EN

ICANN Transcription

IDNs EPDP Community Webinar

Wednesday, 17 May 2023 at 11:00 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/HIdXDg

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 community webinar taking place on Wednesday, 17 May 2023 at 11:00 UTC.

Please note this session is being recorded and is governed by the ICANN expected standards of behavior. IDNs EPDP members and participants will be promoted to panelists for today's session. If you would like to speak during this session, please raise your hand in Zoom. When called upon, attendees will be given permission to unmute in Zoom. We have Arabic and Chinese interpretation on today's call via the Adigo audio bridge. You can find the login details on the wiki page posted in chat. Please remember to state your name for the recording and speak clearly and slowly for interpretation purposes. Please keep your microphones muted when not speaking to avoid any background noise.

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.

Recordings will be posted on the public wiki page shortly after the meeting. All are welcome to use the chat. Please note that private chats are only possible among the panelists in Zoom webinar format. Any message sent by a panelist or standard attendee to another standard attendee will also be seen by the session's hosts, co-hosts, and other panelists. With that, I hand the floor over to Donna Austin.

DONNA AUSTIN:
Thanks very much, Devan, and thanks to everybody that has joined the call today. Myself and Justine Chew, who Justine is the vice chair of the IDN EPDP, I'm the chair of the IDN EPDP, and we're hosting this webinar today to take you through some of the recommendations in the IDN EPDP phase one initial report, which was recently posted for public comment. We thought it might be helpful to just go through some of the background and the recommendations and provide an opportunity for the community to ask us any questions that you might have that will help with providing any comments you might have to the public comment period.

What we'll do today, we'll go through a little bit of background to the introduction of IDNs and understanding variants, and then go through the IDN-related GNSO policy activities, the phase one initial reports, preliminary recommendations, and we're just going to give an overview and a highlight of some of the recommendations that we think might be helpful for us to expand on a little bit, and then just a bit of a reminder about the public comment process. We'll have a Q&A session at the end, but if there's any questions that you have along the way, please raise

your hand. Hopefully, we'll be able to see that, and we can take those questions at the time, or as I think Terri has put in chat, that if you want to put the question into chat, then please preface it with <question>, or <comment> if it's just a comment.

All right, so with that, we'll get into it. So next slide, please, Emily. And next slide, please. Okay, so just a bit of a timeline of the introduction of IDNs. So in 2000, IDN registrations began at the second level, and there was a document which is IDN implementation guidelines that was developed for contracted parties to follow in the management of second level IDN names. And in 2009, there was the IDN ccTLD fast track process. That was a process that was done predominantly by the GAC and the ccNSO, and it was in recognition that ccTLD operators would really benefit from having IDNs at the top level. So that was a fast-track process that was done back in 2009.

In the lead up to the new gTLD expansion in 2012, in thinking about variants at the top level, the ICANN board basically put a hold on variants for the 2012 new gTLD program, because there wasn't really an appropriate variant management solution available at that time. So I guess out of an abundance of caution, the board just decided best not to proceed with variants at that time.

So in 2012, there were IDN gTLD applications through the new gTLD program. So 116 IDN applications were received. 92 of those have been delegated. And in that application process, the applicants could self-identify variants for information purposes. But because of the board resolution, it wasn't possible to proceed with those at that time. So next slide, please, Emily.

Okay, so in 2013, the ICANN board endorsed a procedure for developing the root zone label generation rules. And this has become a pretty important process for IDNs in particular, and actually enables the introduction or is a large part in enabling the introduction of variants at the top level for gTLDs.

So in 2019, the ICANN board approved ICANN Org's recommendations for variant TLD management. And in order to, I guess, verify those or confirm those, the GNSO and ccNSO kicked off respective policy efforts on IDNs. And that's what we're doing here today with the GNSO council approved to charter for an expedited PDP on IDNs in May 2021. And we started our work in August 2021.

And just if you're wondering, what's the difference between an expedited PDP and a normal PDP, it just means that an issue report wasn't necessary, because there was a body of work that had been completed. So we didn't think—the council decided it wasn't necessary to proceed with an issues report. So the charter was developed and the EPDP team was set up as a result of that.

So in 2022, ICANN published the root zone LGR version four and IDN implementation guidelines version 4.1. So the label generation rules have gone through a number of iterations. And that's why we're at version five. And that I'll go through a little bit about the root zone LGR later.

And the IDN implementation guidelines version 4.1. The guidelines are a document that was developed by the community. And I think in the current registry agreement, the registry operators abide by IDN implementation guidelines, it's either

EN

version three or version four. But there's no formal mechanism to ratify the guidelines. So that's part of our charter questions as well.

But there were a number of parts of the implementation guidelines that the board did actually approve. But they've deferred a portion of those for consideration of our work. And that's in phase two of our work. Next slide, please, Emily.

Okay, so understanding variants, the basics. So I'll be honest. Variants, even after chairing this effort for about the last 18 months, variants are a little bit of a challenging concept for me. But basically variant labels are considered the same by the respective script communities.

So if you're looking at and the examples that are on the screen, so the simplified Chinese script and the traditional Chinese script, if you look at those strings visually, they look very the same or similar. But it's the code points that are able to differentiate between what is the simplified Chinese script and what is the traditional Chinese script.

but what hopefully IDN variants at the top level will allow is for those for those top levels that—for a top-level script, a top-level, sorry—So the top-level domain. So in addition to having a primary IDN, you'll also be able to have the variants of that. So simplified Chinese and traditional Chinese, essentially, in theory, operate as one TLD. But in practicality, it's actually the case that they will be individual TLDs in the root. But from a script community and language community perspective, the variant labels are the same.

