ICANN75 | AGM – Celebrating the Multistakeholder Model in the Development of RZ-LGR
Wednesday, September 21, 2022 – 16:30 to 18:15 KUL

SARMAD HUSSAIN: Okay. So we'll start the session. Could we request for the recording to start?

Thank you all for joining today's session in which we celebrate the multistakeholder model in the development of Root Zone Label Generation Rules. Basically, these rules are defined to allow for top-level domains to be supported in the root zone of the Internet. It has been a multi-year project starting from 2011. We'll get more input and details about this work during this session. But the work has been driven by script communities from across the world where many, many volunteers from all across the world have contributed towards developing the rules which should govern what is the unique label and valid label in that particular script and the languages which are written in those scripts.

The work has been done, as I said, with many different script communities, which have included Arabic, Armenian, Bangla, Chinese, Cyrillic, Devanagari, Ethiopic, Georgian, Greek, Gujarati, Gurmukhi, Hebrew, Japanese, Kannada, Khmer, Korean, Lao, Latin, Malayalam, Myanmar, Orya, Sinhala, Tamil, Telugu, and Thai. We now, thanks to these communities and their excellent work over the years, have very concrete rules to move forward for defining the valid domain names, especially top-level domain names, in these scripts.

So with that, let me introduce the agenda for this session. We'll have a two-part session where the first part is a panel discussion which goes through just an overview of what the Root Zone LGR project has been, why RZ-LGR is needed, how it was developed, and how the community has contributed, and what are the next steps as far as acceptance and taking forward the RZ-LGR work by the community forward through the policy development process. And then we'll have some questions or answers. We'll take the question and answers after the panelists have had a chance to make their statements.

In the second part, we will recognize all the community members who have participated in this project over almost a decade and made those rules. We will have Göran Marby, the president and the CEO of ICANN, and Maarten Botterman, who is the chairman of the ICANN Board, come in to share the gratitude with the community for the contribution to this work.

We will also have a picture with all the members of the Generation Panels, the community members who have contributed to this work, many of whom you can see in green T-shirts here, but also many of whom are actually online and they'll join us in Zoom for the photo session. It will be good for you to also interact with the script community from your own region and socialize afterwards. So that's the plan for this session. Next slide, please.

So as far as the panelist is concerned, we have a very experienced set of panelists here. We have Edmon Chung who's the chair of the ICANN Board IDN-UA Working Group. We have Marc Blanchet who represents the Integration Panel for this work. We also have community members

from the different script communities who've worked towards development of RZ-LGR, which includes Wang Wei who's the co-chair for the Chinese Generation Panel. He may not be able to attend because he's not here. In that case, Kenny Huang, who's also the co-chair of the Chinese Generation Panel, he will be sharing their experiences. We have Michael Bauland who is a Latin Generation Panel member, and Dr. Ajay Data who's the co-chair for the Neo-Brahmi Generation Panel. So they'll share their experiences from the communities' work perspective.

Then as we move forward with RZ-LGR, the technical work's done, so now as the next step, it has to be adopted by the policy. We have invited Donna Austin who's the GNSO IDN EPDP (Expedited Policy Development Process) chair. Unfortunately, she's not been able to come in today. So her presentation will be done by Ariel from ICANN. Also, Kenny Huang who is the chair of the ccNSO's IDN Policy Development Process. So Ken, he will be presenting how RZ-LGR work, and Ariel as well, how RZ-LGR's work is being integrated into the policy. My name is Sarmad Hussain. I'll be helping. I'll work with Pitinan to help moderate this session.

So with that, let's get started. Just to sort of show you a timeline of where we started from, the initial homework for this project was done in what is called the Variant Issues Project which ran from 2010 to 2012, and we'll request Edmon Chung to cover that era of the project. Then once the whole homework was done, we launched this project in 2013. For the last nine years or so, we've had the community work on it. But before it was launched, we set up an Integration Panel which

had its own role in the process. The process then was an interactive collaboration between the Generation Panels and the Integration Panel. The work now had at a stage basically in 2022, early this year in May, we published the fifth version of Root Zone LGR which integrates work of all the Active Generation Panels, 26 of the 28 scripts initially planned. The work is completed for at least those Generation Panels which have been active. We are here to celebrate their contribution, celebrate the contribution of the community in this very extended work and important work.

With that, let's move on to our first panelist, Edmon Chung, who's been involved in this project from very early days, its planning days. We will request you to basically share your thoughts on the need for multilingual Internet, which we know is something very close to your heart, and maybe share a background on what was the motivation for starting the Root Zone LGR project. Thank you.

EDMON CHUNG: Thank you, Sarmad. Welcome, everyone. I'm glad to be talking about this topic. For many of you in the room already know that this is definitely a project of passion for me, all the way back to 1999 when it brought me to ICANN. I was just sitting here talking to Marc, you're next to me. First time I met Marc was to talk about IDNs, and that was in an airport.

So I think it's a very big question you put before me, what is the need for multilingual Internet. But one of the interesting stories, I guess, it's been like 23 years of development of IDN. But for the first few years, it

was not easy to convince people that there is an actual need. In fact, I still remember very vividly at one of the sessions that someone went to the mic and said, "I only know two languages: English and the computer language C." So it took a little while to actually get people to understand that there is the importance of multilingual on the Internet. The reality is that the majority of the world does not speak English as a first language or use the English alphanumeric characters only. Local businesses use local languages in their names. Names of people, I mean, obviously use local languages. So why identities online cannot be using their own native language? I think that is really the motivation of IDNs and part of internationalized e-mail addresses as well. Today maybe you think not a lot of people use Internationalized Domain Names still. Of course, there are challenges. One of the challenges, of course, is the universal acceptance of different IDNs and e-mail addresses.

