
MARRAKECH – IDN Root Zone LGR Workshop
Wednesday, March 09, 2016 – 13:00 to 14:15 WET
ICANN55 | Marrakech, Morocco

UNIDENTIFIED MALE: This is the ICANN 55 IDN Root Zone LGR Workshop on March 9th, 2016 at 13:00 WET in the Amatyste Room.

UNIDENTIFIED MALE: I think there is a slight confusion because there was a change in time in this workshop, so I will just start in a couple of minutes.

Okay, so thank you for waiting and we're very sorry for starting 10 minutes late. There's apparently been some confusion with the start time of this meeting. So let's get the session started. Go to the next slide, please.

Welcome to the IDN Root Zone LGR Workshop. Next slide, please. We will be talking about the LGR Tool Set. Marc Blanchet will take us through what is currently being developed and what the status of this work is. We have some feedback from Integration Panel on the work being done by the LGR by different communities and then we have updates from different panels. We have presentations from Thai Generation Panel, Korean Generation Panel, Japanese Generation Panel, and Chinese Generation Panel, after which we'll have a Q&A session.

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.

Again, just because of the confusion in time we apologize for the inconvenience. And let's move to the first presentation on the LGR Tool Set. So over to Marc. Blanchet.

MARC BLANCHET:

Good afternoon. Next slide, please. Okay.

So agenda recap of the project – new features we've implemented since last time in conclusion. Next slide.

So project is web-based LGR editing tool. We defined three phases. There is actually another one coming, a fourth one. So first is enable LGR creation, so you can create an LGR. Second phase of implementation is to actually validate the label or generate [foreign] labels. Third is some LGR management functions such as conversion of language tables. I'm sorry...

UNIDENTIFIED MALE:

Can you speak just...we just want to make sure that you're on a call. Could you please?

UNIDENTIFIED MALE:

Yes, I'm on the call.

UNIDENTIFIED MALE:

Great. Sorry.

MARC BLANCHET:

No problem. So the phase three is about management of the LGR so you can convert the old language tables into the new XML format comparing two LGRs. For example, if you're working on multiple versions of LGRs then you can compare two versions. You could also do operations such as union or intersection. So for example, in those two cases you will get a new LGR which is the result of the operation. And this is the name of the implementation team below. Next slide.

So again, four phases – three listed there. There's a new one that was briefly described by [Sarma] this morning. So we essentially did the whole set of phases and the tool has been available on our own server since then, and now is being available through ICANN, hosted by ICANN. Next slide.

So a few details updates since phase two, mostly from last time we presented this. So now you can edit the WLE Rules. You can edit/define/delete classes, rules, and actions. You can define tags in code point view to be referenced. You can compare LGRs. You can do a union intersection. You can expand ranges to single code points so it expands all the ranges that you can see in your LGR. You can save the LGR summary that usually you would have seen in a model text version so you can now save it for further reference. You can add variance when you import code

points from a file in manual mode which this was not supported before.

The update of the name space – there was a change of name space was done during the [specification] of the in the ITF Working Group to a more standard name space, therefore right now what the tool does is when you have a LGR with the old name space it will just convert to the new name space automatically for you so you don't have to care about that.

ICANN contracted the security firm to do a security audit of the code or the software, I would say more generally, and roughly no issues except one on restricting the capability to inside the XML and we added this enhancement to the code. ICANN is actually using Docker – I don't know if you know about Docker. It's a way to do containers, it's more easy to deploy software in data center, and we provided Docker-ready configuration, I would say, for the tool.

This slide is no more appropriate because it's now available on ICANN as lgrtool.icann.org. IDN team is now also, “Don't go to our site anymore.”

There was a question last time on Unicode version, so the editor itself is Unicode version independent but it actually applies IDNA 2008 rules so it's kind of independent and [not] but any LGR processing will actually depend on the specific Unicode version

and we're aligned with the 6.3 right now which is the current IANA IDNA tables. I'm sorry, next slide. That's what I've been saying. Next slide.

New features. Next slide.