So defining a variant for security purposes, so the example that we have there, the AAA, so the Latin script and the Cyrillic script, as you can see, visually, they are the same. And that's part of the challenge that we've—well, they're not really because the code points differentiate between the Latin script and the Cyrillic script. But for all intents and purposes, when you look at the AAA as a top level domain, visually, you wouldn't be able to tell the difference. So that's one of the challenges that we've been looking at in developing the policy. Next slide, please, Emily.

So understanding the variants and their impact. So variants exist in many scripts to serve language communities globally, potentially impacting billions of users. So while in 2012, IDNs were introduced at the top level, it's still a little bit of failing to some of the language communities because the variants of the TLD is not available. So that's one of the things that we're trying to address here. A single script can be used in multiple languages, and may be subject to variations due to how the languages work.

The DNS makes distinctions between variant labels with different code points, the script communities recognize them as being equivalent. So that if you're just looking at these, visually, those that use the script every day can basically—it might be traditional Chinese or simplified Chinese, but it's considered one and the same.

Variants may exacerbate confusion risks among labels that may or may not be visually similar, potentially causing security and stability issues in the DNS. So as I said, this is one of the challenges that we've been looking at in the work that we've done

in phase one of our work, which was primarily based on charter questions related to the top level. Next slide, please, Emily.

So why have variant gTLDs not been delegated to date? So there's no definition of a variant, but the root zone label generation rules have offered a way to have a consistent definition for ID and variant labels. So in 2012, with the last round of new gTLDs, that wasn't available. And the second gap is that there's no variant management mechanism.

So ICANN developed preliminary recommendations, so that was work done by ICANN staff, but to have it ratified by policy, that's something that the board has requested the GNSO and the ccNSO to do, and that's the work we're doing here. Next slide, please, Emily.

So the root zone label generation rules, which has become pretty important in the context of IDN discussions. So the total number of script communities or generation panels to date has been 17, and the total number of label generation rules developed is 25.

I'm not going to go into how the root zone LGRs are developed, enough to say that it's a community effort that involves not just what we would consider as ICANN community representatives, but it goes into script communities and people with additional expertise in the area. And that's been a separate effort that kicked off probably 2013, 2014. Next slide, please. Thanks, Emily.

DONNA AUSTIN: So if you want to know the variants of an IDN, the label generation rules has a tool that you can use, and this is basically the output.

So if you type in a label, as you can see here, that's identified on the top row, and then it identifies whether it's valid or not, and then it identifies the disposition. So it's either allocatable, which means it can be a TLD or it's blocked, which means that it can't. So this has been pretty important in our work as well to understand the disposition values and the impact of being allocatable and blocked. Without this, because this wasn't available in 2012, it's one of the reasons that the board didn't go ahead with variants. So next slide, please, Emily.

Okay. So another important data point for us in doing our work is understanding that not all scripts have variants. And importantly for us, identified in red, if you can see those, are the scripts that don't have variants at all. There are 22 scripts that have variants, but not allocatable variants, and that's a decision that has been made by the generation panels.

And there's only seven scripts, which is Arabic, Bengali, Chinese, Greek, Latin, Myanmar, and Tamil, that have allocatable variant scripts. And that became a pretty important data point for us when we were considering the charter question of whether to have a ceiling on the number of variants that an applicant could apply for. Next slide, please, Emily.

DONNA AUSTIN: Okay. So IDN-related GNSO policy activities, SubPro PDP, so subsequent procedures policy development process that was conducted a number of years ago now, did do some work on IDNs at the top level. And we have referred to that or affirmed some of

those recommendations, and then the work that we're doing is the EPDP IDNs. So next slide, please, Emily.

So as I said, SubPro did some work on IDNs. And then there's some, the work that we're doing which is we're looking at not just policies that could relate to the introduction of variants in the next round, but also how to factor in those IDN gTLD registry operators from 2012 and finding a path for them to apply for variants in the next round, whatever that next round might be.

The other thing that we've done is we made a decision to split our charter questions into two parts. So the first part was on IDNs at the top level. And the second phase of our work is more related to second level domains. So that's that work that we have started while the public comment period is open on phase one. Next slide, please, Emily.

So SubPro, what was discussed and what wasn't discussed, partially adopted high level ICANN Org variant management recommendations for future gTLD. So that work that was done by staff and the SubPro PDP considered that work and verified some of those.

So the important one is that the root zone LGR is the sole source of validating future gTLDs and calculating variant labels. Variant gTLDs must be managed by the same registry operator and supported by backend registry service provider. So you'll hear us talk about the same entity principle. And that's primarily what that relates to. And second level variant labels registered to the same registrant. So again, that's a same entity principle, which we've

given quite a bit of thought to about how that will work in practicality.

What SubPro does not address is whether ICANN Org recommendation should apply to existing gTLDs and second-level IDN variant domains. So that's work that sits firmly with us, how to operationalize ICANN Org recommendations and other recommendations, studies, advice related to IDNs. Okay. Next slide, please, Emily.

Okay. So it's a bit of an overview of who we are. So this IDN EPDP, the GNSO has a number of different compositions, I suppose, for policy development processes. So we're a representative plus open model. What that means is we have representatives from the different parts of ICANN and they make up representatives from the Registries Stakeholder Group, Registrars, we have ALAC, Business Constituency, the GAC. We don't, unfortunately, have representation from the SSAC, but we have had a conversation or two with the SSAC during our work and we'll have a conversation with them about our initial report in the next week or two.

