JIG
TRANSCRIPTION
Tuesday 10 August 2010 at 1200 UTC

Note: The following is the output of transcribing from an audio recording of the JIG meeting on Tuesday 10 August 2010 at 1200 UTC. 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. The audio is also available at:
http://audio.icann.org/gnso/gnso-jig-20100810.mp3

On page:
http://gnso.icann.org/calendar/#aug
(transcripts and recordings are found on the calendar page)

Participants on the Call:
Edmon Chung - Chair
Jian Zhang - co-chair
Avri Doria
Rafik Dammak
Fahd Batayneh
Sarmad Hussain

ICANN Staff
Olof Nordling
Gabi Schittek
Gisella Gruber-White
Glen de Saint Gery

Apologies:
Bart Boswinkel
Kristina Nordstrom

Gisella Gruber-White: Good morning, good afternoon to everyone. On today's JIG Call on
Tuesday the 10th of August we have Jian Zhang, Edmon Chung, Sarmad Hussain, Fahd Batayneh, Avri Doria. From staff - Rafik Dammak - sorry. From staff we have Gabi Shittek, Olof Nordling, myself, Gisella Gruber-White. We have apologies from Bart Boswinkel and Kristina Nordstrom.

If I could please remind everyone to state their names when speaking for transcript purposes. Thank you. Over to you Edmon.

Edmon Chung: Thank you Gisella and thank you everyone for joining again. And as mentioned, a couple people are on holiday today - on vacation today - so - and this is a - an adjusted timeframe. Instead of a two week timeframe we had the meeting last week because next week is the ccNSO conference call.

And so I think our next one would be three weeks from now in any case. In terms of the session, we're, I think today we're hoping to carry on the discussion last week which was focused on - we started our discussion on IDN variance.

Based on the discussion last week and as I mentioned in emails just earlier today and some of my own thoughts, put together a document for - I guess for discussion.

I think last week we left off by, you know, after making a relatively unstructured discussion, we left off with the (session) that we would try to identify a few aspects of IDN variant policies. So that's really what the document tries to encapsulate at this point.

So perhaps I can start by seeing, you know, if every - anyone has any questions or had a chance to take a look at the document. I guess the plan is to sort of walk through the document. But before I do that, anyone have any urgent thoughts or comments or whether you've taken a look at the document?

Hearing none, I guess it's...

Olof Nordling: This is Olof here. Well actually since it's at the very end of the document, I've got one little comment on the very last bullet point. Having read it in a cursory fashion but the very last one I think here, where you mentioned that for blocked variance to allow overlap. Just a concern that I have being very keen on string similarity assessment and such that all right, if that happens and we would allow blocked variance, and blocked variance for some reason, then in

a future (variation) would become permissible, I think we pave our way for some future problems.

So I would like to see it at least at the outset in order to have it (picture) proof for certain that we regard a blocked variance from that perspective in a similar vein that's preferred (in a) (served) variance.

Edmon Chung: That's a very good point. I have no idea why I described it this way. I was thinking in a strange manner at that time probably. But that's a very valiant - valid point if. It's a better approach to be a bit more conservative to start with.

Olof Nordling: Thanks. I also want to mention that since - well, in one hour maybe we don't reach that far in the document.

Edmon Chung: Right.

Gabriella Shittek: Yes, this is (Gabby). I'm just reading it now.

Edmon Chung: That's okay. Actually I apologize for sending it out so late. It's - it was a challenge to put something together within a week. At least I thought, you know, I'm glad that I was able to do that and thank (Jane) for providing some comments and adjustment before I sent it out as well.

So I guess most people haven't really had the chance to really read through it but I - perhaps I would start by sort of walking through it and, you know, we'll discuss each part as we go along.

And the - again, this is really a very early draft. You will see a lot of place holders just, you know, saying that more should be written here what - generally what (could) be written (and everything). But the - I guess the - there - the main focus is really to think about identifying the various policy aspects.

So I'll - I'm thinking of starting with that. So we'll skip, you know, the first two sessions, really, the background and what I identified as the - a glossary of terms and start with the policy aspects. And I think last week's call and some of my creative I guess consolidation came up with four areas but before talking about those four, I think last week we mentioned this as well.

