DEVAN REED: Good morning, good afternoon, and good evening for the recording. This is Devan Reed. Welcome to the Latin Script Diacritics PDP Call taking place on Wednesday, the 2nd of April, 2025 at 13:00 UTC. We do have apologies from Sylvia today. Statements of interest must be kept up to date. If anyone has any updates to share, please raise your hand or speak up now. If you need assistance updating your statements of interest, please email the GNSO Secretariat. All documentation and information can be found on the Latin Script Diacritics wiki space. Recordings will be posted shortly after the end of the call.

Please remember to state your name before speaking for the transcript. And please note, all chat sessions are being archived. As a reminder, participation in ICANN, including this session, is governed by the ICANN Expected Standards of Behavior and the ICANN Community Anti-Harassment Policy. Thank you, and over to our Chair, Michael Bauland, please begin.

MICHAEL BAULAND: Thanks. Could you share this slide, please? Yeah. So, welcome to the second meeting of our LD PDP. Next slide, please. We'll first go through a brief summary of what we achieved or did last week and the established scope to date. Then we'll continue with the scope discussion, in particular regarding the similarity question. There were quite a few comments last time. And if we reach that far, we will then continue with the charter question one, and then some AOB in case anybody has any comments to add. So, next slide, please.

So, summary of the first meeting. We sent out the early input request to the SO/AC/SG/Cs. And the response is due by 24th of April. As mentioned several times, this is a voluntary response. The groups do not need to respond. In particular, if there are representatives here, they can easily also voice their opinions during the discussions. But in case they would like to make a more official statement or input, this is a great way to do this.

We then reviewed our charter and discussed the variant sets to help clarify what is not in scope for this PDP. In particular, we took a look at the root zone LGR variant sets and showed you where you can find them to take a look in more detail. And we looked at the Unicode tables for help to decide what diacritic means and what is actually meant by a diacritic character. We also determined that we don't want to expand the scope of this PDP. That would take much longer and would unnecessarily complicate things.

We are aware that there are several cases that are not covered by this PDP, but which most likely are equally valid and should be covered. But that should probably be done in a different process. And most likely it's easier to do that once we have created a policy for the very narrow scope that can then be extended for a larger scope. We established the project plan and submitted to Council or going to submitted it to Council.

JOHN EMERY: It has been submitted to Council, and Prudence, our liaison, will be presenting it at the Council meeting this Thursday.

MICHAEL BAULAND: Okay, thanks. And there was a brief discussion about limiting the scope or the recommendations to grants or geographic areas. But this was right at AOB and we decided to continue this discussion at a later point. So there were no decisions taken at that point. Next slide, please. Next slide.

Just a quick recap. We decided that it's out of scope to include letters like the four shown here. I can't pronounce them. I don't know if you pronounce them. Might also be different pronunciations depending on the language where they are used, because those are Latin script IDN letters. They are part of the repertoire, but they are not diacritic letters. The Unicode gives a clear delineation of the scope by stating very clearly what is a diacritic and what is not. And the outcome was a tentative agreement to adhere to the charter at this point and define the Latin script diacritics via the Unicode definitions.

We also talked about how to deal with these other cases that are also important. And the idea was that we could probably give a recommendation for future work on that topic. So not do any decisions for that, but state that there are cases we discovered and we suggest that they should be dealt with afterwards. Next slide, please.

So here's a quick recap of this Unicode definition. We have out of scope characters on the left. For example, the Latin small letter ae, which is also called Latin small ligature æ or Ash from old English. Also, the Latin small letter o with a stroke or o slash or the Latin small letter eth. None

of these letters are a combination of some other character combined with a diacritic. So that's what makes them out of scope.

On the right side, you see cases which are in scope. For example, the Latin small letter c with some sedilia. As you can see, the Unicode table clearly states this is a composition of the letter c, 0063, with the sedilia below 0327. And similarly, the Latin small letter i with a circumflex, which is a combination of the i and the circumflex above, or the Latin small letter y with acute. This is defined as a combination of the y with the diacritic acute.

Whenever the Unicode actually decomposes or gives the possible decomposition of such a character, and one part of the decomposition is a diacritic, then this character is considered to be a diacritic character and thus in scope. And if no such decomposition is given in the Unicode, then it's out of scope. We also put the link for the Unicode chart here. As said last week, this is just one of the charts. There are more charts that have similar cases and they work the same way, of course. Please, if you're talking, we can't hear you, Sebastean. To Tapani, yes, we include also cases with multiple diacritics. Okay, if you can't speak, Sebastien, maybe you put your comment in the comment section and I can read it. Next slide, please.

