

---

SINGAPORE - IDN Program Update  
Wednesday, February 11, 2015 – 09:00 to 10:15  
ICANN – Singapore, Singapore

SARMAD HUSSAIN:

Basically, one of the main other projects is running the IDN fast track or implementing the IDN fast track process for the ccTLDs. We are now also starting a new project on IDN tables for the second level for testing – for enabling pre-delegation and testing in RSEP.

Also working on IDN implementation guidelines. These are the guidelines which are used for the second level for new gTLDs. And then we are obviously very actively outreaching to community, keeping the community informed on the projects we are undertaking, and also reaching out to the community to enable them, to get their input into the program, get them involved as volunteers to some of the work, which is obviously needed for IDNs.

So I'll not go into details for all of these projects, but I will share some details for some of the main projects which we are running.

As far as the IDN TLD program is concerned, the project obviously started in 2011 where there was some initial case studies done to determine what are different variants in different languages in different scripts and how these variant definitions may actually change or be defined across different scripts.

There were six case studies with six different scripts undertaken, and based on that there was an integrated issues report which was

---

*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.*

---

developed in 2012. Based on that integrated issues, next steps were chartered. Basically, the community agreed that there needs to be a consistent machine, [readable] definition of this data, through which the LGR XML Specification Project got started.

Also, the community agreed that there has to be a single root zone definition of LGR and developed a detailed process on how to develop this LGR, which involves integration panel and generation panels from community. And there was a study on user experience on the [inaudible] variants, which was undertaken.

Based on the process which was developed by the community for developing the LGR for the root zone, the follow-up projects were started in 2013. We are currently undertaking that project – those projects.

So the current projects actually include what we call P2.2, the LGR Development Project; and P1, which is the LGR specification and the LGR Tool project and P7, which is the LGR Implementation Project.

So just to give you an overview of where we are as far as the LGR TLD program is concerned, basically the starting point of this whole process is what we call the maximal starting repertoire (or MSR) and this is developed by integration panel, which is an independent panel of experts who are actually reviewing this work. They've already done an initial definition of MSR, and we have members of integration panel here who will go in more detail of where they are with this particular step.

---

Based on MSR, we have generation panels. These are community based volunteer panels, experts from linguistics, from DNS, community members who are giving their time to define which characters or code points for their particular scripts should be included to determine or define labels for root zone, and if any, which of these code points are variants of each other. And if there are any other rules or constraints, to form the root zone labels in those scripts.

We have three panels, which are seated and actively working. They include Arabic, Armenian, and Chinese. We have many other panels which are active and working towards their initial proposals to be seated and [inaudible] their work. So these include Cyrillic, Japanese, Korean, Myanmar, and Neo-Brahmi scripts. Then we have also obviously initial interest from many other script communities including Greek, Khmer, Latin, and Sinhala and Tibetan. So there's a wide variety of communities which are already active and starting to participate in this process. Some of those community members are here, representatives of those community members are here with us, to provide us feedback on the progress they've made so far and the challenges they're facing in the process as well.

Then once these community members finish their proposals, those proposals are handed over to the integration panel for review, and eventual integration into what is the label generation rule set for the root zone.

Just before the Singapore meeting today, the Arabic Generation Panel is the first panel to finish their work and they currently now are viewing it and it's close to being handed over to integration panel. It will first go



---

through a public comment period, and after the public comment period, it will be given to integration panel for its final review and eventual integration.

So much more on that by integration panel and community based generation panels later in this session. Moving on, we are also developing an LGR specification and toolset. The LGR specifications are being developed as an RFC. A [reasonably mature] draft of that is available at the link which is here on the screen, and it basically gives the [inaudible] format to represent this data and rules for different scripts.

The tool which we are developing is focused towards community use. Basically, we are looking at three broad use cases for this tool. One of the first use case is that this tool should be available to the generation panels, so that they can input the data which they're producing and it automatically creates the XML structure at the back for the corresponding LGR, so that the LGR doesn't need to be made manually to assist this whole process.

Once the LGRs are developed, this tool will allow the general user community to enter any label they want in the relevant scripts which are supported by the LGR. What this tool will do is basically, first of all, tell whether the label which has been entered by the user is valid or not based on the rules of the LGR. In addition to that, it will also tell what are other variants of that label in that script. So it will allow a general user to determine based on a particular label how many different other labels are variants of that label.



---

And then for each label, it will basically tell whether that label is allocateable or blocked. So those are the red and the green boxes you see at the bottom of the slide.

