MARRAKECH – IDN Implementation Guideline WG Meeting
Wednesday, March 09, 2016 – 17:15 to 18:30 WET
ICANN55 | Marrakech, Morocco

UNIDENTIFIED MALE: This is the ICANN55 IDN Implementation Guideline WG Meeting on March 9th, 2016, at 17:15 WET in the Amethyst room.

EDMON CHUNG: I guess the recording started and I see people continue to trickle in, but I guess there'll be a little bit of a starting introduction anyway. Welcome everyone, this is the IDN Implementation Guidelines working group. We have an open meeting here in Marrakech. Maybe first of all, just go around the table, just tell us who you are and kind of like a roll call. Can I try to get the attention from the end of the table? Just to introduce who you are and where you're from.

ZUAN ZHANG: Sorry. My name is Zuan Zhang and also AKA Peter Green from CONAC, based in Beijing, China.

[GENE]: Gene [inaudible], network consultant for [inaudible] Networks.

Note: The following is the output resulting from transcribing an audio file into a word/text document. Although the transcription is largely accurate, in some cases may be incomplete or inaccurate due to inaudible passages and grammatical corrections. It is posted as an aid to the original audio file, but should not be treated as an authoritative record.

[RADU]: [Radu] [inaudible] from [inaudible] India.

[NIKI]: Oh yes, this is [Niki] speaking, I am from the [I-group] based in China.

UNIDENTIFIED MALE: Good afternoon, my name is [inaudible]…

CHRIS DILLON: …I'm from University College London, I'm a member of the Chinese Generation Panel and I'm co-chair of the Latin Generation Panel.

MICHEL SUIGNARD: Hello, my name is Michel Suignard, I'm a member of the integration panel.

BARRY LEIBA: Hi, I'm Barry Leiba, I'm currently the applications area director in the IETF and participated in the IDNA specs.

EDMON CHUNG: Edmon Chung from DotAsia and currently the co-chair from the GNSO for this working group.

UNIDENTIFIED MALE: [inaudible] member of the integration panel.

FRANCISCO ARIAS: Francisco Arias, ICANN staff.

SARMAD HUSSAIN: Sarmad Hussain with the IDN program, ICANN staff.

ABDESLAM NASRI: Abdeslam Nasri from Algeria, member of the Arabic Generation Panel.

EDMON CHUNG: It's okay, if you're just visiting, we'd like to know too.

[VIA PESQUE]:   My name is [Via Pesque]. I'm from France but I'm representing a company in South Korea. Thank you.

EDMON CHUNG: Thank you, and those who are connecting online?

MATS DUFBERG: I am Mats Dufberg from IS, and I participate in this working group as a co-chair.

EDMON CHUNG: I see. Dennis and Kal, apologies, you won't be on the audio but I will try my best to look at the chat and someone would help alert me if there's a chat. For the recording when you speak, please just state who you are.

UNIDENTIFIED MALE: [Gal Fahar] just put in his introduction, he's from New Star at Melbourne, Australia.

EDMON CHUNG: Okay. Welcome Kal, Dennis as well. Dennis from VeriSign, also in the working group. First of all, I apologize for my absence in the last meeting from the working group, but in this meeting we're hoping that we would introduce some of the work that we have done in the last little while and have a little bit of input from the community here. And then if we conclude that and we still have time, we'll plan a little bit ahead on what next steps to take.

On the screen, you can see that a very brief document that outlines a bit of the background and we you are, and then there are six topics actually we've identified from the working group

that we would spend a little bit more time on moving forward. If whoever is controlling it can scroll – okay. Someone of you can help scroll down to at least reveal the six. Oh, perfect, it fits perfectly on the screen.

In the earlier session if you were here, apologies that it's a little bit repetitive, but we've identified six areas: transition terminology, format of the IDN tables, consistency of IDN tables, IDN variance, similarity and confusability of labels, and registration data. Essentially, that contains IDN's related information.

These six areas are what we think the next IDN implementation guidelines should include. Some of them have been included in previous versions. Maybe not the same envisioned detail and some of them have moved some of the implementation practices and experiences have evolved, so these are the six areas.

I guess I will spend a little bit of time explaining each of them so that we can get a little bit of feedback from those who are just joining us or just from the community. The first one is really about transition and terminology. The IDNA standards were updated between 2003 and 2008. There are implementations out there. I understand that it's still not completely transitioned over so there's some lingering transition issues. That continues

to be an area that we think is important to point out. And related to that a little bit is also to continue to emphasize the importance of keeping up and also being compliant with IDN standards and RCs, so that's part of number one.

The other part of number one is terminology. The terminology, including what we call IDN variance, including what is now called LGR, label generation ruleset, it used to be more about IDN tables or IDN policies. The terminology has evolved a little bit as well, and there are multiple terminology. One somewhat ambitious attempt would be to try and bring some order to that and maybe suggest a way to use certain terminologies when we talk about IDNs in the context of registries and registrars at ICANN.

That's number one, the first area that we hope to spend some time around. I guess, Mats, do you think – I'm trying to ask my co-chair here – do you think we should stop every six point and ask for input, or it's better to do all six? Mats, do you have a...

MATS DUFBERG: I think that it's better to go through each bullet.