So it's just a few screen shots and it's probably better for you to look at the actual pdf of the presentation because it's kind of detailed inside. So this is an example of where you can list the different WLE rules so if you have a previous LGR that you already have rules, then it will show you the rules and if you see – I don't have a laser right now – but if you see on the right there is a button that says “edit” so you can edit the rules each of them and action and classes. Next slide.

This is an example of a class that you add. It will essentially open a text window and then you can change the content of the text or the class if alter the rule or the action. When you do “save” what happens is we do the tool actually do a basic validation, but it doesn't try to relate if you're referencing something, a tag for example, which is not yet defined. So it's kind of again the idea of a free editor you could do whatever you want and it's only at the end when you are ready then you can validate the whole thing and then the validation will actually make sure that your reference – you don't have any stale reference of

something. So it's really on a two-step process. More free-editing and then validation if something is not correct. Next slide.

This is just to show you the tag addition when you're modifying a code point and tags are actually with auto completion so if you already have defined a tag then it will offer you the choice, so just a simple, easy way for people. Next slide.

This is an example of the [dif]. So what you need to have is two LGRs and in the work space as you may have for those who may or may not have used the tool, you can have multiple as many as you want LGR in your work space and then you can edit any of them. So the tool actually, when you do the [dif] you click "compare" and then you tell "I want the [dif]," then you choose two LGRs in your work space. Therefore, you have to import the two LGR you want to compare before doing that operation. And if it's a [dif] then it will show the result so here is an example of their LGR two versions and shows the difference between the two. Next slide.

In the context of union and intersection, actually the result will create a new LGR, so it will create the LGR, you won't see anything happening, instead you will just be presented with the result of the action. So the new LGR will be put into the user work space and we changed the title so there's no issue on

identifying it and it's automatically created for you so you just got the union of both or the intersection of both. Next slide.

You can expand the ranges so that will convert the range to a list of code points. You can convert a specific range or all ranges in your LGR, and obviously the generated code points will [inaudible] all the properties of the ranges, tags, and stuff like that. Next slide.

The button on the top is to expand all the ranges and the button on the right is in a row of a specific range so you can choose one or the other. Next slide.

We did many fixes, actually, many bug fixes. This is one about a case when and not when definition with an anchor element. I don't want to go through these but we've been improving the software. Next slide.

That's what I was saying previously about the security audit of the code, the software, so that's the quote of the auditor that says, "[Appear] well-designed thing, typical web application of text such as cross-site-scripting and URL manipulation." There was one issue about testing of existence of a file in external entities, this XML stuff that has been disabled since then. Next slide.

So phase three is release. We're currently working on an addition in the tool to enable you to input file of labels and then get the tool process, the labels so you can see if what the effect of the LGR on the list of labels. So collisions or anything. So that's the current work and we should release this probably within this month. It's going pretty well in terms of development at the moment. Thank you.

UNIDENTIFIED MALE:

Thank you. So we'll move onwards. The next presentation is by Michel on LGR considerations. He's here in the capacity of a member of Integration Panel and this is just feedback from Integration Panel as a whole on the current work which has been going on and some of the observations they have made. So over to Michel.

MICHEL SUIGNARD:

Okay, good afternoon. So we're going to use experience we gain in reviewing LGRs from Arabic GP and as well as from the [inaudible] that are coming now on the [inaudible]. So we are reviewing all those LGRs, we are accumulating some hints, if you want, of feedback on all those LGRs that were submitted to us. So this is kind of going over some of the points that are important to us when we're reviewing those LGRs. And so we go in some different consideration, first about the repertoire then

about the rules about the repertoire and as well as some, I would say, advice or hints on how you can get LGR accepted or basically considered by us on... Next slide, please.

So first the repertoire, that's probably the more visible one because that's define your code point set for your LGR and that's probably where you're going to spend the first part of the development is to establish your set of code points that you need for a given script. Obviously, one of the first point is you better make sure that it does fit within the MSR – the Maximal Starting Repertoire which is currently in version 2 so it's always useful that you really consider carefully at the MSR. On MSR we see available on the ICANN website and it's basically a set of code points that is a subset of Unicode 6.3 at this point. Those define some sort of a sandbox if you want to define the envelope for defining your own repertoire.

