ASIA-PACIFIC REGION INTERNET GOVERNANCE FORUM
TAIPEI 2016
A NEW INTERNET ERA

28 JULY 2016
ROOM 401
116:00
WS 81
INTERNATIONALIZED DOMAIN NAMES
AND IMPLICATIONS FOR THE NEXT BILLION

***

This is being provided in a rough-draft format. Communication Access Realtime Translation (CART) is provided in Order to facilitate communication accessibility and may not be a totally verbatim record of the proceedings.

***

>> All right. I think it's 4:00 so let's start okay. Thank you, everyone for coming in on time. Usually this topic, we've done this a few times at APrIGF. Usually it is a small group. I'm glad that we have more than a small group. Especially if you joined the earlier session that was pretty fun, I did a little advertising for this session, it's really talking about how the Domain Name System is changing and expanding. Before I get ahead of myself, my name is Jia Rong, I cover the Asia-Pacific reasonable or for ICANN. Now I shall introduce our speakers for today. There after then I'll run through the order of proceedings and our programme and what we are going to discuss.

So on my left here is Rinalia, she is also a member of the Arabic Generation Panel. Later on you will understand what the Generation Panel means.

On my left here is Hiro Hotta. He is Director of Japan Registry Services, JPRS. And to my extreme left is Dr. Kenny Wang, Taiwan Network Centre, TWNIC. On my right here is Kangana Ranaut, Executive Director for the Taiwan Electronics Development Agency. And joining us remotely is Mr. Rapid-san of Post Communications Telecommunications also known as NIPTICT.

The speakers just introduce yourselves a little bit so people know where you're from and why you're looking at the various language Generation Panels. So that it helps the newcomers in the room.

Now, as I mentioned, this session is about Internationalized Domain Names. And really the key here is that the Domain Name System is expanding and changing and you are no longer restricted to having to use English characters. The technical term for English characters for domain names is called ASCII: So you won't be restricted to using ASCII only but you can use other scripts. And this doesn't happen in a vacuum because even though this programme called the IGM programme is initiated by ICANN, ICANN staff alone cannot manage this. Why? Because I speak three languages, but if you ask me can I contribute to the Arabic panel? There's no way I can contribute to that panel.

We need to have community members from that language community come together and decide what characters or letterings can become a domain name. And from there we take things forward to fully implemented and eventually this would help entire communities come online because right now I use one reason, for example, my father would doesn't speak English, he said the Internet is very good. Oh it is very interesting thing. You can see a lot of things on it. But there's one serious problem is it. What's the problem? You have to know how to spell. And I realized that there was a serious problem. A but being able to type in his own language would empower him to come online. That's what we're all hoping to do through the various label Generation Panels that either the work is completed or the work is in progress.

So I shall now hand over the time to Rinalia who will give an introduction of the IGM programme and why it is important. So I hand it over to you, please.

>> RINALIA ABDUL RAHIM: Hello, everyone. First of all an apology, I have a bit of a cold and it's affecting my voice so I don't sound like my normal self and in case the volume goes down, please raise your hand and I'll try to exert myself further. I'd like to start by setting a bit of context first in terms of why IDNs matter for our world.

I'm sure that most of you know this already, we have in the world about 3 billion Internet users. Nearly half of the users are from our region, the Asia-Pacific region. However, the Asian Internet penetration is still considered to be low. And we expect the majority of the next billion Internet users to be from Asia.

Next slide. And here is the context for Internationalized Domain Names or IDNs. So we live in a multilingual world with various scripts. There are 7,000 living languages in the world and about 50 percent spoken in the Asia-Pacific region itself. Now, the number of writing systems and languages is smaller than number of languages because some languages follow oral tradition; they are not written

down. And because some languages actually share the use of a script.
For example, the language of Chinese, Korean and Japanese share the
use of a script called the Han script. The Arabic script is shared
by more than 50 languages across Africa, the Middle East and Asia.
     IDNs is meant to go across scripts and possible at all levels,
what I mean is third level, top level of domain names after the
engineering taskforce developed the IDN NA protocol in 2003. We have
them to thank for it.   They have been supported and coordinated through
ICANN through the fast track process for country code TLD, what we
normally refer to as the Top Level Domain programme in 2013, what
we refer to as the IDNgTLD programme. Prior to this, all TLDs have
been in the form of ASCII, which is a subset of the Latin script as
mentioned earlier.
     So International Domain Names, or IDNs are names that can be
represented in scripts or writing systems in addition to the Latin
script. By enabling IDNs, local communities can come online by their
own languages and scripts. And here I think it would be useful to
briefly why domain names matter to international users. Forgive me
if it's basic but I think we have new participants to the IGF and
I think it's worthwhile to run through it.
     Domains are useful when you use your email. After the @ sign
there are domain name labels. That's where they are relevant. They
are also used when users access content by web browsers, we typically
type in the URL. They are accessed when users use mobile apps.
Sometimes you don't see it. It is hidden by the application developer
but it's used. They are also used for signing up for online services
and for authentication and by application providers to serve users.
So they're actually quite important.
     Without domain names, using the Internet becomes almost
impossible unless you can remember long sequences of numbers.
Without IDs that go on with various scripts and keyboards will not
fully benefit from having access to the Internet. And even if you
have multiple keyboards, some people have difficulties toggling
between Latin script and non-Latin script keyboards. Next, please.
     So earlier we talked about ASCII and IDNs. Here's the form.
Normally ASCII www.cafe.com with third level domain being www, cafe
being second level and com being Top Level Domain. On the right you
will see a script, beautiful script in Thai. And IDN second level
domain and Top Level Domain. What does that mean? Can you say?
Sample or test or something like that, yes. So this is just to show
you the difference between ASCII and IDN Top Level Domains. Next?
     So I will just you an overview of the IDN programme in ICANN.
There is a lot going on. I won't run through this one by one. I will
give you brief highlights. ICANN's IDN programme promotes on
focusing on engaging script communities around the world to develop
rules and these rules are used to define valid and unique top level
domains. So on this slide you will see a sub-bullet that says label

generation rules. Remember that. That's where the communities come together to develop the rules.