We also have participants and observers. So that means that participants that are part of the representatives from various ICANN Organizations can join and be part of the discussion.

We also have observers. That basically means you sign up to the mailing list and we have liaisons across the ICANN community and board. So we have two liaisons from the ICANN board,

Edmon Chung and Alan Barrett, and we have a couple of ICANN Org liaisons. So Michael Karakash from—actually, I don't know the name of his team, but the team that's working on SubPro and SubPro implementation. Michael is part of that team, but he's assigned to look at the work or maintain an overview and an understanding of the work that we're doing.

Our role is to determine the approach for a consistent definition of variant gTLDs and the utilization of the root zone LGR. Thanks, Justine. Global domains and strategy. So GDS is the team that Michael Karakash is part of. And also part of our role is to develop policy that will allow for the introduction of variant gTLDs.

So our work is to apply SubPro recommendations to existing gTLDs and second level domains, operationalize SubPro recommendations for existing and future gTLDs and address topics not discussed by SubPro. Next slide, please, Emily.

So some of our challenges in the work that we're doing is to permit delegation of variant gTLDs that meet user needs while maintaining DNS security and stability. Allocatable and blocked variant labels introduce complexity due to their permutation. So that just means that it's really a numbers game because you've got allocatable and blocked variant labels. We're not just talking about one label. We're talking about many that are connected to one.

Charter requires coordination with the SubPro implementation review team to address overlapping issues. But unfortunately, the SubPro implementation review team, I think, is standing up for the

first time this week. So we haven't had that opportunity to discuss some of the issues with them.

So in considering our work, we've had to assume that the process for the next new gTLD program is going to be substantially similar to that of 2012. And that's what we've based much of our discussion and consideration on.

Most difficult discussions so far have been around limiting the number of variant gTLDs that can be delegated. So that's the charter question related to numbers, whether to have a ceiling for the number of variants that an applicant can apply for. And we'll come to that later. The process by which existing IDN registry operators could apply for variant gTLDs and adapting the string similarity review, which is a test of whether a string is visually confusingly similar to another, to address the introduction of variant gTLDs. So that string similarity review discussion comes back to the permutation issue that I mentioned previously.

We also have a requirement to coordinate with the ccPDP4. So the ccNSO started a policy process on IDNs around the same time that we did. Some of the questions and recommendations are similar, but some aren't because of the nature of the operation of ccTLDs and gTLDs are different. But we have met with the ccPDP4 on a number of occasions. And I think you'll find in our initial report, we identify a number of recommendations where we think we're consistent. And I think there's probably only one where we think there is a difference, but it's a difference that doesn't have a substantive impact on the policy recommendations. So I think from a board perspective, what they were looking for is to ensure that the recommendations that come out of our work and

EN

ccPDP4 are not inconsistent or cause any significant issue because the policy turns out to be significantly different. So our assessment of that, the EPDP team's assessment of recommendations to date is that they're on a pretty good path or a compatible path. And we've also got liaisons between ccPDP4 and our work. Next slide, please, Emily.

Okay, so as I mentioned, we have adopted a two-phase approach to our work. So phase one, and the initial report that we put on the 24th of April is about top level IDN variant management. Phase two, which we've actually started only last week, but we've started, is about second-level IDN variant management. At the final report for phase one, we hope to have done by November 2023. That in large part will be determined by the public comments that were receive through the initial report. So if we don't need to do a substantive rework of the recommendations that are in the initial report as a result of the public comment, then we're pretty confident we'll meet that November 2023 deadline, or what we've identified as a deadline. But if we do have to do substantive rework, then that's going to push that timeline out.

With phase two, and I don't want to spend too much time on phase two, but we do have, as I said, we've started that work. It's been identified as a dependency on kicking off the next round of new gTLDs. And we are in the process of preparing a report for council for them to share with the ICANN board at ICANN 77 about the timeline for getting phase two of that work done. Next slide, please, Emily.

Okay, so finally, we get to the reason why we're here. So what this part of the webinar will be is pretty much there's a couple of

EN

recommendations that we'll draw down on because we think they're probably the most, I won't say contentious, but probably the ones of most interest. And then we'll quickly do an overview of the others. So next slide, please, Emily.

Okay, so this is pretty important. So some underlying principles of our work. And kudos to Ariel, who's on leave, but has joined this call. So just the layout of these slides, Ariel is very good at presenting things visually to make it easier to understand.

So underlying principles for our work, so the root zone LGR is the sole source. So we've confirmed that at the second level as well. But as you can see, Ariel's identified where those, what sections relate to that. So section 4.1 and recommendation 1.1. Same entity principles. So I mentioned this previously. So this is at the top level of the DNS, the same registry operator must manage the approved labels from the variant label set of a primary gTLDs from the application, legal and operational standpoints. So just ensures that you don't have one operator that operates the primary IDN gTLDs and you have others that operates the variants, they must be held together.

And that comes down to the next principle, which is the integrity of the set. So that's a relationship between the primary label and its allocatable and blocked variant labels shall not be infringed upon as long as the primary label exists.

The conservatism. So we're very aware that the introduction of variants at the second level hasn't been done in the gTLD space before. So we have adopted a more cautious approach to the policy development as a way to limit potential security and stability

EN

risks associated with the variant label delegation. And that really came into play in the conversation around the string similarity review, which Justine will take you through a little bit later. Next slide, please, Emily.