EDMON CHUNG: Okay, so that was number one. Any thoughts or questions? Oh, Sarmad, please.

SARMAD HUSSAIN: I think what we would also like to emphasize is that we actually have a wiki page, and what we've done is we started listing much of this new terminology which we're discovering on that wiki page. So there's actually an ongoing list which we are currently developing, so we actually encourage you to come and visit the wiki page, look at the list of terminology which is currently there. Please feel free to contribute further to it, and as we mature this list, we will also start adding definitions. So please come to the wiki page and you can actually contribute directly there as well.

EDMON CHUNG: Thank you, Sarmad. Any questions about the first topic area? Mats, anything you wanted to add?

MATS DUFBERG: I have been working with PDT for the last two years, or actually more than two years, with the IDN implementations. We can see that it's really necessary to straighten out the terminology, to have a common language between registries and other stakeholders of IDN.

EDMON CHUNG: Thank you, Mats. And, yes, that is a very important piece of work I think, but it is also somewhat ambitious. There have been attempts in the past, and this continues to evolve. But I agree that we should put a stake in the ground and try to come to some agreement in terms of terminology and what we really mean when we say them, because this is also going to have some implications on – because this is part of the contract actually with new gTLDs, so this has contractual implications as well. This terminology and how we really define it hopefully will be helpful, and if we don't do it well, it will be disastrous in some ways, but that shouldn't happen. That's the topic area.

Seeing no other questions on that. I see a number of people have just come in. Just to summarize, from the working group we've identified six areas that we want to have in the IDN implementation guidelines. We're going through it one-by-one to try and get some early input from the community.

We were just going through number one. Number one was about some lingering transition issues, emphasis on staying with the RFCs and the standards, and also an attempt to bring some order to the terminology we use at IDN. I keep trying to use the word "standardize," but I feel that might be too strong. But at least bring some consensus around the terminology we use. That's number one.

EN

I will move to number two. The number two area that we identified was the IDN tables, or the format of the IDN tables themselves. This has evolved quite a bit from the last version of the IDN implementation guidelines. And I guess just as a background for people, in the very, very beginning, the first IDN implementation guidelines, the publishing of the guidelines actually in essence created the IANA-IDN table database.

This is talking about moving from that more fluid format to a more standardized format that more recently has been discussed at IETF, and also I think it's in working group last call. It's on its way to become an RFC, and in order to reflect that, this should be included in the IDN implementation guidelines in the next version. That's the thinking behind it, and this is basically number two. Any addition?

UNIDENTIFIED MALE: Mark, is it possible to just give an overview of where the process is as far as regarding of this format is concerned at IDF?

EDMON CHUNG: Where we are, and what's the anticipated…

ICANN|55
MARRAKECH
5 – 10 MARCH 2016

EN

MARC BLANCHET: I can respond since I'm the co-chair of the lager working group. Yesterday – we did working group last call about a month ago. We got some comments that were integrated into the document. There were no people against the document, so yesterday we pushed a button for the IESG to review, so it's called a request for publication. It will go through IETF last call. It might be short before the IETF meeting, which is in about three weeks, so it may happen after. And then if everything goes fine, the IESG could approve it. I don't know, maybe May time frame.

BARRY LEIBA: I can answer that, because I know the schedule. I'm the one Marc sent it to.

MARC BLANCHET: Yes, actually it's now your problem.

BARRY LEIBA: It is now my problem so I have to review it and I don't expect that to be an issue. I'll issue the last call. The next telechat that it could be on is April 21st, so the IESG could approve it on April 21st and that's likely to be when it gets approved. I don't expect to see any problems with it. Then it goes into the RFC editor queue and takes some number of weeks before they get to it, and that's it.

EDMON CHUNG: Thank you, and it's great to have all these people in the room. Definitely you will be faster than this working group. This working group will probably take a little bit more than April to come up with our reports. But anyway, this is the situation for the format of the IDN tables in the background. Any questions? Seeing none – Sarmad.

SARMAD HUSSAIN: Just a comment. Just to inform everybody here that while this machine-readable format for IDN tables becomes RFC, gets converted into an RFC, we're also doing an implementation of this, and that implementation is also available through the ICANN website. You can go to lgrtool.icann.org. And it does require authentication, and I'm happy to provide that to you if you want to use it or test it. It actually provides you a platform to create LGRs in this format, using through a user friendly interface, and then use those tables to actually validate data as well.

EDMON CHUNG: Thank you, Sarmad. It's lgrtool.icann.org. If you go there, you'll see an authentication. I just did, and I couldn't get in, so we need to ask Sarmad for a login and password.

SARMAD HUSSAIN: Sorry to interrupt, but you can e-mail to idnprogram@icann.org to get that authentication information.

EDMON CHUNG: Thank you, Sarmad. Mats, did you want to add anything before we close on number two and move forward?

MATS DUFBERG: Yes, I think this is an important thing to make all the registries to change from the old format to the LGR, but at the same time we have to realize that the LGR format is quite complex. The old format – at least on the surface – is very simple, so this migration will probably take a long time. I don't know how we would be able to convince all the registries to do this migration.