In addition, ICANN continues to support the evaluation and Delegation of IDN Top Level Domain labels which remain in Generic Top Level Domain in diverse scripts. ICANN is also undertaking projects to promote better user experience, including developing guidelines and reference IDN tables for the second level and working on general accessibility of domain names through the community-led universal acceptance initiative. And here you see a map of the world that shows you the IDN top level domains. These are examples of the IDN CCLTD labels that have been successful evaluated by ICANN so far in the Internet root zone that allows users around the globe to access scripts in their languages. Here I'd like to draw your attention to India where there are more than 10 scripts representing their country's name and you can actually apply for this actually in your country if you have use of multiple scripts for your population.

To pave the way for different scripts to be used in top level domains, communities using each script or writing system must come together to decide what characters or letters can be used and the rules for them. I think it's fair to say without these rules, the TLDs that have varying forms cannot exist and that's going to be a problem for users and we need the community to do the work in order for the labels to be available.

Next.

Here is an example of the work that I have done together with the Arabic script community just to show you what variants look like. You see the letter K in Arabic in two different colors circled? The Arabic red and the green have the same shape when they are joined. And it's a problem because you cannot tell what the original character is. Rules are needed to vary these when they can be used or when they can not be used. Similarly rules need to determine the number of variants which can be used. Sometimes languages that have a lot of variants have too many variants. And not all of them can be allowed. So there is work to be done by language and script communities. Next?

Here we come back to the root zone label generation rules, development status and this is the project that brings script community together. ICANN supports this work strongly by working with local communities to come together to establish what we call Generation Panels. And we call them Generation Panels because the group generates the rules that are used to form IDN top level domains. This work is on going. We are continuously encouraging different languages communities and script communities to come together because we know the world is diverse there are lots of scripts and if you want them to be available for domain names, the work needs to be done. So this highlights more than 12 panels that are working on various writing. For the most recently established in the past year. And the Khmer panel has finished its work.

Here's a brief label generations rules procedure. You have the different scripts working to generate their results. But because we have only one single Internet root zone, the label generation rules for all the scripts must be integrated into a single root zone label generation Rule set. If you are familiar with Lord of the Rings, there is one ring to Rule them all, this is it. This means there will be some coordination challenges. And we need communities of scripts and different scripts to work together to make sure there are no collisions or conflicts. These are my own words so don't take it too seriously.

In terms of coordination challenges, some writing systems share the same script. So there is a need to coordinate for variants such as Chinese, Japanese and Korean using the Han script.

And some writing systems have cross-script variant letters so there's a need to coordinator identify them. These are letters that look similar to each other and may be confusing to users. For example, some letters among Khmer Lao and Thai. And some among Armenian, Syriac, Greek and Latin look the same as other scripts in the set. And speakers on the panel will talk about scripts and will provide cross community coordination work which can be quite challenging and you will get a better idea of what the work entails. Thank you.

>> MODERATOR: Thank you so much, Rinalia. What I'll hand the time over to the other speakers. Each speaker will spend about 10 minutes updating about the work of their scripts Generation Panel. Then after so I hand the time to hero first after that Kenny, then after that both speakers will talk about the cross community workforce, Chinese, Japanese and Korean. But I'll make sure I will be your timekeeper. Therefore then we will hold the time to Rapid who will join us remotely and they will talk about their own label generation followed by the Khmer Lao Thai community work. So we will do Japanese and Chinese work and hand it to Hiro, please. Thank you.

>> HIRO HOTTA: Hi, yes. My name is Hiro Hotta from JPRS. JPRS is registry for .jp. I'm the Chair of Japanese Generation Panel, which is one of the Generation Panels, language and Generation Panels of the area introduced. Today I'll talk about the brief activity history and the main issues for us.

So, this is our list of JGP meetings in the day-to-day events. So we've already have two years experience for meeting. Japanese Top Level Domain should be ruled. These are the members of the Japanese Generation Panel. And I'm a Chair and Akinori, he is a Vice Chair. And he will be succeeding as a ICANN board member. And we have a few professors from Internet world and from trademark and domain name experts. And one is from gTLD registry expert and character code, especially the unicode. And Yoshiro is IDN experts. So we have a variety of expertise in our panel.

So this is one of the requirements to have the barriers expertise panel.