The next point is to make good references when you're going to justify code points for your given LGR, we expect to see references establishing the modern use [because running] to what you will need for domain names applications. So we expect to see some at the station of modern use. So, for example, typically the MSR already took care of that but we try to be not totally too restrictive, so we let some opening if you want to each generation panel to make their own decision so we didn't want to be totally – when the IP did the MSR we wanted to leave

some leeway on the different LGR so they can make their own decision on the repertoires.

Obviously, an important point each application is done with two documents. There is a proposal which is typically a text document with background on all sort of description of what the LGR is about, but it does use some preference or technical details on the repertoire and we want those to be exactly the same that are provided on the XML. The XML will have basically a formal description of the LGR and we expect the same terms or the same terminology to be used in both documents, especially when you do subsets through a tag or references as you know different elements you can create within the XML to subset your own repertoire on those also appear in the same way on the XML, so consistency between the proposal and the XML file is very important to us and we have been doing that quite often.

Again, coming to the repertoire, which is something that people not necessarily realize is that a repertoire is not just [done] of code points, it can also be done of sequence of code points. Sometimes it could be useful. Especially you could see...let's say you define at some point...I'm thinking of the [Latin] Generation Panel... Probably they may find that they need characters beyond the pre-composed characters and if you need characters by accents that are not existing as pre-composed characters, you can also do both ways. You can also have in the repertoire a

sequence of those combining sequences which is based of the base character plus the [inaudible]. Or you could say, “I’m going to have the base character, I’m going to allow that combining [my icons] then I will have established rules about when I can use combining mark with base characters.

So there’s multiple ways of doing that on one of the easiest way typically to make your life easier is just to put in the repertoire sequences instead of just single code point which is different in IDN table. So far we didn’t receive that but LGR format does allow you to create sequences.

And then they need to be classified typically is very common, especially if you have some complex rules in your LGR. Most of the characters are going to be themselves classified in subsets like, for example, you could have an example of [inaudible] we have seen that. You would have consonants, you would have vowels. Is there [independent] vowel or you would have a vowel that have attached to a consonant. You have all different classification of characters and those do show in the repertoire as basically subset on the used mechanism in the XML to do that which is [inaudible] by tag that you attach to a given code point.

Like I said, sequence allows specification combination of code points and that would be very useful for combining marks. We

will see that. We would expect [that] to see that especially on Latin and possibly Cyrillic and few others.

We don't expect them to be used too much in South Asian scripts or Indian script because typically the use of combining marks is it's quite exhaustive and is very difficult to just restrict the usage through given sequences because it would be just too many. You would have infinite number of sequences and that's not exactly the right way to do that.

A sequence cannot have a tag. That's a bit of limitation but in fact there is ways to go around that as an IP. We've been working on those limitation. Next slide.

So that's just an example of a repertoire that we were using Arabic LGR for that, so here you see the names, the code points and then there's references on the right side and you can have multiple references and typically references give... One example that you say once the character was encoded Unicode [in] conversion when this was encoded. That's a minor point for typically LGR. LGR are more concerned about the making evidence of use and those evidence of use are done by typically documents or website or web references, and so then get added as reference and then those reference have a numeric value and that numeric value does appear on the [inaudible] for the references.

So that's how you create basically evidence of use on the IP then we will review those evidence of use on basically negotiate with the Generation Panel make sure that they are all correct. And again, it's very important that you document the same way those references between the XML document on the proposal itself. Next slide.

So when you define labels on the LGR, we are not only trying to make spelling decision, we are trying to make [well form]. And the distinction sometime can be a bit subtle, but the idea is that we are not really after representing a dictionary because dictionary do evolve anyway, they get new terms added on what is correct from a spelling point of view and they may not be the same in two or three years from now. So we are not really going after spelling rules, we are going after [well form]. In some writing system, it's not obvious. Think about Latin again. You could have [inaudible] because that is totally meaningless in any languages frankly, but still if someone want to register [inaudible], they could do it. There is nothing wrong. It's totally wrong from a spelling point of view, but it's [well formed]. So that's kind of the idea.