There's the open question whether we would want to allow more than one diacritic at the same time. So in scope is definitely the dot stop and dot stop, for example, then this would be a case we would consider. Out-of-scope on the left side is also something where we have dot stop and dot step, even though stop is an ASCII-only TLD and step is certainly the A, it's a diacritic. Still, this is out of scope because it's not the

corresponding version. So for A, the corresponding ASCII character would be the A and not the O. But most likely it's not a problem that this is out of scope because it's quite unlikely that the string similarity review panel would make this the case where they say they are confusingly similar. And if they do this, it's then similar to other cases like stop with O and stop with A. If they were to be considered similarly, then this is also not something we would want to deal with.

I just read a comment here. We unfortunately didn't hear you. Not questioning the path you suggested on definition scope and proposing paths for what is out. I'm willing to do the legwork, but can I suggest we gather the Unicode tables to be used? Maybe only the one in the link. Yeah, I guess we can compile a list with all Unicode tables. This would be restricted to the repertoire in the Latin LGR, and we can check what Unicode tables need to be looked at to cover all those characters. Of course, the tables contain more characters. For example, they also have uppercase versions which are always out of scope, and there are other characters in the tables which are out of scope. But yeah, we can try to compile such a list.

The question, I think it was raised by Tapani, is what happens if we have three TLDs on the right column? Stop and then the stop versions with the two accents in the one or other direction. Would we allow this and then say all three could become available, or would only the ASCII only be available and the other two would go into a contention set? This is definitely something we need to talk about. I'm not sure if this is the right time for this. John, do you have something?

JOHN EMERY: Yes, this is a good time to talk about three or more diacritics. I'll also note Tapani's note in the chat about Dieppe with multiple diacritics on one letter, and also Satish asking about if we're proposing general principles or enumerating the full list.

MICHAEL BAULAND: To answer Tapani, the cases where multiple diacritics are on one letter, this is basically handled the same way as if only one character is there, one diacritic is there. If you strip the letter of all diacritics, then just the same in scope situation as if the letter only had one diacritic. If you consider the case where you just want to strip one of the diacritics and leave the other, then basically we have the same case, which is on the right-hand column, where we have multiple different diacritic versions for the same base ASCII version. That's something we need to discuss.

Just reading Satish's comment. Are we only proposing the general principle or are we enumerating the full list of diacritics in scope? I think we can list the full list of diacritics. It's not such a big list. The idea was just to mention the rule with the Unicode tables, but I also don't see a big problem with listing this, except that we must make clear that this is not a fixed restriction. It's just the current examples and this list would have to be adjusted if the LGR were to be adjusted. The recommendation, I think, should be based on the more general rule of diacritics in the Unicode, but for convenience, I think we can list existing diacritics based on LGR version 6 or whatever is the current version, I think 6. Amadeu, please.

AMADEU ABRIL: What I wanted to reply to your last part, I would hate to need to restart the PDP because we have added or raised one concrete character in the LGR. If we want to give that as an example, that's okay, but please never say, these are the list to which this solution applies, because then we are leading into a disaster.

MICHAEL BAULAND: Yeah, fully agree. The recommendation should cover the general case with the Unicode decomposition, but for convenience, it might be okay to list the characters which are available based on LGR version 6. Any other comments in particular to the question what we do with more than one diacritic version of the TLD, the right column? Should we allow this? Should we forbid this? No opinion. Anybody?

AMADEU ABRIL: I don't have my hand raised, but may I intervene? It's Amadeu again.

MICHAEL BAULAND: Yes.

AMADEU ABRIL: I think we should. I mean, that's clearly not in the language we have, but I think this is the one exception in which we should apply the same thing. Otherwise, I mean, there are very rare occasions which this would apply, but I can imagine occasions which would apply. Take, for instance, a brand that has different spellings, and as a trademark attorney, I have registered trademarks with different diacritics for

different markets and without a diacritic for the global market for the same trademark.

So, it's an extreme case, but this could happen, and the solution is exactly the same. So, if we have a good solution, why not allowing the base and two diacritics, or just two diacritics, or just three different strings with different diacritics? Provided that, again, we apply the same principles, the same registry, and they are diacritics in the sense that we are defining here. But the relation to do, even if it's very realistic and would cover 99.99% of the cases, is not the only possible situation.

MICHAEL BAULAND: Yeah, thanks for the comment, Amadeu. I see in the chat that several are also in favor to allow more than two, and I personally tend also to agree to that. In particular, due to the fact that in the IDN EPDP, where we set up rules for variant TLDs, we also explicitly allowed more than two variants to exist. And for that reason, if it's possible and already a case has been made for the real variants, I don't see a big problem if we also make a similar case here. Unless anybody is objecting to that, then please raise your hand or write something in the comment section.