Okay. So the root zone LGR must be the sole source to calculate variant labels and disposition values for existing delegated gTLDs from the 2012 round. So in the 2012 application process, IDN gTLD registry operators could identify their own variants. But in considering the charter question, we have decided that the root zone LGR must be the sole source to calculate those variant labels. So if an IDN gTLD operator from 2012 wants to apply for variants in a future round, they'll have to use the root zone LGR as the sole source to calculate their variant labels. Next slide, please, Emily.

Overview of the same entity principles. So this recommendation 2.1, any allocatable variant label of an existing IDN gTLD from the 2012 round as calculated by the root zone LGR can only be allocated to the registry operator of the existing IDN gTLD or withheld for possible allocation only to that registry operator. So basically anyone from 2012 that applied for and is delegated an IDN gTLD, it's only that entity that can apply for variants of that TLD. Next slide, please, Emily.

So one of the challenges with our work was trying to work out how all this would fit together in the application process. So this is pretty much an overview of the application submission, the administrative check and the initial evaluation processes of the new gTLD process and how variants will be taken into account.

So recommendations 3.1, 3.2, 3.3, 3.4, and 3.15. So you can apply for a variant label, you can apply during an application round only. You cannot precede a primary label application. So that means that you have to identify the primary and apply for the primary before you can apply for a variant. So I guess I should explain that the primary label is equivalent to the source label that is in the root zone LGR. So for the purposes of our work, what we identified is the primary label is the source label and it's the primary label that determines what the variant labels are.

So in the application process, we've said one application can cover the primary label and the variant labels and one application covers the variant labels only after the primary label has been delegated. So that accounts for future new gTLD processes but it also works for the existing IDN gTLD registry operators from 2012.

As a one-time exception for priority processing in the next round, existing registry operators that already have an IDN gTLD, our recommendation, our policy recommendation is that they have priority in the processing which means that if the existing registry operators decide to apply for the variants in the next round, they will be afforded priority in processing above all other applications. So I think SubPro gave some form of priority to IDN applications but this goes over and above that and applies only to existing registry operators that applied for an IDN gTLD in 2012.

So what to include in variant label application? So recommendations 3.5, 3.7, 3.16, implementation guidance 3.6, 3.8 and 3.9. So the applicant will have to, in the application, state why the variant labels are being sought. So that has to be a reason and explanation, something about maybe the community you want

to serve or why you think variant labels are required for whether it's commercial need or whatever.

They also have to include information about their ability to manage the primary label and the variant labels. And one of the challenges that we had in coming up with these recommendations is understanding that this hasn't been done before in an application process. So what additional information would be required by evaluators to account for a variant label application? So this is kind of over and above what a new gTLD applicant would have to provide.

Community TLDs, Geo-TLDs and .brand-gTLDs will have to provide the same documentation requirements as the primary label. Justine, you might have to help me out here. So a community TLD needs to satisfy a community criteria. Geo-TLDs have to provide documentation to support the application. In some cases that means a letter of support or nonobjection from a government and I'm just trying to remember whether that applies to the primary plus the variant labels. Yep. So it has to be support for the primary and the applied for variant labels.

And the brand TLD, so I think the recommendation that we have for the brand TLD is we acknowledge that it's unlikely that a variant would be allowed for with a brand because the primary label—because a brand is usually just the single word or whatever it is, that doesn't have variants. So it's likely that for a brand TLD, it's only the primary label that's applicable.

So in terms of cost for the application, so the cost for a variant label, we've recommended to follow the cost recovery principle

EN

that applies for the new gTLD but we have applied some conditions where there's a base application fee and discounts. So we've kind of—embellished is the wrong word but we have some additional conditions as it relates to the costs or fees. And I think there's a separate slide on that. Next slide please, Emily.

So the application submission, we're still in the admin check and initial evaluation. So what can be applied for, so recommendation 3.17, 3.18, 3.19, 3.20, 3.21, 3.22. So single-character labels cannot be applied for. However, single-character labels in Han script can be applied for after relevant guidelines are implemented.

So what we've done is we've reached out to some of the script communities and sought some guidance on that. So we're waiting on input back from them so that we can develop guidelines around the single character labels.

You can't apply for variant labels of reserved names, variant labels of strings ineligible for delegation but relevant protected organizations are allowed to apply. So strings ineligible for delegation is related to the PDP on—I'm going to get this wrong, it was Red Cross and other IGOs. So that was a specific PDP that was conducted some time ago and that's where the term strings ineligible for the delegation come from. IGOs and NGOs, thanks Justine, and labels not conforming to mandatory string requirements in the root zone LGR.

Further adjustments due to root zone LGR implementation. So this gets a little bit tricky from an administrative or evaluation process.

So if you don't really understand the dot points here, it might be best to look at this part in the initial report.

So the system—and this is the application system—issues a disqualification warning when a label is found invalid or blocked but the application can continue. And that's on the understanding that there could be a misrepresentation of the root zone LGR in the application system.

The applicant can invoke a challenge mechanism for DNS stability review to challenge the root zone LGR implementation. A label correctly assessed as invalid or blocked is disqualified. So that probably doesn't do that section justice. And if it's unclear to anyone, I would recommend that you read that section of the initial report. Next slide please.