In South Asian script, it's a bit more subtle. It's a bit of a fuzzy line what is [well from] what is correct from a spelling point of view and as a guiding factor there it could be also complexity. You don't want the rules to be too complex and so you try to

make rules that make sense and so you basically correct yourself again [inaudible] label that are completely badly [formed].

How we define that, you can also use BNF in your proposal to define what is the structure of your writing system, or you can use any form of regular expression if you want depending of what is the expertise of the [Generation] Panel. We would use anything that you provide us to define what is [well formed] rules of your writing system and then we will negotiate if you want with the Generation Panel to come to a common understanding again between complexity and [coverage].

It's not the intent to fully represent a [well formed] syntax for a given writing system because some of them are so complex that it may not be possible to represent them without going too far in the complexity. So we are just trying to capture at least things that should never happen in a language [that say no]. You have a consonant followed by a sign that is never going to happen. It's not supposed to happen that that sign should follow is a consonant or a vowel or a vowel has to be preceded by some consonant. There are all those things that are inherent to each languages or script and we want to capture the one that makes sense.

It's too bad when I was doing that presentation I did not yet have the [Korean], Lao example. I would have used them otherwise, but we learned quite a bit in fact in doing those two how to work with Generation Panels to come up with basically a mutual understanding on how to get the LGR that was acceptable.

Like I said, the good news for the future LGRs and GPs that create them that they will be able to use what was done as very recently for [Korean], Lao because they could use them as templates for their own work and they can look for inspiration on what was done for those two writing system. So we expect in fact to work to become a bit easier as we goes because we can use... First, the IP itself can use past experience and also the new GPs can use experience from previous LGRs to do their own work. Okay. Next slide.

Yes, obviously the rules, that's an important part of LGRs, which is kind of new compared to what you saw before with the IDN tables. If you remember] in those IDN tables that this kind of free format can constrain you or putting on but they are not really enforced by any mechanism. Now in LGRs we have in fact mechanism to enforce rules. We see those rules are there to enforce [well formed] like I said before and typically we prefer context rules compared to [all label] rules. I'm not trying to go in detail how is XML and the full LGR works, but basically you can

[insert] in the repertoire itself in the repertoire declaration you can typically say that the character can... What is can be before the character, what is after the character, and that's done by the rules that we [inaudible] none... So you basically say, "This character can only appear if it is preceded by these [classic] characters or followed by this class of characters. And that's a very powerful mechanism to restrict labels and make sure that they are [well formed]. So you can also do that on code point base and again, by this contextual rules or you can also do it on the [whole] label, and that's done by the action on the different parts of the LGR.

I give a small example. For example, you say the character must follow a consonant, and so in the repertoire the creation of the XML file you'll have the [inaudible]. X would be the code point of the character.

Okay. And then you say follow consonants and follow consonants would be basically a contextual rule that would be on the [rule] section of the XML file. Rules can define sequence of code points and be used in other rules. That's another thing. I'm keeping it simple so we can understand the rules so you'll have less errors on creating them.

Obviously, it's very important that you validate them with [test] labels, so when LGR is created it's very important for us that we

get a large [copies] of data so we can validate and also we want to get invalid labels so we can validate the fact that those should fail – the LGR processing and so that they are valid we make sure they do succeed when we process them. Next slide please.

This I can't read even from here, but those will give you an example of rules you can go on the slides and they do explain some of the contextual rules and [they are] just given here as examples how they look. In that slide we use [reg ex] on the middle column. Again, I can't read from here. But it's basically what is an XML where it's expressed in a bit friendlier term. And then there's a comment on the right that do explain in plain English what the rule does. Next slide.

