the registry agreement in our analysis. Yes, and I think that's the difference in terms of expectation for the next round. Yeah, thanks, Nigel. Yes.

Any other questions or comments? And I'm not seeing any, and I guess we could move on to the next point, which is regarding the project plan development. So based on our analysis, we try to chunk our phase two work into two parts. And then the chunk one includes the questions that are foundational questions and also the ones that will have impact on AGB as well as substantive impact on the registry agreement. So we listed these questions on the left, and we also reordered it in the way we believe it may be logical or easier for the group to tackle.

So C4, C5, C6, all three questions are regarding the IDN table, and we already found the relevant requirements in the 2012 round AGB. There's a whole section about IDN table, and we believe if this group recommends anything regarding the harmonization requirement or the format requirement, then we expect that will go into the AGB. So we put C4, C5, C6, that's the first three questions to tackle. And then C1, C2, C3, and C3a, these are all related to the same entity requirements. And then there are also

foundational questions. And then also part of it is the mechanism to identify the same registrants that we have to discuss ROID or alternative mechanism. And then we put this as the second segment in chunk one for the group to discuss. And then also we will need some data, even just anecdotal evidence from registry operators regarding how variant labels are managed and how the registrants are identified.

And then after the second segment, C2 part 2, C4, C4a. So these are kind of related, but they're mainly about the domain name lifecycle of variant labels. So C2 part 2 is actually specifically about the activation of the second level variant domains, whether any change needs to be made to the current policy. And then C4, C4a, they're related to whether the lifecycle has to be the same for each domain from the variant label set. So this is also related, and we think this is the third segment that seems to make sense to tackle one after another.

Oh, actually, I should mention C2 part 1 is actually kind of more logical to tackle together with C1 because they're talking about the same entity and using registrants as the definition for same entity at the second level. So I just wanted to quickly mention we kind of parsed C2 into two parts.

Following that is D8, that's the catch-all question, but that's something Edmon brought up. It's regarding the WHOIS, the registry WHOIS and IANA WHOIS, whether any change needs to be made to include the data related to the variant label set. I think it probably stands by itself as a segment to tackle, and there's quite a lot of foundational information we need to catch up on.

So D8 is the next segment. And then D5, that's what I mentioned regarding the fees regarding each transaction. So we do believe it's [inaudible] in case any substantive change needs to be made to the RA to deal with this in chunk one. And then G1, G1a about implementation guidelines, that's similarly as what I noted before. If we make any foundational changes, then there will be a change to the RA most likely, and then we also need to deal with it in chunk one.

So as you see, the majority of the questions in phase two belong to chunk one in our analysis. And this is a quick reminder how many meetings we preliminarily allocated to tackle each question. And the total is 38 meetings to deal with chunk one questions. And we also allocated 11 meetings as contingency buffer. And also because we have the majority that belong to chunk one, we allocated the majority of the contingency meetings to chunk one as well. So nine meetings as contingency buffer. So that's a total of 47 meetings for chunk one deliberation. And it's only about the deliberation part.

So on the right side is about chunk two, that's the ones we believe that shouldn't have a direct impact on the AGB or even the registry agreement. And just a quick summary of what they are. It's D6 about transfer and D7, F1, D6A, D7A, and F2, they're all about— oh sorry, D7 is suspension policy procedure. And then F1, D6A, D7A, and F2, they're all about RPM, especially UDRP or other RPMs.

And the reason is that when we check the AGB in the 2012 round, those policies and procedures were referenced in certain sections, but the meat of the content resides elsewhere. They're not

EN

included in the AGB. They're separate documents. So if there's any change, it will be change directly applied to those documents. And even if the registry agreement did mention, for example, RPMs—but again, it's not in the registry agreement itself, it's another reference or link to a separate document or procedure.

And then also, they're mostly about the registry operator's obligation, like post-launch obligation or pre-launch obligation. So it's like post-delegation issues. And then we believe maybe in the application, the applicant may not need to elaborate too much. They just basically need to comply with these obligations. That's why we believe this chunk can be dealt after chunk one. And then we put it the total of 11 meetings for addressing all these questions. And then we give two meetings that's contingency buffer. That's chunk two. And then I see Sarmad has his hand up.

SARMAD HUSSAIN: Thank you. Just a quick, I guess, point on D6, which perhaps if I'm not wrong, deals with UDRP or sorry, D6A as well. So I guess the question is when eventually we go into operation and there is a dispute and because of that dispute or for some other reason, one of the domain names in a variant set is transferred. Then, of course, to keep the set the same or to not break the set, the other variants need to be transferred as well.

I understand what you're saying is that that should perhaps be dealt with in other policy. But from a registry agreement point of view, or I guess just how it needs to be operationalized, I guess the question is, will that then be addressed timely through the other policy and IDN EPDP will not, I guess, stay quiet on that? Or

EN

is IDN EPDP going to, I guess, take that up at a later time? Just trying to clarify where this work will eventually need to be done. Thank you.

ARIEL LIANG: Thanks. So the other PDP that deals with RPM is the RPM PDP. And then they do have a phase two being planned focusing on UDRP. But then at the same time, the GNSO council just passed a resolution to defer the initiation of phase two for at least 18 months. I believe that's to wait for a phase one implementation to complete. So if this group gets to the UDRP question specifically related to variant labels, I believe it's within our scope and remit to go ahead and deliberate on that particular piece. And then that's kind of consistent with how we kind of deliberate on the phase one questions, because a lot are kind of required coordination with the SubPro IRT. But obviously, it didn't happen until now. And we still went ahead and was the deliberation and based on assumption how certain SubPro recommendations are going to be implemented. So I hope I'm answering your question here.

SARMAD HUSSAIN: Just a quick, I guess, follow-up question. In case it is something which IDN EPDP should or would do, should that be also then part of the orange list here rather than the green list? Thank you.

ARIEL LIANG: I'm sorry, I'm not sure I'm following. So basically, the orange and green talk about how we sequence our deliberation. So we deal with orange ones first as the chunk one, and then we will deal with

chunk two next after completion of chunk one. So that's just a proposed sequence how we address the charter questions. They all need to be addressed.

