

---

BARCELONA – ICANN GDD: IDN Program Update  
Wednesday, October 24, 2018 – 08:30 to 09:45 CEST  
ICANN63 | Barcelona, Spain

SARMAD HUSSAIN:

Thank you for coming early in the morning. We'll start the IDN Program Update Session in just half a minute and let everybody settle down. There are plenty of seats up front, so if you'd like to join us at the tables, please come forward.

Very good morning to all of you. We'll start the session on IDN Program Update. Today we have a few presenters. We'll start with an overview and progress on the IDN Program. Following that, Marc Blanchet, who is a member of the Integration Panel, will provide an update on the work being done by the Integration Panel.

Then, we have updates from the community-based generation panels. We have chairs and co-chairs from Latin GP, New Brahmi GP, Sinhala GP, and Myanmar GP, who will be presenting the updates for the work these community panels are undertaking. We have Mirjana Tasic who is representing the Latin GP, Ajay Data who is representing the New Brahmi GP, Harsha Wijayawardhana representing the Sinhala GP, and Thin Zar Phyoo representing the Myanmar GP. Then, we'll follow-up with questions and answers.

Moving onwards, let's start with the overview of the work being done at the IDN program. We'll start with the objectives. Basically, the IDN Program is focused on enabling the deployment of domain names in local languages and scripts used by communities globally. We

---

**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.**

---

obviously try to focus the work in such a way that these domain names, their use remains secure and stable.

We undertake quite a few different projects. This gives you an overview of those projects. We largely divide those projects in two different sets which are focused at the top-level domains and then we have a couple of projects which are focused at the second level. So, most of our work is actually focused top-level domains. One of the large projects we are undertaking is called IDN TLD Program which has other subcomponents in it. One large component which we have been supporting for a few years now is the development of root zone LGR. You will hear more about that from the community members here as well.

Then, now we are actually starting to embark on a project on deploying, or possibly deploying IDN variant TLDs, and then all this is possible through data and the LGR toolset, actually, the linguistic data being produced and the LGR toolset is actually being developed to manipulate that data and allow community to use that data.

We also, in parallel, run the IDN ccTLD Fast Track Process. We undertake the evaluation stage for that process before going on to the delegation stage at IANA.

At the second level, our focus has been on maintaining the IDN implementation guidelines and also, at this time, we are now starting to develop reference second-level LGRs for registries to use. And of course since all of our work is very strongly and deeply based on community participation and leadership, we have an ongoing program

---

to outreach to the community and get them involved in the work we do. So, we'll overview some of that.

So, starting from the root zone LGR. The IDNA 2008 which is the basic protocol for internationalized domain names, expects all registries at all levels to introduce additional constraints on top of IDNA 2008 to address any consumer confusion possibilities and that could be done as per the mechanisms suggested in IDNA 2008. That can be done by, for example, registering characters which are used in domain names or by using variant techniques.

So, root zone LGR work which we are undertaking is a basis of such a mechanism for the root zone, and of course the aim of the root zone LGR is, on one hand, to make sure that IDN TLDs are available in all the different scripts and languages for the communities globally. However, it must be done in such a way that these TLDs remain secure and stable. Then, as already noted, we are also working on determining whether any of these TLDs have variants and we'll talk about some of these in more detail in the presentation.

So, root zone LGR basically tells us what is a valid TLD, what would be a possible valid TLD in a particular script. Of course, ICANN does not have the expertise to determine what a valid label would be in all the different scripts around the world.

So, the process which was developed by the community to define the root zone LGR was that we need to go to all the different scripts communities, organize them or help them, support them in organizing into what we call the generation panels. So, there's a generation panel

---

for each script we want to support in the root zone. Then, each generation panel works on their script and this generation panel includes experts from the particular scripts themselves, so we for example, could have a Chinese generation panel which has experts from the Chinese script. We have a generation panel on Arabic script which has experts from Arabic which are fluent in Arabic script and so on.

So, the idea was that we have these community-based panels which then set the parameters of work done in the top-level domains in their particular script and also define what are possible variant labels for those scripts.

Each generation panel does their work, develops a proposal, and sends it to ICANN and then ICANN then hands that proposal over to an independent panel of experts which is called the Integration Panel. And this Integration Panel makes sure that the proposal by each of the generation panels meets certain security criteria, stability criteria, and other criteria defined in the LGR procedure.

Once Integration Panel is satisfied that those criteria are met, then the proposal is integrated into what is called the root zone LGR. So, root zone LGR actually contains multiple scripts. This is not coming out correctly for some reason.

So, in the first phase, we identified 28 different scripts which basically we will be working on. And those 28 scripts cover most of the languages which are actively spoken and used across the globe. So, I'm very sorry. This is not showing up. The script names basically are listed I guess in

---

this screen at the top. As you can see, we have about six scripts which have those arrows at the top which are already integrated into the root zone LGR. Those communities have already finished.