EDMON CHUNG: Thanks, Mats. Actually, that brings me to kind of a question. If I'm not mistaken, when we were going through the engineering for the LGR, at the lager group at IETF, the previous tables were somewhat used as a starting point to make sure that the end product would be able to at least describe those tables. I don't think backwards compatibility was sought as a requirement, but

at least there was some discussion and we did look at that. I wonder if Marc can enlighten us.

MARC BLANCHET: The current specifications do much more than the current tables, and we know that we can convert all the current tables automatically to the new format. So in that sense, there's nothing to do for the registries. And I'm not IANA but I understand that IANA intend to essentially convert everything to the new format, but I'm not IANA, so...

EDMON CHUNG: Thank you, Marc.

MATS DUFBERG: I'm sorry that it will not be possible to do such an automatic conversion, because the registries have interpreted this format in little bit different ways. Secondly, the contextual rules are not included in the tables, and those have to be coded into the LGR, so it's not that simple.

EDMON CHUNG: Thank you, Mats. I note this is slightly out of scope of the IDN implementation working group, but I think it's of interest. I do

see that Francisco wanted to add to this. Can you shed some light on…

FRANCISCO ARIAS: I'm not IANA, but the other thing to consider is that some of the tables have had contextual rules and all their IDN policy in comments, which will be of course not possible to convert.

EDMON CHUNG: But the contextual rules can be described in the lager format. There are some comments that may not be completely –

UNIDENTIFIED FEMALE: [inaudible]

MATS DUFBERG: Yes, but if the contextual rules today are written in plain language, then you cannot convert that into LGR in any simple way. Then you have to code it in.

EDMON CHUNG: That's correct, but again, this is a little bit outside of the scope of this working group. It is of interest to this working group though. I'm not sure, Mats, whether you want to take this further or how

you want this working group to keep an eye on that development or not.

MATS DUFBERG: No, this is more a comment that we shouldn't think that the LGRs would just simply be implemented by the registries. It will take time.

EDMON CHUNG: Okay, noted. I guess that brings us to the end of number two, unless there are other comments or questions.

Seeing none, I will move to number three, third topic item. Somewhat related to number two, which is the consistency of the IDN tables across different registries and even identifying across levels. So top level, second level and, might I venture to say, the third level as well. This is the area that we want to explore how much we can talk about in the IDN implementation guidelines – I'll use LGR now going forward, that's label generation rules – across registries and reference IDN tables that ICANN is working on.

I think Sarmad mentioned briefly when we talked about number two, so number three is really focusing a little bit more effort, more attention on how much should we recommend or – not really prescribe, but at least speak to the desired level of

consistency or non-desired level of consistency across TLDs and across levels for IDN tables. That's number three. Mats, did you want to add to that?

MATS DUFBERG: I think that from a user perspective, it would be very helpful if all implementations of, say, Japanese are the same, so that you expect to be able to register the same labels under all TLDs that support Japanese tables. Today we see that different registries have different strategies. Some registries remove all code points that require contextual rules because they don't want to have that, and others include such code points, so that we see differences in, for example, Japanese tables between different registries.

EDMON CHUNG: Thank you, Mats. I have a queue here with Michel and Chris. Michel?

MICHEL SUIGNARD: Yes, I just want to mention that this project to do second level LGRs at this point, among those, 29 [language bases] in fact Japan is one of the LGRs we are working on. I also want to mention that when we did the conversion from some of the tables at that conversion in plaintext – obviously, we did

conversion from plaintext to actual XML LGR rules, for example for Hebrew and Arabic, which has some of those, or even for some of the – I can't remember now, but there were some IDN tables that we found contextual rules and obviously we converted them.

That would be examples that could be provided by ICANN as reference points. I'm not saying that everyone has to follow them, but at least it would be a good inspiration or reference model for anyone that wants to use those languages in whatever level they want to use them.

EDMON CHUNG: Thank you, Michel. Chris? I'll put myself in the queue as well.

CHRIS DILLON: Thank you, Edmon. This is something I've often mentioned during our calls, so just repeating, it's highly undesirable that there would be differences among registries, so the poor users having to remember "Which registry am I using at the moment? It actually fits my typing." A thought that I'd never actually had – and it's quite a frightening thought, I don't know that any of you want to comment on this – is that there could be differences between, say, the second and the third level.

I was presuming that the likelihood is that there will be differences between the first level and the second. But it might be that the first level ends up being very conservative, and then the second and other levels perhaps – I don't know whether I should use the word "liberal," but they have things that you can do which you can't do at the top level. Those might be things like, "Oh, you can use a hyphen," or you can use Arabic numbers, or even you can use the Roman alphabet or other systematical, logical things.

This was my sort of earlier thinking about this. And the more I thought about it, the more I thought that actually this may be quite naïve, and if you actually look at the differences between the second and the top level, it may be very difficult to describe those even if you have one set of differences between top and second and no differences between registries and no differences between gTLDs and ccTLDs or any of these other possibilities.

The more I think about it, the more you look at the differences and you think it may not be possible to do this logically. And if it isn't, you're then saying "Okay, how do we do this in a way which is as user friendly as we can possibly make it?" I ended up thinking perhaps you could have a situation where you start off with your top level rules in one file, and then you add a file which then adds the second, third level all together. I don't know, I just wanted to throw that out there. Perhaps it's a total non-starter.