SARMAD HUSSAIN: Sure, I guess I was perhaps looking at how the green is labeled as that it has no AGB or RA impact. I'm, I guess, saying that there may actually be an AGB or an RA impact in that context. I'm not sure, but I guess that's something I'm trying to raise. Thank you.

ARIEL LIANG: Okay, yeah. Thanks, Sarmad. And yeah, maybe it's not precise, but we circulated our analysis in the list. So what we mean by no AGB impact is that there's we don't expect substantive edits to the actual content in the AGB or the RA because they're mainly just referencing these policies and procedures that are separate from RA and AGB and they provide a link, for example, to the RPMs and then we don't think the change will be applied directly to the AGB because the content can remain the same. And it's just if there's any change, it will be applied directly to those policies and procedures. So that's what we meant here. So there may be no direct impact, but there may be indirect impact because, yeah the registry operators have to comply with these policies and rules and procedures post-delegation. So, yeah, but disregard the thing. And we're just trying to summarize what we mean by chunk one, chunk two, and why the sequence makes sense.

DONNA AUSTIN:	So, just a reminder that the reason we are doing this chunking exercise is to determine whether we think there would be value in doing a phase 2 A and a phase 2 B. So the intent with phase 2 A would be to be able to get through the work so that we're not holding up the next round. And then maybe we could do the second part sometime after. But I think my personal feeling here is that given what Ariel laid out in front of us, I don't see any value in doing a phase 2 A and a phase 2 B and having two initial reports and two public comment periods, because that will just—to Nigel's point about this is a lot of meetings. Yes, it is. There are some ways that we can potentially make the timeframe shorter. As I mentioned last week, we've been talking about whether we should do a face to face meeting. But we know that it will take us six months to get the agreement from ICANN to do that. So there are tools available to us to try to get through the meetings quicker. But I think for the purpose of this exercise, what we're trying to understand is whether it makes sense to do this as a phase 2 A and a phase 2 B or whether we should do it as one lot.

Based on Ariel's analysis, I'm leaning towards we should do this as one lot. And when we respond to the board requests that's come through the council about give us a timeline and a project plan, what that will look like is this is the order in which we will do the charter questions. And this is the timeline for how long we think it will take us to get there.

And when we do that timeline, we will do it on the assumption that we are meeting once a week for two hours a week. We can put caveats around it that there may be ways for us to increase, to get through this process quicker if we do a face to face meeting or if

we start meeting twice a week. But for the purposes of this exercise, this is what we think it will be.

Now, some of you may remember that when we did the updated project plan to council with where we chunked phase one and phase two, phase two timeline was looking at the end of 2025. And I think when the board saw that they thought that's not tenable, is it possible to get done quicker?

I would absolutely think that we can get this done by the end of 2024. But what you're seeing in front of you now on the screen doesn't take into account that we still need to finalize the work that we've done on phase one. That will take us quite a bit of time to get through. Well, depending on the comments that we get back on the initial report, that will take us some time to get there.

So we need to factor that in as well. But for the purpose of this exercise, what we're trying to do is get a better understanding of how long it's going to take us to get through those questions that we think will impact the AGB or, as we just learned, thus the registry agreement as well.

Given that most of it is going to sit in phase 2A, I don't see any reason to have a phase 2B. I think we just do this as a, it's just going to be phase two. We'll get through it as quickly as we can, but the timeline and project plan that we provide to the board, or the council, they will provide it to the Board, is going to be based on a phase two complete set. Unless anybody has any different thoughts on that. But I think as chair of this group I'm very cognizant of the fact that we're getting into the third year of our work. So it's already been a pretty heavy commitment from this

team. And to ask for another 12 months or two years is personally I think unreasonable to think that we can keep the team together for that long. So if we can try to provide a timeline that takes us through to the end of 2024, even that I think is, is a little bit too long, but that we should have a target date.

So what Ariel's laid out in front of us takes us past 12 months, because I think it's 38 weeks and 11 weeks. But we should perhaps think about setting ourselves a drop dead target date for getting this stuff done as well, and see if we can compress the work in some way that we can meet that target date.

So, any thoughts from folks on this? Is anybody intending to retire in the next six months and we're going to lose you? Or what are the red flags for folks with this, the concerns? Ariel?

ARIEL LIANG: Yeah, I raised my hand not to retire from the group. I want to quickly mention about—when we did this more detailed analysis or estimation of time needed for deliberating each question, it actually didn't deviate much from our initial estimation. When we submitted the project change request to the council, we actually allocated 60 meetings for deliberating the phase two trial questions altogether.

So, at that time, we said perhaps the final report for phase two will be submitted to the council in November 2025. And now we did this more detailed analysis. We calculated, and then the final report submission would possibly be in October 2025. So it's only

one month difference there. So I just want to quickly note, our estimation didn't change that much after our exercise.

As to what Donna said, if we do want to put a drop that date, or deadline for us to complete the work earlier than we need to, factoring other factors, for example, more frequency in terms of meeting, and then also face to face meeting, if we can tackle 12 meetings in three days, then that will help speed up a big chunk of the work. So if we could do that, then that's possible that we can finish our work earlier than what we originally expected.

DONNA AUSTIN: Yeah. Thanks, Ariel. And something that I'm also very mindful of is, even if we decided to have two meetings a week, that becomes an extremely heavy lift for Ariel and the staff support to be able to get the information ready to have two meetings a week. So, I just see that Nigel plans to retire next week.

So while we can say, well, we'll just have two meetings a week, sometimes that doesn't become an efficiency. It becomes a drag because we don't have the information that we need to continue with the meeting and we end up deferring.

So my preferred MO is that we use the ICANN meetings as opportunities to meet face to face and we have been accommodated for the DC meeting and that we will have a meeting every day. So we have four meetings and hopefully we'll get through some substantive discussion around some of the early phase two questions.

Yeah, that's right, Dennis, getting stakeholder input is time consuming and it's hard. But I suspect that—and I'd be interested to hear from Edmon whether there—I know that there's a meeting going on in Brussels at the moment, but if there is a drop dead date that we should work towards.

Now, having said that, what I want to lay out for the council is pretty much the analysis that Ariel has done. So it's X number of meetings. We're doing one meeting a week. So it's going to take us this number of weeks. That's what it looks like. This is the order of the charter questions.