In addition, we have another 12 scripts which have finished and in the process of going through public comment or review by Integration Panel. So, we have a total of 18 scripts already finished out of the 28 scripts which we were covering. In addition, we have another 10 or so scripts which we are actively working on. Actually, the communities are actively working on and those scripts will eventually finalize their proposals and get integrated. You can probably access a presentation online available through the schedule to see these names of scripts. They're not showing very well on the screen here.

So, that's the status of the root zone LGR. Now, moving on to variants. So, one of the things which – at least a couple of things which these root zone LGR definitions provide. First, of course, which characters can be used for top-level domains. So, forming a top-level domain label.

Then, in some cases, when we use these characters, some of these characters are “same” within a script or across a script. So, those characters also need to be identified so that we can anticipate what a variant label could be. The reason we want to understand, and in a way, predict what variant labels are is to ensure that those labels are managed in a way that, for example, they should not be delegated to do different entities because, if they get delegated to two different entities, they can potentially create user confusion in the use of domain name system.

---

These are some examples of variant TLDs. As you can see, there are two kinds of examples here. There are examples which can potentially cause security issues, and therefore one label should block the other label. Then, there's the second category of variant labels where it's basically motivated not by just security but also usability which means more than one label actually need to be activated or turned on so that people from different regions can access the same domain name effectively.

Just to go through this very quickly, at the top you see the two strings, Arabic and English and then you'll see a very similar-looking Arabic and Latin script, and you'll see a very similar string which looks like Arabic but actually is a different string, in Cyrillic script. Obviously, visually you can't tell the difference.

At the bottom of each string, what we've given is the Unicode code points. So, if you look at the Unicode points, you'll see that these two strings are technically different, even though visually identical.

So, therefore, if one is registered, the other should be blocked. You don't want both of them to be registered as top-level domains and given two different entities.

If you look at the right, you have examples in Chinese, Han script, where you have a simplified version and traditional Chinese version. In this case, both of them need to be in a way turned on for effective use of this TLD because simplified version is the way of writing on mainland China, whereas the traditional version of the same label is used in Taiwan, in Hong Kong, Macaw, some of these other geographical regions, so really

---

to reach out to all the Chinese community. You would need both versions to be enabled, not just blocked.

Similarly, at the bottom, you'll see example from Arabic script which is also similar in different geographic areas, you will need different strings to be turned on, so that people can actually access them through their own local keyboards.

So, you need variant labels either to address security but also, in many cases, to address usability of those top-level domains across different geographical regions.

Basically, because of this, the community had very early identified that there was need for these variant labels. However, when this work was initially identified back in 2009 and 2010, when IDN top-level domains were starting to get delegated, basically it was not really clear what variants are, how they're defined across different scripts. So, at that time, the board actually had said that we need to ... We cannot delegate a new gTLD, variance of new gTLDs until we really understand how variants work and how they actually should be managed.

So, there was some subsequent work early on in 2011 and 2012 which led to two problems, identified two problems which need to be solved. First, that we need to have a definition of what variants are in different scripts, and then once we have actually defined them, we then – we need to also have a mechanism on how they really need to be managed.

So, since then, we've actually been working on those two problems, as I've already discussed. The solution to the first problem on how should

---

variance be defined is through this LGR procedure which we developed, where communities of different scripts, these generation panels, are actually defining those for us. So, with the now availability of root zone LGR, we actually have a definition of variants for at least the scripts which are integrated.

Then, now, we've also recently launched a community-based study group on root zone LGR which is now studying how technically the root zone LGR should be applied into the current and future processes for determining top-level domains within, obviously, the context of gTLDs and ccTLDs. So now that group is active and looking at how the root zone LGR should actually be technically used.

In addition, there was that second problem which was to look at once we know what variant TLDs are, how do we actually manage them? ICANN Org has actually been working on developing some recommendations on it. Those recommendations were released for public comment in July this year. It's a set of documents. They're available online in case you want to go and take a look at it. We actually had a session on this earlier this week as well where we presented those recommendations. It's closed public comment, and at this point, we are reviewing those public comments to finalize those documents to present to the board where the board will then consider them for subsequent steps.

In addition to those recommendations, we're also developing an LGR toolset, so this linguistic data which has been produced now by all the different 28 generation panels, this linguistic data has to be organized



---

in a way that machines can process them, so that when you put in, for example, label, the machine can tell you whether the label is valid or not for a particular script and also tell you what are the possible variants of that label.

So, we've actually been developing a tool. Before that, we worked with IETF and developed a specification which is called RFC 7940. It's a standard [strike] specification to, in a way, store this information in a machine processable format and then we are now developing tool which allows those LGRs to be used by the community to create LGRs, to view LGRs, to use LGRs, manage LGRs. For example, compare two different LGRs to see the [inaudible] between them and so on. And many different functions. This tool is actually now available at ICANN website. You can go and use it at [lgrtool.icann.org](http://lgrtool.icann.org). There is a very detailed user guide available online as well, in case you want to learn how to use it and what it does. Then, complete, in a way, code for this tool is also available as open source at GitHub, so in case you want to deploy it at your own servers for your own use, you can actually just go online, download the code, and use it separately as well. So, you have both options.