So as mentioned, Japanese, Chinese and -- share a lot of

characters in common. So these are called Han characters in Chinese.
And we call them Kanji and Korean call them Hanja.
    As this picture shows, the Han characters are shared by languages.
So this is one of the big issues to be solved. So I will explain this
later. And Kenny here is a Chair, a co-chair of Chinese panel. So
we overtalk, fight, right?
    [Laughter]
    No. We are friends. Okay.
    As this figure shows, that Chinese Generation Panel needs to
define the Rule for Han character. Only Han character. But Japanese
Generation Panel has to define the rules for Han character
and -- because in Japanese language. We use hankatagana kiragana(?).
All at the same time in a document. And for Korean, Han and hang all,
they use in a mixed way.
    So three have to collaborate to define at least the Han character
rules.
    So this is the framework for CJK LGR integration. LGR mean
language generation Rule. For Han characters. And this procedure
is already agreed by all three Generation Panels. This means that
the Chinese GP, Chinese Generation Panel defines their Rule first.
And Japanese defines our Rule first. And Korean theirs first. And
merge them. And if it can be merged in a good way, you don't have
to care about that, about the result. But it is not easy to define.
I will explain this later. So we have to communicate, consult whether
this character should be usable as a TLD or not or this character
and that character should be the variant or not. So we have to think
about that. I will explain this later. And Kenny will explain that
later, as well.
    And these are the activities of JGP. We established -- formally
established in March 2015 and detailed task description of our panel
is almost done because we are still working. And new issues are coming
up. Not every day but new issues are coming up. So this is always
on the way.
    And some issues or text may be modified on the development through
discussion with ICANN or integration panel and through discussion
with Chinese GP and Korean GP. And through investigation inside
JGP and through consultation with the community, with Japanese
community.
    And development of Japanese LGR. So CJK LGR integration
procedure was developed, as you saw in the figure, agreed and revised
by CGP, JGP and KGP as a framework. We agreed that as a base.
    As an input to the procedure, preliminarily Japanese LGR which
is called LGR-alpha was developed.
    And this is a discussion status for Japanese LGR. Scopes of the
character codes. As I said, can gee, here an Ghana cat an can an should
be used. And it is fortunate that we have Japanese industry standard
for can gee, so we borrowed that definition. And variants, so Chinese

language or Chinese script has a lot of variants. So this means that this character and that character means the same. And they should be regarded as the same character. Or regarded as one character.

But Japanese LGR define no variants, so that means that all the characters in Japanese language are different, are separate, are independent. And integrated Japanese LGR or Chinese and Korean. This means that if Chinese language definition defines some variants, we import that definition, the variant definition. This means that if Chinese language definition defines this character and that character as the same, Japanese language do import that definition.

And the other way I will skip this because it is not easy to explain.

And overview of Japanese LGR. Rap an tour -- Repertoire. More than 6,000 characters. And here an, and can an, we have about 80. In total 6,500 characters can be used as a TLD label. And no variance, as I said.

And we do consultation with our community. Of course, we post our activities on JGP web site public website. And presentation and discussion about TLD strings with Japanese trademark association. And with various stakeholdering in Internet Governance conference which is a kind of national IGF events.

And further communication is planned including public comments within Japanese community and global community through CJK coordination and ICANN process.

And these days, ICANN is thinking about the second level domain labels LGR Rule, second level, something has to have a kind of Rule. Submission of public comments to guidelines.

Then key outstanding issues for CJK coordination. So I bring up two issues. One is the definition of variants must be coordinated among CJK., for example, these two characters are variants of each other in Chinese, although they are totally different characters in Korean and Japanese language. So then for the TLD string is just one. So there's no language context in the TLD string. So users cannot identify that, whether that string is Chinese or Japanese or Korean. So there should be one Rule. But Chinese language wants to have these characters as a variants. But Japanese and Korean language wants to have them separate characters, or independent characters.

In this case, more than 200 groups of Chinese variants are not acceptable in Korean context, so this is a problem. So we have to coordinate and discuss and needs to have one solution.

I think Kenny discuss about this, right?

[Laughter]

So we decide that we will import anything. So it's easy for me.

And second one of the location of all allocatable labels., for example, any combination of characters is allowed in Japanese strings, as I said. And it may make the number of variant strings very huge, as many variability groups are defined by importing Chinese variants.

For example, this is KO university. Has three variant strings.
Among four characters, among them, two characters has two variants
each. So two times two is four. Four variability strings. So if
this string is long, long, long, the number of variant strings may
become huge number. So actually in Japanese context, these four
strings we use all of them. But -- to reduce the number of labels.
Because they don't want to have a huge number of reserved variant
strings which will belong to one registrant of a domain name.
     And JGP is trying to solve it by defining a new for Japanese
and proposing a new application/evaluation procedure. We are talking
about this new procedure in IP. I think this is the end. Thank you.
     >> MR.JIA RONG LOW: 10 minutes. More than that. Slightly more
than 10 minutes but you already covered the cross-community
coordination which will then take some time away for that so that
helps a lot. Thank you so much and I'll hand the time over to Kenny.
     >> KENNY HUANG: Okay. My name is Kenny Huang, co-chair of
Chinese language panel and APNIC. So can you move to the next slide?
     Okay. This slide introduced evolution of Chinese character in
Korean form such as arikopong and we have different stages. So the
character actually evolved from from time to time. And right now we
try to understand how sort of Chinese characters we need to put it
in our repertoire to make sure that the characters belong to our
digital assets.
     Next slide. Unless Chinese script not only evolve but also
spread to other regions other countries such as spread to Japan and
spread to Korea. And I just realized that Chinese characters spread
to Korea starting from the second Century and spread to Japan starting
from fifth Century. Okay, any, that's history.
     Next slide. Countries and regions using Chinese script.
Maturity of country actually using simplified Chinese character. And
Taiwan and also very selective area using traditional Chinese scripts.
     Next slide. Already mentioned a little bit, the Han character
not only used by Chinese only but also used by Korean and used by
Japanese. So there are significant amount of characters are overlap
among Chinese, Japanese and Korean. That's why if we're registered
under second level domain name and you can have your own CCTL Rule
to regulate the policies. But if we want to register the Han character
in root zone because the root zone is competing area, we only allow
register one character. And who can classify? Belong to Chinese or
Japanese or Korean. That's why we're starting to work on this
programme and eventually try to find out if there's any solution that
goes on. Next slide.
     Okay. The variant solution in different scripts, for example,
like a CDNC and Chinese domain name consortium that's the major
organisation working on Chinese language and technology and promote
Chinese domain name and registration. And we have different kind of
standard published including the first one I344 and C3343 and we have

ITF 4713 and regulate how we register Chinese character into different forms such as how we register traditional Chinese character and how we register simplified Chinese character. How these two characters can be effective at the same time. That's that kind of concept was starting from 15 or 10 years ago.