Yeah, I only see positive comments to that. So I think we can summarize that the group agrees we should not restrict this to two TLDs. Even if, as Amadeu said, in 99.99% of the cases, it most likely will just be two TLDs, but there's no harm in allowing these corner cases should they happen. Yeah, then I think we can continue with the next agenda item. Next slide, please.

The big other question here is, what is similarity? This is a very difficult question to decide, so it's hard to have a clear-cut definition of what similarity means, whether two characters or two labels are too similar. It's always a subjective decision that there's no 100% objective decision possible. And a presumption of this PDP is that the string similarity panel already exists, so there's nothing we can do about this. This is a part of the application process, it's a given fact, and this panel also has a process which we cannot and don't want to adjust.

So the panel is working purely on a visual similarity level, their evaluation, and the question is, are there any measures that this PDP can take to provide more certainty to applicants when considering whether to apply? The implications here is that this PDP is to identify an exception process for an already existing process that seeks to prevent user confusion. So the idea of the string similarity panel, why it's there, is to avoid situations where two TLDs exist in the root zone and they are too similar to each other that users could confuse domain names like example.tld1 and example.tld2.

And this certainly must be avoided that users get confused and the easy solution therefore was to just allow one of them and reject the other. But as we know, and that's why we are here, this is not a good thing if the entity operating both of the TLDs is the same and it implements some kind of rules which we are going to define and discuss that prevents the user confusion for those two TLDs. Then the main reason to reject this in the string similarity panel is removed and therefore we can create an exception process for this string similarity process. So yeah, I guess we agree that we won't change or adjust the existing string

similarity panel process or anybody trying to change that. And if they do, if they want, I guess this is the wrong place to do it.

I just read some comments. One second. Tapani is writing, could we propose rules saying that the registry should only allocate same level domain to both ASCII and diacritic gTLD? Yeah, I guess this would be something that we would do. I think the idea is to follow rules that have been created by the IDN EPDP because that PDP already had to implement special rules recommendations that make sure that even though the variant TLDs by definition are confusingly similar, that the registry operating them has some kind of rules that make user confusion impossible or very unlikely. Amadeu, do you agree?

AMADEU ABRIL: Sorry, the process of finding the buttons to mute and unmute is long for me. I don't want to discuss about similarity and user confusion even though I love to do that. You are completely right. This is a given. It's only if there is similarity and determination of probability, not just possibility but probability of user confusion, these exceptions would enter into place.

Now, what I think we should do is be careful not to invent new terms. And in the applicant guidebook there is extreme similarity. And then there is user confusion. And I'm a little bit surprised that in many documents for this PDP, the staff has always used visual confusion, which is something that appears nowhere in the applicant guidebook. And that's confusing in itself because it equates visual similarity to user confusion. And it's not. There are two different steps and two different

evaluations. Indeed, for us, we are at the end. So it must be similarity that's deemed confusing.   But we should not use terms in our documents that introduce new categories because then the confusion will be even higher. Thanks.

MICHAEL BAULAND: I think you confused me. No, I understand what you mean. Yeah, that's right. Similarity and confusing is not necessarily the same. It's just that the string similarity panel, this is looking for strings that are confusingly similar on a visual scope.   So they are not looking at meaning or language translations or something like that. But just if you look at one label and look at the second label, whether users could be confused by this. And in this context, similarity and confusability is kind of the same. Of course, confusability only happens if the usage of similar strings is in such a way that users can be confused. And otherwise, it's just similar but not confusing.

And this is something we want to discuss here and make rules to have the strings which are similar, but due to their nature and their usage are not confusing because the rules we create avoid this confusion. So they are just similar, but not confusing. I hope I was not confusing too many people now.   Ariel suggests to look at the AGP language to align definition of similarity. This is the draft language as for now. Yeah, we should certainly take a look and do that.

So the question back to the scope is we are to look at diacritics which are similar. But I think it's not up to us to define what this similarity means.   But just to have something in the way that our

recommendations just use the outcome of the string similarity panel. So whenever the string similarity panel decides two labels to be similar or confusingly similar, then our recommendations can kick in and depending on further circumstances, avoid this rejection.

But if the string similarity panel does not object to those two strings and says, well, they are not confusingly similar, then I think also our PDP should not make any rules for those labels because they can already coexist without any additional restrictions. So it's no benefit to make restrictions if this is not the case. Any thoughts here? Comments? It's a quiet group today. But that's good. We're fast. No objections, no comments, no objections.