Then the final use case is for backend use where people may want to compare two LGRs or convert LGR from an existing format of IDN tables. So those are basically utilities for use by registries and other people who are working with LGRs at the back end.

All this technology we're developing is going to be open source. It is going to be accessible online so that the whole community can use it, not only directly, but also integrate it in their own systems.

Then, as I said, as far as the MSR and community work is concerned, we have members of integration and generation panel with us this morning and they will provide you further update on the work they are undertaking.

Moving on then, one of the other major work which IDN program undertakes is the implementation of IDN ccTLD fast track process. At this time, since 2009, since the program was launched, 45 strings have been evaluated, IDN ccTLDs have been evaluated from 35 different countries and territories. Basically, it covers 17 different scripts and 26 different languages so far.

Recently we've implemented the second panel for the two panel string similarity process. This was a revision of the IDN ccTLD process which was done in December of 2013. The second panel has come back and we posted the results of the review in October of 2014, and those have obviously been shared with public.

---

Fast track process also goes through an annual review process that's part of the process itself. Due to that, to get feedback on multiple points, we have actually recently released an IDN ccTLD fast track process for public comments. It's currently available online for public comments. There are a few questions which ICANN is asking the community to respond to as well, but the community is free to respond on any aspect of IDN ccTLD fast track process, so please participate in the process. Go through it and please raise any questions or comments you have on the process through the public comment.

One more new project with IDN program is currently aiming to start is the IDN – basically, there was a request from GNSO to look into the possibility of having some reference table for pre-delegation testing and RSEP because when registries apply for IDNs they're supposed to submit – the gTLDs apply for IDN support, they're supposed to submit IDN tables which are tested. However, there were no reference tables available, so registries – applicants – do not know what is the reference against which their IDN table submissions would be testing against.

So, as a first step, the tables have been shared through .SE which is the provider who tested these tables on behalf of ICANN. ICANN, however, we are now taking the next step forward and actually going to see if we can develop these tables in a more formal fashion, so that we can actually post these tables as reference for these processes for the community.

The process which we are currently undertaking will require that first we develop very detailed guidelines on how these tables will be developed. Then a provider will basically develop these tables, not only



---

develop data, but also list down all the authoritative sources for language tables, which includes national standards, dictionaries, and other references which are being used to develop these tables and provide analysis of the data which is eventually developed.

Once the IDN tables are developed, they will go through two expert reviews – one by a linguistic expert for that particular language and also a DNS security and stability expert to make sure that these tables do not raise any security and stability issues.

Once all these expert reports, documentation, and data is defined, they will be released for public comment for the community to ensure that the data and analysis is correct, and based on any feedback, the tables and documentation will be revised and all this information will eventually get published for the use of the community.

So moving forward, we are also very actively engaged as IDN program to reach to the community and tell the community what we are up to and obviously get their involvement to guide us as well as participate in the programs we are undertaking so that we can make progress in those programs.

We obviously provide these sessions at ICANN meetings. We also reach out to SOs and ACs to provide the community feedback and also go to other relevant fora around the world to actually get the communities involved and keep them updated on the IDN program at ICANN.

We have active mailing lists. Please join these mailing lists if you're interested in knowing about the program further. I will stop here, and



---

we move on to the update from Integration Panel. This will be done by Wil Tan, who is a member of the Integration Panel.

WIL TAN:

Thank you, Sarmad. Good morning, everyone. The Integration Panel would like to give some updates on what we've been up to since the last meeting in LA.

Very briefly, we've published a new version of the Maximal Starting Repertoire version 2. It is now in public comment. It contains six new scripts and associated documents along with it.

We've also published a few documents that hope will help the generation panels in their work, and also we've been involved in a few e-mail exchanges and document reviews with various GPs in their stages of development.

So MSR-2, we released it in December and there are 34 days left in the public comment period. So we would like to actually encourage anyone who is interested to submit your comment before it ends. You're more than welcome to.

MSR-2 is a complete [word] compatible replacement for MSR-1. It contains six new scripts. But since it's a complete replacement, you don't necessarily only have to comment on the new scripts. You can also comment on the existing scripts that are in MSR-1 and we're more than happy to receive those comments.

The six new scripts are actually Armenian, Ethiopic, Khmer, Myanmar, Thaana, and Tibetan. We have not made any changes to the MSR-1





---

scripts. Even though the MSR-2 was developed with Unicode 7.0, we've kept it to just a 6.3 subset. As some of you may well know, 7.0 poses some challenges, unsettled issues in the technical community, so we left it at 6.3.