I just want to make sure that we're sort of on the same wavelength at least to start with. I think a very fundamental - an important and fundamental part is to understand that an IDN variant is really a two different domains on - in the DNS and especially with existing standards and technologies, they are just basically two different strings.

What we're looking at is an IDN variant policy - policy framework - that would define essentially by policies that they are same domain. So this is, I think, a very important understanding or a very important starting point. If anyone I guess feels uncomfortable with this starting point, I guess it's - need to bring it up and need to address it right away and talk about it. Any thoughts, questions, comments?

Hearing none, I think another important starting point for the group, I think we mentioned it last time as well, is that the discussion is really focused on the TLD level. We understand (our) registries would implement, you know, the different language policies and all that.

That's, you know up to the TLD operate especially for ccTLD side of things. That is, I think, an important understanding and what we - what we're focused on with - at least within this group is what to do at the root zone at the top level for the top level strings. So that's, I think, another important starting point.

Avri Doria: This is Avri. Can I ask a question about that one?

Edmon Chung: Please go (ahead).

Avri Doria: Yes, I totally understand when we say that we're creating, you know, or suggesting or what have you policies related to the TLDs. But when you look at - and I'm not exactly sure I understand what the phrase means - policy operatives for implementation which is your fourth bullet. Might there be included within that constraint for what happens at other levels, and second levels and such?

So while the policy making is for the first level, it could have implications at other layers or you say, "No, it's only about how it gets stuck in the roots and has nothing to do with how second level and perhaps others would be treated."

Edmon Chung: Thank you Avri. In fact, that you for the sharp eye and in fact I was going to sort of bring it up later once everyone bought into that that it's - yes, I think you're right. I think, you know, there are certain implications and we probably won't be able to completely sort of de- how should I say it, detach the two.

But the focus, you know, the way that you put it is precisely, you know, sort of what - where one starting point. And there will be some, I guess, almost implication or incidental implications to - on - for those DV language policy that is being used for the second level. You're quite right.

And what I - what I'm trying to mean in terms of the policy operative is actually operatable policy, so what to do rather then a grand general statement about a policy. I think in the - I guess the experience at the new (DTLD) process we realized certain policy recommendations, once it gets down to - gets over to staff for implementation, sometimes that it's not operational then, you know, it really presents a challenge for staff to implement them.

So what I'm trying to say is really I guess trying to be more aggressive in setting a - policies that are operatable that's what I meant. I was trying to me-mean by there but (you) can probably, you know, suggest different wording.

Okay, so I guess that answers, Avri, your question and thank you for the observation actually. So I have so far identified four areas or four aspects of the whole discussion about IDN variance. And I'll just read them out basically. The first one which we sort of discussion - well, the requirements for strings to consider and IDN variance, I sort of avoided saying defining what an IDN variance is.

I mentioned the reason, trying to come up with a descriptive definition of an IDN variant probably going to be proved futile. We've tried it many times and - because there're different - different languages have different variant or flavors of variant. So that's - I think it's very difficult to try to do that so instead of that I think a probably more fruitful approach might be to try to identify a set of properties or attributes for something to be called an IDN variant.

So this is the first one. The second one is a follow on of that. In fact, the second aspect is, you know, once we try to talk about it that way, I think we would probably need to flesh out a little bit more a framework or in - I guess a better way to describe it might be a super set of the language policies that we have seen in the major languages that have been discussing this issue.

Once we, you know, we try to select a set of the type of attributes then, you know, we have essentially a framework in saying, you know, these are some of the elements that are common. These are some of the elements that optional for different types of languages and we might be able to come up with a - you know, a framework that says what a - an IDN language policy looks like and how it would produce IDN variance.

And if we have a sketch - a skeleton - of that, we can much better define a policy for it. So - and this leads into number thir- the third item which are the

types of IDN variance. I want to note that in the IDN working team final report there is - there was a concept about type one and type two variance. I sort of find that probably not as useful operatively and that's why (invest) comes back to Avri's question about why I mentioned operative.

It - the IDN working team report talked about IDN variance type one being visually similar one then type two being none visually similar but meaning - similar meaning one - identical meaning or something like that.