And then in addition to that, we can say that there are things that we can do to try to beat that October 2025 final report deadline, such as face to face meetings and having dedicated working sessions at ICANN meetings. So we can push that there. But what we will lay out is pretty much it's going to look like October 2025 from what Ariel's just told us.

So as the leadership team, we will push the envelope on getting as much meeting time as we can during ICANN meetings. You know, what are we doing there if it's not to progress policy work? So we will push for that. We've been able to do it in DC, but it is a policy forum. So I think there was probably a bit more leniency to get us that time. I'm not sure it's going to happen for the October meeting.

So, Edmon, any input from you on whether there's any revised board thinking on the request or just does anybody know if is there any scuttlebutt on when they're hoping to have the AGB published?

EN

EDMON CHUNG: Thank you, Donna. Edmon here. I'm actually here at the board workshop. Nothing in terms of a particular update, but as I listen in, I think a lot of things make sense. The plan, as laid out in Cancun, is to obtain the community input for the timelines and then ICANN Org will help build a critical path that would lead us to that timeline.

So the input from here and also some of the other projects that were identified would come—hopefully we will receive that by the DC meeting, and then thereafter, if we get everything, all the timelines that were asked for, then the target is to have an integrated timeline with the critical plan by August. So I think that is still the plan.

DONNA AUSTIN: Thanks, Edmon. All right. So what I propose we do is that based on the analysis that Ariel has laid out for us, and I don't see any value whatsoever in doing a Phase 2a and a Phase 2b, given Phase 2b at the moment is only identified as 11 meetings. I don't see any reason to split this into a 2a and 2b. So we'll go forward with developing a Phase 2 timeline and project plan based on Ariel's analysis. We'll bring that back to the group. And we will talk about it again before we get to ICANN 77, although that time is running out too. And then so that everybody's on board and we can have that ready to provide it to the council at ICANN 77.

So if folks are okay with that as a plan, I think we'll go ahead and do that. All right. And it looks like option one is the way to go.

Yeah, Edmon, I understand the challenges with ICANN meetings, so we are looking at the possibility of a face-to-face meeting as well, but the intel we have on that is that it will take six months to get agreement to do that.

Ariel, was there anything else that you wanted to add or were you in good shape?

ARIEL LIANG: I think we're good to go. And in the leadership meeting, I'll just confirm approach how we're going to provide that update to the council in terms of the project plan and timeline. And I don't know whether it will help if we, for example, just putting our calculation, like reduce the number of meeting deliberation just to see how it takes us, like what day it will take us. I don't know, but we can discuss the approach, because I was just thinking like if we're building the project plans, requesting face-to-face, and then we say that will cut down our timeline by like 30% for example, what would be the initial report delivery date and final report delivery date. And maybe we can provide some estimate as our best effort. And then maybe that will help the council understand the value of having a face-to-face and reduce the overall length of our deliberation. But we can discuss the detail with leadership team.

DONNA AUSTIN: Okay. Thanks, Ariel. So I know we have a couple of people who need to leave the call now. So Dennis and Nigel, what we were going to do in the second half of this call was to have Sarmad run

through the IDN tables at the second level, because that's going to be the first thing that we focus on.

I think we still have critical mass here that we can continue with that, but Dennis, I'm pretty confident that you're okay with the IDN tables at the second level, but is there any reason why we shouldn't do that? We can just go ahead with that now and we can have that available. The recording will be available to folks who couldn't join the call. It is a pretty light attendance today though. Justine, any thoughts from you on whether we go ahead now? Okay. Dennis can listen to the recording. All right, let's go.

Samad, if you're okay, let's go ahead and do the run through I suppose. We are very light on in attendance.

SARMAD HUSSAIN: Sure, Donna. Can we share the presentation? I hope it's okay, we'll go through the same presentation which we used earlier. Most of the content is actually the same. All right, so let's get started with this.

So we'll just take you through the concept of IDN tables. Because, of course, some of the charter questions which are coming up will be discussing some topics which are around IDN tables. From a terminology point of view, we've actually been referring to IDN tables in multiple ways. Over time, this terminology has evolved.

Initially, IDN tables and variant tables was used as a means to refer to this collection of data, which allows I guess anyone, registry operator or a registrant or anyone who's relevant and interested to see actually how a domain name label actually can

be formed and what are some of the conditions which are put on them.

Eventually, we developed a much more formal mechanism. Actually, even the earlier mechanisms were reasonably formal, but then we actually developed even a more formal mechanism. We'll get into those in a little more detail later. And then that more formal mechanism we started calling label generation rules.

So when you see these terms being used, please consider them as interchangeable. There's actually no real difference in usage. There may be some implication, for example, when you call an IDN table label generation rule in how those rules or that data in IDN table is encoded, but that's an internal detail. So these are equivalent terms.

So what do IDN tables contain? What are IDN tables? There have been rules to develop domain names and those rules have existed from very early on. Initially, many of those rules were actually encoded as part of the RFCs themselves. So for example, in ASCII, there are actually rules which say that domain names can only be formed using letters A through Z, capital or small, digits 0 through 9 and hyphen. And there is also a rule which says the domain name can be no longer than 63. And then there are also rules that, for example, a domain name cannot start with a hyphen or end with a hyphen and so on. But for ASCII, this has sort of been just part of the—somehow defined in different RFCs.

So there are two kinds of things which are getting defined here, right. There is actually a list of characters, which should be allowed. And then there are some rules which apply to those

characters. And as you can see that both these things have existed even before IDNs, even for ASCII, these actually existed.

And then for top level domain, they are actually a more special category of a label. And so, for example, there's been a expectation that top level domains really should be limited to letters and digits and hyphens should be used in top level domains and so on. And so, again, that was also part of an RFC definition. So these rules were really not explicitly listed in any IDN table as such.

So when we're talking about IDN tables, we're talking about the repertoire. Repertoire means what should be included. And it's best to see, for example, this in the context of ASCII, which I guess we are much more familiar with. The domain names which we currently use also define our repertoire. As I said when we're using ASCII, ASCII standard actually is much larger. It contains all these other characters we are familiar with and we use every day. So like a percent sign and a plus sign or a bracket and so on.