So I just take the silence as agreement to the statement that we won't define what similarity means. We just go with whatever the string similarity panel does and then work with the decisions they do. If they decide it's not similar, then we are out. Our PDP is not applicable. But if the string similarity panel says, oh, those two TLDs are similar, then our PDP kicks in and we check for further prepositions. And if they are all met, then the policies should become active and define how these TLDs could possibly coexist. Asteway, please.

ASTEWAY NEGASH: Thank you, Asteway Negash, for the record. I'm not so familiar with the details of the string similarity review process, so I apologize beforehand if I say something stupid. But I feel like one of the requirements for the string similarity review to evaluate two strings, possibly a diacritic and

an ASCII version, is that those two labels are not bi-ends of each other. So can someone explain? Thank you.

MICHAEL BAULAND: Exactly. First of all, there are no stupid questions at all. We all have our expertise in some areas and we don't know everything in other areas. And I just encourage everybody to ask whenever something is unclear or you don't understand something, because something which is very obvious to me might be quite difficult to understand or is not known to other people. So we should try to get everybody on the same level. So please, if you have any questions, don't hesitate. I will also ask if I don't know. So you can do the same.

And yeah, this process, you're right.   If strings are or labels are considered to be variant, then they are not considered in the string similarity review process because they are already caught beforehand by the root zone AGR mechanism.   This is a very clearly defined mechanism. Everybody can use this. There's a tool, this root zone AGR tool provided by ICANN. And this is a clear-cut answer to the question whether two labels are variants of each other or not.

So unlike the string similarity review, which is to some point subjective, this variant question is 100% objective and there's no discussion here. And just those labels are out of scope because there are already mechanisms to take care how these labels will be handled, depending on whether the variant is an allocatable variant. Then the IDN EPDP rules kick in, or if it's a blocked variant, then this means that those two labels will go into contention set if they are applied for at the same time.

Or if one already existed prior to the application, then the second one will automatically be rejected due to being a blocked variant.

But yeah, it's a good point you raised. I think we ought to also discuss this internally. It might be a good idea to invite someone from ICANN who is more knowledgeable in this whole string similarity panel process and give a quick presentation. How this process looks like, how is it working and possibly answer any further questions. So we will look into this to invite these to our calls in one of the next calls. Yeah, it's a good idea.

But independent of how it works, we take it as a given that it works somehow. And we will take their results as a fact for our PDP and just work with that outcome and not deal with similarity questions ourselves. If there are no objections, I think we can go on with the next. Yeah, Samad is probably the right person, Satish. We'll check that internally and get that to this group as some background information, which are definitely good to have. So with this, I would like to hand over to John to explain the scope and charter question one.

JOHN EMERY: Yeah, so today we want to have a bit more of a deep dive discussion into charter question one. So last week, just as a reminder, we went through the scope of this PDP, which is limited to examining a single issue that we all know quite well. And basically, charter question one comes out of this part of the scope. What mechanism is needed in order to allow a single registry operator to simultaneously operate both gTLDs?

So in charter question one, we basically want to go through now and have a broad discussion broken down about potential considerations, potential rules that we might want to establish in this working group. And so breaking it down, we have the objective of our working group to determine the limited circumstances in which a base ASCII gTLD and the Latin script diacritic version of the gTLD can be simultaneously delegated.

So breaking this down, I think now would be a good time kind of start a more open discussion about charter question one, especially part one, under what circumstances would a base ASCII gTLD and the Latin script diacritic version of the gTLD be simultaneously delegated, if any? And so with that, I'll hand it back to Michael to lead discussion. And I hope we have an open and robust discussion on this. But I'm happy to see so much agreement already. This is fantastic.

MICHAEL BAULAND: Thanks, John, for the quick explanation. Gave me some time to drink a bit and read some of the comments and write a comment myself. Since when I'm talking, I'm not that multitasking capable to actually look at the chat and respond to that. Sorry for that. So I already see Amadeu, you're turned up, so please start the discussion.

AMADEU ABRIL: Okay, I don't think this will be a surprise. But first of all, I should apologize. It's very difficult for me to answer this question without touching somehow on some of the subsequent questions. Trying to be

sure. This is about preventing user confusion, not how to allow an applicant to move forward only.

So from the user confusion perspective, for three reasons, I think that the conditions are that the TLD follows the two TLDs, sorry, or multiple TLDs follow the so-called same entity or R3 principle. That is, that a domain without Michael.quebec with accent and Quebec without accent, or Mark.saopaulo with the first A with a tilde and the other one without the tilde. Both domains should be registered to the same registrant, the same domain name holder, owner, if you prefer, the same registrar, the entity managing it, and the same registry.