That, however, doesn't really quite inform a - I guess a policy operatable by ICANN which it seems to me at least like what the problem is previously. And instead of that, what I'm sort of proposing is define sort of types of IDN variance with respect to their allocation and delegation properties not by the linguistic policies.

Because, you know, my feeling is that if we go down the path of trying to identify it based on linguistic properties, because we won't have all the experts of all languages in the world, we probably won't be able to set up a policy that would be robust.

But instead of that, if we look at, you know, what giv- what type of variant instead of - in the light of whether they should be allocated, whether they should be delegated in the zone, that probably would allow us to end up with a more robust policy that can be used for the root zone.

And then number four is really just again the follow on to three. It's almost like four steps that I guess the - we need to go through. And the fourth one is really if we identified the types of IDN variance, what exactly policy wise ICANN should do when it receives an application that - an ATLD application that has - that requires - IDN variance.

Those are the four aspects that I have sort of consolidated from our previous discussion and then also added - I guess added some thoughts myself. It

might not be exhaustive so I guess looking for comments and input and what other aspects of this IDN variance this group needs to look into or any comments or thoughts on these four.

Olof Nordling: Edmon, Olof here.

Edmon Chung: Yes.

Olof Nordling: Reading it I was pretty confused between one and two which I almost find to mean the same thing. So probably should be clarified that it is some kind of high level requirement and more detailed requirements on number two because you can read both as being more or less equivalent.

Edmon Chung: I actually sort of agree with you. I was writing this and as I was fle- trying to flesh it out I realized that they're quite similar. The only thing is that one is a slightly higher level view of two. If people think it makes sense to actually put it as one so that it's less confusing, I'm - I actually am agreeable to that.

Olof Nordling: I wouldn't mind.

Edmon Chung: Okay yes. I think - I think it probably makes sense because when I was trying to put out points in the next section I realized that they're, you know, really there's one point about number one and most of the thing is in number two. So yes, I think putting them together makes sense.

                         Any other thoughts on other angles or, you know, aspects of this IDN variant thing that we think should be included in our discussion?

Sarmad Hussain: Hi, this is Sarmad.

Edmon Chung: Hi. Please go ahead.

Sarmad Hussain:	I just wanted to comment on I'm not sure of basically when will these variants be proposed? Is it only that the variants will be considered at application time or could they be added later on after (litigation) of the primary string?

Edmon Chung:	That's actually a very good point. At least in - for the time being, the - well when I was preparing the document, the - their thinking was that it would be - it would need to be included in the application especially for the ones that you want delegated. And it has to be defined based on the IDN language table - IDN language policy that is included also in the application.

	The question about, you know, I guess whether at a later date you might want to apply for it, that I think I included a short note about that for sort of reserved variants to then apply for them.

	I had - in that little note I think it's in - on which page - under Number 4, to have it seen - you know, consider it as separate application, I had a - I had I think somewhere - oh yes, the third last point at the end of the document, the view - I guess the general view right now is that you should have your IDN language table set up and have something that's, you know, well - robust before applying for a TLD, okay so there isn't an expectation that it would change all the time.

	But I think what you said is valid if it changes in the future what happens. But on the issue of reserved variance, because the preferred variance (or) the ones that for example that you would automatically include that - conceptually that's the thing, but reserved variance there - they might be considered a different application altogether. And for that one...

((Crosstalk))

Sarmad Hussain:	I was also suggesting the possibility where at the time of application there's no variant suggestion and then later on, maybe a year down the road when the TLDs are really dedicated, there's an (eviction) application to include the

variant (foreign), you know, (really) dedicated string. Is that going to be possible?

Edmon Chung: I think that's a good point. I think it's probably, from where I see it, seems sensible. As long as this is a, you know, it's a different - it's - you pr- you will have to go through a different application process. I think that makes a lot of sense.

What I'm co- what I - the question is whether we want a special application process for these types of things or this is an application process that needs to be included into the, for example, the IDN and ccTL - the process. And the new gTLD process, like just using the existing - well, not existing - but using the then current process to apply for a variant.

I think both would work. Do you think there needs to be a special consideration or do you think, you know, just going through the regular application process as a different application would probably work?

Sarmad Hussain: Actually the reason is because the current application process perhaps as used that there's going to be (assigned) (restring) which is being proposed. So I'm actually just contemplating that scenario where primary string is already approved and then already dedicated and being used.