And then for JPIs such as already mentioned, they don't have variant issue although they are using Chinese character as Japanese character. But you don't have that same concern. But can have different meaning. So you don't have that kind of variant issues.

Korean basically no such variant issue but concern some core point of Han character that with cause some kind of confusion from Korean community point of view.

I suppose even from Japanese community point of view, they have a similar concern, but they are willing to collaborate. So that takes away that kind of concern. So next slide?

Okay. The coordination principle was it is not created buy the Generation Panel. It is created buy the coordination panel. The principle you have two type of cooperating principle. Once you register some certain character and the certain character you need to identify whether it will be allocate in DNS root long or you will block in the DNS root zone. Because the character different kind of variant. So we need to define how DNS system dealing with that kind. Do we need to allocate in the root zone or block in the root zone? Next slide, please?

Probably have already speak. For Generation Panel and Japanese panel, how will we indicate together? We have one table from the one character love in Chinese and simplify love and traditional love depends on who you're talking to.

[Laughter]

And we also have love from Japanese community JGP. And because we have different kind for the character and also different variant, when you put it together, you combine what be generated, the right-hand side. That will be integrated in the table on the right-hand side. Okay. Next slide.

That is also Rule created by integration panel. Our task will be complete once we finish repertoire, Chinese generation repertoire and we generate the variant set and go through integration panel consultation, we will evaluation and LGR. So eventually finally we can generate a CJK generation Rule, put it in the Rule zone to the DNS registration.

Next slide. So far we are already giving from our proposal, we propose an almost near 20,000 character to propose level belong to the range of Chinese character. And if the character user allow to register the specific character. Because of the first new gTLD so we did not apply this Rule. But for the next round if we can come out with integrated solution, that will be effectively for the next open. Next slide.

Okay. Here comes more detail about how we are going to solve the kind of conflict because if we have Chinese character, Japanese character and Korean character and there's a number of overlap such as JGP and overlap between KGP is around 6,000. And overlap between Chinese character and KGP is 4,147 but right now go through the communication and consultation effort, right now we dealing with most critical characters only between CGP and KGP. The most critical reduce down to 200 to 600 character along after two-year consultation, debating, fighting. Okay, it's done. Next slide.

It's very technical but anyway in official registration form would be something like this. You have first is original registration form. Second simplified to traditional to simplify automatically activate. The third would be the Chinese character.

Already introduced so I'm not going to repeat the discussion. Next slide.

So, what's next? The next year coordinate between the CJK community and we will adopt guiding principles to seeking rough consensus as mentioned. Right now going through two-year effort, we already reduce a second amount of characters down to 260 character overlap that we are still debating how to deal with the 200 character. And we suppose we going through technical feasibility study and also put it in the consideration in terms of cultural or social consideration because some character in different community they have very, very sincere concern. So we need to evaluate that kind of issue, as well. So we are now working from computer science point of view also from social science point of view. And eventually a harmonized solution is anticipated before the end of 2016. I wish Hotto-San can agree with the timeframe.

>> MR.JIA RONG LOW: Thank you very much. Based upon the programme we are supposed to have cross community work which they already shared. But maybe, let's just spend five minutes maybe from each, both Hiro and Kenny, if you could share just one comment like either based upon what you think the progress is and do you think 2016 would be the good timing. That's one question.

Another question could be like what do you see going forward after this work is done?

>> KENNY HUANG: I thought Hotto-son would do that. Right now reduce the size of overlays is most severe overlap especially the kind of pressure not only members' point of view but also the pressure from the community because we have language expert community. And they quite insist some characters should remain the Rule they prefer. Such as they are also using the character in their own society.

So we're going through right now trying to do some sort of statistic going through based on the current registration from Korean DNS database and also doing statistics from other database and Hong Kong database, trying to find out the frequency of the characters being used for the registry domain name and also the other progress,

working on the newspaper statistic, what kind of characters been used. We calculate the frequency used in different media an current to current most fashionable language. So once we come out with frequency statistic, we try to orient with Korean or Japanese or Chinese community. So we try to compare the different frequency. And for high ranking, for example, if a character is higher ranking use in Chinese community, probably low level agree to use Chinese Rule. And if some character were more higher frequency use in Korean community, probably we will just adopt the Rule. Highly used in Korean community.

So basically we are heading to this approach. And because the Rule is quite simple. So we believe after a few weeks' computation, the results should come out pretty quickly. So I believe some significant outcome would be generated in September. And so we probably will have a big announcement from end of this year if everything goes through. That's my suggestion.

>> HIRO HOTTA: We have issues because as I said, we are requested from IP that we should reduce the number of allocatable labels. For example, we already have IDN.jp registration around 120,000 registrations. And observing the registration, there are many labels registered. And which have a lot of variant labels if we import the Chinese variant definition. So we do have to reduce that. So we haven't decided how to reduce that. So we need to think about that seriously from now on. Yes. And that's the biggest thing.