For three reasons. Experience has shown that this prevents confusion. Second, because this is coherent with the solution for variants, and we know that we are not talking about variants, but we know that this problem of the diacritics has certain similarities, limited but certain similarities with variants. And third, because also the approach that the ccTLDs have taken.

Now, this is something that Tapani also said, I only want to make a friendly amendment. The question is not that the domains are registered to the same people, that we prevent being registered from different registrants. That is, it may well happen that Mark only wants to use Mark.saopaulo with the tilde because he believes that this is the correct answer. If he prefers not to use, and even not to allocate, not to register the other one, but still then this should be blocked. But at the end it means the same, that the entities having the domains in both TLDs or three TLDs should be the same all along the chain, registrant, registrar, registry. Thanks.

MICHAEL BAULAND: Thanks, Amadeu. This is an important point. The same entity principle, which is also part of the IDN EPDP solution for the variants. And you've been talking about the same entity for the second level registration. I think this is important, but maybe we should first take a look at the TLD application process.   So also there, the TLD should most likely be belonging to the same--

AMADEU ABRIL: Sorry, Michael, but I think that is already in the conditions of our scope. It's how they can be delegated to the same registry. So I think this is out of discussion.

MICHAEL BAULAND: Yes, Tapani raises this question that what happens if the applications are for labels where the meaning is different and the entity is different. But I agree that this is not something that we can cover. John, can you help me with the scope here? Is it really, as Amadeu said, out of scope? Or is it something we have to decide to not make possible? I'm not sure right now.

JOHN EMERY: Yeah, I think I'll have to check if anyone else from staff is clear on that. I think that essentially this working group could make the rules for it.

AMADEU ABRIL: Sorry, but if you read the questions, it says how they can be delegated to the same registry.

JOHN EMERY: So Steve comes in. Steve.

STEVE CHAN: Thanks, Amadeu. This is Steve from staff. And I may be a bit of a miracle, but I think the charter probably could have been a little more precise because it essentially implies the same entity, but then it also asks the question whether or not this is the right rule. So maybe what a middle ground could be is that while this group may not need to specifically recommend it should be same entity, since there seems to be broad agreement that that should be the case, it probably would not hurt to make an explicit recommendation nevertheless, if that makes sense. So as I know, there's probably a bit of ambiguity in how the charter is drafted, because it really does imply that it should be same entity and doesn't really leave it up for question. So I hope that helps. Thanks.

AMADEU ABRIL: Sorry, for personal reasons, I have discussed that in real terms in the past, how to do that in the past. And the summary was, after all the discussion we had with many people a long time ago, is well, perhaps that, I mean, the applicant, as I say here, could be different in terms of legal person, but should be one operation of the two TLDs this together. Otherwise, the rest of the solution doesn't work. So formally, like

Donald's applied for 300 TLDs with 300 different entities, from a legal point of view could be a different entity, legally speaking, but operationally speaking has to be the same. Otherwise, the rest of the solution falls apart.

MICHAEL BAULAND: Yeah, thanks, Amadeu. I tend to agree here that if we were to allow these two TLDs to be operated separately, then there's no way to avoid user confusion, because then the one registry wouldn't just know what labels have been registered in the other registry. And the goal here is to avoid confusion. And this can only happen if we also enforce like the same entity on the second level. And like it's been done in the IDN EPDP. Just reading the chat quickly.

Anil says, I think we should restrict to the principle of the same entity concept wherever applicable. And yes, I agree. And Justine said, agree, we should avoid establishing different rules which contradict existing policy from the IPDP. And Tapani has clarified same entity as in the registry or as a registrant for the same second level domain.

So we need to have the same entity for the second level registration. So if you register example.tld1, then it must be enforced that example.tld2 must be either blocked for any other person or it must be available for the same entity that registered example.tld1. If that answers your question. Sebastien, please. Still having audio problems? We can't hear you. Yeah, seems like you're having sound issues. So Tapani still seems to wonder why I'm thinking of a rule that in Quebec case, we could require same second level domain only to same registrant. But in the

Sjoberg case, that would not be a good rule. But if you don't do this and you have different registrations, example.sjoberg and example.sjöberg once with an o and once with an o, then these two domains could be owned by different entities.