EDMON CHUNG: Thank you, Chris. I will add to it and then move on. Mats, stop me if you think I'm going too much into details, but responding to what you said and also building on what Chris mentioned, my view is that there are going to be cases where legitimately they should be handled differently. Japanese is a very good example that you put out. In a ccTLD context – because say for example .jp – then you have context that you kind of – if you see an IDN and you see it looks like Japanese .jp, you would have a reasonable expectation as a user that that is Japanese.

But even if you see two Kanji characters, for example two Han characters or three, or a string of Han characters .jp, you would imagine given the context that that would be Japanese. In the gTLD situation, you may lose that context. A situation for, let's say, .asia or even .com, where you have a string of characters that are all Han characters .asia, then you don't know whether you are dealing with a Chinese string or a Japanese string. At the registration situation when a registrant registers, they may pick Japanese or they may pick Chinese. The resulting variance and the resulting handling in the user experience for that particular string would become different.

In my view, when you look at a gTLD situation, if you care about the user experience on a gTLD kind of scope, and also given that

the user doesn't have a contextual indicator to go by, you may have to have different language policies for a gTLD versus a ccTLD. Also, if you look at a gTLD situation, if you have a TLD that is in Japanese, let's say in Katakana, then the second level, at least that gives you a little bit more of an indicator of what that whole domain may be.

So that registry may decide that they want to follow the .jp rules as a ccTLD. But a more generic term, either in ASCII or other form, you may want to just treat it like a gTLD mode. And also when the situation comes where it's a whole string of Kanji, then you may want to include variants to avoid – in terms of the philosophy for going about this idea, and treat it that way. That's how I think it may be desirable to actually have different types of implementation.

I don't know whether you will be addressing it, but Arabic may have that kind of situation as well. Mats, I'm guessing you will probably want to respond, but I'll go to [inaudible] and then Sarmad.

UNIDENTIFIED MALE: Hello, my name is [inaudible] I'm from SaudiNIC from Saudi Arabia. Actually, this point is a very important point from the user perspective. The problem is that the end user or the registrant doesn't know variants well. It's something new to him.

So he or she doesn't know the differences between for example Arabic Kaph and Urdu Kaph or Arabic Ayin and [inaudible]. For him, this is something new. It was new for me when I entered this area maybe 12 years ago, and I'm operating a registry.

I believe we should have something baseline for all registries, and if a registry wanted to something in a different way, they should say or state what is the difference, so we have some baseline for all registries. For example, if you want to do more secure or having less variants, we can add to it, so that it would be simple for the user to know. The problem is that – I know many of my friends tried to register a domain name under [inaudible] and under other Arabic TLDs, new gTLDs, and it wasn't the same experience.

For us, you can enable, for example, six variants; for others, only four. For the rest, some of them, there are no variants. This is a problem actually. Some of the registries are advanced. For example, they considered the variants that are needed for you in order to reach it all over the world. Others, no, just only Arabic and that's it. If you go try the domain name in countries like Pakistan, Afghanistan, or Iran, you will not be able to reach the domain name, and this is a problem. So that's why again – back to this point – there should be – first of all, the user is naïve.

He doesn't know the complexity of variants and what he should enable or what he shouldn't enable. We need to have a baseline between all registries, and then each registry state clearly in the variants section, "We follow the standard, plus or minus these things," to keep it simple to the user. The user doesn't know how to open a language table or to see the difference between language tables or run a script to show him the difference. No, in simple words, for the registry. I hope this will be somehow addressed in a way to make the life of the end user easy. Thank you.

EDMON CHUNG: Thank you, very good comments. Just to note in terms of the scope, this is not the working group that would create those variants. We're talking about certain policies and rules that are kind of a level above, but that's very important. Sarmad, you wanted to…

SARMAD HUSSAIN: Yes, I just wanted to read a comment posted by Dennis Tan. He says, "Regarding consistency of IDN tables, while it is desirable that end users have consistent experience, this comes down to business decisions of registries and registrars relative to how they want to serve their customers in its markets. Competition

should take care of this problem that is consistency. [It's good to] follow the best practice. Thank you."

EDMON CHUNG: Thank you. Thank you Dennis. Mats, did you want to respond to all that?

MATS DUFBERG: I think that Dennis said it quite well, at least that we don't have the power to enforce anything, even if we would like to do it. So I don't think we can avoid inconsistencies between different registries. But on the other hand, we have the chance, together with other parts of ICANN, to point to some kind of standard, and I think that many registries will use standard tables if they're available, because they don't want to invent their own tables if they don't have to.

EDMON CHUNG: That's probably a very correct observation. With that, any further comments? Please.

ABDESLAM NASRI: I just want to underline the word "consistency." I have an idea to put behind this word, kind of process, kind of how to do this, especially when we talk about all 29 scripts for instance. Maybe I

think within a language or a script like Arabic script, this may be the community work and taskforce work. That's when we go beyond, or even for a long run when Unicode is evolving, how to keep this consistency? Just give one, two, three [inaudible] how is the consistency, measure of assessment or calculation process? Thank you.