But there is another aspect which is what we want to talk about today. Another challenge is that of policy. What kind of policies? It's just domain names like English domain names that technology is there for other languages. So what's the problem? Well, there is. Very interestingly, early on we found that when you add different languages onto the Internet, especially the unique identifier system, especially domain names and e-mail addresses, there are situations where different characters have different kind of natures.

If you're very strict on definition, this is not a good example, but I still use it because this is the easiest way to explain to people who doesn't know a lot of languages. The best way to think about it is that your

English domain names right now, capital letters and small letters, you can use it the same way. When you type in a particular domain name, different capital letters or small letters, you go to the same place. The same is not true for Internationalized Domain Names, and one of the classic examples is Chinese, for example. We use traditional Chinese and simplified Chinese, and they are different code points for domain names technically. What is then needed? However, people use them interchangeably, especially, for example, in Hong Kong, increasingly, people use simplify Chinese and traditional Chinese interchangeably.

So what we need is a system to map them together. Think about the situation where you want to register a domain name, you have to register all the variations of capital letters and small letters, you're not registering one domain, you'll end up registering tens of domain names. That's the policy. That is one part of the policy that we are talking about. It is also tied into something that is really important in my mind, which is to reduce abuse and to increase the trust for the DNS by end users. So what people see as the domain name or believe they're typing in the domain name, the policy should allow it to, as best as possible, map the user experience with the expectation. That is an important part of what is called the Root Zone LGR, the Label Generation Rule set.

But talking a little bit, I guess, about the history. First of all, I am now explaining that is so important. Very early on, I used to think that what's the problem? If these different languages—it's sort of like the word colour that is spelled in British English with the U, and color in American English without the You, and so what? You registered two

different names, and that's that. What I realized is that it's not like that. Different languages really have different characteristics, and this is what the Root Zone LGR tries to address in terms of a policy that is consistent and also technically able to be implemented consistently for the root zone and for IDNs to work consistently.

So with that, actually, just as a little bit of a background, Sarmad mentioned that the Root Zone LGR work has been ongoing since 2010. In fact, that was when the IDN ccTLD Fast Track first came in. So the first IDN TLDs were starting to be put in place in 2010. That was also when a group between the GNSO and ccNSO was pulled together. I was honored to serve to help chair the group, which is called the Joint IDN Working Group between the ccNSO and GNSO. For those of you who like to know ICANN history, that represents the first time that we brought the ccs and the Gs back together. You look at the trivia of ICANN, you would see that initially the ccTLDs and the gTLDs are all together in what is called the DNSO. But this group brought it back together because there were a number of common interest items, including single character IDN TLDs, Universal Acceptance of IDN TLDs, and of course, IDN Variant TLDs, which is a lot of the part which the Root Zone LGR addresses. So that was back in 2010.

Then, of course, I guess as we move on through the years, different issues have been dealt with policy-wise. But the reason why we're celebrating today, and looking at here, I think next time if we really come out of the pandemic, we should have probably some wine and beer to actually celebrate. But goalposts coming to 2022, I think what is really exciting is to see that most of the active languages around the

world and the scripts that are actively used are now able to be used in the root zone.

I will end by saying that there are many different debates here at ICANN, IDN has been one of the things, as I mentioned, since 1999. The other thing that has been discussed and debated since 1999 is probably WHOIS. We won't go there, but what we are looking at is the final stretch of making IDN real. I'm really excited to be serving on the Board IDN-UA Working Group right now as the chair to kind of see the rubber hit the road and see IDNs in real action. I think this is one of the issues that probably represents one of the strongest community consensus that it is in the global public interest to make IDNs work to make them work universally and to policies to prevent abuse and support IDNs.

That was I guess the little bit of a story that I think is important for the program. And from there, the many years of hard work is built what we now have, I think, as a very usable and consistent, technically and policy-wise, strong Root Zone LGR, which I think my other panelists will touch on. Thank you.

PITINAN KOOARMORNPATANA: After Edmon already shared about the history of this project, I think we will move on a little bit on inviting the relevant panels to share a little bit of how they experienced how the work has been done. So at the beginning, I will start off by Marc Blanchet from the Integration Panel. He will give a little bit of explanation of how we organize the work so that we can gather the knowledge from the script

communities, so many of them, into the single set of rule set. So over to you, Marc.

MARC BLANCHET: Thank you, Pitinan. Good afternoon for those of you here, and good morning, evening, good night for people remote. It's an honor to be here. So kind of celebrating 10 years of work which is quite an undertaking, especially for the colleagues on my right that did most of the work, which is the Generation Panels.

To start with, I want to get a top-down understanding of what we have been doing, starting with Unicode. Unicode is the standard for encoding every possible glyphs used in any script and language in the world since the beginning of the writing. Unicode is actually, as has been recently encoded [inaudible] Euro glyphs. But those glyphs are not good for Internet in the [fire]. One of the most important criteria is security. We don't want to have phishing based on look-alike glyphs.