But for domain names, it is defined, the repertoire defined for domain names includes only the letters, digits and hyphen, which are of course listed here. And we, as we talked about, this is normally for general domain names, but for top level domain names, the repertoire is even more restricted. It basically says that it's only the alphabet and not the digits and hyphen, as far as ASCII domain names are concerned.

So that's where we're coming from. And when we get into internationalized domain names, we have to define the same kind of things for internationalized domain names as we have defined

for ASCII. So we have this large encodings in Unicode, which encodes all the different letters of all the different scripts.

And the question, it's the same question, set of questions, what are the letters which should be allowed inside a domain name? And like ASCII, we also have to define those letters or shortlist those letters and we call that the repertoire. And that, of course, work need to be done. For ASCII, it's available, but for other languages, it needs to be defined. And of course, that's what is one of the things which would go into an IDN table.

And then, in addition to that, there are also these other rules which would need to be defined. So in ASCII, for example, there are rules around hyphen. In other languages, are there any additional rules around any other characters which need to be defined so that the labels which are formed are sort of intact?

And when we go to these other scripts, another set of rules or type of information which is also needed is whether any of those characters which are shortlisted in the repertoire, whether any of those are variants of each other, meaning whether they may be actually considered same by the end-user community. And if they're considered the same, then those need to be identified, because if they're not identified and they just go into the repertoire and then create different domain names that may actually potentially cause an end-user security issue. Because you will actually have two different domain names, which will be seen by the end-user as the same domain names and it can potentially cause people to click on links which they don't intend to go to.

EN

So IDNs for the second level would need to—and also for the top level—would need to define these rules like they're done for ASCII. The rules for the second level need to match the letter-digit-hyphen scheme and the rules for the top level need to follow the letter principle as is done in the ASCII space as well. Root zone LGR or label generation rules define the rules for the top level, and then I guess what registries do, and also I guess ICANN has been supporting that process by defining reference label generation rules, is to define their own rules to use or determine what are good and valid domain name labels at the second level. So basically, IDN tables are ways of determining what is a good label at a second level.

Okay, so objective of IDN tables is to enable second level domain names in local languages and scripts used by communities globally, but to make sure that it is done in a secure and stable manner, so as to not put end-users and registrants in any harm's way.

And then eventually once the IDN table is defined, what happens is that if an end-user wants to register a label T1, they would put it somewhere in a system with a registrar or some registry or somewhere else, where they would want to check whether that particular label T1 can be registered. It will be for under the relevant TLD, there will be actually an IDN table for that particular script and language, and the system will run the label through that IDN table and come up with some responses.

One response could be that based on the IDN table, the response will say that your label T1 is not valid and cannot be registered under that particular TLD. And that's perfectly fine, depends on the

EN

TLD policy and also other things, for example IDNA 2008 standard and so on.

Also, the other answer could be that the label you're applying for T1 is perfectly valid under the TLD through the IDN table, and it says you're good to go as far as registering it is concerned, for example. It can also provide you additional input as that this label actually can create—it has some variant labels. And so it can give you actually a list of variant labels and also perhaps tell you whether some of those variants are registrable, I guess, so they're allocatable variants, quote unquote, the terminology we've been using, and some of them actually are blocked. Donna, I see your hand up, let me stop.

DONNA AUSTIN: So in terms of the IDN table and whether the name can be registered. So who decides that? Is it the registry or is it the registrar or is the IDN table independent of the registry?

SARMAD HUSSAIN: For the second level, it is decided by the registry, the TLD.

DONNA AUSTIN: And at the moment, some registry operators don't have the same IDN table.

DONNA AUSTIN: They could or could not, and there could be a good reason for it as well. So, that should be possible, that's driven by the business

EN

need. So there can be two kinds of differences between one TLD and another TLD. One difference is that some TLD is offering X languages and another TLD is offering Y languages, so they can choose a different set of languages and scripts they're offering. That's one difference.

The second difference could be that within the same script or language, they could actually have a slightly different IDN table. And that could be perfectly fine as well. So, for example, suppose you're in a country which actually has an immigrant population which comes from a certain other country and it actually extends the writing system in certain ways for that country, then having just a table which not only includes local language, but also language of the immigrants which are coming in, there could actually be a potential extended IDN table for that country, which may not, for example, be applicable to another country.

So I'm saying it's sort of a business decision for each registry to decide who their target audience is, and based on that, they can design a table. The only, I guess, requirement is that IDN tables can be different as long as they are secure and stable for the end users.

DONNA AUSTIN: Okay, thanks.

SARMAD HUSSAIN: Okay, so any other questions till here? Otherwise we move forward. So IDN table is eventually telling whether a label is valid under that TLD and if it is valid, whether it has any variants.

So, as far as the decision on whether it's valid or not, it will look at repertoire and rules and then whether the label form is within that repertoire and follows those rules. We will look at that in a little bit as well.

So, an IDN table defines which labels can be registered for a particular language or a script at the second level under a particular top level TLD. So we just discussed that. And the purpose of an IDN table is to manage security, stability and usability of IDNs. Security, stability, of course, and then we were also talking about usability that from a usability point of view, obviously that's a business decision. If a TLD is targeting a particular population, they would design an IDN table which is usable for that particular population. So they can actually, since target audience can change, the design for an IDN table could also change.

So IDN tables include Unicode. So this is sort of what an IDN table includes at a high level. So IDN table includes Unicode code points. So, of course, we're using local scripts and languages. So it's no longer ASCII. We're using Unicode, which encodes all these different scripts.

And it includes a few things which are listed in blue. So it includes a language and script tag. So when an IDN table is being defined by the registry operator or for a TLD, they will say that, okay, they can put a tag which can contain two kinds of information. They can make a language-based table or a script-based table. That's up to them. And then they actually have to specify that. So they can say that this is a table for German language in Latin script, or they can say this is an Arabic language table for Arabic script.

EN

They can even choose and say that we're going to make a generic table for all Arabic script, irrespective of language. So they will skip the language information and say, this is a table for script. Which can then support all the different languages which are used to write—which are written in Arabic script.