In addition to the general TLD program, as I had mentioned earlier, we also are running or supporting the IDN ccTLD fast track process. Through that process, we've already delegated 58 different strings belonging to 40 different countries and territories across the world. Of these, 56 have already been delegated, representing 38 different countries and territories.

---

IDN fast track process also goes through a review process and it's actually currently going through a review process and being ... The second string similarity review guidelines are currently being reconsidered based on previous use by ccNSO and we're anticipating once the ccNSO concludes that review, we will update a fast track process to include that feedback as well.

So, that's some of the work we're doing at the top level. At the second level, we do at least a couple of projects, one on IDN implementation guidelines and the other on reference LGRs.

So, we'll talk briefly on IDN implement guidelines. These guidelines are business second-level IDN registration policies and practices which we recommend to address consumer confusion. These are binding through contracts on most registries and registrars offering gTLDs. They're also encouraged to be used by ccTLDs.

Basically, on the request of community, we start ... The previous version was actually developed back in 2011. And based on community feedback, we started a review process in July 2015. We worked over last few years, couple of years, and finalized through a community-based IDN guidelines working group with nominees from GNSO, ccNSO, ALAC, and SSAC. That group eventually finalized these guidelines in May of this year. These were published as IDN guidelines version 4.0 in May. Currently, we are internally reviewing these guidelines to determine what are these additional requirements with these guidelines posed, and understanding those requirements we aimed to eventually bring

---

this back to ICANN board for approval and implementation early next year.

The new guidelines cover seven topics and have actually 19 different guidelines, focusing on transition from IDNA 2003 to 2008, format of IDN tables, consistency of IDN tables and practices, IDN variant labels, similarity and confusability of labels, and then publishing registration policies. Also, terminology. So, it covers a wide variety of areas and then multiple guidelines available.

In addition, obviously, to all that work, we support the community work which the community undertakes for the IDN program. These are some of the outreach activities which we have actually been involved in, in many cases supporting the community-based generation panels and then we also try to reach out to the community at ICANN meetings and otherwise at the regional forums to let the community know of the work which actually is undergoing.

That's a very brief overview of what we do at the IDN program and we'll continue. Basically, next I'll invite Marc Blanchet to give you an overview of what is being done by the Integration Panel in the context of development of root zone LGR. So, Marc?

MARC BLANCHET:

Thank you, Sarmad. Thank you very much. So, the agenda, my presentation is essentially the scope of our work and activities since the last time we did an update which is ICANN 61 in March in San Juan.

---

The scope of the Integration Panel. The Integration Panel itself is a panel of independent experts. Independent here means that we're not part of the generation panels, for example. Tasked with reviewing proposals presented by generation panels. So, we're reviewing. We're not doing any script proposals ourselves. We're not determining code points or variance by ourselves. This is done by generation panels. And if accepted, integrating them into a consistent set of label generation rules for the root zone.

The decisions of the IP are required to be anonymous. That's our scope work. We also ... The procedure says that we must take into account any public comments submitted in response to the posting of the generation pane's output. Therefore, if there's any comment from the community about the generation panel LGR proposal, then we have to take it into account.

Since seven months, it's been a bit busy. We did these slides three weeks ago, and since then, we have additional things that are not in the screen. So, it looks that we're busy. Well, actually, we're busy because the generation panels are providing LGRs and work.

So, since last March, we had [inaudible] proposal from Myanmar and we also had since three weeks ago the Hebrew one. [inaudible]. We reviewed the draft LGRs of Chinese, Cyrillic, Japanese, Latin, those Neo Brahmi scripts which there's an error on the slide, which is not Bengali but should be written [Bengala]. Devanagari, Gurmukhi, Gujarati, Kannada, Malayalam, Oriya, Tamil, Telegu, and Sinhala. We also reviewed the Myanmar LGR.

---

The Cyrillic has been reviewed and has been submitted after the public comment period. It's been deferred to do concurrent integration with Latin, Green, Armenian all together.

Korean has not been submitted after public comment. So, we're waiting for the next steps on this side.

We also released the MSR-3 shortly after the ICANN 61. That included three CGK and three Latin code points. MSR-4 is actually going to be scheduled soon. So, again, that slide is no more current. We did receive a formal request for some additional Latin code points and also Myanmar code points. So, this is something we should schedule to be done soon, and future MSR whenever it's appropriate for additional code points of scripts.

We plan to produce a new root zone LGR. This would be the third version of the LGR. The target was planned to be around this time. However, it's been delayed. Essentially, the reason is we are depending on the completion and delivery of the script LGRs from the generation panel.