>> RINALIA ABDUL RAHIM: Thank you, Hiro. I just wanted to clarify something that Hiro meant. He said the IP requested the Japanese community to reduce the number of their variants. I just wanted to mention that IP here does not stand for Internet Protocol. It stands for the Integration Panel. This is a small group of experts that actually oversee all the review all the submissions from the community script panels and actually say, yes, this is okay or no, this is not okay, go back and do some more work. So that's who he means by IP.

>> MR. JIA RONG LOW: Okay. Thank you very much. In the interest of time we will move on. So we'll invite Rapid to join me. Can he control his own slides or do I have to do it from my side. The remote speaker? Thank you. So Rapid over to you.

>> Thank you very much for joining us. Please go ahead we can hear you loud and clear.

>> RAPID: Good afternoon, everybody. My name is Rapid Sun from National Institute of post telecom and ICT from ministry of post telecom communication and I'm also member of the Khmer Generation Panel in the root zone also. And next slide, please? So I'm going to present you about our progress of the proposal. And one is about the introduction of our Khmer language and about our Khmer script April member of Khmer Generation Panel and the project schedule that we had conclude and our methodology that we are working with small group with technical person and also the feedback. Next slide, please.

So for the Khmer language it was in the early seventh centuries. And
our script is based on an Abugida type originating in the south India.
It borrows some word from Sanskrit and Pali. And also borrowed and
found in Thai and Lao and Kuay, Stieng Samre Cham and others. The
official language Cambodia 15 million people and also some
1.3 million people in Thailand and Vietnam.

Next slide, please. So our script is within the ISO 15924. And
we are writing from left to right. And we have a lot of script, about
146 characters, include the consonant, the vowel and dependent
consonant and subconsonant also. And from use of FF we also have some
special character from U190 to U19 FF. Next slide, please?

So here is the member of our Khmer Generation Panel. We have
Dr. Sopheap he's supposed to be a speaker for today, but unfortunately
he is not well for the past few days with a big flu. So I'm also a
member of the Generation Panel and also acting as secretary. We also
have a member from telecom Cambodia who also ran the DNS, the .ka Top
Level Domain and some member from the e government department and
some ex pier I don't know unicode and some from the linguistic also,
the two lecturer from the royal University of --

Next slide, please.

So here is our project that we had been working so far that we
first we developed the principle that we submit to the ICANN. And
we also work on the code points and the variants, the variant between
the Khmer language and the Thai and Khmer and Burmese and also work
on the label Rule, the whole label Rule. And we also working to host
the public consultation after we have the first draft of the proposal.
And also last from the proposal, we develop the XML. And then we submit
to ICANN for public comment. Next slide, please?

So our methodology that because of all of the members in our
Generation Panel, all of them, some is from linguistics and some from
the technical DNS or some from the e-government department. So then
we decide to work to build a small Working Group, especially because
these identify the consonant, the vowel, the independent vowel or
other script. Mainly related to linguistics. So we build a small
team within the linguists mainly and then after we have the first
proposal, then we call for the meeting with other members of the group,
of the Generation Panel that we present them the proposal. And then
ask for their comment or their feedback regarding on this. And we
also develop the variant that in Khmer language we also have some
variant especially some consonant when we form a label, then some
consonant is local the same. So then we also identify some variant
in Khmer language. And next slide?

And also we also develop some variant between the Khmer and
Thai and Lao and Burmese. There are about 17 characters that we think
is a variant. And then so have a discussion with the Thai Generation
Panel on this. And we have developed the label Rule that we develop.
For example, some special sign is not follow, can not be followed

with some consonant. Their vowel cannot be independent, for example, and then based on these label rules, then we develop the XML file.

Next slide, please. So this is the feedback. We have our proposal from 0 to 1, about 10 editions proposal. And we developed and get the feedback within the member, the Khmer Generation Panel. And also the proposal with the integration panel that we also get the feedback from them. And after the final proposal, we have the public workshop which we invite for larger community, from the academy, from the civil society, from NGO and from the users, that they can give the feedback of our proposal.

And the proposal we submit to ICANN and then due by the integration panel and also get the public comment and it is end by 28 May 2016. And then we got a few comments, one from Thai Generation Panel and one also from integration panel. With the Thai Generation Panel, we had discussed about the variant. And then after the discussion, there is no variant between the Khmer and Thai. There are some regarding in term of the consonant. Some consonants look similar, but when it's formed as a label, then it is not considered as a variant. This is our conclusion after we had the discussion with the Thai Generation Panel.

So next slide. Okay. I come to the end of my presentation. And then if there's any questions, I'm happy to answer.

>> MR.JIA RONG LOW: Thank you so much, Rapid. We'll leave the questions to the end. We will hand over to talk about the Thai panel. And I think Rapid raised a really good point about the challenge in cross-community work for the Chinese, Japanese, Korean one is very different from the one for Lao Thai Khmer, but I'll leave more details to --

>> I will try to simplify a lot of people. We're getting lost here what we're all talking about. It's about language. In general what I speak on capacity of the Chair of the LGF panel. I don't know much the linguistic person at all. But we have the members coming from university, blah, blah, blah.

When we start to work, we just wonder what it is. And there are experts sitting in the room. Edmon. He's the one talking, shouting about the email in local language. And that inspire me of what it is. Why we have to care about having an email in local language, the address and all this stuff. And we start to study. Thailand is coming late in this. But we try to look into holistic approach other than focus only on the LGRs alone. Because it is only the root zone. Which only engineers working. But we try to look into the broader sense. So our panel's working. And then what we found is that Thailand have 35 scripts, I don't even know who they are. 35 scripts that we have to take care of. Including Bali in India is surprising us because it is spoken language not written form. But because of religious, the Thai scripts was used to construct the Bali language in religious. And that is something. That's what we get of running the LGR. We

found out two things. One is that it is easy for us because nobody use our characters, thank God.