And this application only includes one or two variants which people realized that may actually be beneficial from user's perspective once they go and apply the IDNs. So there will be no primary string in the application for delegation. And I'm not sure...

Edmon Chung: I think it's (distinctive)...

Sarmad Hussain: I'm not sure it's going to - I'm not suggestion a different process I'm just (unintelligible) something. And it may be, you know, catered - it could be catered to an existing application process. I'm not sure.

Edmon Chung: Okay. No, I think that's a very valid question and definitely we should, you know, include it in - at least in the initial report, so throw it out and see if other people have thoughts about it.

And in terms of addressing it I think, you know, either having a special process for this type of application or just going through the regular process, those are two possibilities I guess and we can put all this into the initial report and see if we get any further comments back.

Avri Doria: This is Avri. If I could add something on that theme.

Edmon Chung: Please.

Avri Doria: On the ccTLD application process it sounds like it's almost a small tweak to the application process to add something to say if the primary string has already been allocated what is it and go on from there.

In the new GTLD process, we would need to do something completely different in that there really is no (recourse) strong primary string, secondary string, what have you, variance as far as I understand it (all). So for something like this to work in IDN and GTLD as I said, there would need to be (serious) thinking on how to rework part of the process just to do it in the first place, let alone to add one later.

Both of those need to be rethought and I think while rethinking how you would do it in the initial (incident), you could take care of the how you would do it in the later incidence of adding one.

Edmon Chung: Probably. But I think - no, Avri, I think you're right. I was thinking that if (beams) first thought that they're quite different. But, yes, if we, you know, deal with it so for the initial then we will probably be able to use a similar way to deal with it for the sequence. Yes.

Olof Nordling: Olof here if I may. I think that well given that you already have the primary string and then you would come with the request I think is a very valid question really to bring up because things will change and we're just in the beginning of IDN top levels, the main deployment. So probably revelations will come and people will realize that, yes, okay we have this variant and we would like to introduce that a few years down the line for example.

And I think it's less cumbersome then full application process. And I think it's likely more cumbersome then the (R sep) process which is another way of changing services of course for an existing registry. So I just realized that this is worthwhile thinking about and the question is very well put I think.

Avri Doria: Hi this is Avri. I wasn't trying to say that the question wasn't put, although I tend to think that you adding the variant may be more cumbersome then people might allow for. And this is given my concern about what happens at the second level and beyond in variance and what happens, you know, in terms of how one deals with similar strings in that.

So you're right. The process of getting it into the root may be simpler but the process of setting up the behavior of how it will work below that may be more difficult.

Olof Nordling: That's very correct Avri.

Edmon Chung: Right. Yes, I think the - those are very valid points. I will definitely need to include it in our consideration. Given that we're on this topic, do we want to just jump into, (something on the list) or just go back to the flow of the document?

It's probably better to go back with the flow of the document because it's (really) important to start off with what an IN variant is before we jump into

this whole discussion about what to do, you know, in terms of allocation and delegation.

So yes, adding variants later. That's - I would add that as number five in terms of the list of policy aspects that should be covered. Anything else on people's minds?

Again, I apologize this document came out so late so perhaps, you know, when you think about - you know, when you come up with something please feel free to throw it to the main list as we continue this discussion.

So why don't we move on in terms of walking through the document? I think the next section, the idea is to collect some few points for each of the aspects. So I'll start through the four of them again I guess.

The first one was - I guess we can talk about the first two together really as mentioned by Olof. The first two are the requirements for a string to be considered an IDN variance and the framework of sort of the attributes that compose this - an IDN language policy such that through which an IDN variant is produced.

So as I mentioned, it's - the - I was looking to this issue and looking at all the - you know, a lot of the background materials. The - it seems to me that it's quite difficult to try to come up with a what I would call a universal definition of an IDN variant absent an actual structure that says, you know, these are the things that is - that will produce an IDN variant.

And this is where - this is sort of basis of number one and two. And just quickly going through it, in terms of number one, again I'm - I think one important thing is that based on the IDN language policy these IDN variants should be consistently produced.