Next, just some quick statistics. On the table on the left, we've tabulated the new scripts in MSR-2, the number of code points that are PVALID under Unicode 6.3 for that script, and then how many of those code points made it into MSR-2. So in total, we've added 702 code points to MSR-2. For rationale and why we made certain decisions, please consult the Overview and Rationale document that gets published together with the MSR-2. That brings us to 33,492 code points in MSR-2, basically.

We would like to just basically lay out – categorize – the documents that we have available on the root zone LGR project Wiki. All the documents are published there, and there are just three categories of documents, like the overview documents are also actually in public comment. We released them in January. The three documents are really good starting points for any GPs or anyone interested in the project to look at and review.

For anyone who wants more in-depth technical discussions, we have technical documents – one on the variant rules, the whole label valuation rules, as well as the XML, how to use the XML machine [inaudible] format, how to craft your LGR and code it into XML format. Of course we have the foundational documents such as the procedure as well as the MSR, too, in there as well.



---

So all of these can be found on the project Wiki, and from it you can find the document repository and all the documents are there. That's it.

We have another session from the IP – two sessions from the IP – this afternoon. I just wanted to point out that. I'll be talking through some guidelines and [Michelle] will be talking about some – he's going to go in depth into designing variants and [inaudible] rules.

SARMAD HUSSAIN:

Thank you, Wil. We'll move right along, going to community updates. The first up is Nabil Benamar of the Arabic Generation Panel to give an update on the panel's work.

NABIL BENAMAR:

Thank you so much, Sarmad. So I will present an update on Arabic GP on behalf of the Taskforce of Arabic Script IDNs. This is a community-driven based as a taskforce on Arabic script IDNs. It has been created under the framework of the Middle East Strategy Working Group, and there is a link for the taskforce.

We have many objectives. The first one is to establish this Arabic script label generation rule set for root zone, and then to tackle the second level LGR for the Arabic script. Then we can go to the Arabic script and [inaudible] data, universally [inaudible] for Arabic script IDNs so that we can tackle also the issue of the e-mail address using the Arabic script, WHOIS, and many other issues that we are thinking of moving forward.



---

Making technical challenge around registration of Arabic script IDNs, operational software for registry and registrar operations, and all DNS-related issues as well as the DNS security as well.

So currently in this taskforce we are 33 members, and applications are still being received. The members are from 18 countries all over the world and members have nine language communities using Arabic script, which means that we using different languages but under the Arabic script, such as Arabic, Malay, Saraiki, Sindhi, Pashto, etc. And with further expertise in use of Arabic script from East Asia, South Asia, Maghreb countries with different specificities related to these different languages.

The members are from diverse disciplines, so we can have linguistics – we have linguistics, we have technical background members, registries [DNS], registration [to DNS operation] as well.

The membership is open, so it's community based. All the details and interests of members are posted on the website of the Middle East Strategy Working Group. Also the discussions are publicly archived. All the details are in this link. Background and introduction for the Taskforce for Arabic Script IDNs is also available for everyone. It's publicly published.

Any conference call that we have done are registered and archived in the website. Anyone can listen to our previous conference calls, and also all the face-to-face meetings that we have done.



---

This is an overview of the Arabic script gTLDs assigned and delegated, so that you can see. We have already 20 TLDs, like [inaudible] .IRAN, .PAKISTAN, etc.

This taskforce has launched during the third – or second – Arab IGF meeting that was held in Algiers back in 2013. They did a presentation during the IGF in Bali. The first steps was to outreach during the Middle East DNS Forum, and this taskforce presented many presentations during ICANN meetings as well as IGF meetings. The last one was an organized workshop during the third Arab IGF in Lebanon.

Setting up things that [we] have done. The first one was the formation of Arabic Script Generation Panel and establishing principles for inclusion, exclusion, and deferral of Unicode code points. These code points should have an analysis of Maximum Starting Repertoire and feedback to Integration Panel.

After that, analysis of code points for label generation rules so that we have also established principles for inclusion, exclusion, and deferral of Unicode code points regarding the LGR.

Second step is the code point variants. This concept to be used in our work so that we can establish principle. We have established the principle for code point variants. We have analyzed the code points for code point variants, also code point variant dispositions as we have seen with the presentation of [Dr.] Sarmad. So the concept of allocatable versus blocked.



---

Then the whole label evaluation rules. So that's establishing principles for whole level evaluation rules and determining this whole level evaluation rules.

The final step is we have issued a document, so a document in our work that has been actually sent to everybody. It's publicly seen that we can have the feedback of the Integration Panel and the community. Text document in summarizing these label generation rules for Arabic, as well as its XML format.