[Laughter]

It is too complex. And you don't need to know anything in detail. Let's assume that there is one language that you have. Let's say three level of the language. You have upper and you have lower. And one word comprise of the beginning words, ending, upper lower, the way you type on your keyboards to construct one words. Because you can start by anything. A lower first, the first character, lower and later characters or upper one. So we have four positions in your language. This become a disaster. Because when I was a kid before the government, you have to go to typewriting. Typewriter class examination. You have to know how to put all this I don't know what they're called. The typing, the minutes, the words per minute, that kind of stuff. You're not old enough. You have to go to the typewriter exam. At least you have to pass 60 words per minute to qualify as a government official. And the way they put the upper and lower is fixed. You have to put upper first because mechanic will not work. Because of the typewriter work like that.

Then come to the Internet. The kids don't even know the keyboards. They type whatever they like. Then we put the technology people in Googles and Microsoft sitting there and they just say it's waste of time. And we demonstrate that it's not working at all. It's depend on the typing of these four letters and that's what we learned. That is a big asset we get of these LGIs. So with the root zone, but it's the local Han text that Thailand found that we know now to go and fix that problem. We get this from the work that we've done. So we know exactly that this is a problem. There are a lot of problems used as example to doing this exercise is help locally to talk to the technology provider because, especially to mobile phone. This will lead to -- they will start to look or this will lead to -- because the combination of the character, user look the same but we have what's called sequencing problem. And that's what we learn out of this work in the LGR. And working with in Marrakech, we also have the networking. If you look into -- Hiro mentioned a bit about a second level that public comment already close now in February, we all do see that the problem of cross scripting of our language in Thai and Lao will happen more on the second level because that is not in the hand of IANA or because of the way that the reseller is putting the second level if they're not controlling this, this might lead to the fitting site. We already demonstrated that we can put the Lao in combination with Thai. Normally the user is using this kind of automatic QR codes or whatever, it's automatically input the URL for you and the user will never know that they're entering to the wrong site. And those are the things that we worked. That is about our LGR in common. You don't need to know the Thai language. But there are a lot of exercise that will be coming out late a bit because we try to look at holistic approach.

We try to enforce the email to international. It is also related to this work in LGR. How you put the email in Thai language with the Top Level Domain. And for the second level domain, we did not see the urgency because Thai only have one ID. And it's different case than in Japan, Korea and China. They have a lot of idea. No Arabic. Only one idea and Lao do not have. So we are not in the groups. If you know the detail, we are not in the groups to work as either group 1 or 2. So we are not working in the second level domain yet because there is no urgency there. It should not be any problem. But one last thing is with the 35 language I think you know about this story where we kick off, when you kick off me for the LGR. Remember kick off. You kick me off.

     [Laughter]

     Yeah, we've -- ICANN has come to the point of kickoff meeting, kickoff event in Thailand and we found out that one of the language in Thailand is speaking by 3.5 million people. Then I visit the location and I found that everywhere they use to promote that local character, local string. All the hotel try to put that. And I have a speech to invite them to be in. And a lot of people is really crying, the old people, because they thought that in the net well destroy all the language. Because then -- they never thought that would be possible. And that is important for IDN and I think the Internet should not destroy who they are because they use that language. 3.5 million people are using it. They couldn't have a sign in front of the office, in front of the hotel. But they never thought they could have an Internet website that can put the characters in. It's not difficult for all engineers. But that, another side story of what you can do with LGR. You can bring this thing in. So don't try to look around. You may find something the same as what I found. Even though you're using Roman and try to look around at whether there is some minority, some people that still use other scripts even though you using the Romanisation already. But I be sure that we're going to find out because that language that I mentioned is also used by Burmese. It's a cross scripting between Burmese and Thai, in the northern Thai regions. So I do hope that we have energy enough next year to start that LGR panel for northern Thai. We also have the same we call Hans, but in the other, in other format like Burmese character that is what I want to share with you that which LGR brings it to preserve because the language is a culture, it's arts, it's people that you're dealing with. So it's important that ICANN outreach and it's important that the local community aware that it's there and we need to work together.

     Okay, thank you.

     >> MR.JIA RONG LOW: I think that's a good segue for me to say as conclusions before we go into the Q & A. I think the second point that was made was with regard to we think that we know all the languages there is existed. And we think that it might be represented somewhere.

But what's very clear is that there are many languages around the world that are becoming extinct. And some of them were a result of us relying on computers.

From ICANN what we hope to do is that we can capture the script and allow it to be used on the Top Level Domain. And this is where you can volunteer your scripts or the script that you know may be some minority community from your country might be using it, but they are not actually being used on the Internet. So if you know that, please write to us on this email, on the first part. IDN TLDs @ICANN.org. We will show you how to set up the labour Generation Panel, who and which kind of experts should join that panel. We can give you some advice. And then we help you in terms of the scope, how do you start the panel and so on. So we can do the support work and we hope we can work with you. So if you know about this, please help us spread the word, as well, so we can preserve some of the important scripts that are around and so that we can actually make the Internet inclusiveness and diversity on the Internet.

The other events where you can get involved with regard to IDN work is, well, always you can review all the work through public comments. The various speakers share their work after the completed, they have to put it up for public comment. So if you speak that particular language or you use that particular script, you can always go in to see the work in the proposals. And you can comment on it. And your comments will always be valued. So we hope that you can give your public comments to that. And then the third one is keep yourself updated to various updates. So you can join the mailing list which is VIP@ICANN.org. I find this very interesting. To get updates specifically on the IDN programme. So these are the three areas.