Metadata in description. So this is not part of the formal definition of the—actually, normative definition, but this is more like anybody can put in any kind of textual description of the table in there. And that's normally just I guess for consumption of people who are just reading it. So it's more for human consumption, I guess.

Then come the more formative parts or more normative parts of the IDN table. One is the repertoire. So this is the list of characters which are allowed. We talked about that earlier. That that has to be clearly specified what can create a valid domain name. So if you list a list set of characters here, then only those characters could be used to create a domain name under the TLD for that script or for that through that IDN table.

And if you put in any additional character, then it will create an invalid label. So it's very important that we list all the different characters which could be used. We also identify whether there are any variant characters. Variant characters can be motivated by visual similarity, but also, for example, semantic similarity as in Chinese. So they could be different reasons why a community wants to do some variants, but normally that is defined by the community because it's what the community thinks are the same characters. And then they could actually be rules. And rules can vary significantly from script to script. There are sometimes more

significant rules in complex writing systems. We've seen in South Asia or East Asia, for example. But there can be all kinds of rules.

And then I guess what we've highlighted here is that the rules can come from two sources. There's sometimes technical kind of rules and then sometimes they're linguistically oriented or motivated. An example of a technical rule is that RFC 5891 says that Unicode string must not contain hyphen in third and fourth position. That's because you have this X and dash dash and so on. So they're saying that a Unicode string shouldn't have it. And that's sort of a technical restriction with IDNA 2008 standard has put in.

But from a linguistic point of view, there are different characters, for example, can be used in different contexts. And if they're not using those contexts, they can potentially create rendering issues and therefore end user security problems and so on. So there are those kind of rules which communities can describe on top of the technical rules and they have described through the process we've been following

So IDN tables. So we have now to put in all this information in a file. And I guess the question is, then how do we put this information in a file. There have been earlier formats which are described in RFC 3743 and then RFC 4290. So they were slightly different formats which created different Slightly different formats which created text files where you could just go and put this data in, the data which we listed here in blue. The language and script tag metadata, which can, for example, include the date of publication and those kind of things. Then the more normative part which is the repertoire variants and rules.

And In 3743 and 4290, the repertoire, they were normally listed as a list of code points. So it was in certain ways machine readable. Similarly, variants were also listed as pairs. So they were also in some ways machine readable, but the rules, interestingly, which were articulated in the earlier formats, RFC 3743 and 4290, they were normally English description of the rule, which meant that a machine could not really read and automatically parse that rule and therefore largely these IDN tables were for human consumption. They really could not be read automatically by machines and humans would read them and then code software to actually implement these rules and so on.

And that was, of course, okay, but eventually RFC 7940, the new standard was developed in which actually this last bit where the rules were also encoded in a proper systematic fashion in a way that it actually can be processed by machines, rather than just describing the rule as it's done on the screen in English. And that's the main advantage of moving to RFC 7940, is that now IDN tables are completely processable by machines. Somebody defines an IDN table. We can write a program which can process it without really having a human to interpret it to see what the rules for example are.

So moving on, then. This is what we talked about RFC 3743 which was an informational RFC encoded the whole repertoire variants and rules in text format. The repertoire and variants, one could still process by machine but rules were not machine readable.

RFC 4290 again did the same thing, rules were normally in a comment section and not readable by machines because again

they were explained or described in English. Or in a human language, not formal syntax.

And RFC 7940 which is now a standards track RFC allows everything to be done in XML format and it's completely machine readable and machine possible and that's a significant advantage.

Other reason here is because when you're writing a rule in English, if two different human beings actually read it, they can actually interpret that rule in a different way. But when you actually encode such rules in machine readable formats, then it becomes much more consistent to interpret and implement these rules and it's no longer left to the interpretation of the reader and that makes it a bit more consistent as well.

This is just visually sharing some of the examples. If you look at the bottom comment, the code points you can see are written in machine—is a list of code points and then it also gives multiple columns of these code points are actually variants. So you could actually read code points and their variants from this list.

But if you look at the rule part in the comments up above them, it says code point 002D hyphen minus based on a reference from RFC rules. The rule is the label must neither start nor end with 002D. So basically a machine cannot parse this English text. It doesn't really know what to do with it. It will just ignore it and therefore a human will need to read it and then code this in software. This is the earlier format.

Again, repertoire and variants in this next format is also machine readable to some extent, but the rule, as you can see, is again

explained in text and really needs to be interpreted by human reader and put into software code. And it's again open to interpretation.

Whereas when we move to RFC 7940, the code points, the variants and the rules are all very well defined in form of a formal specification which are completely machine processable.

Again, these are some of the same things which I just said, why XML format? XML format is used in the LGR format, which is in RFC 7940. Rules are more precise than writing them in plain English, prevents misinterpretation and promotes consistent use of rules by different parties. The LGR can be developed, reviewed and compared automatically.

So when we are comparing two tables, for example, so if you have one IDN table and another party is using the same IDN table, you could compare those tables, but the rule, you will never be able to compare the rules. But in this case, if they're two different parties or if you're two different tables, we can actually compare them if it's an XML format and say whether they are exactly the same or not.

So because of those properties, the LGRs are much easier to test. Basically, one could say that maybe the rules—it's just too [stringent] a format and not everything in a language then can be encoded and we really need open or flexible English language to really encode the rule, but we've actually now encoded 26 scripts in root zone LGR with all the different kinds of rules coming in from the different communities and we actually have been able to encode it. So that's an example that this more formal mechanism

is also flexible enough to encode all the different kind of rules required by the different kind of scripts.

Of course, one criticism is that this format becomes much more complex. This is not really very readable by humans. And we agree to that. And because of that, what we've done is we've actually developed HTML versions of these XML files. What that does is that it automatically creates an HTML version of an XML, which becomes very human readable. And what we've been doing is when we create these now LGRs, we release both the HTML version which is automatically generated and the XML version. And XML versions for machines, but the HTML version is for the humans to read. It's just as any other normal HTML file which you can actually just read through your browser and reads well. You don't really see all this XML kind of stuff in there anymore.

So who develops IDN tables? This is a question which was asked just by Donna recently. So as we discussed, IDN tables for the second level are developed and used by the registries for the second level. So they are actually developed by the registries themselves based on their own business requirements.