Okay, so existing registry operators applying for variant labels. So the EPDP discussed the possibility of a standalone process for existing registry operators to apply for their variant labels prior to the next new gTLD application process. And to understand the feasibility, the EPDP team examined the new gTLD process flow. And assuming future rounds will have a similar application evaluation elements as 2012 and anticipating new elements based on SubPro PDP and this PDP, the EPDP team observed that variant label application must go through the same steps and stages as any other application. And the new gTLD program will require modification to accommodate gTLD variant applications.

We had a number of conversations around this. I think because we do understand that IDN gTLD operators from 2012 have been disadvantaged in some respects because they can't apply for

variants. The position of the IDN EPDP team was that the most expedient and cost-effective path to move forward for existing registry operators was to apply for variant labels through an application round. And as I noted on a previous recommendation, in some part to—I can't think of the word right now, but I guess recognizing that disadvantage, we have recommended that one-off prioritization for existing registry operators from 2012. Next slide please, Emily.

So this might look a little bit complicated but it goes to the application fee structure for IDN gTLD applicants. So if you're applying for a primary label only, in the next round you'll pay the base application fee. If you apply for a primary label and no more than four variant labels, you will still just pay the base application fee. If you apply for a primary and more than four variant labels, then you may incur an additional fee to account for any additional evaluation or costs incurred by ICANN in processing. And that fee would be determined by ICANN.

For existing registry operators in the 2012 round, for the next round, if you apply for four or less variant labels, then the base application fee will be waived. If you apply for more than four variants, then you may incur that additional fee. So the base application fee is waived but you may incur an additional fee for a larger number of variant labels being applied for.

Future registry operators—I don't understand that, not allowed or not allowed. Justine?

JUSTINE CHEW: Yes, because if the applicant applies for only a primary label in the next round, then in the future round they will be a future registry operator. That's the reason why it's not allowed, because of the timing of it.

DONNA AUSTIN: Right, okay, thank you. I won't go future round after next round but it's except for the existing registry operators from 2012. So it's a discounted base application fee and a discounted base application fee plus additional fees. So I guess within our recommendations, there is some consolation, for lack of a better word, for existing registry operators for from the 2012 round. Next slide please.

So I'm going to hand the next few slides over to Justine. String similarity review was one of the most difficult discussions we had and we had a small team look at that and Justine oversaw that effort. So with that I'll hand this over to Justine. Thank you.

JUSTINE CHEW: Thanks, Donna. So with the string similarity review, as Donna said, this took a while. We had a small team looking into generating examples to see what might happen with string similarity.

Now, if you understand from the 2012 round, string similarity review is basically a process of comparing and applied for string against a group of other strings or other TLDs, I should say, and a group of other strings. And what I mean would be in the case of— if someone applies for say string A in the last round, 2012 round, then that string A would have had to be compared with other

EN

existing TLDs, existing ccTLDs, strings that are requested as IDN ccTLDs, other applied for strings, right. So that's a group of, a pool of strings and TLDs and a string that is being applied for has to be compared against.

And with the string similarity review, it is basically a visual similarity check. So there's a panel, the string similarity review panel, that looks at the strings in comparison visually. So you imagine that if you have to check one string against another string, and you do that thousands and thousands of times, depending on how many strings you need to compare, it can get quite complicated.

So in the 2012 round, we were only dealing with primary, so one string against other strings. The question for the group, the small team, as well as this particular EPDP is, what would happen if we introduce variants for each primary and each existing TLDs, if they were to have variants?

So you can imagine that it can get pretty complicated. And in fact, the small team actually took slightly more than a couple of months to come up with good examples of how to demonstrate what might happen. But essentially—and this is a simplified version of what took place in terms of the deliberation of the EPDP, and we're trying to dumb it down to a conceptual level so that it's easy for folks to understand. So you have to bear with me if I start kind of jumping around, because I think we may have to change the sequence a little bit in order to help understanding.

But at the end of the day, what the small team and this EPDP settled on is that we needed to do certain things. One is to take

EN

into consideration the need to mitigate the potential risks of string similarity incidences. And the two risks that were identified were denial of service, and that's not DDoS. It's denial of service, which basically is a no connection. In simplified terms, it's a 404 error.

And the other risk, which is the more important risk, is the risk of misconnection. Because with the denial of service, you basically just end, if you hit a 404, that's the end of your journey with that particular URL or the domain name. So there is not necessarily going to be harm caused by that experience.

But if it's a misconnection risk, then presumably, one of the situations would be the user would be misled and taken to a different site than what he or she anticipated to have landed on. So when that kind of scenario happens, then obviously there's going to be potential for abuse, DNS abuse, phishing in particular. But we will demonstrate that in a little while.

And then the part of the expanding to examine the role of variants in string similarity review would be to try and detect more combinations of visually confusable labels. Because when you start introducing variants, and earlier on, we explained that variants are typically the same word, but possibly look different, or could be they look similar, but actually completely different words. So there are those kind of complications involved.

But at the end of the day, if we don't put the variants into play with string similarity, then you're basically cutting off that possibility of detecting more combinations of confusable labels.

But at the same time, because if you imagine that one particular string has got 10 variants, and another string has got 10 variants, then you could be potentially looking at hundreds of comparisons, depending on how that permutation runs.

So we also were very wary that we didn't want to impose unnecessary complexity on the panel to do their comparison. So we only refrained from requesting that blocked variants be compared against blocked variants. And I'll go to that, I'll expand on that concept a little bit more when we get to the next slide. In fact, let's go to the next slide now, and I'll come back to the last two points here.