Now, let's move to Q & A. Any questions or any thoughts on what the speakers have presented or thoughts about while you're mulling over, I shall just start with one which I think is very interesting because I didn't really think about it until mentioned on the whole issue of phishing.

You know, sometimes you get this email which looks like it's your friend, but actually it's somebody else because, for example, at ICANN, it can be someone who poses as me, but instead of like -- with threeNings.org -- and pretends to be me and I give them my information. "Jia Rong asked for your birthday." And then you get my birthday is this and then you get phished.

And the situation of IDNs where mentioned, if somebody decided to change scripts. So use two letterings for Lao and suddenly switch two letters for Thai. The computer reads it as code so it cannot tell the difference. But the computer sees the two things as different. So Lao script and Thai scripts but the user would read it as the same thing. So actually they are two different addresses but this actually has more implications for the user than anything else because it creates that issue of security. And I realise it is very good point

that's been flagged. And we will need to come up with a solution to think about how to resolve this. So that's just one of the points I think is very relevant. I'm not sure whether our speakers have any comments with regard to this issue.

Okay. What about any questions from the floor? We are open to that. Please.

>> My name is Rakir. I'm from Pakistan. I have a question about this IGN programme. I mean there are certain languages which are using the same script. For example, Arabic script or Persian or other language of the Pakistan or Afghani language, they are using the same script. So how is it this problem tackled within the IDN programme or in the IDN system?

If, for example, I write I ran in Urdu language or other languages the same number of characters there. For example, if I write Iran.com in Arabic script, so how exactly is this thing tackled? This is Iranian, the Persian Iran or the Urdu language or Pashtun language Iran?

>> RINALIA ABDUL RAHIM: So when the repertoire were developed and there's all these different languages that use the same script, in terms of identifying it for review, you place a tag on it, meaning that this character or letter, whatever it is, comes from let's say Iranian or Persian or so forth. But when you are on the Internet and with a label, it's almost impossible to tell. Like hero was trying to mention that. And that is why there has to be a Rule that's agreed to by all the Arabic-using script community that says if this label is being used, then something else that looks similar to that has either to be blocked or not allocated. And that's why we need the community to come up with the rules, to propose it and to get agreement.

>> MR.JIA RONG LOW: Thank you. Over there? Microphone over there, please.

>> My name is Ungjun Jong from Korea. I think trademark is very important for setting up domain name policy. And Hoto-San mentioned before discuss Wednesday the trademark association about TRD strings. My question is that what was the main topic or any controversies of the discussion with the trademark association?

As far as I know, maybe I am wrong, Korean community has not consulted with the trademark owners about specific things. That's why I question this. Thank you.

>> MR.JIA RONG LOW: Thank you very much about the question.

>> So the first thing is that surprisingly do they want IDNTLD? That's the first issue we discussed because of course in Japan, many trademarks are written in Japanese characters. But on the Internet, they just want to have English names as a TLD. So I think that may be a lack of Emails capability or something. But at least at this moment they don't want IDNTLD very much. That's the first issue discussed with them.

And second one is as you said, if they want to have TLD, IDN TLD,

they may -- that TLD string may have variant. So, for example, even in Japan, the two characters may be interpreted as the same character. But it is not a deep issue for them because their current basis whether IDN necessary or not.

>> Just to add a few. Actually it is not a new question. It's a very old question. Back to country 20 years ago when we talk about IDN protocol. Older IGF engineering team that went out interesting that kind of consider because they consider we already have domain name and domain name is identifier so they can represent any kind of -- modeling or name in language in web content instead of put it in domain name. Even right now still a lot of people criticize why do we have to pay attention on the domain name? Because right now everything was using APP app. So you can just go to the destination by itself. Also you need to pay attention on the domain name. So we probably went to the wrong industry. Just kidding.

So back to the principle of communication. Data communication is a multilayer from layer 1 to layer 7 from TCP is five layer. The major characteristic of the labeling is every layer is -- you don't need to worry about what the protocol will be run on top of labeling. When you deploy TCP IP, what solo would be wrong on top of TIP. We are doing a specific layer application. If you consider as application, it does knots have any commercial interest at all but if you consider we are enabling the possibility of an infrastructure software infrastructure, then we are multiplying the future potential for application. So if you consider trademark definitely something that relate to the IDN, but it's not closely related, it's partially related. But how we do, how we're trying to do just maximizing the potential for application. But alone doesn't consider basic. Still it is considered as an infrastructure. So we are developing the infrastructure prepare for the future.

>> Hi, my name is Sayeed from Afghanistan. I come from a country where there's tons of languages, 52 registered languages. And there's many more. And we use Arabic script.

My question is that if we have identified some languages that are part of the IDN Working Group and there are some languages where we cannot find some experts who can contribute into the IDN Working Group, what is the process of this variants and LGR development and review process?

When the email and the domain system, IDN domain name system is fully functional, would it be a problem to add more languages to it? Is there a process established for that? Or no?

>> RINALIA ABDUL RAHIM: So the label generation rules is going to be released on a cyclical basis. So LGL1 I think has been released. And there will be a period where there will be acceptance of new submissions and the experts will look at the submissions to see whether it's appropriate to add to it. So it's not the end of the line. You certainly have an opportunity to participate.

>> Hello. My name is Casey K from academia. I pretty much appreciate all the work you have done for the various technology development but I'm interested in what happen after all the technological development is completed? So how do you perceive the takeup of all IDN top level and second level domain names will look like? So that's my question.