And then there's the user confusion because if you have one registration with one of the TLDs and you do not make sure that the same domain name in the other TLD is blocked for the same entity, someone else could come and register it and thereby try to hijack the domain. To pretend it's the other domain. And this would lead to user confusion. And for that reason, it's not a good idea to have that. Yeah, well, you're right. There already are sjoberg.se and sjöberg.se, one with an ö with dots and one with a normal o. And they don't seem to cause much confusion.

First of all, we do not know whether the string similarity panel would decide those two letters to be confusingly similar. Most likely they would. But also, what a TLD, especially a ccTLD does at the second level is something we can't change here. So I think those two labels might be confusing. But the Swedish registry obviously thinks it's not. But this is also then limited to the Swedish language. So I think so far the most comments I read here were also in the sense that we need to have the same entity also for the TLD level in order to make that confusing. To make this not confusing.

So another open point we can talk about is whether we would like to limit the applied for IDN strings of existing ASP-gTLDs where the existing ASP-gTLD is a workaround for the proper IDN string and/or vice versa, limit the applied for ASP-gTLD string that acts as a workaround for the

existing IDN gTLD. This was also one of the questions, whether we want to limit this. Amadeu, please?

AMADEU ABRIL: Hi Michael. Can you explain to me what a workaround is? I don't see how we can operationalize that workaround. In which perspective? Linguistic, social, technical, financial? I think we need to have a recommendation that's fairly objective. And the problem is a similarity to the user confusion. Not the reason why that registry may want or not that second TLD.

MICHAEL BAULAND: Bill, do you want to respond?

BILL JOURIS: I have much the same point that I think Amadeu was making. If we're going to create a policy for ASCII versus ASCII with diacritic, I think it needs to apply going forward, not just retrospective for somebody who once had an ASCII version previously. I think if somebody wants an ASCII version as well as a version with diacritic going forward, I think the same criteria ought to apply there as well. Thank you.

MICHAEL BAULAND: Yeah, I agree there. And to the question, what is a workaround? It should be considered in the way that you have your label, which is an IDN label, i.e. has at least one non-ASCII character. And the workaround is a label which only consists of ASCII characters. So, if you don't have

any special characters at your keyboard, you would write it this way. Steve, please.

STEVE CHAN:

Thanks, Michael. This is Steve from staff. And perhaps we have good news, at least from the staff perspective, which is I think we don't actually need to define this word workaround because the first part of this working group call actually essentially defined the scope, which I think is what this question that originated from the issue report is getting at, essentially. And so this group has, as far as we can understand, seem to define it as via Unicode and that table is illustrative. I don't know if I should really try to pronounce it, but stop and stop and different versions essentially is, I think, providing a more precise definition than "workaround." So, it seems that this question is probably already answered by the first parts of the discussion today.

MICHAEL BAULAND:

Thanks for the input. But it still remains the question that what if you have a label with two diacritic letters? Would we only want to make a policy where this letter, where this label and the ASCII only label can be allocated at the same time, or would there be some kind of intermediate? So, just right in the chat, the Deja name, which has a diacritic E and a diacritic A, would this only be allowed to coexist with the Deja label, which is ASCII only or should it also be possible to have something like Deja where the E is still the diacritic, but the A is ASCII only? Is this something that we want to have? We don't want to have? Any thoughts about this?

I'm nearby. There should be at least one diacritic character, but we may not have upper limit on number of diacritics in the label. You mean, as a starting point, if you don't have diacritics at all, of course, we are out of scope. But the question is, should you be able to have the Deja in the correct spelling with two diacritics? And should it be possible to apply for the Deja with just one diacritic and one ASCII? If they were to be rejected by the string similarity, all of this discussion is, of course, based on the fact that the labels were to be rejected. If they are not, then, as discussed earlier, it's out of scope and they can just coexist.

AMADEU ABRIL: Michael, it's Amadeu again. Why should we limit that? I don't think we need to discuss whether we want this. We should discuss whether we should prevent that.   We are not in the business of grammatical correctness. What does this BIZ mean? Really, nothing. It's just a diminutive of something else. So, we will not tell somebody, oh no, the correct spelling of your TLD should be that one in German or in English or in Vietnamese or wherever. So, in that case, if somebody wants something that only, for any reason, they want that, the question is whether that creates user confusion, not whether that's a good TLD. I don't think that we should be in that business.

MICHAEL BAULAND: Yes. The reason why this is asked is that this PDP was about cases where the ASCII-only version is a workaround for the diacritic version. And if you just remove one of the diacritics and leave the other, this is most likely not a workaround anymore. But I agree to some of the things

written in chat. We don't know if the spelling with only one diacritic is actually a valid label in some language. And if we decide not to limit this PDP to two labels, to two TLDs, namely the only ASCII one and the correctly spelled one, then I don't see a reason why not also allow this case where one of the diacritics is allowed.