So again, test labels. That's very important that we get label to validate LGR. That's the only way we can verify that the rules on the repertoire are correct and it does typically out of those we have a negotiation with the various [NGPs] and we sometimes have to change the rules to make sure that, in fact, labels that are supposed to fail do fail and labels that are supposed to succeed do succeed. And we always get surprises like sometimes we get a large [copies] of data that would fail the LGR and that's we see has to be fixed. Also we need to know those really a bad spelling. Okay, so next slide.

Variance. Okay, I need to go fast here. Basically, variance it's something that's been covered otherwise but obviously the LGR do cover variance, variance tables for [in script] or other script and so you would have to look at those and look obviously for is it within your own repertoire from repertoires that have some relationship. We do expect LGRs to do that. Next slide.

This is an example of a slide that is in fact, part of the Arabic LGR and you can look at that one. Next slide.

The example for the Arabic [element] LGR, you can look at what was done for the Arabic LGR so you have at least an idea from a true LGR that was put use on being now is available. Next slide.

And again, [use our own] expertise we know LGR we even know XML, we know all to do those things and don't hesitate to ask our opinion and work with us when you're creating your own LGR. Next slide.

Okay. Thank you. Sorry....

UNIDENTIFIED MALE: We'll move on to community updates. We have first update from Thai Script Generation Panels. So over to [Piti Nan].

[PITI NAN]: Thank you. Good afternoon, everyone. So next slide please.

Okay, for the [Thai GP] we found October last year actually, so this is just a roughly of the language is about 70 million speakers and [all this is just like we] moving forward a lot of trying to localization so on this we also working on the e-mail address, so this is just a roughly. Next slide please.

Okay, so we already defined the code points and in entire script there are a number of speaking but in the writing is actually only one. Next slide please.

Okay, this is a component of our GP so is comprised of the DNS, the policy maker, the linguistic, and also the registrar. Okay. Next slides.

This is the original plan. We actually unfortunately we cannot meet the plan. We can do it halfway on the red highlight. So right now we already defined the label rules, but still in the middle of trying to hold a public consultations. Now the plan has changed and probably by the next meeting we will be able to submit. Okay. Next slide please.

Okay, so this is just to summarize that from the [inaudible] log.com. We [inaudible] one is only the [Thai], the rest are speaking but is actually writing the same. Okay. Next slide please.

We finalized the code point repertoire. We found that two of the character is no longer in use in modern writing system, so that will be in the proposal. Okay. Next slide please. Another one.

So for Thai script we [similarity] because the rules from [Sanskrit] so it's [similarity] to Laos, Khmer, Burmese, and some script in India. What we are working as a group is now Laos and Khmer and Thai, so in the next will be the variance definition. Next please.

This is too small, but from the PowerPoint you can download. Within the Thai script itself there is some of the script that could define the variance but after yesterday we have consultation among us, Thai Khmer and Laos, and find out that it's not that confusing. So we are agreed that there is no variance to be defined.

This next one is the Laos and Thai. If we are Thai and Laos, we looking at the characters we definitely know which one is which, but there was a question that if there is a font difference it can cause the similarities but that's not the case. Okay. The next one Khmer script. And then the next one is Burma which is not active yet. Can we go to the next one, Burma?

This one, unfortunately, you cannot see because the font problem. Okay, the next one.

These are the rules. We have some specific rules of the writing system that need to consider the sequence of writing. Another one please.

So the sequence of code point we already done. Certain code point already done. Now is the sequence that we are still have to consult with the public consultation. They are the writings. Our writing system have the consonant, vowel, and the tone mark which come in a certain sequence but is never really defined as a standard. So that's the point that we will have to hold a consultation and hopefully we can find a final within the next meeting. To the next slide please.

So this is the change of the plan. After this supposed to have the consultations within March or April, and by the end of April there should be a vote to submit the proposal. So that will be the summarize of it. Thank you.

UNIDENTIFIED MALE: Thank you. And we'll move on to the next presentation by Professor Kim on Korean GP.

SEUNG KEON KIM: Thank you. I like all the slides but we focus on changes made since the last time. Next please.