The IDN protocol called IDNA that was defined by the IETF restrict the whole Unicode set to a smaller set based on the properties of the characters, not their appearance or their visuals. For example, IETF didn't look at every single pair of characters to verify their visual similarity. LGR work further restrict the set of characters and define rules to make a string in any script a candidate for a DNS label or domain name.

To create the set for a script, Generation Panels were formed as a group of experts for their specific script. They spent months and years to actually define the restricted repertoire of characters, variants, and

related set of rules in their script that should be used in TLD zones for the purpose of this work but was actually useful for second level. Part of this work actually excludes some of the characters such as the one that are historic or not in use. The formal output of this is an LGR, Label Generation for the script, and that's actually an XML file for those who are technical.

The work is based on some principles. For example, conservatism, simplicity, security. Example of the use of those principles is we don't include characters that are not in use today. We don't want spammers, for example, to use some obscure character that is not in use but is similar to one that is in use. So for the purpose of the process, the Generation Panels have to document their characters and identify if they are really in use. And we the Integration Panel, we're here to verify those claims.

We want stability and longevity. For example, if Unicode encode a new character recently, then that character is actually not a good candidate to be included in the LGR. Why we're saying this is because, like everybody else in software development, sometime you make changes because the actual property of that character was not the right one and the change. Well, in the context of identifiers, we don't want to be in a situation where such a label, for example, a TLD string contains a character that was valid, and then later on becomes invalid. That is not good news.

Another example of design principle is simplicity. During the work, sometimes we receive LGRs from Generation Panels that were looking

like language grammar. That's not the place to put some language grammar. It's really restricted to how you write the script.

The Integration Panel was formed to review those scripts LGRs to verify the principles and also to integrate all of those LGRs together in a single Root Zone LGR. By integrating, the IP also look for cross-script issues that may arise that obviously may not be seen by a single script Generation Panel, so the Integration Panel as the purview of all the LGRs. The IP reviewed in detail all the script LGRs received enough and was playing devil's advocate by asking the GPs to justify the inclusion of each every single character in the proposal. As you may know, some scripts have many, many characters. The IP is composed of five experts from different domains, Unicode, IDNA, DNS, and languages and scripts. I am a member of that Integration Panel.

The Integration Panel, the integration of all the LGRs, brought to IP some design choices. For example, cross-script variants sometimes create an avalanche effect, too many scripts. So a variant of one is related to a character, another LGR, which is, as a variant, another one go all over the place. So for simplicity, we sometimes made design choices that seemed the best for the whole community but may have impacted more or less some LGR files. But obviously, we never ever impacted the work of the Generation Panel. The core Generation Panel as LGR work—repertoire and rules were never impacted on our work.

For some scripts, the actual line between including a character or not in the LGR is pretty thin. Moreover, the similarities between some characters, especially between scripts are often significant. So there

are many instances where the final choices made by the Generation Panel but agreed by the IP were tough. We didn't have to fight in the street but certainly over a beer or two. So if you could come back to the other slide—thank you—you see down in the slide a little arrow that says, "Needs more work." We sometimes end up like being the professor at the high school receiving homework from a student, and the professor says, "Not good enough. Continue working." Then the student comes back with a new version and the professor says, "Not good enough." So, sorry, my colleagues of the Generation Panels, we were not trying to do that way.

So during all those years, we were often in a situation where we could be bribed by the GP for the purpose of getting their characters included in the LGR. Well, only very few of you know but I have a big announcement to make. We have been bribed. Am I getting your attention? So for the interest of the community, now it's pretty done, I need to disclose, right? We received one gift from one GP and I'll show you. Do I get your attention?

UNIDENTIFIED MALE: Is it a pile of cash?

MARC BLANCHET: Oh no, sorry. That's not it.

UNIDENTIFIED MALE: Is it a check?

MARC BLANCHET: Here's the gift. We received from a GP a T-shirt with their script on it. That was really nice. I would have loved to be bribed by receiving a T-shirt for every single script of the Root Zone LGR, even if for some script, it might have been more challenging, given the number of characters. To avoid any confusion, that was a joke.

So in summary, this fantastic work would not have been possible without the great work of the Generation Panels. The Integration Panel was there to help and facilitate. I would like to thank my colleagues on the Integration Panel: Asmus Freytag, Michel Suignard, Will Tan, and Nicholas Ostler. I would also like to thank the ICANN staff involved since the beginning of the journey, naming Naela Sarras, Nicoleta Munteanu, Sarmad Hussain, Alireza Saleh, and Pitinan Kooarmornpatanai. I asked her today to help me speak her last name and it's still difficult for me, so I hope I'm not too bad. I would also like to thank the Generation Panel members but I'm not going to list all of them because there's probably a few hundreds. Long life to the queen—sorry—the king—sorry—to the LGR and to the open and diverse multilingual Internet. Thank you.

PITINAN KOOARMORNPATANA: Thank you, Marc. Now you know why we took many years because we were busy playing jokes. All right. Then let's move on.

So now we move on a little bit on some share of the experience from the GPs. So we actually have more than 270 volunteers to this work from seven different GPs. We obviously cannot host everybody here,

unfortunately. We have representatives from three Generation Panels here. We have the members from Chinese GP, Latin GP, and Neo-Brahmi GP. So I have just one minute, just go around first for Kenny, for Michael, and for Dr. Ajay to just briefly talk about what is your GP like, what are the components of your members, and what is the characteristic of your script? So just one minute each first, and then we go to the question. Thank you.