Registries may refer to reference LGRs when developing the writing and tables. Reference LGS are being published by ICANN for the different scripts. And they're being published by us, but they're not developed by us. We actually go to the relevant script community, seek their input and based on that we develop it and then we obviously take it through a public comment process, address any input from the community at large. And then based on community input, finalize them and publish them.

So again, reference LGRs are only a reference point. It doesn't mean that everybody has to follow them based on their own business decisions. Registry may still decide to design their or develop their IDN tables differently from reference LGRs. That's perfectly okay as long as the security and stability aspects are not compromised.

So that's how the IDN tables work. Of course, the concept of IDN tables, quote unquote, also applies for root zone. The root zone, we have one big IDN table, which is called the root zone LGR. It is also developed by the community process. And of course, currently now policy is being developed to use that as a mechanism to develop, define the valid top level domains and identify their variants. And the procedure link is provided here. As I said, this has really been a community based process.

So here are some examples of what an IDN table may have. So this is a very mini IDN table which says that it allows five codepoints. I can design an IDN table as smaller or as large as I would like.

So this IDN table says that only hyphen, B, C, I and dotless I are allowed. So that's a repertoire. And then in addition to the repertoire, it's defining that I and dotless I are variants and going from I to dotless I is a blocked variant, whereas going from dotless I to I is an allocatable variant. We'll see how that plays. And then there's a rule which says hyphen cannot be at the start position of a label. So it's a very simple IDN table.

If we give the label BBC as an input to this IDN table, when we run it, BBC's label is allowed through this IDN table, because there are

no rules rejecting it. And there is, of course, both B and C are part of the repertoire. When we give it CIO as an input, that would not go through because C and I are there, but O is not included in the repertoire. So the IDN table will reject this label saying that O is not a code point that is supported.

IF we give it a hyphen BBB as an input label, even though hyphen BBB are all part of the repertoire, this label is still rejected because it violates the rule that hyphen cannot be at the start position of the label. So because of that, this particular label is rejected by the IDN table. So this is actually how the IDN table works. It allows some labels and it disallows some other labels.

Looking at this example a little more, we now have some labels. We've already looked at BBC. Let's look at CBI. So CBI is also allowed. The repertoire covers all the characters in this label. As well as there are no rules which are being violated. So it is valid. But in addition to that, what is interesting to note is that I maps to dotless I and when it maps that, it creates a CB dotless I variant. And when we are going in that direction from dotted I to dotless I, if you look at the variant definition, it says that it would create a blocked variant. So CB dotless I is therefore a blocked variant.

Similarly, if you take C I C dotless I, you can work through this example, it will say that the original label which we normally now call primary, I guess, is valid and is allowed. It also creates one allocatable variant and two block variants and you can actually work that example out through this set of rules here. Let me stop here and see if you have any questions so far. I don't hear. So let's move on.

So IDN tables may generate allocatable and blocked variant labels. We are all now very familiar with that. Allocatable variants can be activated if allowed by the registry policy and requested by the registrant. And IDN guidelines 4.0 suggests a mechanism for automatically activating variants as well by registries without the request for registrant as long as the script community allows for it. Again, IDN guidelines are now currently getting implemented, part of them through as IDN guidelines 4.1 and then some part of those guidelines have now been deferred to IDN EPDP to look at and see whether you those are good guidelines to move forward.

But one more thing is that IDN guidelines 4.0 suggests that all activated variant labels should be registered by the same registrant, the same entity to prevent user confusion. This is a similar principle as the same entity principle for the top-level domains, I guess applicable to the second level.

One of the questions which has been raised, and this is one of the recommendations from IDN or guidelines from the IDN guidelines, which is now going to be considered by IDN EPDP so it's no longer part of IDN implementation guidelines 4.1, is whether these IDN tables would need to be harmonized.

And I guess we'll spend some time to understand what harmonization is because eventually this is something which this group will need to decide on whether that is something which you will need to be done.

Okay, so a TLD can implement more than one IDN table. So that's one another thing to note, that under a TLD, there can be one IDN table or two or 20 or 50 or 100. And we actually have examples of

TLDs which have tens of IDN tables being implemented under the same TLD.

So there's an IDN guideline in version 4.0 which says IDN tables under one particular TLD should be harmonized. And it's also one of the recommendations in the IDN variant TLD recommendation that IDN tables under TLD and also its variant TLDs should be harmonized. So, so those are two, I guess, levels of harmonization, which is needed.

Let's try to first understand what harmonization would mean. So basically IDN guidelines version 4.0 explains harmonization as two measures. These two measures are suggested to prevent cases of IDN variant labels being generated by different IDN tables under the same TLD to be allocated to different registrants.

So, point is they can—we just looked at or we just discussed, I guess, what I've said is that more than one IDN table can be implemented by an IDN under a TLD. So a TLD could, for example, be implemented, for example, a German table and French table and Spanish table. These are three different IDN tables it may be implementing, and you can, of course, understand that all of these three tables are going to be in Latin script.

And I guess one of the things which this is saying is the two IDN variant codepoints or IDN variant codepoint sequences. So if you have two different IDN codepoints in one IDN table, the two codepoints in one IDN table, which are considered variants in that IDN table, then they should not be considered non-variants through another IDN table.

So suppose I apply for string T1, which contains some characters and suppose I run it through, let's say a German table. And it creates two strings saying that you have T1 and then T1V1 and these two are variants of each other. Under the same TLD, if I run it through a German table, but under the same TLD, if I run it through the, let's say, the French table or the Spanish table, and if the table says that, okay, T1 is fine and T1V1 is fine, but they're no longer variants, then it creates a challenge that if I apply it through the German table under the same TLD, the same two labels are variants, but under the same TLD, if I apply through the Spanish table, let's say they're not variants.

And that actually can be a little problematic because once the registration is done, the end user doesn't really know whether the registration was done through the Spanish table or the German table. They will just look at the labels. And in some cases, they could be registered to the same people. In other cases, they could be registered to different people because they're no longer variants. And the difference then can be arbitrary and not predictable and then can degrade the end user experience. There's a hand up by Ariel. Let me stop here.