Okay, we have at the bottom of the slide we have [48119] characters, [hundred] characters of [0.3]. It is quite stable. There won't be much deletion or addition I expect. Next please.

We analyzed Chinese variant groups. There's about 3000. Out of them, we concluded that only 303 variant groups can affect Korea. So we analyzed only one tenth of the 3000 Chinese variant groups. Next please.

For 303 variant groups, we concluded that 44 of them were acceptable and that the remaining 259 are hard to accept at this point. Next please.

We analyzed the unacceptable 259 variant groups once again, and we subdivided it into two. 97 variant groups have similar meanings and the remaining 162 Chinese variant groups have different meanings, and we are in the process of translating the meanings of 162 variant groups. Next please.

On February 1st Korean LGR has been formally [seated] by ICANN. Actually we started our working since the end of 2014 but now Korea is formally [seated]. Next please.

We expect that the Korean LGR for the [inaudible] will be published hopefully in May of this year. Okay, that's all. Thank you.

UNIDENTIFIED MALE: Thank you. We'll move on to the update by Japanese Generation Panel [Hotta san].

HIRO HOTTA: Thank you, [Sarma]. My name is Hiro Hotta from JGP. I'm the Chair of the JGP. The slide deck itself here contains almost the same as the have in the Dublin meeting. I'll go very fast. This is the meeting events. Next please.

We have eight members from various stakeholders of various expert types. Next please.

And this is the [inaudible] so this means that the Japanese, Chinese and Korean language script has the common characters in their label, so we have to coordinate. Next please.

And this is the framework of [CJK] [inaudible] integration for Hang characters. This has been already agreed by three GPs. Next please.

Activities. So we changed the [inaudible] or something like that so this is very historical. Next please.

Discussion status for Japanese LGR [one], the LGR are [inaudible]. Scopes of the character code. We borrowed that from Japanese Industrial standard. About the variance. Definition of variance for Kanji, the Japanese LGR doesn't define

any variance, but as I said, we share the characters with Chinese and Korean characters or labels. We import the variance from them and think about how to define variance as an imported variance. So we are thinking about that. And for the WLE we may have very small number of rules even if defined but is not fully discussed yet. Next please.

Overview of Japanese LGR. Yes. We have around 6000 Hang characters and Hiragana Katakana, they have 85 and 89. We usually use the Hang, Hiragana and the Katakana in a mixed way in writing a word, so these three scripts will be mixed in a label. And variance, as I said, no variance are defined. Next please.

So the last two slides are newly created, so I will explain it a little bit deeply. So [each] reduction of the number of allocatable labels is very necessary. So as I said, we don't have variance. But we do import the variance from C N or K so that makes us the variant label group bigger and bigger. So for example, logically it may be more than 10,000 allocatable labels if we don't define any rule for that. Because in Japanese case, every combination of characters is allowed in a word, so logically the number of variance will be big and big for the labels. So we analyzed [Kanji] domain names are current registered under [JP] and the biggest size of a set of calculated variant labels which follows the variant definition from Chinese language table, so it's around 20,000. But it is big or not big but I know that the IP Integration Panel

think that it is a big number. An IP has just requested it last week but the IP had just requested JGP to reconsider the reduction of allocatable labels farther. So we are thinking about the further seeking ways to reduce the allocated labels.

And in addition to that, we request ICANN some ways that is complement to LGR which can be used to serve the issues, because JP want to have any combination of a label and it is not foreseeable in advance. So how should be solve that? So if the reduction cannot be done by inside the LGR, we need to ask ICANN whether such a procedure can be made as a panel. Okay. Last slide please.

This is not technical one. Our communication with the Japanese community we are doing communication whether to educate what the LGR is and to seek their opinion on our definition of the LGR. So presentation and discussion with Japan Trademark Association it was in October and no objection against draft of our LGR. And presentation and discussion with various stakeholders in Internet Governance Conference Japan, it's kind of localized here in the event in November. Again, no objection was raised against our draft.

In other related activity, the JGP submitted public comments to guidelines for second level LGR. Thank you.