And for the actual next version, it will depend on which set of LGR we are going to receive first, either the Neo Brahmi scripts and Sinhala, CGK or Latin, Cyrillic, Greek, Armenian. Probably the Neo Brahmi ones, but we'll see how they come back to us as a group. So, we'll do only one set at a time, one release of LGR for one set. So, those, for example, three sets will be over three different versions of LGRs just to manage the work load and the number of code points and variants and all the integration needed. That's it for me.

---

SARMAD HUSSAIN: Okay. Thank you.

UNIDENTIFIED MALE: Hi, [inaudible]. I think I misunderstood the thing about Hebrew. What did you say about Hebrew?

MARC BLANCHET: So, part of the process, the first step of the process, is to get generation panel proposal which says what is the plan of the generation panel in terms of the work, what script they will be looking at. People, the individuals involved and all that stuff. So, we do receive the proposal and review it to make sure that the right people are there, the coverage of the script is fine and all that stuff. That's what we just received in the last three weeks.

UNIDENTIFIED MALE: Yeah. Okay. So, you've just received the proposal. Thanks for clarifying. I misunderstood.

MARC BLANCHET: Of the generation panel. No problem.

SARMAD HUSSAIN: Alright. We'll take one more question and then we'll move on.

---

DON HOLLANDER: Thanks. Don Hollander. Marc, what are the implications for delaying any of these things, given that there's no new top-level domains being proposed at the moment?

MARC BLANCHET: So, the root zone LGR is not only applicable to the future TLDs, but also to the current TLDs, because there are current TLDs in the root which also possibly have variant domain names.

So, obviously, as far as use of root zone LGR is concerned, there is possibly [immediate] use. Just to give you an example, we actually have Arabic script already integrated into the root zone LGR and we also had a request from Arabic ccTLDs to delegate, for example, variance of those labels which are already currently delegated.

But, to determine which variance – what are the variants of those ccTLDs or gTLDs and which could possibly be delegated, we would need that root zone LGR definition.

SARMAD HUSSAIN: Okay. Let's move on to the community updates from our generation panel members. We have brief updates from four of these panels. We'll start with the Latin generation panel. We have Mirjana Tasic who is chair of the GP and she'll take the floor.

MIRJANA TASIC: Which way? That way, I suppose. Oh, another way. Okay. Thank you. Generation panel in this composition is working the last two-and-a-half

---

years. This is the brief overview what we have been achieved during these two-and-a-half years.

This is the short agenda. Let's start as soon as possible. We started in Summer 2016 after a new call for volunteers and GP proposal for inclusion and exclusion principles were sent to ... This is just a short history of our panel.

We first created some principles and then we sent it for an informal public review. After that, we collected ... We started preparing our future task. We found using mainly online sources. Some 455 languages. But we choose to not process all of them. Then, we were collecting ... We were verifying code points from these languages. Are they in MSR, those days MSR-2, now it is MSR-3. Then we found some code points not in the current MSR and we proposed them to be included.

Then, we submitted first LGR proposal with ... Something is wrong there. No. During this year, we started to work on variance. We have both types of variance in Latin script. We have in-script variance and cross-script variance with different – with scripts which are familiar with Latin. Now we are focusing mainly last few months on producing in-script variance because we have finished with cross-script variance.

As we see, Latin script is spread all over the world and there are a lot of languages using Latin script. So, that's just one map to see how Latin script is used all over the world.



---

The scope of work of Latin Generation Panel was we have started with maximum MSR number two. Now we are working with MSR-3. Those are subset of code points allowed in IDNA 2008.

You can see on the screen – I won't read all of them – Unicode ranges which has been processed. Then, we started with a list of 455 languages in scope, in total, all sources where ... We used all sources online and I think we didn't use any textbooks or something like this.

Because it is a lot of languages, we selected to process only languages with EGIDS 1-5. First, we processed the languages with EGIDS 1-4. It contains 181 languages. And from those languages, with EGIDS 5, we processed those languages which has more than one million speakers.

About variant analysis. I said at the beginning that we have two kinds of variance. In-script variance and cross-script variance. So, we started first to work on cross-script variance and we identified somehow that those scripts have some connections with Latin scripts, so we looked at Armenian script, Cyrillic script, and Greek script.

Initially, Latin GP had 14 members and three observers, although seven or eight people are working regularly. We had people ... We had diversity in our group. We had people from all over the world. Also, we had different kinds of expertise in the group. So, we think all the necessary expertise is in the group. The workload is very, very big, so we needed so much time to finish it.

Challenges. We had many languages, many code points to process and not enough members to cover the workload. So, we found some kind of

---

solution that we first process the EGIDS 1-4 languages. Then we processed EGIDS 5 languages. Only 29 of them. And we developed some kind of very simple and fast procedure to analyze languages and to verify code points. And we divided workload in two groups, Repertoire Working Group and Variance Working Group. Repertoire Working Group was verifying code points using the languages selected in previous step and Variance Working Group was looking how to organize and find a methodology to process variant and cross-script and in-script situation.

So, that's a short statistics about our working groups. We have finished some of the stuff. This is also ... Those are some numbers which are showing what we have done up to now. Number of processing languages, number of code points found, attested number of new code points found in [inaudible] languages.