AMADEU ABRIL: Okay, this is Amadeu again. And I promise it's the last time. Again, I repeat, we should find a reason why to deny this. I repeat, I think that the right thing should be to be able to apply for even a stupid. My God, good luck with that. But they wanted that. Why deny that? Too long? No ASCII workaround? Well, it's their choice. Now, the question is whether there may be reasons for that registry or in the language, I think, that we don't know today here for doing this admittedly strange things. But I don't think this should be the criteria that we don't know now, unless we have a good reason to say these things are dangerous from this perspective.

Second, going back to the workaround, it still makes me nervous. Give you a concrete example. There is a city very close to Barcelona, it's Valencia. Its name has a grave accent on the E in Catalan. In Spanish, they don't use that. Both languages are official then. It has two names. So, would they be told in your model that Valencia without an accent was not a workaround? Because, in fact, it's one of the two official languages? It's certainly not a workaround. It's a different way of using that, but it's not a workaround for the other with the accent. But they may want to apply for both. No, they don't. Imagine that. So, sometimes you have two ways of writing things, and it's not that one is

the workaround for the other. They simply have two ways because they are expressed in two languages.

MICHAEL BAULAND: Yeah, thanks, Amadeu. Sebastien made a good point in the chat. He said a brand could call itself Deja with just the E being a diacritic and the A being the ASCII, just to look different. I agree that we should not make policies based on language or what is the real word because TLDs and domain names in particular are not restricted to dictionary words or whatever. So, yeah, I agree that we should just look at the Unicode fact that whenever the base character is the same, then the diacritic versions and the base character would be covered by the policies we develop and not restricted to just the two cases.

Tapani says, even though TLDs can be non-words, if they are words in some language, it does have real-world implications. Yeah, I think that's true, yeah. One other question that was raised last time was whether we want to restrict this to certain application types, for example, brands or geographic names. We stopped that discussion last time because we were out of time, but since we have some more time now, we might want to come back to that now. Any thoughts, suggestions here? Bill, yes, please.

BILL JOURIS: Yes, I think we should definitely not try and restrict it to geographic names any more than we restricted to real words. I think that way lies madness, not to put too fine a point on it. I think if we're making a policy, we should make a general policy and not try to slice things into

various special cases that will cause confusion for the people trying to implement the policy. Thank you.

MICHAEL BAULAND: Okay. Thanks, Bill. Tapani says he agrees with Bill. Any other thoughts, agreement or non-agreement? I personally also tend to agree here because with my technical background, at least from a technical perspective, it makes no difference whether the TLD is a city name or a brand name or just some fantasy name or some other word from the dictionary. Technically, I don't see a reason to restrict it, and therefore, why should we? It only makes the policy more complicated, more complex, and if we just state it's for all TLDs, it's easier to understand. Satish also agrees to not restrict it, so unless there are any other opinions in favor of restriction, then please make yourself heard. This issue was raised by Philippe. We may take his views. Philippe, do you want to add something here, please?

PHILIPPE FOUQUART: This is Philippe here. If that's referring to my comment last time about the proper nouns, I was trying to find, I guess, a shortcut. It seems that we're trying to avoid the ocean, and I think that's what Bill was saying, essentially, in trying to address all specific corner cases in all languages, and if we say we don't make a difference between a geographic name and any string, for that matter, because that string, as Sebastien said, could actually be a brand name.

My sense is that there's a lot that we'll be relying on the panel, if you see what I mean. To me, the issue is very much whether it will be much

simplified if a panel decides that two strings are actually not similar because they're an expert in one particular language. So that the registries can be different, but I'm not sure that, well, that's not an answer to Anil's comment, but I'm not sure we should try and address each and every corner case, and if that's what Bill was saying, I would agree with that. I hope that's helpful.

MICHAEL BAULAND: Yes, thanks. I think it is helpful. And for the panel, I also think that the string similarity review panel, it does not look at the type of the TLD. They just compare the labels, both of existing ones and of applied for TLDs, and the comparisons are between all the types of TLDs, and all of those can be found confusingly similar. So, Ariel, please.

ARIEL LIANG: Thanks, Michael. And actually, I think Amadeu made a really good comment last meeting, reminding everybody about the process flow of evaluating new gTLDs. What he mentioned is that the string similarity review comes first before the application evaluation. The application evaluation includes GeoName, brand eligibility, etc. So still, the similarity usually needs to be resolved first before there's the subsequent evaluation of brand and Geo, etc. And I'm also going to post the applicant journey document from the AGB in the chat, and everybody can take a quick look at the process flow there.