And what I mean by that is that I guess there is always a concern that a variance is arbitrarily created for application and certain abuses might be - might result in that case. So in terms of that I think it's important that, you know, you have IDN language policy, you put in the primary IDN always comes out with the variant strings. And that's really the sort of overarching concept.

And then in terms of framework this is where I used a lot of my, I guess, creativity and - I'll bring it up and see what people think. Talked a little bit about it but I was looking at all the variant policies - all the IDN language policies, you know, that have been put in place so far.

And these are some of the common elements. We would have a set, you know, at least one or a set of character tables. We would have - if there are variants there would be at least one that contains a character variance mapping table like which character map - which characters.

And then there would be a set of rules describing how the IDN variants are generated, you know, how - by permutation or whatever rules. And then there would be a set of rules that would describe what type of variance, which ones created would be put in the zone, which ones would not be put in the zone.

And then finally there would be a - might be a set of rules describing any allowed or disallowed combination of characters or placement of characters to add to that as well.

These elements I think describe at least all the language policies that I've seen so far. I wonder what people think - do people understand what I'm trying to say, in Number 1 and 2? Or if anyone knows of a IDN language policy that has elements that are beyond what is being described?

Sarmad Hussain: All right can I give a couple of examples? This is Sarmad again.

Edmon Chung: Sure.

Sarmad Hussain: And I'm not sure if they are - these examples are captured by these statements. So for example in Arabic script one could write a TLD with the optional vowel marks or without those vowel marks. And so the vowel marks are really optional, so they're not really - they don't have a variant as such. But there are two strings, one which is specified with the vowel marks and one which is not specified with the vowel marks.

And the two are actually considered equal by the users, because users can, you know, while reading they type those vowel marks automatically. So that's one concern. And then there are others, you know, for example languages the - even Indic languages there's this concept of writing this mark, it's called (hallan) which sort of closes or mutes inherent vowels and that is sometimes also optionally written and readers can - so these are combining marks which are sometimes optionally used. And I'm not sure whether that's - that is captured here.

I was also thinking about abbreviations. There are sometimes used and could be potentially considered variants of longer strings. And I'm not sure, you know, a mapping technique would not be able to map the two onto to each other. So I'm not sure if something like that could also be considered a variant in TLD context.

And if it is considered a variant, then how that would be catered to as well so there's a couple of examples.

Edmon Chung: Right, for the first two yes I'm quite aware of that. And I think usually the - I guess that's where I use the operative framework. You know, in terms of describing those - the situation with Arabic and some Indic languages what is usually used is Point Number 2 and 3 the first one being in terms of - it's described in terms of a sequence of characters.

So usually the mark is for a particular character. And if these - if a string of these two characters you would map to a particular character, right, which means that it omits the - the mark. And that is how the character variance table would be set up. And then the set of rules would basically describe in terms of how the primary IDN would be created, I mean, how the IDN variant would be created from the primary, that's the set of rules basically describing it.

But there would still be a character table which would identify the sequence of sort of characters and its variants, being that the sequence of characters being the character with the mark and then the variant being the one without the mark. Does that make sense?

Sarmad Hussain: Actually if you for example map there are - as you are assuming there are two different vowels each accented by a different mark. And essentially I think from what I understand you're saying, that each of those marks is mapped onto a null as a variance.

Edmon Chung: Right.

Sarmad Hussain: So if they don't appear the string is still compatible or, you know, the same. But...

Edmon Chung: Right.

Sarmad Hussain: ...that could create a problem when you're coming back. So if you have let's say a string without any marks and you want to see what it's a variant of, then you will actually over-produce because null - you would have all nulls and then you'll create all the three vowels at each possible place.

I hope you understand what I'm trying to say.

Edmon Chung: Correct, yes. I do understand and that is the situation. And therefore there needs to be a categorization of those variants. So certain - a number of them you would you over-creating but then those would be considered a special kind of variance where you just don't - you don't allocate them or delegate them.

Sarmad Hussain: I just need to think it over.

Edmon Chung: Sure. But operationally that's really what happens. Like you over generate, but then you basically okay these ones are really over-generate and this why we consider these as a type of variant that was over-generated essentially.

                 But, you know, again I think there might be some issue there; I'm not saying that there might not be. But that is definitely - at least in terms of considering this model that was my original thinking. And I think it makes sense to specifically talk about this issue in the write-up so that people can be more clear. And if there are problems with it we can more readily identify it.