KENNY HUANG: Thank you. Kenny Huang, co-chair of Chinese Label Generation Panel. It's my pleasure to give some basic introduction regarding Chinese Generation Panel. Basically, a member of Chinese Label Generation Panel comes from—before we joined the Generation Panel, actually, we were already starting working on Internationalized Domain Names together, for example, for speaking territory, Taiwan, China, Macau, Hong Kong, and also Singapore as well. So we already collaborate these few speaking territories to join the alignment we call CDNC, Chinese Domain Name Consortium. In addition to that, also joint effort with the other Han character using territory, for example, like Japan, like Korean. So basically, the members of Label Generation Panel majority cover in Taiwan, Macau, Hong Kong, and Singapore, this region. But regarding to the coverage of basically Chinese, it was the second largest language using in the world. So regarding to the coverage, basically we have written in our policy proposal, basically you can see a lot of places. Actually, they are really using the Chinese character. It's quite often, even for Malay, like [inaudible], walk down the tree, you can see a lot of Chinese label and using in [inaudible] like

Singapore, the Chinese label is everywhere. So basically, the coverage of Chinese character actually is almost worldwide coverage. The characteristic we cover traditional Chinese and simplified Chinese altogether. So including a lot of language, we have a lot of characters we haven't used long ago. Through that kind of effort, we try to integrate that kind of effort, put it together into a Chinese Label Generation Rule and put it in our repertoire. Sorry, more than one minute. Thank you.

PITINAN KOOARMORNPATANA: Thank you. Michael, go ahead, please.

MICHAEL BAULAND: Thanks. I'm a member of the Latin Generation Panel. We've had quite a lot of part-time members, so to say, because many people joined and then left again. In the end, we had a core team of seven people: Bill, Dennis, Hazem, Mats, Meikal, and our chair, Mirjana. And we have experts from different areas, linguistic experts, some Unicode experts, and some experts of the registry and registrar topics.

The coverage of Latin script is known by everybody because we all speak English and thereby use a Latin script. But it essentially there are several hundred languages using the Latin script, and our challenge was to cover most of them because all of them would have been impossible. So, the main characteristic is of course diversity of the languages. Thanks.

PITINAN KOOARMORNPATANA: Thank you, Michael. Dr. Ajay? We are aware that you are the man of the hour so you have to also leave to other UA-related session. So after you introduce the Neo-Brahmi Generation Panel, could you also share the magnificent work to work on that script altogether in one panel and share something about that altogether? Thank you.

AJAY DATA: Thank you, Pitinan, for consideration. This is Ajay Data for the record, co-chair for Neo-Brahmi Generation Panel. We had four countries covered, including India, Nepal, Sri Lanka, and Bangladesh. Obviously, we all know Neo-Brahmi Generation Panel was a successful example of multistakeholder process where people came together to achieve a common goal with consensus. What GP included people from academia, civil society, government, technical, and language communities and industry who work together to achieve rules for all nine scripts and 22 official languages of India, and along with obviously all the countries which I mentioned.

The challenges were many from that perspective because it is not just the language. You land up dealing with the political influence at some point where people would like to remain in their rules, but ultimately the challenge is to be sorted with consensus and discussions in the community. We went in each country, physically had meetings in each country, and discussed with the authorities, discussed with the concerned stakeholders, and also seek their doubts, whatever they have in their mind and clear them of. At the end of the day, this was empowering their country in that region.

Simply, we followed a traffic light process, a very interesting one. We used the green, red, and yellow light. So if anybody has a green symbol, that means that is a distinguishable label. You can have that green symbol and nobody worries about them. Everybody says, "Put your green, red, or yellow," which was very easy to discuss what we need to discuss and focus on. These rules continued during the discussion and we keep moving the characters from one color to another. This gave us an opportunity to focus only on the problem areas where we were discussing nine scripts. As our other colleagues and members of Integration Panel will appreciate, we were not dealing one script. We were dealing with nine scripts. Hence, we need to also standardize the rules. We standardized the [inaudible] rules from that perspective so that we do not have conflict among the scripts within the country. Hence, the rules were standardized among the scripts. And with almost 70+ volunteers across all the stakeholder groups, we generated those panels. I think within two years of time, we worked with all the nine scripts, and we did our task. For one script, we took one year more extra, that is Bangla. Now I think we are sorted with that too. With that note, please excuse me to join a session and I will join you in 10 minutes back. Thank you very much.

PITINAN KOOARMORNPATANA: Thank you, Dr. Ajay. Then let's move on to some questions to Kenny and also Michael. So let's start from Chinese GP because you use the Han script, which is obviously used in Korean and Japanese as well. So could you share how did the Chinese GP coordinate with other GPs on this? Thank you.

KENNY HUANG: Okay. Thank you. Basically, as I mentioned, we already established technical engineering group before the Generation Panel. So we're familiar with each other with our engineering team in Japan and also from Korea, not only engineering but also language experts from Japan and also from Korean. So before we joined the Generation Panel, basically we have set up a coordination among Chinese, Japanese, Korean to see how we're going to work together. Because before having Root Zone Generation Panel, basically all the Internationalized Domain Name was registered under ccTLD so there is no any overlap code point issue. But if we want to put the Han character in a single space in a root zone, we definitely need to coordinate. So before that, we already foresee we need to work together.