So this is an example that the small team came up with to show the impact of not introducing variant labels, or the way the model that we've introduced would work. So what do we mean? So as I said before, if we're just looking at the 2012 round, you are just going to be comparing A1 and B1. So I'm going to use the word labels, and not confuse strings and labels. We kind of use it interchangeably.

So with the 2012 round, because there's no variants involved, it would only be a comparison of A1 and B1. End of story. If they're not visibly confusable, then they will both get through. And this example only looks at two labels. So imagine if there are three labels that needs to be, or 10 labels or whatever, and each label could end up having hundreds and hundreds of variants. So therefore, you see the permutation becoming quite exacerbated.

So just focusing back on this particular example of two labels. So B1 has got variants, B2 to B32, what you see in the pink box on

EN

the right-hand side. And label A has got also variants A to A14. So I need to explain that the green box means that it's allocatable, and the pink box means there are block variants. So that's the difference between why it's green and pink. And obviously, the blue ones are the primary.

The distinction that we are trying to introduce now, when we are talking about variants, is that we have allocatable, which are the green ones, and the blocked labels, which are the pink ones. And what the model that we've come up with, affectionately named it hybrid model, is that we would compare the primary with the allocatable. We would, on one side, compare the primary against blocked labels. And then we would compare the primary against the allocatable on the other side, and the primary against the blocked on the other side.

So what I've just described is, we compare A1 and B1, which is route 1, as you see here. Then we compare B1 with A2 and A3, which is route 3. Then we compare B1 with A4 to A24, which is route 5. So that's one side of it.

The other side of it would be, we compare A1 to B1. Obviously, that's done earlier. And then we compare A1 to B2 to B23, and that would be route 2. Then we also compare A2 and A3 with B2 to B23, and that is route 4. So that's how we come up with the five different routes. And you notice that there is no line between the two pink boxes. So that's what we mean by, we refrain from requesting that the block be compared with blocked.

And you will see the reason for this is because, if we had not done this, then with the yellow routes, you would see that... Okay, let

EN

me reel back a little bit. By doing this, you will see that with the yellow routes, we have actually identified a few combinations of confusable labels. So if we didn't do this with the hybrid model, we wouldn't have picked this up, or these would have been ignored. And the reason for doing this is because we want to avoid having two labels, which are actually confusingly similar, being delegated to different parties, or one or both being delegated, because they will end up confusing people at the ending. So that is essentially why we have proposed the hybrid model.

So I also mentioned earlier that the permutations could be very, very large. And we also were concerned that we didn't want to place undue, unnecessary complexity in the comparison model. So Emily, could you go back to the earlier slide?

So one of the things that we did as an exception to the hybrid model, and to the adaptation with the string similarity review, is that we would allow the string similarity panel some leeway to omit having to compare labels of scripts which have very low level of confusability between the scripts.

The easiest way to explain this would be like, if you look at the Latin script, and say the Chinese script, the scripts are very different. So there is a very low chance that there would be labels that would look similar, and would be confusingly similar in that manner, for that matter.

So for those kind of situations where invariably the scripts won't look confusingly similar, we will give the string similarity panel some leeway to omit those kind of comparisons. But we will also have guidelines and criteria, objective guidelines and criteria, as to

EN

how these things could be omitted, and why these things would be omitted. And we're also requesting for some additional research or study to identify the nature of the scripts.

The last bullet here is that it says all labels from the variant label set share the same outcome, which basically means that if you think of the integrity of the set principle, the set must go together. So the set, which includes the primary and all the allocatable and the blocked, they must stick together, you can't separate them. That's the integrity of the set principle.

So if an applied-for string, or any of its variants, is found to be confusingly similar to an existing TLD, whether it's gTLD or ccTLD, doesn't matter, or any of their variant labels, then the entire set of labels that is being applied for would be ineligible to proceed, because they would invariably conflict with something that's already existing and would introduce confusability if they were to be allowed to proceed.

If, in the case of where we find confusing similarity between an applied-for set, whether it's the primary or the labels, sorry, the variant labels, against another applied-for primary and variant label set, whether there's some combination of confusability, then what would happen is both the sets would go into a contention process. And upon the resolution of the contention, then the set that prevails would be able to proceed. So that's the explanation of shared outcomes. Okay, Emily, can we jump two slides down, please?

All right, so this is a scenario that we came up with to kind of illustrate the harm, the possible harm that a misconnection would

bring about. Essentially, it is what I tried to say earlier, is because variants tend to have this thing about being similar, whether it's similar in sight or similar in kind of understanding, really, in some cases. It is about trying to avoid confusion.

So in this situation where someone sees a particular label, right, thinks it is something else, goes and looks for it, and then gets led to some other website or space that they thought were incorrect, they weren't looking for that, really. They were looking for something else, but invariably got to somewhere else.

And in this situation, there isn't harm done per se unless obviously the entity that's running the green label at the end on the right corner is an abuser. So in this situation, there may not be harm. But what we're trying to say is, because there is this risk of someone being diverted to another space, to another destination that they didn't anticipate, then there's always going to be potential for DNS abuse, phishing, for example. Okay, so moving on, let's move to the next slide.

Okay, so objection process. Okay. Well, in essence, each label is a label. They should be treated the same from a technical perspective. So there isn't any reason why all the labels shouldn't be subjected to objection process. So if one label is subject to objection process, then every other label should have the same treatment. So that is essentially the principle behind Recommendation 5.1.