So this is a sum-up of our work that we have done. So we can see all the labels, all the code points that have been used. So we can see the purple is for inclusion. So now we have proposed 135 code points, and excluded [inaudible] LGR by Arabic Generation Panel. The number was 86. This is a sum-up of our work.

Looking forward to continue our work is to finalize the Arabic script LGR proposal for root zone so that now we are in the final internal review and going to next step, which is a release for public comment and submission to ICANN.

Then we're going to go forward for Arabic script LGR for the second level toward universal acceptance of Arabic IDNs. Thank you, [inaudible].

SARMAD HUSSAIN:

Thank you, Nabil. Let's keep going forward and go through the rest of the other community updates. The next one is by Hiro Hotta, update on Japanese Generation Panel.

HIRO HOTTA:

Thank you, Sarmad. My name is Hiro Hotta from JPRS, which is a ccTLD registry of Japan, .JP. So I'll briefly update the data of the Japanese Generation Panel.

We have meetings, nine meetings so far. So we informally established our meetings, [so] Generation Panel actually of August 2014 and we did eight meetings. We discussed two things. One is how to establish this generation panel, and the second one of course is the content of generation rule.

We have at this moment eight members as a first tier. And of course we have second tier who helps these guys. I am the chair of the JGP, and I cannot remember who is from JPNIC. He's the vice chair of the JGP. And we have several persons who are experts of IDNs, Internet, and language and so on.

Our activities. We have been talking about how to [consist] the JGP, and it's done and we've just submitted our proposal to ICANN on Friday last week. We also describe what JGP should do, and it's almost done. Of course some more taskforce issues may come out during discussion with ICANN Generation Panel [and ways] CGP and KGP, Chinese and Korean, because we do share many characters with Chinese and Korean, so we have to coordinate and cooperate among us. And of course, inside JGP we may have come up with some new issues, which is not known now.

---

We are developing our Japanese rules. We are still in the basic study stage. What are intrinsic issues and what are the basic solutions for the issues?

This is the status of the development. The first one, scopes of the character codes, repertoire – Kanji, Hiragana, Katakana and maybe ASCII. When we submitted – before this Singapore meeting, we thought that the ASCII, which means English alphabet characters, should be able to mixed with Kanji, Hiragana, Katakana which are the local characters, but in the discussion with the [IP] and [inaudible] maybe we will give up the mixture of ASCII.

And for Kanji, which is a Japanese [inaudible] level one and two, that has around 6,000 characters, so we almost have reached consensus about this, 6,000 characters for Japanese local languages.

Variants and [inaudible] still under discussion, so all Kanji characters will be independent or may have variants. We are still in discussion. A language LGR passively adopts the variant definitions, if any, of other language LGRs. I put “if any” in parenthesis, but definitely Chinese rules have variants I think.

In addition, a language LGR defines all the characters in its scope to be alloctable. And Japanese LGR actively defines variants for its own? Or not. It’s under discussion.

The whole label evaluation, our rule may have a very limited number of tiny, straightforward rules even if defined. Thank you.



---

SARMAD HUSSAIN: Thank you, Hiro Hotta. We'll move on to the next update on Korean Generation Panel work. It will be given by Professor Kim.

KYONGSOK KIM: Thank you. Oh, the first page is missing. My name is Kim Kyongsok [inaudible] university and I'll give you updates within Korean Generation Panel. Five issues will be presented today.

The first one is whether or not to include Hanja in KLGR, and second whether to include [LDH and KLGR]. Thirdly, what kind of list can we have for KLGR, especially which Hanja characters will be included. The fourth issue is Hanja variants. Lastly, reconsider recommendation 7 in SAC060 in Han variants.

Whether to include Hanja in KLGR or not? We discussed this issue and there is rough consensus that we'll probably include Hanja in KLGR. It is not the final decision yet. Probably [we need] to make a [poll] and/or public hearing to reach a consensus among Korean language community.

Secondly, LGH issue was mentioned already by Japan. There is nearly the same situation. I'll skip this issue.

Thirdly, what Hanja characters are [we] going to include in KLGR? We have not decided yet, but we considered a few candidates. The first one is K Hanja in IICORE. IICORE is defined in ISO/IEC 10646, and it is an international core subset of up to 10,000 Han characters out of about 76,000 Han characters in ISO/IEC 10646. Out of 10,000 [inaudible] Han characters, about half – 4,700 characters – are in [K Korea].





---

You can see example here. In the first line, KR is marked as blank. And in the second and third line, KR column is marked as K0. So we considered including those KR characters in IICORE.