Sarmad Hussain: Yeah. And there was a second abbreviation...

Edmon Chung: Yeah the abbreviation part is an interesting discussion. Off the top of my mind I would think that on the top level it might not be appropriate. The way that I see it is that unless it's an - is a result of, you know, going through the language table and abbreviation probably won't be it's very difficult for ICANN to make a determination of whether an abbreviation is a variant or I guess quote-unquote variant.

                 Because if you add that into sort of the definition of IDN variance for the root zone management then I think, you know, it really opens up a much bigger concern. I don't know if others want to add to that.

Olof Nordling: It's Olof here. If I can chime in. I tend to agree that - I mean, we're bordering to the situation where we have translations almost and an abbreviation while

you have done something else with this string in a very clear sense. And I think that ends up with a lot more problems than it solves, really, to accept that as a variant.

I want to go back to thinking about your Arabic example and if you want to have a string which then normally would be let's say three consonants long and then you have intermediate vowels and that would be a variant. But if I understood Edmon right the idea would then be to have a primary in the sense vocalized.

So you had that one and then you could have the non-vocalized as the variant or in order to perhaps avoid all the mixtures of all the other vowel combinations that could occur.

I don't know if that made any sense but I was trying to reflect upon it myself and I found it - well it's becoming a bit difficult to have well vocalized and non-vocalized as variants of each other. But it would probably be simpler if you had a vocalized as the starting point and the non-vocalized as a variant of that one. Just very, very - a thought from my side. Thanks.

Edmon Chung: Thanks Olof. I think it makes a lot of sense in terms of the concept. In total abbreviation (unintelligible) just jumping onto it, we probably should include some description about it. I guess the question is, Sarmad, whether, you know, you feel comfortable that at least in the point saying that abbreviation really is not, you know, to be considered as an IDN variant.

Woman: (Unintelligible). Please hold on.

Olof Nordling: Okay we...

((Crosstalk))

Edmon Chung: Hello?

Sarmad Hussain: Sorry, yes I am okay with that. But so what I was actually trying to bring up was that it's probably important not only to define what variants are, but also to sort of the more specific of what variants are not.

Edmon Chung: Right.

Sarmad Hussain: And I think we need to sort of try to spell out as many things on both sides rather than just trying to define what variants are.

Edmon Chung: Right that's very good point. And, yeah, as Olof mentioned translation and also you mentioned abbreviation, I think, both of this really should not be considered IDN variants at least that would be a starting point. And we'll throw this - again we'll try to create a document and throw it out for comments and see what the larger community thinks.

But in terms of - going back to the Arabic and Indic examples I think Olof talked about it in a more clear way than I did. Does that sort of address your concern?

Sarmad Hussain: Actually the primary string is -- from my perspective since we use Arabic script -- would be the one which is most used by the users so that is - which is going to be most likely typed by the user when they're accessing a TLD. So I think...

Edmon Chung: Right.

Sarmad Hussain: ...perhaps looking at it from a vocalization point of view or some other linguistic point of view is also a good thing. But I think from a usability perspective it's probably best way to so, you know, that's my opinion that I think that would be the best way to determine what should be the timely TLD string.

Edmon Chung: Right.

Olof Nordling: Sarmad and Edmon, Olof here. I think it's almost like if you would need an additional concept; something of having like you say the non-vocalized as the primary string because that's mostly used, I mean, you don't use vocalized text very much...

Edmon Chung: Right.

((Crosstalk))

Olof Nordling: ...I think.

Sarmad Hussain: Yeah.

Olof Nordling: But then perhaps something called an additional concept may be needed so call the base string which is actually...

Edmon Chung: Right.

Olof Nordling: ...some kind of notion that this is the basis from which we actually start linguistically speaking but the primary string for the IDN purposes is the non-vocalized. If that helps, I don't know but I just see a need for two different concepts.

Edmon Chung: I think it makes a lot of sense. In fact just before you said it I typed it on my notes saying primary string and then base string question mark. So we're completely in the same wavelength there. So you would have the primary string - primary IDN which is the one that most users use. And then you would basically use a base string to create the variance.