Another issue is we know each other quite a long time. So we're happy to work on in this issue together. And different organizations, they're willing to commit certain resource to get a problem solved. By the way, because they are so many character overlap between, for example, like a Chinese and Japanese or Korean with Chinese Han character, there are so many code points actually overlap. So in order to resolve the overlap, we have quite intensive meeting, not only within the ICANN but outside the ICANN community we also set up several independent meeting with the Korean Han character community, and also with Japanese Han character community as well.

So I remember we even have one day trip go to Seoul because the other meeting room has been booked. So eventually, we have a Han character discussion in a Seoul train station that day. After the meeting at almost midnight, whatever your midnight, fly back to Taipei. So actually, it's very, very comprehensive, a lot of engineering discussion, a lot of Han character discussion. So I think that we created a very good chemistry to working with community members from Japan and also from Korea as well. Due to that kind of harmonized chemistry, I think that's the best way to work in the community, and eventually we can deliver a foreseeable outcome to generate that repertoire we desire. Thank you.

PITINAN KOOARMORNPATANA: Thank you, Kenny. We also have Generation Panel from Japan and Korea in the room as well. So we should have some photo after this. Okay. Let's move on. Next, I would like to ask Michael, our Latin GP member, because your script is being used in so many, many languages. So how can the GP manage to come to the conclusion? Thank you.

MICHAEL BAULAND: Yes, that's true. Latin script is used all over the world, not just in Europe or America but in Africa, Asia, Australia, everywhere, there are languages using the Latin script. So our first difficulty was to choose the languages we were able to analyze, and for that we used the so-called, EGIDS scale, that scale which assigns every language number, which states how much the language is used. For example, English is

scored zero because it's used almost everywhere. Then it goes down to scores five, for example, which says like the language is in vigorous use with literature in a standardized form being used by some, though this is not yet widespread or sustainable, or six, which means the language is used for face-to-face communication by all generations and the situation is sustainable. But since we were working for domain names, of course, only those scripts, those languages with stable writing system could be of use for us. And that's why we decided to only take those languages with an EGIDS scale up to four into consideration, plus the ones within scale of five that had at least one million users. This is of course somehow arbitrary but we had to make some decision what to include.

For those languages which were still more than 200 languages, we looked through the Internet for sources which characters are actually used in those languages. All of these characters have been included into our repertoire with some exceptions. For example, if a character is used for a click sound also or if they look like an exclamation mark, then this was not possible to include then. Even though we just consider those languages, it does not mean that all other languages with an EGIDS scale lower than five are not included at all. It just means that we didn't look at their characters. Most likely 99% of their characters are already included, because what we found is that most of the characters we included have not been used just in a single language but they have been used in many languages.

Another problem was then to decide on the variants whether one character should be considered to be a variant of another character.

And for that, we introduced a scoring system and we compared each character with other character, and then assign a number to that relation, like one for being a homoglyph, two for being nearly identical, three for distinguishable, and four for distinct. And all seven of our core members assigned such a number to each relationship. To consider two characters to be a variant, we decided that at least five of the seven members had to give this relationship a score of one or two. There were quite some heated discussions in there, whether some characters should be a variant or not, because it's sometimes really not easy and it depends on the font that is used to decide whether they are so similar that a normal user is unable to distinguish them or whether they have still some distinct features. Thanks.

PITINAN KOOARMORNPATANA: Thank you, Michael. So I think we have a little bit of the tour around the world. I guess before we move on to the next section, any last word from IP or GP or anyone on this? All right, so then I'll hand it back to Sarmad. Thank you.

SARMAD HUSSAIN: Thank you, Pitinan. Thank you, Kenny, Dr. Data, Michael, and Marc for giving us a bit of peek into the journey of developing the Root Zone LGR. This effort has, of course, taken thousands of volunteer hours from all the different volunteers from all the different script communities.

As the work now concludes, it was considered by the ICANN Board. The Board resolved and requested the ccNSO and GNSO to consider

Root Zone LGR in the policy development process in their respective policy development processes in the context of IDNs. We will now go to the work being done by GNSO and by ccNSO to get a little more insight on how that work product which came out of this effort by the community is now being used as part of the policy development process.

So we'll first go to—as I shared earlier, Donna or Justine, unfortunately, are not able to come. They are currently chair and vice chair for the IDN Expedited Policy Development Process at GNSO. We request Ariel from ICANN staff who will share some details about how GNSO is integrating Root Zone LGR in its policy development process. Ariel, please.

ARIEL LIANG: Thank you very much, Sarmad. This is Ariel Liang. I am supporting the GNSO Council, specifically the IDN EPDP. I'm rather a latecomer or newcomer in this long-term project but I feel very privileged to be part of the process contributing to the adoption of RZ-LGR and the broadened access to the Internationalized Domain Names.

So when I joined this project, I had this question. Since one of the most difficult issues is defining the gTLDs and variants, and then RZ-LGR already did that, they already have figured this out, then why do we still need policy development process? The reason is that to operationalize the RZ-LGR and then make the variants at the top level into a reality, you really need to have consensus policy to adopt the RZ-LGR and also adopt relevant policies and rules related to the

management of variants. In that way, you can make variant at the top level into reality. So that's why we have PDPs come in to play.