And we know from the 2012 round, there were four types of objections. String confusion objection, limited public interest objection, legal rights objection, and community objections. Now,

string confusion objection is singled out as different to the other three here because it deals with confusion. So visual confusion. Again, it's still the visual aspect of confusion.

And basically, we say that objections with Recommendation 5.2 and 5.3, 5.2 at least, the objections may be filed based on confusing similarity between the labels that are established by the hybrid model. Okay, so if we use the hybrid model, the adaptation of string similarity model, and it comes up with combinations of confusable labels, then you can use that as a basis for a string confusion objection. And you can use the primary, or you can use the allocatable, or you can use the block if it appears in that group of confusable labels. And the outcomes of it would be consistent with the 2012 applicant guidebook, so nothing changes per se. Except the fact that there's an introduction of variants.

With the limited public interest and legal rights and community objections, the objection may only be filed against the primary label and the applied-for variant labels. So it would omit non-applied-for variant labels, and it would omit the blocked labels. And because it's not about confusability anymore. It's about that particular string or the label.

If the objection against the primary prevails, then the application in its entirety, whether it comes with just the primary or the variants, would be ineligible to proceed. If the objection against the variant prevails, then only the variant would be withheld from being allowed to proceed further. So the rest of the set that is being applied for may still proceed. Next slide, please.

Right, and string contention. So the idea behind string contention, obviously, is if things are found to be similar, then they get put into a contention set. So obviously, applied-for labels that are variant labels would be placed in the contention set. And the entire variant label set would be processed together in a contention set. And one of the outcomes, as I said before, is if one is found to be problematic, then it would affect whether the whole set gets through or just parts of the sets would get through.

Okay, I think that is all we have for string contention. And then moving on, I think we'll go back to contractual requirements. So that's it for me. I'll hand the talking stick back to Donna.

DONNA AUSTIN: Thanks very much, Justine. Okay, so overview of the contractual requirements. So this is similar in some respects to how we've treated the applications and the ability to apply for primary variant in the one application.

So to apply the integrity of the set principle in the registry agreement, so future IDN gTLD operators, the primary label and the approved variant labels will be subject to one registry agreement. So the way it works at the moment, you have a single TLD and one registry agreement. But what we're recommending here is that the set be kept together under the banner of one registry agreement.

For existing IDN gTLD registry operators, so approved variant labels, subject to a separate registry agreement, but linked to the existing registry agreement. And that is really because we

understand that there are some challenges to potentially bring an existing registry operator, and if they apply for a variant to bring that in under one registry agreement. So we understand that there are some challenges. So for simplicity, we're just recommending that if they do, if an existing registry operator does apply for variants, that the variants are under a separate registry agreement, but they are linked to the existing registry agreement for the primary.

Same registry fixed fee applies to a registry that manages variant gTLDs and one that manages a single TLD. So at the moment, registry operators incur a fixed fee. I think at the moment, it's $25,000 a year. But there is a calculation if you go above, I think, 50,000 domains in a 12-month period. But what we're saying is that the single registry fee will apply for the primary and the variant. So one registry agreement, one fixed registry fee, and then the registry level transaction fee, which is calculated based on the cumulative number of—no, sorry, I'll take a step back.

There is a registry level transaction fee, which is what I was referring to. So once you reach above, I think it's 50,000 transactions, that's for a single TLD registry operator, you will incur additional registry fee. What we're saying here is that that fee will be incurred based on the cumulative number of registrations from the variant label set. So it won't be calculated on the single TLD, it will be based on the label set. So we're acknowledging that a gTLD and a variant gTLD coupled together, that registry level transaction fee will kick in based on the cumulative number of registrations, not based on just the number

of registration in one of the TLDs, whether it's a primary or the variants.

The community TLD, brand TLD, gTLDs, category one safeguards, variant labels will be bound by the same restrictions as the primary label. So apply the same entity principle in registry agreements, so the same registry service provider for each critical function for an existing IDN gTLD and its variant labels has to be the same.

So what that means is particularly for an existing IDN gTLD operator, they have certain providers at the moment that will provide critical functions for the TLD. And if they apply for variants, then the variant has to come under the management of the same vendors or service providers that are currently in use for the existing IDN gTLD.

All registry transition processes encompass the primary label and the variant labels and the same data escrow provider has to be contracted for the primary label and the variant labels. So basically, if you're using one operator for the primary, you must use them for the variants as well. Next slide, please.

Delegation and removal, so delegating variant labels, so no ceiling value for delegated variant gTLDs was considered necessary. And I'll talk about that separately. We also recommend the creation of a framework for developing guidelines for the management of variant labels by registries and registrars. So acknowledging that variants don't currently operate for gTLDs, we do think it would be worthwhile creating guidelines, but acknowledging that that can't happen until variants have been in operation for some time, what

we're recommending is that a framework be created to enable the development of those guidelines in the future.

The primary label and approved variant labels will be subject to the same delegation timeframe. So what that means is that I think currently registry operators, once they sign the contract, they have 12 months to delegate. And what we're saying is that the primary label and the approved variant labels will need to be delegated within that 12 month timeframe. I think there is provision within the registry agreement that with agreement of ICANN, you can extend that delegation period, but that's something that would have to be done with ICANN. So what we're saying is that same delegation timeline will apply.