EDMON CHUNG: Thank you. I kind of am jumping ahead, but when we actually deliberate on this particular issue, perhaps it may be useful to think about whether some guidelines – if we know that there will be some inconsistencies, and the so-called inconsistency may be good or bad depending on how you see the entire lay of the land. But how do we obtain – if there are more standard ways in relation to the LGR for people to obtain those language tables automatically, that might be something that would be useful. But I don't know whether it's in scope or out of scope for this group to talk about.

To summarize what I was just saying, we may be able to say something, or we should consider whether we should say something like registries should keep their LGR here or somewhere on their website or something, or some standardized area which would allow other people to get that easily. I see a hand up. Francisco.

FRANCISCO ARIAS: Thank you, Edmon. In regards to what you said, will the requirement to publish the tables in the IANA repository serve that purpose?

EDMON CHUNG: That's a good question. Currently, the guidelines do not dictate that you have to – Sarmad, you want to...

FRANCISCO ARIAS: I think the current version of the guidelines recommends that you publish the tables, and for the gTLDs, which is what Sarmad was trying to say, the contract requires them to publish.

EDMON CHUNG: Okay, thank you for the clarification. You can also refer to others. You don't have to have your own. But that would still work if we require that. That's a good point. Maybe that's already the right area to put it. But then we're going to the issue of whether that whole database is going to be updated with LGR, which goes back to what Mats mentioned earlier. Anyway, apologies, this is going a little bit more detailed into when we actually talk about the issues. Please.

UNIDENTIFIED MALE: If you'll allow me to clarify my previous point, I think for example for the Arabic, we have the Arabic task force working for the LGR for the second level. What I recommend – or at least this document should address that, say, a recommendation for gTLDs or TLDs in general, if you want to change something different for the Arabic script at least, different than what the task force has finalized, please state it in a simple word for the end user. Don't expect them to run it or check the differences. This is just a recommendation, if you agree on it. That's it.

EDMON CHUNG: Again, I think specifically on this topic, that might be a little bit more on the implementation side. The reference LGR, and you say if there's some variants – sorry for using that word – a different set of variants, a different set of goal points, then you should make it simple and justify it. I'm actually curious whether this is inside the scope of this working group to consider, or is this part of the reference LGR discussion? I'm thinking through this. Mats, do you have any thoughts on actually providing some recommendation specifically targeted in a way to the reference LGRs, and any alternative LGRs from that?

MATS DUFBERG: Since we don't have any resources to create any reference tables within the group and there will not be reference tables covering

the entire need, I don't think that we can point at reference tables and say that you recommend it or you should use such tables. The only thing that we can talk about is the need of having consistency between TLDs.

EDMON CHUNG: Thank you, Mats. I can't imagine us saying something like if you propose a table that is not the reference table, here are some things that you should consider. If you publish something other than the reference tables, you should at least state that you understood the reference tables and the rationale for moving beyond that. The question is whether it is in scope or out of scope for this working group to include in the IDN implementation guidelines, I think. I don't think we need to –

MATS DUFBERG: I think that we could recommend such a writing, if there are reference tables or when there are reference tables. But there will not be reference tables for matching all languages and scripts that are used by the gTLDs.

EDMON CHUNG: Sarmad?

SARMAD HUSSAIN: Yes, at this time, the reference tables being considered are for 29 languages only, so obviously if there is a point to do those reference tables or do reference tables, it will be limited to those 29 languages. If there's work being done beyond those 29 languages or at script level, then I think that's the case what Mats is also suggesting, that then the working group also needs to resolve those cases and make some suggestions on how those cases may actually also be addressed in this context.

EDMON CHUNG: Thank you, Sarmad. Michel.

MICHEL SUIGNARD: Yes, to add a bit to the Japanese case, because that was an interesting case, in fact the current reference table for example is purely for Japanese, so it has no variant because you're not thinking of adding variants for Japan. But you could in fact perfectly imagine another reference table that would be basically done with the purpose to be used within the context of Han, if you want, not knowing that it's Kanji or not. That then would mean that instead of just being the same table, instead of just containing the pure Japanese Kanji, it would also have the variant they will declare on it.

The same day you will do the coordination in the root when you integrate the Japanese LGR with those of the Chinese LGR. You could do the same thing on the reference table, even for second level, it's just a different piece of work you do. You just basically have to do the integration if you want at the second level with a Chinese table, and then you get the variant set for Japan as well. You could do that, but I don't see an infinity of versions that you can basically make your version that is not considering coordination – or integration if you want – on the version that is basically isolated.

So you have basically two versions: one that is standing alone, which is used basically by – like your first example, when you know you have Japanese context. That one doesn't have deal with variants. And the other one that may be used in a context where you don't know if it is Kanji or [inaudible] or whatever. So in that one, yes, you may have to address variants, but it's basically a binary thing. It's not a multiplicity of versions, just like two.

It would be the same for Arabic. I think you could have a situation where you do basically an integrated version or you do basically a standalone version.

EDMON CHUNG:	Thank you, Michel. I think that's a pretty good point. Whether you have context or no context, that may be two ways to do it, and it might not be a multiplicity of things. That's I think a good point.

With that, I think I'll draw a line on this item and go to number four. Number four is about IDN variants. In fact, from the very beginning, I think even for the first IDN implementation guidelines, there was talk about IDN variants, but definitely not as well formed as we are now. In the last few years, a lot of work has been done on the IDN variants to formalize it and to make it more standardized in many ways.