So I just want to give you a quick background about the GNSO EPDP, how that works. The GNSO Council has initiated this PDP in last year and the group started its deliberation August 2021. It's functioning as a hybrid model, so it means it has representatives and participants from across the community, as well as liaisons from the ICANN Board and ICANN Org. In fact, on this panel, Michael and also Edmon, they're a big part of this group. They have been actively contributing to the work. So it's very privileged to work with them on the IDN EPDP. As Sarmad mentioned, this EPDP is the under the leadership of Donna Austin from GNSO Registry Stakeholder Group, she's the chair, and Justine Chew, she is the vice chair and she's from the ALAC. So you can see the multistakeholder representation in the IDN EPDP.

Also in terms of the scope of the work, they are focusing on two areas. One is the definitions of all gTLDs and the variant management mechanism. Second is the IDN Implementation Guideline and how it should be updated in the future. So the IDN Implementation Guideline is related to second level IDN registrations and the contracted parties need to comply with that.

The charter includes seven topics and just 48 questions in total, and RZ-LGR is one of the seven topics. So when the group is taking into consideration the RZ-LGR, one thing they noticed that they're already a big body of work that already currently exists on the IDN subjects, and also the new gTLD Subsequent Procedures PDP has developed a variety of recommendations related to IDNs. So basically, this IDN

EPDP has to build on the existing work and then take those recommendations into account.

Just to give you a bit of flavor regarding the consideration of RZ-LGR and the draft recommendations that this group has developed. SubPro actually already recommended compliance with RZ-LGR for the generation of future new gTLD and also calculating their variant labels. So in practice, this means that in the application submission system for the future rounds, any applied-for strings as well as their variants need to go through an algorithmic checking against the RZ-LGR to ensure they're valid. The EPDP team has developed a recommendation that's consistent with what SubPro has recommended. So it's basically recommending compliance with RZ-LGR for existing gTLDs with respect to variants. This means that if an existing gTLD registry operator such as Arabic or Chinese TLD or registry wishes to activate a variant of their existing gTLD, such label requested needs to also be checked against RZ-LGR to make sure it is valid and also allocatable.

Under this example, the IDN EPDP team also developed some additional guidance regarding this RZ-LGR compliance. So for example, in the new gTLD application process, there will be a DNS Stability Panel to evaluate the applied-for string and the EPDP team believe that the DNS Stability Panel will be best suited to review and determine whether applied-for string is indeed valid according to RZ-LGR. If an applicant disagrees with the panel's determination, it can actually use the limited challenge mechanism to challenge that determination. So that's inconsistent with what SubPro has proposed.

Some other examples I want to showcase is related to the technical aspects of the RZ-LGR operation. So, the second example is regarding grandfathering. So, the RZ-LGR is expected to be updated throughout its life cycle, either as a result of new script LGR being integrated or the updates or revision to existing script LGR. There may be a very extremely rare case that a proposed RZ-LGR update does not support existing IDN TLD. So, in such case, the IDN EPDP team has recommended grandfathering the delegated gTLD and to maintain the stability of the root zone, but they also come with some other conditions and implementation guidance. One of the implementation guidance is expectation for the Generation Panels to make best efforts to retain backward compatibility with existing gTLDs and their delegated and allocated variant labels to maintain the stability in the root zone.

The third example I want to showcase is about single character gTLD. So we know there's already existing Chinese, Japanese, Korean gTLDs, and all these in the root zone but we don't have single character ones. For the folks who know about this language, their idiographic language means that one single character already has the meaning and can stand by itself. So you may be familiar with the geographic locations or family names. They're one character but people know what they mean. However, in the root zone, we don't have them already existing because there's a lot of security, stability, and confusability concerns. That's why the IDN EPDP team is also tackling this particular question. It's consistent with the SubPro PDP that to recommend the single character gTLD be only allowed for limited scripts and languages where a character is an ideograph. So that

specifically means Chinese, Japanese, and Korean languages, and Han script, they can have single character gTLDs.

But there's also additional work that needs to be done because there may need to be additional guidance or restrictions and further understanding to understand which characters should be allowed or should not be allowed to be single character TLDs. That's why the EPDP team would like to work with the Chinese, Japanese, and Korean Generation Panels and get their linguistic expertise to understand this question.

So these are some kind of examples of this IDN EPDP's work. You can see it's very interesting and we will love to have broader community engagement. You're welcome to observe the meetings of the EPDP team. It's open to the public. We have broader participation from the community so you're welcome to join even as a participant. And also in the future, we will publish the initial report that includes preliminary recommendations and will be subject to public comments. So we'll welcome your input and feedback on the work. So with that, I will turn the floor to Sarmad. Thank you.

SARMAD HUSSAIN: Thank you, Ariel, for a comprehensive overview of how the work by GNSO Policy Working Groups are integrating the work of the community on developing Root Zone LGR into the policy for gTLDs.

So with that, let's turn to Kenny who chairs the ccNSO's Policy Development Process Working Group on IDNs and request him to

share how the working group is considering integrating the work on Root Zone LGR into ccNSO policy. Thank you.

KENNY HUANG: Okay. Thank you, Sarmad. Basically, I need to change my hat from co-chair Chinese Label Generation Rule into chair of ccPDP4 Working Group. It's my pleasure to introduce about how the Root Zone Label Generation Rule is being considered by the ccNSO. I can provide you some example. If you consider Root Zone Label Generation Rule, it's a technical requirement for the root zone, then on the ccPDP4 will be considered a policy requirement for other ccTLD operator to how they operate the IDN ccTLD. Basically, it's more related to the policy requirement. In addition to consider the requirement, basically, the working group was split into three sub-working group.