We analyzed the number of characters in IICORE. Most of them are within KS X 1001 and a few are from KS X 1002, 1027-1, and 1027-2. Some are not included in this KS X.

Secondly, you can [inaudible] Han characters in KS X 1001. It has the [inaudible] within Korea about 40 years, and there are 4,888 Han characters in this standard and there are 268 compatibility characters which are not included in KLGR. To be specific, it is 15 characters not included in IICORE and we are [considering] whether we include or exclude those 15 characters in KLGR.

There is rough consensus that follows. The following sets seems to be a good starting list for KLGR. Firstly, 4,700 characters in IICORE and 15 characters in KS X 1001 were not included in IICORE. About 50 Hanja characters marked KP column, but not in K column or by IICORE. And probably there may be some more, but I guess it won't be many, even any.

A list of Han variants. Within Korea, there is officially no variant list, so we need to make a variant list. Once KLGR finalized the Han list itself, KLGP will start working on a variant list.

This is about Recommendation 7 in SAC060. It was mentioned by Mr. Sarmad to me and we considered a relationship between SAC060 and Han variants. It was our understanding that Recommendation 7 seems good in general, especially when there is single variant mapping table.



---

However, in the case of CJK, variant mappings are different among CJK. I'm not sure if SAC060 can be applied to CJK case [directly]. Okay, thank you.

SARMAD HUSSAIN:

Thank you, Professor Kim. Now we move on to the last community update for today's session on the work by Chinese Generation Panel, and then we'll take questions from the floor. We have Kenny Huang who will be making the presentation.

KENNY HUANG:

Good morning. My name is Kenny Huang, and Wang Wei is on my right-hand side. I'm happy to give an update on Chinese Generation Panel.

So far we have up to ICANN 51 a meeting we're running for [inaudible] Chinese Generation Panel meetings, including one CJK joint generation panel meetings on the 8<sup>th</sup> of January.

On the end of last year, December 29<sup>th</sup>, we submitted [inaudible] Chinese Generation Panel proposal to CJK mailing list. So all the CJK members can review on the proposal made by Chinese Generation Panel.

Basically, we just try to highlight what's the philosophy we try to integrate CJK label generation rule. Initially, we need to respect the principle not only defined by Integration Panel, but also reference to various technical reference documents such as SAC060 document and also such as recommendation from IAP.



---

Second, we concentrate on the overlapped code points. Also, that will be a case-based analysis, [inaudible] not try to theorize the abstraction, because when we try from [inaudible] point of view, when we try to theorize all the potential possibilities, I think that's impossible to complete the integration [all the] integration rule set.

So we try to focus on a case what we have and also prioritize the evaluation according to the data availability, such as we have data source from [inaudible] table, and we have data source from [inaudible].

Right now, we focus on the Zone A. That means the overlap between Chinese label and Japanese label. In other words, [inaudible] about 6,181 code points.

[inaudible] one example how we integrate Chinese label integrate with Japanese label, and that's one of the case. Also, we probably have different kind of scenarios and that's just one of the scenarios. Of course we also need to discuss with Japanese community to see how to make more a scenario, how to work in all the algorithms to integrate two labels.

For example, on the first row, we can see [inaudible] code point from [inaudible] traditional or simplified Chinese code point, and [inaudible] Chinese code point language [inaudible]. And we have different [kinds of] Chinese variants. For example, we have Chinese variant one, two. And that could be allocatable disposition. It could be blocked disposition and it depends on the definition rule given by Chinese Generation Panels.



---

The other case is come from Japanese code point. For example, Xj. So far, based on the source, the data source that we have, they don't have any variants. So all the variant [inaudible].

So we're starting to work out a computational algorithm. For example, if you find a Japanese code point among a Chinese code point and we allocate or we start to allocate all the variant tables to a Japanese variant table and we set the value disposition become a block. So [inaudible] integrate Japanese code point will be something like this one, and the Japanese code point will be generated [whole set variant] and with the disposition [inaudible] blocked.

Last one of the scenarios, as I mentioned, and also again we need to discuss with Japanese to obtain consensus in order to move forward.

Okay, [inaudible] real case, according to previous algorithm. For example, the code point [inaudible] in Chinese is [inaudible] and we have traditional code point and simplified code point. That's first row.

Second row is Japanese code point. This character is [inaudible], but they don't have any variants.

Based on the previous computation algorithm [as I suggest] [inaudible] Japanese [word generate] new variant such as 611B and with new variant 7231. This new variant actually did not initially exist in their variant table, but right now [inaudible] new word to the variant table [inaudible] be blocked.