This is a topic area that the working group will spend some time not only to talk about the nomenclature, the kind of terminology again, but the types of states of variants. Whether it's allocated, activated, blocked, reserved, and also what types of relevant policies owned by the same registrant: activation policies, automatically activated, activated later or no activation, and some other types of recommendations specifically on variants, given the depth of experience we now have in the subject. That's number four. Any questions? Sarmad.

SARMAD HUSSAIN:	Just to add to that, the column at the right actually states the current IDNGWG, the working group position. And for this

particular point, the current thinking of the working group is that the guidance is going to be provided at a reasonably higher level compared to the other three points which have been discussed so far.

EDMON CHUNG: Thank you, Sarmad, for the clarification. Mats, do you want to add anything? Please go ahead.

MATS DUFBERG: One question is if variants is a must, if you have a Chinese table, a Han table, is it a must to have variants? There are other contexts where variants are common, but is it a must?

EDMON CHUNG: Mats, apologies, I didn't quite get you. You mean this is a question the working group should ask, or you're asking the question if you're dealing with Chinese registration, is it a must to have a variant?

MATS DUFBERG: I'm putting out the question. What is the view of the working group and the participants in this meeting? Is it okay to have a Chinese table without any variants?

EDMON CHUNG: I see shaking heads, but I'm trying to figure out – for the IDN implementation guidelines, we probably should not be prescribing something along those lines. But rather, within our scope would be we can point to the reference LGRs, we can point to other community efforts and probably recommend that whoever implements IDNs reference them, or at least understand and point to them. But I'm not sure that this is where we should talk about that subject.

Also, this is specific on a particular language. I saw some shaking heads. I wonder if anyone wants to jump in.

MICHEL SUIGNARD: The only thing I know is that every IDN tables I've seen for Chinese that are recent, all are variants, so at least from a usage point of view you would consider that probably every Chinese IDN table would have variants, because of the traditional/simplified issue. It's not impossible to create one without, but it's not common practice. As far as you want to go for a recommendation or not, that's open to you to obviously. But practice seems at least to be saying that Chinese use variants because of traditional and simplified. Ignoring that would be making a major mistake, in my opinion, of anyone trying to use such a table.

While obviously for Japan, Japan is a completely different issue. Japan by default, you probably don't want a variant, because variant in fact in Japan is a complicated subject, because there are no simple rules to say that among the thousand permutations possibly you could have if you push the Chinese model on Japan, you don't know which one is the preferred version or what version a given customer would take among a thousand permutations. Where in fact for obviously China, Chinese context is much easier to determine for the user if you want traditional view of the root label or you want a simplified view of the root label. So it's a complicated subject. I think we should just practice [on that opinion] in that situation.

EDMON CHUNG: Thank you, Michel. Mats, I see your hand up.

MATS DUFBERG: Yes. There are actually, a number of gTLDs with a Chinese table without any variants, and there are gTLDs with Han script tables without any variants. So it's actually out there already.

EDMON CHUNG: That's a good point. Again, I go back to what this group is I guess chartered to do is probably not weigh in on any specific language or specific table, but rather perhaps our

recommendation needs to be at a point where if you consider developing or when you're actually implementing it for your TLD, these are the things that you should look at. And part of it, rather than specify what language should do what, is to go to the community efforts, understand them, or go to the reference tables. That to me still seems to be kind of the level that the IDN implementation guidelines should be at, rather than at specific languages and tables. Does that make sense? But perhaps others, any suggestion... Michel.

MICHEL SUIGNARD: Yes, I would say that the reason why a lot of IDN tables for Chinese don't have variant tables is because they have no clue how to do that. It's a complicated subject, and to do it right takes a lot of work and energy to get it right. If you do it wrong, it's worse than not having anything, so to some degree it's also a question of getting them right. It's still a work in progress, frankly, to get the variant for Chinese done right. Because it depends on repertoire. Even if you're willing to integrate with another part of the CJK universe, it's a very complex subject, so I'm not surprised that people are afraid of touching it, because we're still kind of working on it at this point.

EDMON CHUNG: Thank you. Mats, you have your hand up again.

MATS DUFBERG: Yes, I think that the reason why TLDs or registries have chosen not to have any variants is the same as some of them have chosen not to have any contextual rules, is that it's simpler to implement. It's cheaper to implement, and they don't think that the user community cares or their customers care.

EDMON CHUNG: Thank you, Mats, that may be a correct observation, and I guess that ties a little bit into what Dennis mentioned earlier. His point was to leave it to the market. I support the free market, but I think there needs to be some polices around it to safeguard users as well. I hate to be the one saying that maybe we know better, but I think at least as we look at the idea and the implementation, it's probably better to be more conservative to start with and then becoming more liberal over time.

Mats, I'm going to probably move to the next couple of topics, but what I'm struggling to get a sense is – perhaps you can explain a little bit more clearly what kind of action item this working group can take on, on the particular subject you mentioned, because I don't see us going so detailed into specific languages again and tables. I have Michel's hand up, and then I'm hoping to get a sense from Mats, and then we'll probably draw a line on this topic and move forward.