At this moment, we have finished ... I mean, we have candidates for cross-script variance with Armenian, Cyrillic, and Greek Script. There is a huge job in front of us to finish within script variance. I thought initially that there won't be any script variance but now they are drawing from minute to minute.

We also submitted the second-round proposal for IP in September 2018. We got comments and suggestions and we have reviewed it during our last Brussels meeting.

So, where we are now. Our plan initially was to finish all the stuff by June this year, but you can't predict always when you start something which is completely new to finish it on time. Also, we have a lack of people. So, we are planning somehow to submit our next proposal until

---

the end of this year or maybe the beginning of the next year because December is always difficult to work. That's all from us. Thank you.

SARMAD HUSSAIN:

So, in the interest of time, we'll keep moving forward and we'll take questions at the end. Our next presentation is on Neo Brahmi Generation Panel. It's going to be presented by Dr. Ajay Data who is the co-chair of NBGP.

AJAY DATA:

Thank you, Sarmad, and good morning, everyone. This is Ajay Data for the records, co-chair for Neo Brahmi Generation Panel. We are a little different than all the panels and a kind of exception. We are doing nine scripts. Generally, it's one script, one panel. So, you can imagine the task which is underneath. So, we are three co-chairs, Mr. Kulkarni and Professor Udaya. And 60-plus volunteers from four countries: India, Nepal, Sri Lank, and Bangladesh.

This is [inaudible] gone. It is a kind of history around that in 2010, country code top-level domain names started happening globally first time in Arabic script and in 2014 India also [inaudible]. And in 2015, the panel got seated but the actual work started in 2017. I took over as co-chair and we started formally working on those scripts.

In 2018, we started first Devanagari, Gujarati, and Gurumukhi as a first set of document which goes to public comment. Public comment, out of nine, eight scripts were done. Two of them are still open [inaudible] number seven. Rest of the scripts, public comment period [isn't] over.

---

We have done various obviously cross-script variance and various scripts within our scope and out of scope. With Sinhala, we have done ... Samad's slide did not cover about our visit to Sri Lanka and Nepal. So, this is also a face-to-face meeting we had for covering our scripts much more better and creating a consensus and [inaudible].

And just to give you an example, for those who are not from generation panel, that this is one script which is [inaudible] Devanagari. [inaudible] that example. Another example is in Gurmukhi. So, I'm going back again. It is Devanagari and it is in Gurmukhi.

So, a normal person will see them absolutely same. There is no [inaudible]. So, let me [inaudible] slide which will tell these two labels. They can create confusion. They can create a phishing attack possibility. And this is the problem which LGR [inaudible] are solving through variance, just to give you a more clear perspective around what is the role of a variant in real life.

We have done eight scripts as I have told. One script is close to completion. We actually had a meeting yesterday with Bangladesh also. So, Bangla script is on the way. Actually, you are not wrong, Marc. This ISO code is Bangali. Some people wanted it to be called Bangla. That's fine.

In 18 months, from active work to finish, probably the fastest work which we could do, and 60 volunteers. So, this was an amazing [inaudible] by volunteers working [almost] ... We had [inaudible] 15 days call. So, that is a fixed date the call is going to happen. So, this helped us to get on the call and plan your week, plan your day

---

accordingly. And everybody followed up individually that what we are supposed to update. So, everybody was on task that 15 days [inaudible] update. And that actually worked very, very well. And we, including for the [inaudible] also followed up through [inaudible] that we can get some help to them and get their work completed.

For those who are not familiar with panels again, you can see how the domain names and IDNs look alike. So, just to give you a perspective how the top-level domain name which is after the dot and before the dot as a second-level domain name look like when there are top-level domain names. Actually, these are all valid domain names which are working already on the dot-[inaudible] domain name which is in India. That's it. Thank you very much. Any questions, we will take questions later.

SARMAD HUSSAIN:

Thank you, Dr. Data. We will keep moving forward. The next presentation is by Harsha Wijayawardhana who is the co-chair for Sinhala Generation Panel.

HARSHA WIJAYAWARDHANA:

Good morning, everyone. First of all, I will cover what we have done up to now. So, there are six topics. I am the co-chair of the Sinhala GP and Dr. Ruvan Weerasinghe is the other co-chair from University of Colombo School of Computing.

---

So, the Sinhala script belongs to the Indo-Arya sub-family. So, we use Sinhala for Sinhala, Pali, and Ecclesiastic language and Sanskrit to write Sinhala is used.

So, these are the members of the panel. I just want to also ... I was going through this and I wanted to also mention that we also have a fast track through the fast track process. There were two ccTLD IDNAs which is dot-lanka and dot-[inaudible] for Tamil. So, I was also involved in that process, so therefore, this is my second time I'm doing this work.