UNIDENTIFIED MALE: Thank you. And let's move on to the last presentation for this session. It's an update on Chinese Generation Panel. It is going to be presented by Wong Wei who's going to be doing this remotely, so Wong Wei, over to you.

WONG WEI: Thank you. After the Dublin meeting the CGP rechecked the repertoire we made before the Dublin meeting and we have our current repertoire it's consisted of three parts. The first part is the CDNC table. The CDNC Taiwan meeting last year [inaudible] chairs into CGP repertoire. And the CDNC table 2015 had 41 new chairs requested by Hong Kong community and all of this added into the CGP repertoire.

The second part is from the China State Council [inaudible] common use in 2013. It brings 27 new chairs into the CGP repertoire. And to consider future extension of CGP repertoire, we added 145 character from [ICOR] into the CGP repertoire. But we didn't add the characters in [ICOR] extension B into it because some of them are not fully supported in Chinese [IT] system. The next please.

So among the 172 new characters from the [minimized] table and [ICOR] table, there are 107 chairs that happen to also be the JGP characters, so we reviewed these characters and we create

the variance and relations between them and the original CGP repertoire. The next please.

We get the feedback from the IP, so CGP depicted the flaws in the latter version of CGP LGR XML document. We [completed] the description of [inaudible] ID, we add the type of [reflective] and we corrected some spelling errors and added some comments to the action rules.

Speaking of action rules, though the WLE rules are not finally decided, we hope to keep consistent with the action rules of current CDNC which is besides the – we accepted the original applied or simplified or traditional or the other variant labels where it should be [blocked]. The next one please.

So considering we have to admit that we haven't finished the coordination work with [JNK] besides the 107 KGP chairs, I mentioned about [CP] does not seek to borrow or import more characters or variance from the JGP and we acknowledge that KGP provided thousands of [inaudible] chairs and they also in the range of CGP repertoire. Next please.

So the next step, CGP agreed to generate a merged LGR based on [inaudible] proposed by [inaudible] but I just get the suggestion from [Asmus] that there are still some characters are not covered by the current CGP repertoire including 22 characters in the IDN table proposed by .asia. So in next

meeting, CDNC meeting we will discuss if the CGP repertoire should have the [fourth] part which means whether or not it should cover all the Chinese characters in [all] IDN tables proposed by the [ministries] The next please. The last one. I'm sorry.

So the challenges is still the [variance stats] between CGK because obviously there are some hundreds of characters which are variance in Chinese environment but not variance in Japanese or Korean environments. So we suppose we can reach 100% consensus on all of the Chinese characters, but it seems impossible to. It is impossible for us to reach that goal so probably we need that plan B to try to find out alternative solution to [suppose] the unacceptable variant groups [proposed by K]. [Inaudible] has invited JNK to come to Beijing to hold a face-to-face meeting at 20th and 22nd March, so I hope we could be smart enough to find a ideal solution to solve our difference. Thank you.

UNIDENTIFIED MALE:

Thank you. So that was the several presentations we had for this session and now let's open the floor for any comments or questions anybody may have. Are there any questions/comments?

[HARTMUT]: This is [Hartmut] from Pakistan. I have a question for Michel. Michel, how do you see variance how it will be handle or whether it will be handled by ICANN, by the community, or by the registries because it may create confusion between the language communities. How to handle if it is allowed it may create some difficult situation and if not, it will be not accessible for the communities.

MICHEL SUIGNARD: The only thing that we do is that we put block to allocatable. What happened to allocatable variance is really outside of the IP scope. It's [only] become a decision above us on what happened. We're just saying that it's good to limit allocatable variance because there is no good technical solution to in fact entertain allocatable variance in my opinion. So we're trying to restrict it as much as we can because I don't think it's feasible to have too many allocatable variance when you have a label, but what happened to that it's really beyond our scope as Integration Panel so I can't really answer that.

UNIDENTIFIED MALE: So getting a set of variance and their dispositions is the first step in the process and being [work at] the root zone the intention is to do that as conservatively as possible, which means that the number of allocatable variance is as small [a set] as possible.