MICHEL SUIGNARD: Sorry to be taking time here, but a lot of reason why those contextual [inaudible] is in fact [to save] security. In fact, for example when you start to have some punctuation or punctuation-like characters – in the second level especially, because in the root you can't have them, but in the second level you can have some of them. Or sometimes you have digits in a mix of Arabic. So you have those contexts on there, they're [trying] to protect whoever is using those labels. It's not just for the fun of it. Many of those context rules are meant, in fact, to improve the quality of the labels and the experience of anyone using domain names. So ignoring them is in fact at your own risk, in terms of security of the whole IT system.

EDMON CHUNG: Thank you. I was going to go to Mats for the last word, but I'll let – Francisco, you're going to – okay. So I'll go to Mats for last word on the item, and then we'll move to the next item. Did you want to add, Mats?

MATS DUFBERG: Yes. First, a comment on the contextual rules. The way that the registers have sorted that is to remove the code points that require contextual rules. Going to the variants problem, one

possibility is that these guidelines actually state that the variants is a should or a recommendation for certain scripts tables and language tables. Another possibility is to have more general wording about it. I don't know which way we should go, but I think it's important to have this discussion and to raise this issue, because it looks very different in different corners of the world.

MATS DUFBERG: Thank you, Mats. Now I think I understand much better where you're coming from, and I think, yes, that should be part of it. We should look at the previous versions of the IDN guidelines and see what we have there. I'm pretty sure that we deliberated that in the IDN working group in the GNSO specifically on that topic, to not mandate variants, but for certain cases making sure that it's very important to have. We should probably take a look at those, but I see that Sarmad wants to add, or Francisco. You finally decided that you still wanted to say something about this.

FRANCISCO ARIAS: Sarmad convinced me. I just wanted to point out that there is a document from SSAC, SAC60, that contains recommendations regarding the consistency of tables between first and second level and across the TLDs, that this working group may want to consider.

EDMON CHUNG: Thank you. Noted, and that'll definitely be referenced and also studied I guess a little bit in the working group. With that, let me draw a line on number four and we have about ten minutes to go five and six. I'll probably do five and six – they can't be done together, so separately.

Number five, similarity and confusability of labels. What we were thinking about here is, again, from the very beginning, the IDN implementation guidelines, this was an issue. In fact, this was an issue that kind of created the IDN implementation guidelines. Homoglyphs and attacks perhaps utilizing homoglyphs and homograms. The working group intends to look a little bit further into whether we need more comprehensive recommendation of this. Currently, the idea is to have more guidance at a higher level, pointing to some of the work that has already been done, especially from Unicode and other areas. This remains an important part of the IDN implementation guidelines. That's number five. Mats, did you want to add anything?

MATS DUFBERG: I think it's important to really discuss, should there be strong recommendation – a "should" or a "must" – to have mechanisms to block homoglyphs between different scripts?

Say that there's a TLD that supports both Latin and Cyrillic. Is it okay not to have any mechanisms, or is it a must or is it a should? What level should we be on? I think this is important to discuss.

EDMON CHUNG: Thank you, Mats. Sarmad?

SARMAD HUSSAIN: I would also like to add that there are actually community-driven efforts currently being undertaken for LGR, for the root zone, where these communities are considering cross-script variants which could fall into this kind of confusability discussion. In case one wants to make a more detailed recommendation, one possibility is also to refer to that kind of work, which will actually go formally into LGR for the root zone as well.

EDMON CHUNG: Thank you, Sarmad. That is useful information, and as we consider our timeline for producing the end result for this working group, how it matches with the other timelines may be something we want to take a look at. Not that we should let other areas dictate our speed, but it would be useful to get a

sense when those other work would be ready for this group to take a look at.

I guess that's the number five. I'm not seeing any further discussion. Oh, Sarmad, please.

SARMAD HUSSAIN: Just to come back to the point, I think if there is a generic guideline not specifically to a case, something to the effect that any cross-script work which gets into the root zone should be considered for this context. I'm not sure how strong a wording one would need to make, but if there is community work which ratifies that there is homoglyphic confusability and that is verified through the integration process and goes into the root zone, that is sufficient enough I guess motivation to consider it for the second level as well.

EDMON CHUNG: Thank you, Sarmad. Good points there. With that, seeing no other hands, I'll move to the last item. The last item is about registration data. With the introduction of IDNs and also the developments about registration data, which is the contact information and probably including e-mails as well, that information also being internationalized. The working group has started the discussion on this. We have identified it right now.

We're not entirely sure that this should be within the scope of this working group to weigh in, in terms of recommendation yet. But this is an area that would be related and is something that we will consider providing some recommendations on. Please.

UNIDENTIFIED MALE: Have you reached something, or still it's an open issue? Or have you have consensus or have discussed it in detail, or nothing?

EDMON CHUNG: Sarmad, why don't you add further?

SARMAD HUSSAIN: In the working group, I think the working group was still not decided on whether to include this in the topics they want to consider or not. There has been discussion, but there has not been a decision like on the last five points. And if the community thinks – if you think this is something the working group should consider, I think that would be good feedback to have.

EDMON CHUNG: Yes, so still an open item. So if you feel strongly that this working group should take it on and provide some recommendations, please speak up.