                 I think that is a very useful concept. The question then becomes, you know, what to do with the base string; is the base string also delegated and stuff but

that's also, you know, the base string it's by definition is an IDN variant of the primary IDN. So it would follow the discussion on what to do with these IDN variants in the - in the discussion.

Sarmad, does that make sense? You know, if we have a type of thing that's called a base IDN which may or may not be distinguished from the primary IDN - the primary IDN could be the same as the base IDN but in certain cases the base IDN might be a different one which, for example, in Arabic might be the vocalized ones.

Sarmad Hussain: Actually the introduction of base string as something separate from a primary string I'm just thinking it may cause confusion - further confusion. So, you know, I think we have primary string right now and variants and variants are motivated by many other things and base string is perhaps one of those things to which variants are motivated.

But I'm not sure whether documenting it will make things more clear or make it more confusing. That's my sort of initial thought of it. Another question which I wanted to raise in the similar note is that in many countries there are multiple scripts used.

And so the question is I agree translation is a different TLD. Is transliteration a different TLD as well or not?

Edmon Chung: Okay.

Sarmad Hussain: Or is it a...

((Crosstalk))

Edmon Chung: In terms of your first item there the base string concern, I think, you know, your concern about confusion is valid. Perhaps what we can do is in the document we can identify that this was being discussed sort of on a separate

note and put it on a separate note that this was discussed and, you know, as a possible thing.

In terms of...

Sarmad Hussain: Okay.

Edmon Chung: ...transliteration my first reaction is probably that no it's quite similar to translation in a way because then you talk - there are two issues with transliteration one of which is that you would end up with a language table that has multiple scripts.

And that - that's somewhat a borderline case in terms of the ICANN IDN guidelines where you're really supposed to - in terms of an IDN language table - IDN language policy really should be contained within one script. Unless those scripts are, you know, really, you know, in a country where you would write a word combining the different script elements which is, for example, in the case of Japanese they use various scripts and they combine it within one phrase.

So in - I guess your example that probably isn't the case. It's, you know, you're really writing in one script; the whole thing is one script. But you can write the whole thing again in a different script; sort of transliterate in a different script.

In those situations at least my point of view at this point seems to me like that, you know, they are - they would probably be considered different language policies all together. And different language policies would regulate them and they're two different IDN applications.

Would that make sense? Other people have any thoughts on that?

Olof Nordling: So just add to the mess that I think transliteration creates because, I mean, the transliteration rules - there are transliteration rules but they vary. So sometimes they're not - they're not - they don't remain the same and there are parallel transliterations depending on between which languages.

Just to take one example I was brought up when Beijing was Peking, transliterated to Swedish. Then transliteration has changed to Beijing. So you end up with - you have a very weak basis for making that kind of connection. And it's very likely to change.

So I think it's really a risky path to make some - well to see that as variants of something that is going to stand the test of time.

Sarmad Hussain: Can we sort of - so yes I agree generally. What about when we talk about transliteration in base (LDH) and in other script? Because (LDH) is already part of the - any language table irrespective, right?

Edmon Chung: Not necessarily, most of them do but not all of them do. There are - I'm pretty sure that I've seen language tables where (LDH) is not included in the table.

Avri Doria: This is Avri. I'm about to sign off because I have another call. But this brings up a whole issue that I think we've talked about a little and I get into my sounding like a broken record in terms of what happens to these strings once they're declared variants and what it means at their - at their other level. Because to say that you're associating a transliteration, the (transliteralization) or you're associating an (LDH) are you really associating it as a variant?

And we haven't really discussed what that means because are the second levels going to also be at that transliterated language or character set or (LDH)? So I'm not quite sure what it means to sort of say they're related strings under the same authority and that being something different then they are a variant because they are treated in a certain way.

And so I think as we widen the definition of what it may mean to be a variant we start to get really confusing as to what might happen at the lower layers - or lower levels - or higher levels or whatever is the right direction - of that name.

And so, that's okay I have another call I have to chair at 9 o'clock so I'm signing off now but I start to get really confused with the direction we're taking and understanding what it means to be a variant as opposed to just a string controlled by the same authority. And there may be a difference between those two.

So thanks.