Okay. So, the code points are actually here. MSR-3 repertoire includes two code points for sequences. So, there are some code points, so the characters that we don't use at all right now. They had been used earlier but now it's basically kind of obsolete. So, therefore, we are not included in there.

So, these are in-script variance that we have. If you look at it very carefully, they look very similar. Some of the issues that we have, [inaudible] similar. So, these are some of the issues that we have to actually work on right now for the in-script variance.

So, we also went through the cross-script. These are called confusables right now. Earlier, we thought it was variance but now we have gone in to call it the confusables. So, if you look at that plan, what we found out was Sinhala and Telegu we have a similarities. Then, Kannada and Malayalam. Also, Sinhala and Devanagari, Gujarati. All of these, actually, [inaudible] are very similar looking.

---

Also, we have realized there are some similar-looking confusables [inaudible] Sinhala and Malayalam and Sinhala and Myanmar. These actually could actually create some confusion there.

So, these are the evaluation rules. We don't have that many, except Sinhala has what are called Sannjakas, different from the other Indic scripts. So, these are [nesos, half nesos]. Other than that, mostly it's very similar to Devanagari and the other Indic scripts.

So, this is where we are right now. So, we have now ... Our script has gone now for the public comment and most likely, once it's finished, we should be able to finalize. That's all. Thank you.

SARMAD HUSSAIN:

Thank you, Harsha. The last, but certainly not the least, is the update on Myanmar Generation Panel being presented by the chair of the Myanmar GP, Thin Zar Phyo.

THIN ZAR PHYO:

Good morning. Good morning, everybody. My name is Thin Zar Phyo from Myanmar. This is my short agenda for today. First of all, I would like to introduce my language and my scripts. Myanmar [was formally] known as [Burma] last 20 years ago after one of the government changed. Country names were changed [by this] new government. Nowadays, we call it Myanmar. Myanmar and [Burma] are the same. Burmese is the official national language in Myanmar. We have other several ethnic groups also. The biggest is Shan as you see in my table.

---

Myanmar is [formally] left to right. [inaudible] the same [inaudible] script.

This is our Myanmar GP. We have almost ten persons in my group and we have also three advisories from my GP. As you see, in my GP, there is no professor yet. There is no professor, doctor in my group GP yet. But, all of my members are trying, passionate, trying to promote our language and our script.

We have 87 code points, 50 sequence, and 12 code points including three repertoire created. We have eight in-script variants. The first glyphs are [inaudible] for LGR [inaudible] for them.

In-script variant analysis. First of all, I would like to explain Myanmar and Malayalam. After we decide which character are variants, after we decide, analyze this, Myanmar variance we got two characters only. And Oriya also.

We have confusable character code points. In Myanmar, we have only two characters [inaudible]. From Myanmar and Malayalam, confusable code point as you see in my table. There are only three characters.

As you see, Myanmar character shaped are very [inaudible] circle shaped. Other [inaudible] and other characters are a little [inaudible]. Some scripts are I think [inaudible] or they have a lot of [inaudible].

Confusable code points. We have only four confusable code points. Today, I don't need to explain more about my script because you already are familiar with our Burmese family because of [inaudible]. I will skip some slides for today.



---

This is our evaluation rules. We are a little complex for evaluation just because some Myanmar words are adopted by [Bali]. That is why some characters are fixed in a [word]. So, we have quite complex for evaluation rules, as you see.

This is our future plan. We will outreach for public workshop in [inaudible]. We have to finalize the proposal early 2019. That's all. Thank you for your attention.

SARMAD HUSSAIN: Thank you, Thin Zar, and thank you to all the panelists. Let's open the floor for anymore questions or comments anybody has in the room or online.

UNIDENTIFIED FEMALE: We have one online from [Leanna]. "I would like to know how popular is the IDN domain names in Myanmar. How many domain names do they have?"

THIN ZAR PHYO: It's difficult [inaudible] for IDN. I think around 10 or 15, I think. For second question, how many domains do you in Myanmar? We are still trying that, for IDN in Myanmar.

SARMAD HUSSAIN: So, currently [there are no] second-level domain names offered in Myanmar.

---

THIN ZAR PHYO:                    Yeah. Yes.

UNIDENTIFIED FEMALE:        Okay. There’s another question from [Aziz Modi]. “Thank you, ICANN, for supporting IDN program. My question is very simple, why ICANN doesn’t have their own IDNs.”

SARMAD HUSSAIN:                We’ll have to get back to you on that one. So, that’s actually a good question. We will take that internally and see what we can do.

UNIDENTIFIED MALE:            Just to follow-up, exactly the same feel I had in [inaudible] public forum to the board, and the ball was transferred to David Conrad that he should look at and see if we can do something until Kobe.

SARMAD HUSSAIN:                Okay. So, we’ll connect with David as well. Thank you. Edmon?

EDMON CHUNG:                    Thank you for the updates. As usual, it’s very, very informative and it’s good to see the progress that’s happening. Just offering a little bit of an observation. Personally, obviously, those who know me know that I think this is a very important program. I’d like to offer an observation

---

and also a question to the group. How do we get more people interested in this?