All delegated gTLDs and delegated allocated variant labels be grandfathered despite updates to the root zone LGR. So this is one of the possibilities, is that the root zone LGR will be updated from time to time, which may mean that some delegated IDN gTLDs or their allocated variants may no longer be consistent with the root zone LGR. But what we're recommending is that in that instance, that the delegated gTLDs and the variants be grandfathered. And we have a number of reasons for that that you can read in the initial report.

And also related to any changes in the root zone LGR, so the generation panels and integration panels when they are doing that updating, we're recommending that they make best efforts to retain full backward compatibility.

Okay, we're almost there. Removing variant labels. So removal of a primary label requires removal of its delegated variant labels. So

this goes back to the sanctity of the set and the fact that the primary label is the label that determines the variants. So if you take the primary label out of the root for any reason, then the delegated variants must come as well because they don't have any connection to a primary.

Removal of a delegated variant label may not necessarily require the removal of other labels from the variant set except when it's removed due to breach of contract. So all we're doing is acknowledging that there may be instances where the registry operator has a variant label that's been delegated, but perhaps down the road, the registry operator may decide that there's no need to have that variant delegated anymore. So they want it removed from the root. So they are able to do that. And that won't compromise the delegated primary and other variant labels in the set.

But where a label is removed due to a breach of contract, then that means that all the labels must be removed because if there's a breach of contract for one, the reality is that it's a breach of contract for all given that we've also recommended that the primary and variant labels come under one registry agreement. Next slide, please, Emily.

Okay, so why we've decided that no ceiling value is necessary. So the recommendation reads that no ceiling value for delegated top level variant labels from a variant label set is necessary as existing measures in the root zone LGR to reduce the number of allocatable top level variant labels, as well as economic operational and other factors that may impact the decision to

apply for variant labels will keep the number of delegated top level variant labels conservative.

So in part, this goes back to the slide previously on the root zone LGR and recognizing that there's only seven scripts that actually have allocatable variant labels. So with the exception of Arabic, the other six scripts already have limited the number of allocatable variant labels. So to either one between one and two or up to four. So there's already restrictions placed on the number of allocatable variant labels. So given that, we didn't consider that placing a ceiling value was necessary.

In addition, various factors such as cost operational competence, potential challenges with variant management may result in a conservative approach by applicants anyway. SSAC confirmed that the volume of delegated variant labels doesn't necessarily create security and stability risks. So while I think the conversation we had with SSAC, there was certainly a concern by SSAC that the variants could create some challenges because there's a lot of unknowns. But at the end of the day, a TLD in the root is just a TLD.

Guidelines for the management of variant labels by registries and registrars should help address SSAC's concerns regarding the lack of a common approach in managing the variant TLD. So that's recommendation 8.2 about the developing guidelines for the management of gTLDs and their variants. So we hope that will go some way to providing some kind of consistency in the introduction of variant labels. And that also goes to the framework for developing the guidelines to enable that guideline that we're

EN

recommending be created during implementation. Next slide, please, Emily.

Variant label states. So there are variant label states are delegated, allocated, withheld, same entity, blocked and rejected. So there's a recommendation that ICANN Org records and tracks the variant label states as long as the primary label remains delegated and variant label transition path. So that's in recommendation 9.3 and implementation guideline 9.4.

In the conversation we had around these label states, it became— well, I think it was probably always evident that the terminology that we use in different contexts within ICANN, you will have a word that means different things depending on the context. So we had a number of conversations around that in considering the variant label states.

Particularly allocated and delegated have different meanings in different processes. So that was a little bit of a challenge. But this is for the purposes of tracking the label state for an IDN and its variants. This is the terminology and the path that this goes through. Next slide, please.

Okay. And just on the previous slide and the challenge that we had in some of the terminology, we do have the public comment open at the moment. And one of the things that we have requested input on is the glossary. Because particularly in light of the fact that some of the terms that we use, we've applied a meaning that is appropriate for the context of IDN gTLDs and their variants. And we appreciate that that could be, some of the terms could be used in other contexts. And that creates a bit of

confusion when you're trying to understand the body of work that we've put together. So in addition to public comment on the recommendations themselves, we're also seeking input on the glossary and whether the terms that we are using are cause for confusion in other parts of ICANN.

So just a reminder on the public comment. So it opened on the 24th of April. It will close on the 5th of June. So we have 19 days left apparently. So there's a link there to the phase one initial report. You can submit your input via the provide your input button on the public comment page. And for each preliminary recommendation, the way that we've done this is to indicate a level of support or non-support and explain your rationale.

So we are conveniently at time. So that's the end of the webinar. The webinar and recordings will be made available on the GNSO calendar. And I'm sure, yep, that's already in there where you can find that information. And I guess if you've got any questions, you can take those—if you have a representative group that is part of the EPDP, please bring those questions to your representatives and they can bring it to the EPDP team or contact the GNSO secretariat and any questions that you might have, they can pass those on to us.

So thank you for your time today. I appreciate it's a lot to get through. And I know we have quite a few members of our EPDP team on this call. And I just want to thank them very much for the work that they've done to getting us to the initial report phase and also to our team of Ariel, Steve and Emily and also Devan that kicks us off every week, and Justine, who's been a great support to me in this work.

All right, so we look forward to your comments. And if there's anything that's unclear, then please, there are avenues that you can get those questions to us and we will get back to you as soon as we can. So with that, I think you can end the recording. Thanks, Devan.

DEVAN REED: Thank you, Donna. This call is now adjourned. Thank you all so much for joining. Have a wonderful rest of your day.

[END OF TRANSCRIPTION]