However, in certain cases, it may still create a large set of variance which are allocatable and for that purpose there has to be a subsequent process which looks at those allocatable variance to decide which of those variance will actually be allocated and what will be the process for that. That is work which is subsequent in the sense that it's a separate project. It's called IDN Variant TLD Implementation Project which is we are currently doing homework on it within ICANN and we are aiming to come back to the community with more details on that project very soon, perhaps in the next three to four months, possibly hopefully by the next ICANN meeting. So that's a discussion which needs to be done within the community after some initial analysis of the facts has been done.

UNIDENTIFIED MALE: If you allow I have another question, whether it is possible for the IDNS to handle accessibilities for disabled persons to make it necessary to have in their languages.

UNIDENTIFIED MALE: I think accessibility is probably going to be another layer on top of the IDNs in applications. So for example, speech interfaces which can spell them out or speak them out, whichever is possible or some other mechanisms. So I think that's just another layer on top of it which is an application layer and that

may not be directly in the scope of the work which IDN program undertakes.

[ABDUL SALAM]: This is [Abdul Salam] from Algeria from the [inaudible] script. I have a question about how to accompany the evolution of the Unicode versions. So you have new characters being introduced it's a concern, for example, Arabic script. So how will ICANN plan to move from 6.3 to version 8 or later, version 9 of Unicode. Thank you.

UNIDENTIFIED MALE: I'll request Marc Blanchet to respond to that.

MARC GLANCHET: If moving up to anything above 6.3 will require some work inside IETF and somewhere else, so I'm not sure what to do about anything up after 6.3.

UNIDENTIFIED MALE: So there are some initial discussions going on on how to follow up on that as Mark said this is a discussion which is not inside ICANN, it's a much wider discussion which includes IETF and Unicode and many of the other community members. So that is something which is already initiated and some more work is

anticipated which will start soon on it. So there is no clear answer to that except that the community really needs to get involved and try to understand the problem and now find a feasible solution moving forward.

[HADEF MWAURA]:

I'm [Hedef Mwaura] from Ethiopia. I am a new Fellow and I am a member of the Ethiopic GP. It's a newly formed group. We have two or three meetings, very few people. So my question to the National GPs is what kind of composition does GPs have? What best compositions in terms of expertise do you suggest to me?

Another question is can you explain to me what the variance and tags issues are in this [inaudible]. Thank you.

UNIDENTIFIED MALE:

That's a slightly longer discussion and I'm happy to take that with you after this session offline. But in any case thank you very much for volunteering for Ethiopia GPN. We're very excited to have Ethiopic GP working and coming towards a LGR proposal. So we'll follow up after the session.

NABIL BENAMAR:

This is Nabil Benamar from Morocco from the Task Force for Arabic script IDNs. My question is to Marc. As you know, we have

finished our work for the root zone as Task Force members and we are now tackling the second level and in the meantime the Universal Acceptance. The first comments are we have noticed that, for example, for the e-mail address – even if this question is beside the current topic of the discussion – the @ in the address hasn't been modified, for example, for the addresses. My question is where we can elaborate more so that we can have another character or something like that that replace the @ in the address to something Arabic or which is the same case in the other scripts. Is it in the ITF, in ICANN, as well? Thank you.

MARK BLANCHET:

E-mail it depends if you're talking about messaging within Facebook, but e-mail on the Internet uses various protocols and therefore those protocols are defined in the ITF so that should be the place, but it might be a big endeavor to change the @ sign.

UNIDENTIFIED MALE:

I think @, not a character rather than a symbol for separation. Yes, so if you wanted to change this character to accommodate Arabic EI, Arabic e-mail addresses I think most all the applications should be changed to accommodate this new character.

MARC BLANCHET: Clearly that's within the ITF scope to discuss that.

UNIDENTIFIED MALE: So are there any other questions or comments before we close?
If none, then thank you all very much for attending the session
and now we'll close the session. Thank you.

[END OF TRANSCRIPTION]