I see that the presentations are very informative. A couple of the questions, like whether ICANN uses it and stuff, really points us to what's the relevance of this for the other parts of the community? I'm hoping ... I'm happy to try to work on this for the next time we do this and to make it a little bit more engaging in terms of how these things are relevant to other parts of the community and what challenges we are presently facing.

I point this a little bit to universal acceptance and how they're being used and those kinds of this. I understand this particular update is for the, I guess, technical progress. However, in order to be able to engage more of the community to be part of this and pay attention to this, it might be useful to have that kind of relevance and engagement. So, just offering this as a thought.

MICHAEL CASADEVALL:

Michael Casadevall, ICANN fellow. Coming in from an outsider perspective, one of the large problems with getting involved with the IDN program is it's very much not accessible to the outside world. Most of the table work generation is done by IANA and is done by a select group of people, especially this doesn't help with people like myself who are mono-lingo or speak other languages very poorly.

I mean, from a technical perspective, there is work that can definitely be done. For example, there have been historical, if not current, issues

---

with TLS certificates and IDL names, specifically to [inaudible] that actually [inaudible]. This conversation came up on the universal acceptance mailing list, but there's no real connection between some of the more technical aspects and from the fellowship.

I mostly came to this room because this is the only outlet I could see that was really talking about IDNs in a public forum that I could find, outside of the working groups.

SARMAD HUSSAIN:

So, thank you for your suggestions. They're really appreciated. I think we'll come back to you offline as well. We'll connect with you to get those suggestions and maybe [inaudible] planning for the Kobe sessions and see how we can integrate that input and perhaps make, reorient these presentations more from a community involvement perspective. So, we'll come back to you. And if anybody else also want to contribute, please do let us know and we'll connect with you as well, as we go into Kobe planning. So, we have a couple of comments [inaudible]. Okay, Dr. Data?

AJAY DATA:

So, I would like to talk about the TLS comment because this problem [inaudible] to be solved, how I still would like to know more about what's not working.

I just want to update everyone here. I attended an ICANN strategy, [inaudible] strategic plan, and number four [inaudible] IDN and universal acceptance as a strategy of ICANN.

---

So, this is [inaudible] top of the agenda and I think this will get a lot of more focus [than what] we are looking for in the days to come. As I heard, our board member from Japan is not here. We know that IDN is going to be focused quite heavily in Kobe and this is going to get much more ... [inaudible] this is wonderful. [inaudible] we can all work together to get more participation. That's exactly what we need. Universal acceptance is obviously one of them, but a lot more. And ccNSO also. I saw some discussion about IDNs and policies around that. I think we need to [get more involvement] for sure. Thank you.

SARMAD HUSSAIN: Any more questions, comments? Edmon?

EDMON CHUNG: I thought there was question over there but I don't know who put up their hand.

UNIDENTIFIED MALE: Not really a question. This is [inaudible], Verisign. We're doing some work and I think, Edmon, you know about this, at IGF with a dynamic coalition on DNS issues that involve universal acceptance. So, if anybody is interested in that. Because it's working in a different forum outside of ICANN. Talking to people that know about ICANN doesn't really do a lot of good but going outside of ICANN is a more interesting forum. So, if anybody is interested in that, just chat with us.

---

EDMON CHUNG:

I'll just jump in. I wanted to respond to – sorry, I missed your name. Michael. It's interesting that the point ... It seems like ... You mentioned a point that it's not just that the ... I used to think that the IDN stuff is very dense and hard to decipher and therefore it's hard for people to come in and contribute. You're giving me a different note here in saying that even if you are able to decipher the information, it's kind of hard to actually contribute. I think that's something that perhaps we need to look at as well, in terms of our information provided, which avenues that new people coming in that are already up to speed could actually contribute to the development.

I don't want to miss that particular point, not just about preparing for next meeting and making it more engaging for the community, but between now and then, there may be some ways to look at how new people can be incorporated into what is going on with all the work that's happening here.

MICHAEL CASADEVALL:

Michael Casadevall, ICANN Fellow. As a follow-up to this, I think if you really want to get people involved, you have to make it clear that you don't need to speak all the languages on this list. There is a mental component here. And even though I only speak English fluently, I understand the Unicode coding issues. I mean, the issues you have simple collision are going to actually affect the ecosystem in a larger way that I don't think has been talked about.

Has anyone seriously ... And this is me coming in new. Has it been talked about the collision issues that happened with SSL certificates

---

and the validation path [inaudible] go through Open SSL? The can of worms can go quite a bit deeper than just DNS because of the way DNS names interact further down the pipeline.

EDMON CHUNG:

That's a good point. We should avoid the names collision thing that happened with the new gTLDs. I don't know whether those of you who are aware of names collision would remember early on how the new gTLDs affected the SSL certificates and then further on, name collision issues. Here is a note about IDNs and potential impact there. I actually don't know. Sarmad, has there been work on that?