ARIEL LIANG: Yeah, thanks. Just for my professional curiosity, I mean, maybe curiosity is not the right word, but I'm trying to understand exactly what harmonization means. So based on what you said, does harmonization mean that the variant rules or the variant relations must be consistent based on the script of the IDN table? So at the second level, IDN tables can serve different languages, but as long as the script is the same for these languages, then the

EN

variant relations or the variant label rules must be consistent. Is that what essentially harmonization means?

SARMAD HUSSAIN: That's exactly part of the harmonization. There's another part of it which goes across scripts.

ARIEL LIANG: Okay, I'd love to hear the detail because I think that's a key point for the group to deliberate on the harmonization question. So I appreciate some more elaboration.

SARMAD HUSSAIN: So there is a part which pertains to a single script and that's the example I gave you. And then there is another part of which goes across script. So when you are doing, for example, Latin and Cyrillic, where many characters actually look the same. You can actually have a letter A in ASCII or letter E or some other, which actually have a very almost exact—actually not almost in some cases, exact equivalent in for example Cyrillic script. And that, of course, needs to be managed across script then. And that's sort of another example.

So basically, let me give you actual examples here, which were coded here. So let's suppose there's an Arabic script top level domain and it offers two tables. One for Arabic language and one for Urdu language. And that's perfectly fine, right. If somebody is offering an Arabic table, they can actually say that, okay, I have an audience in Middle East, North Africa. So we'll do an IDN table for

Arabic language and then say that, okay, we have a separate audience and let's say South Asia. And for them, we are actually going to do an Urdu language IDN table.

And they define these two tables differently for their own audiences. And you can see that this last letter in each table, this character is slightly different. There's 064A as far as Arabic is concerned, and it's 06CC for Urdu language and that's perfectly fine. That's how these things are typed in those communities and that's why those Unicodes are used.

So far, so good. But when they actually register a label and these are two independent IDN tables, even the code points are not the same. But when they actually, somebody applies for this 628, 64A, 641 character, top left, that goes through the Arabic table and Arabic language table says this is perfectly fine, move forward.

Similarly, when somebody applies in South Asia with the top, sorry, left bottom examples, 628, 6CC and 641, that perfectly goes through the Urdu language table and that is also registered. And each of them, because of these two tables are not harmonized and they are independent of each other. So this is like the German and Spanish example. Both those labels which are considered or created are actually totally independent of each other, right.

But look at those labels, they are visually identical. And from an end user perspective, both of these are being offered under the same TLD. And there is nothing wrong in the current process where you can actually have an Arabic language table under the TLD and Urdu language table under the TLD and they can

actually be applied. And currently, the process will allow for these two labels to go through and be registered. But when they get registered under the same TLD and they're visually identical, it creates problems for end users, because this can be used for phishing and for other purposes.

So what we are saying is when you're doing these multiple tables, these tables need to be harmonized. So then what happens if you do the harmonization of these two tables? When we do the harmonization of these two tables, then what happens is that we say that we include this extra letter in in this merged table.

There are multiple ways of harmonizing, by the way, there's no one mechanism, you can actually harmonize it in different ways. So this is only one way of harmonizing that you actually create a merged table to see that whether there are any other variants inside from one table to the other table. That's one way of doing it. One can also actually do other ways of harmonization.

So at the end of the day, what we are looking at is the result should be consistent, the actual process of harmonization could vary and there could be more than one process. But what we are saying here is, so one mechanism is that we create a union of all the tables, which is the merged table, right. We have two sets, Arabic set and the Urdu set, we create a union of that set to create a union table, which we call the merged table and then that union table will then have both the characters and we know that those two are actually variants. So we actually add that variant definition and then use the merged table to see whether the new—so if the Arabic version was applied first, when the Urdu table is applied, it will first go through the Urdu table and we'll say it's okay.

But before we say yes, it can be registered, we run it again through the merged table first and we see that it actually creates a variant which is already registered through Arabic and therefore harmonization will block the second label to be registered.

So harmonization prevents registering potential variants, which can, I guess, creep through if the harmonization is not done and two different IDN tables are, for example, used. There was a question in the chat that, "But ICANN when reviewing such IDN tables can eliminate one, correct?"

So the answer to that is that no, we cannot eliminate because when we review, we review one IDN table at a time. Currently, harmonization, unless if harmonization is approved as a policy, then we can approve multiple tables at the same time. What we do is when we are given an Urdu table or an Arabic table and they could be given together or at different times, we will only look at Urdu table or the Arabic table or German table or Spanish table, for example, individually.

Obviously, we understand that there is a problem of harmonization, but the current process and current policy does not allow that kind of review at this time. I hope that answers your question. Okay, thank you.

Okay, so this was an example of within script harmonization coming back to, Ariel, your comment, that is the same script and we need to do harmonization. But then this is the example of cross script harmonization. Suppose a top level domain has audiences in Cyrillic script region as well as Latin region. So they offer two IDN tables, one in Cyrillic and one in Latin and people

can come and register a label, either through the Cyrillic table or through the Latin table.

And if there's no harmonization being done across these IDN tables, in this case, in different scripts, you will see that each of those strings which are listed on the top on the top left and bottom left, one in Latin, the top one and the left bottom one in Cyrillic, they're identical, but they will go through these IDN tables and will be registered independently of each other to different registrants.

And that, of course, you can obviously see that you will then have these two labels under the same TLD exactly visually identical labels under the same TLD, even though they're coming from different scripts and that is also obviously a problem which should be resolved and is a problem which can be resolved through harmonization.

So again, if we create a common merged table, a union of all the tables, we then identify that there are cross script variants here which need to be blocked. And if that's the case, and the process of harmonization is integrated, then if the Latin one is registered first, it will be registered and then when the Cyrillic one comes in, it will first go through the Cyrillic table in this particular case, in this particular implementation and the Cyrillic table will say that it's good to go. But then it will as a second step come through the merged table and it will identify that it has a variant, which is the Latin and the Latin is already registered. So it will block the Cyrillic string, which is equivalent to Latin.

So the harmonization basically stops this within script issue and cross script issue, which can potentially cause security issues for

EN

end users. So let me stop here and see whether you have any questions on harmonization and why it's needed and what it prevents and, for example, how it could be done. Yes, Ariel, please.