Last, Chinese Generation Panel repertoire and variant type. The case was given by previous ICANN meetings, so I'll just skip it. That also has been presented in ICANN London meeting, [inaudible] potential



---

[inaudible]. If we still want to figure out how to integrate together, something we probably need to consider – for example, the usage pattern among Chinese, Korean, or Japanese community such as we collect a code point, a common and modern string from a newspaper and try to computationalize their frequency of use and their usage pattern, and then we can come up with the most [scientific] to identify what kind of code point [they should] keep, what sort of code point we can integrate or harmonize together.

This is another suggestion given by Dr. Joseph [Chung] who is a well-known linguistic expert in Chinese language. He proposed we probably need to redefine the variant in the one-language context.

For example, the modern pronunciation should be the same, and modern meaning and application are the same. Based on that principle, we can redefine CJK variants across [inaudible] language context, because so far we didn't base on that language context to define the variant, so probably we can consider a language context point of view to define the variant.

Finally, some suggestions. We tried to institutionalize all the CJK [working models], such as we need to recognize the principle either given by Integration Panel or given by [inaudible] different kind of technical communities, and we need to develop repertoire variant type mapping. Also [inaudible] CJK panel. And we concentrate on [overlapping] code points, as I mentioned previous. We need to set up a goal schedule and deliver and try to make it happen.

Finally, [inaudible] challenge that we have. For example, the principle, we probably still have – some members have some concern based on



---

the principle and we need to clarify the concern to see what's the best way to move forward. Also, we need to deliver repertoire, variant mapping and variant type among CJK community.

Also, the final suggestion is we probably need to set up time tables and deadlines for every step. We [inaudible] schedule by the ICANN because ICANN actually [inaudible]. For example, [inaudible] is going to launch in the second round of new gTLDs, and the how the new gTLD is going to apply that label generation rule, because if we release the [inaudible] gTLD, that means [the process] cannot reverse once [your] rule [inaudible]. Thank you.

SARMAD HUSSAIN:

Thank you. Those were the updates from different community-based generation panels. We will now open the floor for any questions you may have. So please come up to the mic and let us know if you have any queries to any of the members.

MICHEL SUIGNARD:

Yes, this is Michel Suignard from the Integration Panel. I just wanted to make a clarification on one of the presentations, if you don't mind. It's about the Korean presentation. It was mentioned quite a bit [inaudible]. I just want to mention that the [inaudible] is in fact fully included in the MSR. We made sure it was fully included.

On the same time, I would really like in the future to be compared against the MSR [content] and not [inaudible] because there's quite a bit more characters [on] every generation panel should really look – [inaudible] CJK – should really look if the content is in MSR for sure,

---

because Korea it was kind of [inaudible] in the CJK work to some degree because there is no Hanja reference IDN for us, so we had to kind of guess what would be an IDN content for Hanja.

So we based in fact our work when we did MSR-1 on the information that we could get, which was basically to make a union of the .ASIA set of the .JP [inaudible] because we thought that there were like 15 characters missing otherwise from the very basic case, so we added those.

But I would like in the future that – because there may be a chance that in fact some characters may be missing for Korea in MSR because we don't know. We don't have really a good reference for that, so MSR is currently under review, so it's a good time to verify that all the characters you may need are included in MSR and you better do that soon because the review is taking place now. That was my only . . .

SARMAD HUSSAIN:

Thank you. We have Edmon.

EDMON CHUNG:

Thank you so much, and thank you everyone for a very informative update on this very important issue. I think it seems like there's good progress and that's good to know.

I wanted to go back very, very far into, Sarmad, your earlier presentation. I notice a communications plan. I probably missed it from before. What is envisioned on the communications plan and going



---

forward? That's [inaudible] on that, because I think it's the very first two slides or something. There, you got it.

SARMAD HUSSAIN:

So, basically a communication plan was developed internally for the IDN program to manage and plan and eventually undertake the outreach activities for this program.

Basically, there are – just to add – we are focusing on two distinct stakeholders here. One is the general community, obviously, which is impacted by the IDN program. But also the second set of stakeholders are potential and current Generation Panel member who are directly involved in the work with the IDN program. And there are internal strategies to outreach to both kinds of communities. For example, we have strategies which are geared towards reaching to communities which are [not formed] as Generation Panels right now. We are outreaching to those communities trying to inform them of this work which is needed, try to encourage them to participate in the process and so on. It's a series of strategies aimed at two kinds of stakeholders.

Again, [the aim] of different strategies obviously is to [apprise] the community of the current work, and obviously get the GP members involved in the process.