The first sub-working group working on their variant management, therefore, focus on how you manage your variant, how many variants you have for selected IDN ccTLD. Second, working on the de-selection requirement. The De-selection Sub-working Group, for example, de-selection you can consider, that will be considered as a retirement policy. We identify certain criteria and if the criteria matched then the IDN ccTLD need to activate the retirement policy process to go through the policy development process. That will be in the second subgroup. Third subgroup will be String Similarity Subgroup. We also need to focus on how can we deal with a string similarity.

In addition to the combination work from the three sub-working groups, basically IDN ccPDP4 also need to work on how can we

implement that kind of policy at a given timeframe. Basically, the proposed timeframe, we are given an initial report in January of 2023 that we will give the first initial report. That's pretty much the timeframe we propose to do.

In addition to that, because we just realized also EPDP has the requirement for single character top-level domain, so based on that demand, we'll probably need to bring back all the experts together to consider if any technical consideration we need to propose to that kind of requirement, not only technical requirement but also need to combine with the language experts and the other experts as well to consider what kind of policy recommendation we can offer to the single character top-level domain requirement. So that will be so far the ccTLD ccPDP4 trying to achieve. Most of the tasks will be expected to be complete about an initial of next year. Thank you.

SARMAD HUSSAIN: Thank you, Kenny. The ccNSO work is also on its way to integrating the work done by the Root Zone Label Generation Rule panels into the policy work as well. As one of the aspects identified by the resolution by the ICANN Board, there was a request to both GNSO and ccNSO to collaborate with each other for consistent implementation. May I request for maybe short comments on behalf of GNSO and from ccNSO to share how the two SOs are collaborating in these IDN-related policy development processes?

KENNY HUANG: Thank you. As we mentioned, working on ccNSO, we have assigned a liaison working with EPD process. Anil, our co-chair of ccPDP4 is a liaison working with EPDP. So he will keep updated information from EPDP also to bring the information back to the community as well. In addition to that, basically, we need to harmonize the work with GNSO but that doesn't mean every single policy-wise should be identical. Because ccNSO and GNSO, naturally, they're a little bit different. So it doesn't mean that policy-wise should be identical, but we can harmonize a process in terms of policy requirement in terms of how to integrate Root Zone Label Generation Rule as minimal requirement for policy requests by the technical community as well. I like feedback too.

ARIEL LIANG: Thanks, Kenny. Kenny mentioned the liaison from ccNSO for the GNSO EPDP. And likewise, the GNSO EPDP has a liaison to the ccPDP4. His name is Dennis Tan from the Registry Stakeholder Group. Of course, on the liaison level, they have a lot of coordination and monitoring each of groups' work. Also I want to note that the ccPDP4 and IDN EPDP have met a couple of occasions, I think, at least. The latest is in July. Then these two groups, they have compared and contrast their recommendations under the same topic and look at similarities and differences. I think this kind of session was very helpful for the two groups to understand each other's progress and identify discrepancy and understand the rationale. As Kenny said, it doesn't need to be identical, but at least we need to understand the reason for

differences and then try to harmonize the work. That's all I want to add. Thank you.

SARMAD HUSSAIN: Thank you, Ariel. Let's move on to the next slide, please. We are short on time, but if there is any question, maybe we can take one question from the floor. Please come to the mic. Then we'll move on to the next part of the session.

LEVY SYANSEKE: Good afternoon or good evening. Great presentations. I'm Levy Syanseke from Zambia, ICANN75 Fellow for the record. I think I have one question with regard to how the Latin script managed to get a lot of languages on board. My concern is I come from a country where for certain languages, the English language had to omit certain characters out of the alphabet to allow the language to actually begin to conform to those characters. So how then do you get to a place where you can properly incorporate other languages in the script? If we were to look at, let's say, the same tribes where certain characters in the alphabet were omitted, they decided to bring them back in the language and how they write them. Doesn't that bring a certain level of compromise? Then how does the integration clearly work in that regard? That's my question.

SARMAD HUSSAIN: Michael?

MICHAEL BAULAND: Thank you very much for the question. Actually, it was not so much to which characters we were able to include because there is a so-called maximum starting repertoire. This is a list of characters which in theory are possible to be included in the root zone and we were able to choose from that repertoire.

We have a linguistic expert in our group who is working at the university in language department about African languages. He was very good in helping us to find and decide on [inaudible] cases when it was related to certain characters. But if you think that one of the languages is not represented, you can, of course, approach us and we can maybe talk about that afterwards.

SARMAD HUSSAIN: Thank you, Michael. I think that also points out to a very good feature of Root Zone LGR, which is that it's not static. We have a Root Zone LGR version 5.0. But if the community feels that there are additional characters which need to be added, it is an incremental process and the community can reformulate the Generation Panel and submit an updated process. We continue to support any additional characters needed as languages and scripts also evolve over time.

With that, we'd like to move on to the next part of this session with request to panelists to stay here. But we'd actually like to invite the CEO and president of the ICANN Organization, Göran Marby on the stage, as well as the chairman of ICANN Board, Maarten Botterman. Please come on the stage and we'll give you the floor to thank the community for the great work they've done over the last decade or so

as well as share your views. Thank you. We will first invite Göran Marby to start.