>> MR.JIA RONG LOW: Thank you. I think I've heard Hiro before. So I ask him to answer. So the question is after all of this work is done, how do you think people are going to use it? What do you think the pickup rate might be like going forward?

>> Simple question. I facing this question before. The issue is I think it's really to the team you're talking. Physical space or cyberspace. Nobody can find me with Roman names, the ASCII. It's impossible. If you map between the Thai characters, I can spell 100 ways to represent my name. So you will never find me if you don't know how I spell my name in English. Because like my family name, using both Sanskrit and Bali to represent it is not Thai. It is not the consonants. It's the syntax that represents that this is coming from Sanskrit family. So there are 1,000 exceptional rules in the language. But when you're born you get the TNA character names. So if it is not available, so how could you identify people?

So they did not know that is is available. And people get used to using English. But this will bring in the new thing. So you can write your name in Thai. You can have email in Thai. You can know this government office, you type in, you can find -- that are the major issues that we think that people just wait and they don't understand that this technology is available and is important. And you talk about the physical, the cyber space identity, we talk a lot about e-authentication frameworks. And if you cannot identify a person in their local context, so why we have to use the technology is impossible. Like public infrastructures is not going to work because the key point is legally binding. But there are no legal documents in English. Pass poured is not a legal document -- passport is not a legal document, by the way, we don't have the legal framework in English so it's still in Thai. So anyone in court, everybody, even your western we will spell your name in Thai first. But the assignee have to spell the name that this is the person and then the court order will start to work. And these are the local contexts. I do believe that the acceptance rate will be high.

And I have to mention that in detail even though I use my email address in Thai, can I send to anywhere in the world and they can send me back, it's not a problem at all. Not meaning that I have to each an Thai corrected names. I don't go into detail the company called blah, blah, blah technical works had been used. If seamlessly I send to Google, I did the ID is will be. Is perfectly done. And the problem is who can provide a service at which cost? We are now working with technology supplier like outlook from Microsoft and

Google who offer these kind of services for the Thai citizen. So I think it took away a long time from technical availability all the way to technology suppliers.

>> Thank you. I wonder whether there are other speakers need to add? Kenny.

>> KENNY HUANG: I also agree with you. Basically when you answer the question, it depends on what sort of timeframe you're talking about. If it's about six months, it's difficult. Two years, a challenge. But when you consider 20 years, it's a highly all kind of potential possibility can be generate. If you think 20 years too long let's consider Internet TCP/IP. It took 30 years to commercialised. Even for IPP6 was created in 1996. Right now it took 20 years, getting 10 percent growth rate. So has great potential in the future. That's why I mention if you consider its application definitely hopeless. But if you consider it's a software infrastructure, it's a technology infrastructure, once it create, it's waiting to be materialised so be patient.

>> MR.JIA RONG LOW: Thank you, Kenny. I don't think I can put it more eloquently. That was very good. I hope we can go back and get a transcript. I'll just read it next time I get the same question.

[Laughter]

I think Kenny put it on top of that, I did a tour around with my colleague say mad Hussain who runs this programme for ICANN, when we were going around the different countries and often he would tell me, okay, you lead the Asia-Pacific work. So can you always give introduction to ICANN? So when I talk about ICANN and the IDN programme and I'll say imagine a day when your local community can come online without typing a single word of English. You see the faces change in those people's faces. Why? Like mentioned in Thailand, there's a large community that cannot speak English and cannot use English. If you want to sue someone, you have to write Thai, which means it limits them from using the English system and this application basically allows for then the sky becomes the -- because allows the local language on the Internet.

Another area I think I can highlight, as well. There are many national initiatives that are thinking of driving this forward. For example, even with working in the Thai government and also working with Rapid also from the Cambodian government, they are willing to trying very hard to push this for -- from the application layer, but they also have a vision how can they provide this even better?

So from the Indian government, for example, they are promoting the active use of dot barat which is Hindi script, the script which is Indian script, to encourage people to use their own local script for domain names. So there is a national initiative which drives this.

I think it has to be multistakeholder model in a way where governments drive national initiatives and the software and the businesses also pick this up and the users also are able to pick this

up. But as Kenny mentioned, I don't think I can add anymore to that.
If we look at the landscape of 20 years, I do believe that we will
see many communities come online using their own scripts and
languages.

Any other questions? We have time for one last one.

>> Yes, thank you. Lisa from the Philippines. It's good to hear
that languages and scripts can be preserved and appreciated by
communities online, as well in the Internet. Of course in the
Philippines we use Roman script. And it's because of centuries of
colonization. But we also have our own script. It's called Alibaba.
And I was wondering if there has been initiative already in the
Philippines already to use that script? Otherwise because I know some
groups who are also using -- who are studying and preserving that
script. And it would also be good to see that it's also included in
this new IGR or IDS.

>> MR.JIA RONG LOW: Thank you. I will just take this question.
As far as I know, I don't think that script is being used. So if we
could take this offline, I would really appreciate it if you could
just write to me and also to this email IDNTLDs@ICANN.org and tell
us about this script a little bit more and we'll find a way to work
with you. Thank you so much.

Okay. I think this pretty much wraps it up for this session.
Thank you so much for your time. And can we give a big round of applause
to our speakers? Thank you so much. And thank you very much, Rapid,
for joining us remotely. Thank you so much.

***

This is being provided in a rough-draft format. Communication Access
Realtime Translation (CART) is provided in Order to facilitate
communication accessibility and may not be a totally verbatim record
of the proceedings.
***