Edmon Chung: Thanks Avri. Thanks for being on the call as well. And I think that's a very good point. And if you're still on I think that's really actually what you mentioned is a - Aspect Number 6, you know, really looking into IDN variants and the implications on the second level especially really the zones, what - are there a particular set of requirements for the zones of IDN variance TLDs?

Avri are you still on?

Avri Doria: Yeah, I am.

Edmon Chung: Did I make sense?

Avri Doria: Yes that does make sense but that means there's also implications to what we define as variants or the word is getting broader and broader and broader in its definition that we have variant in the very narrow sense, variant in the narrow sense, etcetera. Bye, I'm sorry.

Edmon Chung: Right.

Olof Nordling: In this sense...

Edmon Chung: Did...

Olof Nordling: Olof here.

Edmon Chung: Please Olof - please.

Olof Nordling: If I may continue on that? Those concerns on the second and other levels well they are - will appear if we cross the sort of language and script order. So, I mean, remember the old fashioned definition of a variant character, I think it was something like autographically equivalent character to something else and then you could use those.

     And apparently that's not enough. But at least we - whether we stay within the same script and language as the original string I think that's a very important aspect.

Edmon Chung: Right. Yeah those are - I think those are very good points. And I think in terms of - I want to come back to just quickly noting this thing. In terms of - at least in terms of transliteration, Sarmad are you comfortable including in that list of things including abbreviation, translation, transliteration as sort of at least in the starting point to not include them as IDN variants?

Sarmad Hussain: Yeah, I think that's where it should go.

Edmon Chung: Okay. Okay.

Sarmad Hussain: But there may be some cases like Japanese, Chinese you were talking about where this may actually not be true.

Edmon Chung: What do you mean by that?

Sarmad Hussain: You know, for example it depends on whether you consider - so in cases where and normally they use a different writing system so different writing systems are used so there may be two ways of writing the same thing. I'm just thinking - I've not really familiar with Chinese/Japanese coding and scripts, but there may be something there.

Edmon Chung: Okay.

Sarmad Hussain: For example...

Edmon Chung: That's interesting.

Sarmad Hussain: ...Chinese and simplified Chinese.

Olof Nordling: (Iragon) and (Katakan) and (Kanji) and such and, I mean, well where a (Kanji) sign can be written in (Iragon)...

Edmon Chung: Right.

Olof Nordling: ...and they would be - well the strings would be in some sense in Japanese autographically equivalent even though the individual signs are not necessarily. Well hard to tell. So, yeah, well there will be special needs but I think we have noticed that.

Edmon Chung: Right. But I guess in terms of identifying as IDN variants it's probably - that's why one of the things I wanted to bring up is that if we want to sort of define this in a way where we can operate it, you know, operate the policy.

And the examples especially with transliteration I think it's probably, you know, more suited to a different type of thing instead of calling it an IDN variant. Because it's a alternative way of writing something rather than, you know, in the case of I guess Arabic where you have the marks or not and in the case of Chinese where it's simplified and traditional Chinese those are situations

where they are linguistically considered really the same within that particular language and writing system.

So we're talking about something that is - it seems to me that is still slightly different. And in the case of Japanese at least so far the Japanese sort of - let's say authorities in a way - have maintained that, you know, in the concept of domain, in the concept of IDN variants, (Iragon), (Katakan) and the (Kanji) not mapping them is the appropriate way to go about it for Japanese, because they - while sometimes considered equivalent or sometimes not considered equivalent is one of the things.

So I think we've exhausted the hour. We've - I think we've covered pretty good ground. I'll try to put all the discussions that we had today and update the document and send it around for discussion further. And again if anyone has any thoughts please feel free to use the mailing list and throw in your ideas and thoughts especially given that this document came out so late.

Hopefully in the next few weeks we'll have further discussions. So the next meeting would be three weeks from now because this meeting was moved up from next week. So I don't have the date with me but I think (unintelligible) will keep us on track. And I thank again (unintelligible) for helping us on that as well.

Before I wrap up I just - any other questions or points people want to raise? (Jane), did you want to add anything before we close?

Jian Zhang: No I'm fine.

Edmon Chung: Okay in that case thank you everyone for joining tonight and good night.

Olof Nordling: Edmon, thank you. Bye.

Sarmad Hussain: Thank you. Bye.

Edmon Chung: Bye.

END