So I think, at the stage of string sim review, the panel wouldn't really know about whether it is a Geo or a brand, and they're not in the

position to determine that. So I think I just want to remind folks about the comment from Amadeu.

MICHAEL BAULAND: Thanks, Ariel, for reminding this point from last week. Yeah, that's certainly also another point in favor of not restricting this to the TLD type, but just cover all of the TLDs independently of their type. Thanks. So, I think we are mostly in agreement that we don't want to do this, the restriction to certain TLD types.

Seeing that we have 10 more minutes, I think there's one more question that was raised, is whether we want to limit this to applied for IDN strings of existing ASCII gTLDs, or whether we want to include new gTLDs as well. So, any thoughts about restricting this to existing ASCII gTLDs?

Anil said that he thinks we should include both existing as well as applied for, and one thumbs up there. I also think that I see no reason to restrict this to cases that already exist now or have been applied for in the previous round. The policy is quite broad in that sense that the problem of these cases can occur also in future rounds, even if none of the TLDs existed. If we go back to the Quebec case, I see no reason to say like, well, Quebec created their TLD in the previous round, so they have the chance to get the correctly spelled version, but anybody applying in the future will have to decide whether they want the correctly spelled version or the ASCII only version.

I think technically there's no reason, and also from a policy and fairness, I think that everybody should have the same opportunity here.

Sebastien says the policy is here to last, not just to solve one case, no reason to limit, and Claude mentioned the Montreal case. Unless there is anybody with objection here, I think we should also make that possible for cases where none of the TLDs exist yet. Okay, I think we are in agreement here, and with just five minutes left, I want to hand it over to Saewon for the outcomes and action items.

SAEWON LEE: Thank you, Michael. So I'll be quick. There seems to be a lot in agreement today, and I might be a bit repetitive because I didn't follow exactly what was the outcome of last week. But just going off of the recap, and again, this AOB is also with the next steps as the action items also serve as the next steps for this group.

So first of all, I think the team is still in agreement to define the Latin script diacritics per Unicode. And I think we agreed that we will not define what similarity means or change the existing string similarity review process. But this working group will use the outcomes of the string similarity panel, and the recommendations that come from this PDP will only kick in when it's not defined under the review process.

And again, I think we have to bear in mind the consistency of terms from the existing body of work, and especially not to be confused with similarity as well as the user confusion. And related to diacritics, we decided to allow for multiple diacritic cases, while not restricting to certain TLD types, bearing in mind the process flow. So, for example, not restricting to real words or geographical brands, brand names. And lastly, the outcome from the last discussion was when discussing the

instances of what to limit, not limiting to only applied for IDN strings of existing ASCII gTLDs, but also accepting new gTLD applications, while limiting it to same entity operating the ASCII and diacritic versions. If there's no objections to this, I'll go straight on to the action items.

Again, this follows from what the outcomes were. First of all, just bearing in mind the consistency of terms and having to make sure that we follow the outcomes of the string similarity process, I think the team, we would like to suggest the team to look at the AGB language that was shared in the chat from Ariel. And again, we'll share this via the mailing list. I think leadership and staff will go back and also try to organize a meeting to invite an ICANN Org expert, maybe Samad, to share with the team on the string similarity process, the criteria and the details and the workflow.

And then lastly, which actually came about in the beginning of the discussion, is for the leadership and staff to try and list a Unicode table to look at all those characters or make a list or compile such a list of what is within or not within scope. If anyone has any objections to this or has anything further that I may have missed, please do raise your hand. Okay, I'll hand it back over to you, Michael.

MICHAEL BAULAND: Thank you. I just read Tapani's comment. Just for the record, I don't like using the Unicode definition for diacritic. It'd be better to use visual confusability. Users should not be expected to know that O with a strikethrough is technically different from O with two dots, but that's not a strong objection. I agree to go on with the Unicode meaning.

Thanks, Tapani. The objection is noted, but I guess the majority agree that we stick with the Unicode case as this is a clear-cut definition. And it's also the one handed to us by the GNSO Council. But we will make some notes, some recommendation that these other characters are not forgotten and that they should be considered in some other way after we have created our policy for the clear-cut diacritics.

Bill says there are some 40 diacritics. If you don't find that number, you have missed some. Thanks. Bill was involved in creating that repertoire, so he should know. Any other last comment? I see we are already at the top or half of the hour, or depending where you are, maybe at the top of the hour. If not, then I thank you all for joining the call, and Devan, you may stop the recording now. Thanks.

[END OF TRANSCRIPTION]