I guess, as we do stuff, we do pay attention a little bit to it, but not directly involving that community.

SARMAD HUSSAIN:

Dennis?

DENNIS:

Michael, there is a name collision analysis project public session this afternoon at 3:15 and I invite all of you to join there and ask your questions there. They are now forming the project plan, so the scope of the project plan is very important. So, please, come.

MICHAEL CASADEVALL:

I intend to, although [inaudible] this is out of scope because this specifically relates to SSL certificates. While SSL certificates are

---

encoded in Puniname on a technical level, people are not going to be typing in this and I can easily see pandemonium happen once we get into a world where name collisions [inaudible] similar and DNS ... Well, let's just say a lot of the infrastructure that underpins PKI never expected [inaudible]. So, I'll leave it at that.

SARMAD HUSSAIN:

Thank you for that input. We'll take that on board. We'll follow-up with you as well and also internally, but yes, please do still raise that at the name collision session. Your point is also taken, Edmon, that we need to obviously provide a channel for those who want to engage with IDN program beyond obviously the scripts.

We actually had ... Maybe I have a couple of initial thoughts on it, but I guess we can take that offline and see how we can probably manage that. Don?

DON HOLLANDER:

Just further to what might help Michael is a simple introduction. Why is this important rather than ... And leave it at this is why it's being done for what I will call us mere mortals and while all you super-powered people here who are doing all this detail analysis, this becomes gray to me. I sort of understand what you're doing and I think the examples that were shown here were very useful to me to see.

But maybe, to help ... Why? Why is this important? Why is it important now? So, we have some TLDs currently in non-ASCII scripts who would



---

like to have variance. So, those sort of examples I think would be useful to us mere mortals at this. Thanks.

SARMAD HUSSAIN: Thank you, Don.

UNIDENTIFIED MALE: The question that I have, I think Edmon maybe can enlighten because [CJK] is doing [inaudible] on the IDNs right now. In Sri Lanka, we have an issue where we had the dot-lanka, dot-[inaudible] for some time. The content had not been ... It was very slow-coming. So, do you have any projects where ... There are a couple of views also I have because with IDN is coming, the content would be in different languages and then there is segmentation.

Have you come up with any tools where ... I know there are translation tools, but any tools that are easy to use where you could integrate all the content? Anybody who is working on that right now?

SARMAD HUSSAIN: Yeah. I think the content layer goes a bit out of scope from our work here because we're particularly focused on the domain name system. But anybody else want to ... Edmon?

EDMON CHUNG: Yeah. As Sarmad says, it's kind of difficult for this community to work on the content, but on that particular question, which has been asked

---

many times in the past, I see it as the reverse. Yes, maybe for IDNs to spark new content is immediately difficult. But, if the existing content that is there can utilize [inaudible] IDN to redirect there as a starting point, that will, in itself, help inspire more content in that language. So, I see it in a slightly reverse way. It's hard for the community to create content because content needs to be created somewhat organically.

In the past, we've tried this path, but content is created organically in the community. But, if you show that IDNs as people see it on billboards or whatever actually jumps to that particular language site or content, that inspires more usage and therefore inspires more content. I kind of see it in that sense rather than us trying to translate everything into a particular [inaudible].

SARMAD HUSSAIN:

Okay. So, we actually have one minute left. We have a comment or question online. Let's take that and then we'll come back to this room.

UNIDENTIFIED FEMALE:

It's a question from [Aziz Modi]. "How can we get more and more people adopting IDN as there are not much use case? What is more ICANN is doing to promote it?" That's the question.

SARMAD HUSSAIN:

So, this would potentially go into the universal acceptance domain. Don, do you want to respond to that at all?

DON HOLLANDER:

I'm happy to respond to it and I think it's a cheeky question from [Ashish Modi] who is a universal acceptance ambassador. So, one of the things that ICANN is doing is supporting the universal acceptance steering group, which in turn is supporting universal acceptance ambassadors to go out into their local community and professional communities to raise awareness of the issue for the developer community.

And just for people who don't know what universal acceptance is, it's the idea that all domain names work in all applications. So, if you have an e-mail address – in my case, [inaudible], and I want to use that to register with my tax department and my tax department says, “No, that's not a valid e-mail address,” then that's a universal acceptance problem. I did try that and I did ask whether I still had to pay taxes now and they said they would put something in the post. So, they always have an answer.

SARMAD HUSSAIN:

Thank you, Don, and thank you, [Ashish]. Probably the same question back to you. So, out of time. Last thing I would like to share with you is that all the information we've presented here and more is available on ICANN website. So, a quick way to get there is to type [icann.org/IDN](http://icann.org/IDN) and you'll get to all the wonderful information around IDNs.

If you want to connect with us, please write individually to us or you can write an e-mail to [idnprogram@icann.org](mailto:idnprogram@icann.org). So, thank you all very much for attending this early morning session and we'll close the session. Thank you.

**[END OF TRANSCRIPTION]**