EDMON CHUNG:

Adding to that – thank you. Thank you so much. The reason why I ask is I guess I hear that this is an internal plan and was not shared. Not asking for transparency on this issue, but I think if the larger community knows





---

a little bit more about it, it may be useful. For example, wearing the At-Large side hat, we could probably help you in getting the word out.

This is something that is, in my mind, not only about formation of the LGR, the generation panels, but eventually also as these LGRs are being published for public comment or getting the community review, we want people to know ahead of time so that we can get ample feedback for the Generation Panel or the Integration Panel products as well.

So this is I think an ongoing thing and I would encourage you to share it with the larger community and perhaps we can chip in to help as well.

SARMAD HUSSAIN: Okay, thank you.

ZAKIR SYED: Excuse me. Question?

SARMAD HUSSAIN: Yes, please go ahead.

ZAKIR SYED: This is Zakir Syed, ICANN fellow from Pakistan. I have a question. It could be very basic. Probably this is my very first session on the ICANN session.

The question is with the introduction of IDNs on ccTLDs maybe on the second level domain names, do you think it's going to create challenges in terms of IP and domain name disputes? If yes, are you – I mean, this

---

IDN program and the team – is in cross coordination with the IP constituency at ICANN? Thank you.

SARMAD HUSSAIN: Yes, Edmon, you want to comment?

EDMON CHUNG: Two things. One, this particular program that is brought up in this particular session is focused on a very narrow scope, which is on what is called IDN variants. So on a linguistic and technical level, certain characters are considered the same. So the intellectual property rights and those issues I guess are not part of this specific program.

However, what you raise is very valid. This is something that is very important for the entire ICANN process. This was specifically discussed quite some years ago before the IDN TLDs were implemented. Also, specifically there is a discussion at the – a very old working group called the IDN Working Group under the GNSO (the Generic Names Supporting Organization). When the new gTLD program was created, part of the GNSO policy recommendations including an IDN Working Group Outcomes report.

In that report, it specifically talks about some of the challenges, including the interactions with UDRP (the Uniform Dispute Resolution Policy) and some of those issues that once you have different languages and characters, how is that handled?

But on that topic, actually, the general consensus is that those intellectual property rights issues is better dealt with in the trademark



---

conventions and experience there, which already has experience handling different languages. Brands that operate in multiple jurisdictions, they have already fought some of those battles and we should take those experience into here rather than create new ways of looking at those issues.

However, back to this particular variant issue which tries to look at two characters and potentially considering them the same, the issue you brought up may have implications later on the actual implementation into the root. That's project 7 and I think it's called LGR Implementation, P7. That is not started yet. Well, that is started, but it hasn't come out for public comment yet. When that comes in, I think that's probably the area specific to this program that is most relevant in terms of the policy of handling confusables or issues in intellectual property rights. Does that . . .?

SARMAD HUSSAIN:

Just to follow-up on that very briefly, so IDN program staff does provide support to any such work if it comes up, but it is directly handled by other teams within ICANN. Any more questions?

UNIDENTIFIED MALE:

We have a question from a remote participant, [Don]: "A question for the Arabic team. Have all the Arabic script TLDs and ccTLDs agreed to apply the same rules to their second and third level name space?"



---

**NABIL BENAMAR:** Thank you. We are now finalizing the work of the top-level domain and we will tackle the second level just after this work. But I think it will be just a recommendation for the community how to use the IDNs and the LGR in the second level. Thank you.

**SARMAD HUSSAIN:** Just to add to that, as far as the Arabic Generation Panel work is concerned, that is limited to the top level. And any work undertaken for the second level is being done by the Taskforce on Arabic IDNs which is a community-based group and it's not through the Arabic Generation Panel work directly.

**EDMON CHUNG:** Adding to that, I think in response to that question, in my mind I would caution against setting that type of expectation. The reason why is because the need for a fully unified set of variants is motivated at the top level because there is no way of identifying the language context or linguistic context of the characters being used.

Case in point, Hanja characters, Kanji characters, and Chinese characters, at the top level there's no way to figure out whether it's supposed to be Japanese, Korean or Chinese. However, in the second level, especially in the context of ccTLDs, the implementation of the rules could very much be different. So I would caution against a general expectation that it's always going to be the same, because the motivation for the unifying, in a way, of variants at the top level is because you lose the context of being able to have the linguistic nuances.



---

However, I would go to say that in the context of gTLDs, that may be more similar to the top level situation. So you could think about the gTLD second level space to be a little bit more akin to the top level, but certainly the ccTLD case may not map nicely to that expectation.