GÖRAN MARBY: I was hoping that he could start, actually.

MAARTEN BOTTERMAN: In fact, I can. I want to thank you very much for giving me this T-shirt. It says "I develop Root Zone Label Generation Rules." I really cherish this shirt, but the "I" is not me, it's you.

On behalf of the ICANN Board, I would like to thank you all for the tremendous efforts you have put in the Root Zone Label Generation Rules, and that's the acronym RZ-LGR. The work to develop proposals on how to use various languages and scripts to domain names in a secure and stable way is incredibly important to ICANN and to the greater Internet community. Over the past nine years, a wide range of experts from 44 different countries have come together to form Generation Panels to diligently develop solutions for international domain names that delicately balance both technical and linguistic needs. The members of these panels have included linguists who provided expertise on how the specific script is used in different languages, technical experts who have broad knowledge on the technical implementation of the scripts, including the Unicode coding, the use of fonts and how to support scripts in different operating systems, domain name industry and computer networking professionals who brought technical insights into Domain Name System and IDNs. This included representatives from both gTLDs,

registries, registrars, and ccTLDs. And not to forget the local communities and end users will help to guide experts on end user requirements for multilingual Internet.

The amount of volunteer work and effort from these Generation Panels was enormous. Since the first Generation Panel was formed in 2014, you've created 17 Generation Panels, engaged 270+ script community volunteers with more than 10,000 volunteer hours, opened 30+ public comment proceedings. In total, you finalized proposal for 26 scripts, analyzing more than the 386 languages and covering many more. This work would not have done without your dedication and expertise, and it's truly a testament to the multistakeholder model at work.

So I'd like to thank and I want to recognize each script Generation Panel for their significant contributions to RZ-LRG over time. I kindly request that the members from each Generation Panels here in the room or in the Zoom Room to raise their hands and I'll call out your panel. The Arabic Script Generation Panel, the Armenian Script Generation Panel. Aren't you going to raise your hands on every script, Sarmad? Chinese Script Generation Panel, Cyrillic Script Generation Panel, Devanagari Script Generation Panel, the Ethiopic Script Generation Panel, the Georgian Script Generation Panel, the Greek Script Generation Panel, the Hebrew Script Generation Panel, the Japanese Script Generation Panel, the Khmer Script Generation Panel, the Lao Script Generation Panel, the Latin Script Generation Panel, the Myanmar Script Generation Panel. I know the Neo-Brahmi Script Generation Panel is in the room. The Sinhala Script Generation Panel

and the Thai Script Generation Panel. Thank you. Please join me in a round of applause to recognize those who have contributed to the work.

I would also like to thank my former and current colleagues on the Board, IDN-UA Working Group, including Edmon, who have provided oversight on this work since 2011. Even though the Root Zone LGR is finished for all the active Generation Panels, I hope this is not the end of your participation in this community work and that you would continue to stay involved in IDN and Universal Acceptance with ICANN. The ICANN Board and the Board's IDN-UA Working Group will continue to promote Internationalized Domain Names and Universal Acceptance to all domain names and e-mail addresses so that we can make the Internet more accessible for users around the world. This is on top of our agenda. Thank you very much.

GÖRAN MARBY: So one of the advantages of following Maarten is that he said most of the things I was supposed to say so I don't have to—

MAARTEN BOTTERMAN: I just said that the last meeting where you spoke first.

GÖRAN MARBY: Oh, okay. Thank you. Maybe we should coordinate the speaking notes a little bit. I got to keep it very short and then say that and then speak for 15 minutes.

But essentially, what you've done is that to create a building, you need to have building blocks, you need to start somewhere. What the work you're doing, the continuous work of this is one of the most important building blocks to build that house. Without it, it wouldn't be a house. It wouldn't be anything that we can do something about.

Next time, can you make bigger T-shirts for people with normal size? This is an extra large. It's probably my problem, not the T-shirt sizing.

But the thing was that it's the sort of unsung heroes. A lot of people are going to get credits when the Universal Acceptance makes it possible for people to actually go online using their own keyboard narrative and all of that. But to think of it that it actually starts here, without putting those things into the actual core of the Internet itself, nothing else would happen.

Once upon a time, Cherine said, "Budget without money is just dreams." And the thing is that I was thinking about when I read my speech, which I'm reading right now in my head, is the fact that we can talk as much as we want about the inclusiveness of the Internet and diversity of the Internet, but if we don't add those building blocks into it, it's just dreams. So what you basically are doing is to make my dreams come true, and for that, I'm very grateful. I will wear this when I've been on a diet. Thank you very much.

SARMAD HUSSAIN: Thank you, Göran, Maarten. We would like to take a picture of all the Generation Panel members. We'd request everybody who's participated, contributed, to please come on the stage. We'll also

request people who are online with us in Zoom to please turn on their cameras. If we can stop maybe sharing the slides and share the video of people who are online, then let's maybe gather at the stage here and we'll all pose for a photograph. Thank you.

After the photograph, we'll have a chance to socialize with each other. We're sorry, because of the COVID constraints, we cannot offer any cocktails for that socialization. But we hope you'll still take some time to meet people from your script communities who've put in all that effort to develop the Root Zone LGR. Let's come on the stage for the photograph. Thank you. We can stop the recording. Face the camera.

[END OF TRANSCRIPTION]