ARIEL LIANG: I don't want to monopolize the questions. I mean, I'm also learning in the process and I'm just wondering in terms of the variant relations, what is allocatable, what is blocked. Is that decided by the registry operators individually? Especially, for example, the cross script variant rules. How is that decided? Is that based on what the RO believe should be blocked or what should be allocatable where there are some the reference table, for example? You mentioned that they can check the suggestions there. I'm just wondering how that's done.

SARMAD HUSSAIN: Yeah. So as far as the variant sets are concerned, we normally consult with the community which uses the script to define the variant sets. Because that pertains to the security issue because if something's a variant, it's considered the same, then it should be identified as variant and perhaps not allocated to different registrants.

But whether that variant should be allocatable or blocked is generally left to the business model of the registry, of course. And, however, obviously, there is this SSAC recommendation that number of allocatable variant labels should be reduced. And that's

EN

something, of course, a registry should take into account when they decide how many variants to make allocatable, of course.

But it is really up to the registry. And Satish is making a comment. So yes, when it is a mixed label, it is not allowed to be allocatable. I was, I guess, talking in the context of same script labels and their variants. That part is decided by the registry. There's actually an IDN guideline which says that mixed script labels are blocked. They shouldn't be allowed for registration. So that part, Satish, you're right. That part is not allocatable. So mixed scripts are automatically blocked, but within the same script, whether they're allocatable or blocked, that's left to the registry to decide. Any other questions? Yes, Ariel.

ARIEL LIANG: I just want to know that Donna had a comment in the chat. It'll be helpful for us to understand the ICANN processes associated with IDN tables. And I know there's an IDN table update project. We only have seven minutes left. If it's not enough time to cover that, but maybe in another occasion, we can learn more about the IDN table update project just to understand how it's reviewed historically, that probably will help for our deliberation.

SARMAD HUSSAIN: Sure. So let's then continue. And so we are actually at ICANN, based on community input, creating reference label generation rules, just for registries to, I guess, understand community requirements as they develop their own IDN tables. And we're making more and more of them available through—you probably

saw a recent public comment process and we'll be now finalizing those and publishing them as well.

IDN table review, the way it's done. This was a question right now. Basically, a generic top level domain registry operator intending to offer registrations in different languages and scripts must be approved to offer IDN service for languages and scripts. And what they do is they develop to do that. What they do is they develop IDN tables and share those with ICANN for review. And then ICANN takes them through a review process, looking at some of these things we talked about from a security stability standpoint. Realizing, of course, and appreciating that IDN tables actually can be different based on business models.

And they are then after review approved and they are actually added to Exhibit A of the registry agreement. So each registry has approved IDN tables and different registries can actually have a different set of approved IDN tables, depending on what they apply for.

IDN tables are reviewed through multiple processes. They are actually reviewed. So when you're applying for a new gTLD, you can actually, with the application, you can submit IDN tables. And those are reviewed during the registry system testing. It's called pre-delegation testing, PDT process. You could actually also apply for additional IDN tables or new IDN tables at a later time after delegation and then that goes through an IDN service, which is sort of a specialized registry service evaluation policy process. And then also when sometimes ROs are changing, backend operators or registry service providers, it also goes through the RST process which looks at—and in that process we review IDN

tables as well because sometimes IDN tables are changing because the registry operator is adopting the tables of the receiving RSP rather than the ones which they were implementing earlier.

All the IDN tables are published online in the IANA repository, whatever is approved and ccTLDs are also encouraged to publish their own IDN tables they implement in addition to the gTLD registries. And what we've now done is to make it a very consistent and transparent process, we've actually made a tool available online. And what it does is, if you're a registry operator, you will develop or design your IDN table. And you can upload it in the tool and choose the reference LGR which we've developed for that particular language or script. And you can, just on a click of a button, it gives you a report on whether the IDN table meets the security stability requirements.

When we check internally within ICANN, we use exactly the same tool and see exactly the same report. So now it is possible for registry operators to check even before they submit the table to ICANN and we also get exactly the same view as they are getting. So it's a now completely open and transparent process.

And the criteria for evaluation is the reference LGR with the rules and the variant relations and the repertoire checked based on what the community said is the right solution. The community which uses a script. So reference LGRs are developed by the community. We are only checking for security stability issues. If they are different in design, but there's no security stability issue, the report you get from this tool will say that this is good to go. So it's not checking that it is exactly the same. It can be different. And

EN

the tool doesn't really say anything about that. But if there is a potential security and stability issue, then the tool will identify that. So that's what is online and the IDN table project actually which you were referring to, we've run that over the last couple of years based on registry stakeholder input that we need to make the IDN table review process consistent and transparent and it is now completely open, consistent and transparent based on of course reference LGR definition provided by the community.

And I think that takes us to the end of the slides. So thank you very much. And let me hand it back to you, Ariel and Donna and Justine.

DONNA AUSTIN: Thank you. I see that Justine has a question in chat. So is the reference LGR the same product of a GP that goes into the root zone LGR?

SARMAD HUSSAIN: So the answer, as Michael said, is no, because the design parameters for second level are slightly different from top level. If you remember when we were looking at ASCII as well, top level is just letters and second level can be letters and digits and hyphen. So, so the rules for second level and top level are slightly different. So when we are actually developing reference LGRs for the second level, we make sure that we use the second level rules, design rules, rather than just the Roots zone level rules. Thank you.

DONNA AUSTIN: Thanks, Sarmad. So just a reminder that there's no call next week because I think we clash with the GNSO Council meeting. So enjoy the week off and thanks to Marika for filling in today. I know that we don't have Emily and Steve. So Marika has joined us to, I assume, take on the note taking responsibilities.

Okay. Thanks, everybody. The leadership team will do some work on the timeline and project plan, which we'll circle back with the group maybe in a couple of weeks. And also we'll put a note out about the changed time for these calls. And if there's no objection, we'll move forward with UTC 12. Great. Thanks, everybody. We'll see you in a couple of weeks.

DEVAN REED: Thank you all for joining. Have a wonderful rest of your day.

[END OF TRANSCRIPTION]