EN

UNIDENTIFIED MALE:	I think there is another group for the WHOIS working on something, am I right?
EDMON CHUNG:	Do I get it that we should not consider this at this moment, especially given the work that is being done elsewhere?
UNIDENTIFIED MALE:	Yes.
EDMON CHUNG:	Okay, that's useful. Francisco?
FRANCISCO ARIAS:	Just very quickly on that, I think you're referring to the RDS PDP working group, the Registration Directory Services PDP? I wonder if we'll be able to check with them if they are indeed considering that issue. I'm not sure.
EDMON CHUNG:	Thank you. I guess I'm going to try to summarize a little bit. And Mats, please add or correct me. It seems to me that what you're saying is that we shouldn't overlap the work and also shouldn't preempt any work that is currently ongoing, and I think Francisco's point is well taken. We should probably go and get a

sense of whether what we may consider here would overlap or preempt discussions over at the RDS.

If there isn't, then we have reason to believe that the IDN, part of it, maybe there's already something, because a number of working groups on the WHOIS side have already completed their work, especially in the internationalization part. So maybe there are certain areas, but if those interact so much with the existing current work, then we should stop and let the RDS complete its work first. I think that's – did I kind of summarize that properly? Mats, did you want to add to that?

MATS DUFBERG: Yes. From my point of view, if we should have any opinion about what is shown in WHOIS or equivalent, is maybe if it should be recommended that the IDN label is shown as a U-label in WHOIS or not. Further than that, I don't think that this working group could go.

EDMON CHUNG: Mats, please finish first.

MATS DUFBERG: Yes. Having a meeting on how internationalization should be done in WHOIS, in other fields, I don't think that that belongs to the IDN group. Maybe the showing of the U-label or not.

EDMON CHUNG: Thank you, Mats. I think there was a different expert working group that did look into the matter. I'm looking to Francisco to see if we can add to that.

FRANCISCO ARIAS: Yes, indeed there is a working group. Not a policy development working group, it's an expert working group, as you call it, derived from the WHOIS review team. It's the Internationalized Registration Data working group. They finalized that work a few months ago, and I believe they have recommendations in this regard. However, they're not policy, so there are still some things that need to be decided in order to find the appropriate way to implement those.

EDMON CHUNG: Thank you, Francisco. It may be useful as this group considers our work on this particular topic to not only reference the work from the expert working group, but also get a sense from your team to see if further work is being done and whether we could actually reference the expert working group product or it's

better not to touch on it yet, back to the point made earlier. If there's still policy work that is following up on that, then maybe we shouldn't touch it yet. If not, or if that's stable enough, then maybe we should include it, or parts of it, in our recommendation. Please.

UNIDENTIFIED MALE: I think the issue raised by the colleague before, it's very important to have the representation of the domain name as U-label, not I-label, and I add to that also to list the variants that are enabled. It's a very important thing for maybe hosting companies, so that they look up the domain name and then they add these things. It's useful information to automate some kind of hosting.

EDMON CHUNG: Thank you. I encourage you to take a look at the expert working group product, as it does touch a little bit on that area as well. Sarmad, I lost connection, sorry.

SARMAD HUSSAIN: Okay. There's a comment by Dennis Tan regarding registration data. "It's not in scope for this working group. We shouldn't segregate registration data processing for a subgroup of domain names."

EDMON CHUNG: Thank you, thank you Dennis for that. There is a last comment from Chris. Please.

CHRIS DILLON: I came into the room actually thinking the same and thinking this isn't relevant, but I can't remember the detail of that report. I know that the report was saying it's important to display U-labels. What I cannot remember reading is whether they had the idea that there could be quite a few of these U-labels in the case of variants. I have no recollection of reading that, and if I didn't read it, if it isn't in the report, then extra work may be needed there if we really want to see the variant labels listed, because that may just be news to them. Maybe I'm wrong, but I don't remember it.

EDMON CHUNG: Thank you, Chris. I guess we'll take that as homework in our next meeting, especially when we go back to this particular topic, we'll take a look at that report and take that as input as well. With that, we're at the end of this session. Please. Keep it short.

UNIDENTIFIED MALE: I have one question. I'm not sure if it is part of this group, but what about the XML standard format for variants and IDNs? Shall we have a recommendation regarding this standard format, or it's something technical and another group should be working on it?

EDMON CHUNG: Thank you for the question, that's exactly number two. You probably missed that. When you came in we were probably after number two.

UNIDENTIFIED MALE: No, I'm sorry, the EPP extensions.

EDMON CHUNG: EPP extensions. Good question. The EPP extensions are – I think I don't have a quick answer and we're at the end, but that is a very good question. We'll definitely take that into the working group and consider that part as well, the EPP extensions, also the variants. Thank you.

With that, I guess this brings us to a close. Thank you for all the inputs. I'll see if Mats wants to have a last sentence, just ten seconds before we close.

MATS DUFBERG: No, it's fine.

EDMON CHUNG: Okay, thank you, and we thank you for the feedback. We'll continue to work on this and as we go along – again, this is not a formal report yet. Our aim is to create a formal record with some formal feedback over public comment period as such. Okay, thank you everyone. Bye.

UNIDENTIFIED MALE: Thank you.

[END OF TRANSCRIPTION]