UNIDENTIFIED MALE: Hi, my name is [Manny] and I'm from Singapore. I've got a question. You're talking about variants of tables on languages. Is there anyone working on [Indic] scripts?

SARMAD HUSSAIN: Yes, we have a group which is being formed. It's called New-Brahmi Generation Panel which is dealing with multiple scripts. If you need more information, please feel free to get in touch with me right after this and I'll pass you—

UNIDENTIFIED MALE: I appreciate it, because I am working on the IDN application for the next round for the Tamil language, which is [inaudible] Singapore, Malaysia, [India], and [inaudible].

There's an extension on the outreach. Is there any activities taken up by IDN as a whole [inaudible] outreach program on IDN application process or implementation of its program

SARMAD HUSSAIN: I guess a short answer is no, and if you have any requirements or needs, please let us know and we are very willing to work with you.

---

UNIDENTIFIED MALE: Yeah. I appreciate it, because [inaudible] Singapore in the month May and I would like to add IDN outreach schedule. We can do it via Adobe Connect or somebody could come and updates.

SARMAD HUSSAIN: Sure. Just please let us know and we'll certainly follow-up.

UNIDENTIFIED MALE: Thank you.

SARMAD HUSSAIN: So we still have time for at least one question if anybody wants to ask a question or make a comment. Please come to the mic.

STEVE SHENG: Thank you. Steve Sheng, ICANN staff, but speaking on my personal capacity here. I think I just want to echo what Edmon said that the root zone is a special shared resource. There is only one root zone that is shared by the Internet community, and therefore I think the policies made for the root zone [inaudible] necessarily more conservative than other root zones. So this may or may not apply to the second level.

One question that I have, I think because the root zone is a shared resource for the Internet community, and therefore by necessity, that would require different language communities within the script to work together to find meaningful compromise to move things forward. So



---

with that, the question is directed to the CJK community. Has there been coordination on your work and what's your plan on that? Thanks.

UNIDENTIFIED MALE:

Thank you, Steve, for the question. Yes. The simple answer is yes. At least every ICANN meetings we meet the CJK, met every time how to coordinate. It seems that the Chinese community goes first and fast, but the J and K are catching up with it.

In this meeting, for example, yesterday we had a [inaudible] meeting inside CJK. We coordinate that plan, how to proceed [in size]. So please be patient to come up with some result. Thank you.

SARMAD HUSSAIN:

And just to add to that, the IDN program does organize coordination meetings between panels which need to have those meetings as regularly as requested by the panels.

WANG WEI:

Wang Wei from CGP. Yeah, just as [inaudible], we realize that the coordination between the CJK is pretty important if we want to reach the final solution, which works for the whole Internet community on the Chinese character.

Actually, before we came to this meeting, we had a discussion in China with some linguistics. For the second level Chinese domain name registration under .CN and [inaudible], we had the size the repertoire is almost 20,000 and we have discussed this issue with Chinese linguistics and I think we are prepared to reduce the size of the repertoire to

---

decrease the complexity of coordination, maybe decrease the size to about 6,000.

Of course I think even with the [inaudible] size, there are still a lot to of overlapping characters and both of us, the three parties, have different definitions, so we raised another issue there.

We think a redefinition of the variants, for Chinese variants, and definition for the variants across the three-language environment is needed, which will make the linguistics of different countries understand this issue better.

Maybe after this Singapore meeting, we will ask [inaudible] from ICANN if it's possible to create an independent Adobe Connect account for the three of us, which will help communicate when we want it.

I think it's a tough job. Anyway, after a painful and torture discussion in Monday and Tuesday, we have reached a rough consensus. We will work together and try to accomplish this work as soon as possible. Yes, thank you.

SARMAD HUSSAIN:

We are over time. We'll just take a quick comment, Andrew, if you still want to make one, but just keep it quick.

ANDREW SULLIVAN:

I'm pleased to hear that people are working together on this CJK case, but I want to remind you that the mechanism that we created doesn't actually require that everybody get exactly the same results for a given



---

character. That is, the same characters all have to be treated, but you can have different results depending on the language.

So for Korean, for instance, you might have a blocked variant; whereas for Chinese, you might have an allocatable one or something like that for the same character and that's okay. If you need any additional help with that, I can talk to you about it after.

SARMAD HUSSAIN:

We have a detailed session on an LGR workshop later today, which will dive deep in many of these issues. For those who are interested to continue this discussion and learn more about it, please come and join us there. It's from 1:00-2:15 in Moor and we hope to see many of you there at that time. So let's close the session and thank the panelists for their contributions. Thank you.

[applause]

[END OF TRANSCRIPTION]