Transcription ICANN London
IDN Variants
Saturday 21 June 2014

Note: The following is the output of transcribing from an audio. Although the transcription is largely accurate, in some cases it is incomplete or inaccurate due to inaudible passages or transcription errors. It is posted as an aid to understanding the proceedings at the meeting, but should not be treated as an authoritative record.
On page: http://gnso.icann.org/en/calendar/#jun
The recordings and transcriptions of the calls are posted on the GNSO Master Calendar page

David Cake: Okay, I'll take the chair for this one. So our next item on our agenda is - is IDN variants. Can we start the recording please? Started?

Yeah, and I would - yeah, like to - could we start introduce us to IDN variant issues please? Who's here to take the call? Yeah.

Sarmad Hussain: Yeah, thank you. My name is Sarmad Hussain. I'm going to be presenting this update on IDN variant TLD program on behalf of ICANN. Next slide please. So just as a background when we talk about IDNs we are actually - in the (unintelligible) domain names we are actually talking about domain names which other - or which are in other scripts like Chinese, Arabic Cyrillic, outside general US ASCII domain names which we already have.

And as for as variants are concerned - I'm sorry, these slides don't show up well in the PDF form - but we are talking about the top level domain names which do different labels as far as the Unicodes are concerned, Unicode points are concerned. But they are either visually or otherwise exchangeable or considered same or confusable by the end users.

Next slide please. So to support the IDN variants in the root zone, the ICANN community, at the direction of the Board, undertook several projects to study and make recommendations on the viability, sustainability and delegation of these variants in the root zone.

And way it has worked is that there have been many projects, which I will give an overview of, over past few years. And these projects have looked at

how it may be possible and currently we are in the process of now implementing these variants.

Next slide please. So this started back in 2001. The first step was to really look at what are the - how variants are actually defined across multiple scripts. And there were case studies done for six scripts as a representative set of scripts on Arabic, Chinese, Cyrillic, (unintelligible), Greek and Latin. And these are case study basically identified what various kind of issues are related to variants.

These issues were eventually, in 2011 and '12, integrated into a consolidated report which was eventually called the Integrated Issues Report for Variants.

And based on the problem definition which came out of the integrated issues report a solution phase was devised which has, in the - which was as Phase III looked at various directions in which the solution had to be implemented, what needed to be defined was a label generation rule set for the root zone which basically included which code points from across the various scripts are allowed in the root zone, which code points are variants of other code points and whether there are any label level restriction rules.

So there was a specification which is currently out and is being used by the community to specify their proposals for LGR. There has also been a process which has been defined so basically all communities have been invited to develop a proposal for their script. I'll talk about that in a little more detail later as well.

And initially in Phase 3 there was a process which was defined to bring these communities together to develop these proposals. And then there was a study which was done to see what are use implications on how - user implications on having TLD variants in the root zone. Next slide please.

So that's - so we've included details of each of these projects but I'm going to skip these - these are there for you to look at and if you have any queries we

can come back to it. So can we skip next slide please? Next slide. Next slide. And two more slides. Yeah.

So at this point we've done the background homework. The issues were identified, the process was identified and we are currently in what we call Project 2.2 which is the implementation of label generation rule set.

What that means is that each of the communities, which want variants in the root zone and otherwise as well, for which there are TLDs they would need to come together and define the set of code points which are allowed for forming the top level domains and define whether they are variants for any of these code points or not.

This project started in June 2013 which call for the integration panel. Integration panel is independent panel which is going to take proposals from community generation panels and integrate them into a single language generation rule set for the root zone.

So integration panel was the first one to be seated in June last year. And in July there was a call out to form script level generation panels. In February this year Arabic generation panels - so these are script level generation panels, not language level generation panels - so Arabic generation panel was formulated.

Integration panels released the first maximal string starting (unintelligible) which is the - in a way a super set of code points which may be allowed and each of the - in the root zone - and each of the generation panels are then - are supposed to select from these - this super set and determine which sub set of the super set relevant for that particular script.

So the MSR was released yesterday well the initial MSR was released in March and there was a public comment period. And after public comment - (integrational) public comments the final version of MSR has been released yesterday.

The Chinese generation panel is now also being formed; the proposal has been received by ICANN. And we are currently in the process of outreaching to other communities to formulate other script generation panels as well. Next slide please.

So this - so this is I think one of the things we do want to extend to the communities. We need the communities to come up and start working towards creating their generation panels and devising the proposals for the generation for their particular scripts so that they can be considered by the integration panel to be included in the root zone to enable the TLD variants.

So as I said, Arabic generation panel has been formulated; Chinese generation panel - script generation panel is also underway and there's also some work already for Korean and Japanese. And (Neo Brahmi) which includes many of the scripts listed there. But formally as far as the proposals are concerned we've only received proposals from two script communities. Next slide please.

This is the - just a summary of maximal string - maximal starting repertoire which has been released through the integration panel work which includes 22 scripts and it's short list 32,790 code points out of the 97,973 code points which are currently available in Unicode 6.3 which are p-valid or (unintelligible) ORJ as - and are candidates - potential candidates for the NGR.

Next slide please. So going forward I think one of the most significant things for us is to get the generation panels together, as I said, get the community involved, get - and assist them in any possible way to develop proposals for the LGR which are eventually considered by the integration panel.

In the first version of the maximal starting repertoire there was some scripts left out and the reason for leaving those scripts out was that they were not really applications for those scripts. So just to get I guess something out as

soon as possible so the generation panels can start working on them. MSR1 focused on scripts which had clearly applications.

We are now going to work on the scripts - other remaining scripts from within the Unicode and release MSR 2 which will then be complete maximum starting repertoire. And in the meantime MSR 1 can be used to develop proposals for the label generation rule by the generation panels - existing generation panels.

We are also aim to release the label generation toolset as soon as we get some proposals in from the generation panels. And as I said, we are continuing to reach out to the community to get these GPs formed and the proposals in so that we can get LGR 1 out.

And in parallel there is some progress on implementation of variants that is some homework which is currently being done internally however the focus has really been on getting the generation panel work going and LGR out. And I'd like to stop here. Thank you very much and I'll respond to any questions if you have.

David Cake: Thank you. Anyone have any questions?

Ching Chiao: Thank you, David and thank you for the presentation. This is Ching Chiao from the Registry Stakeholder Group. Actually I - and thanks again for the presentation. I think a good update. And at this point it seems that the community has had an expectation that this meeting is something for the IDN variant policy or - sorry the MSR and the LGR rules can be somehow finalized.

So there seems to be a little bit kind of delay here than the original - I mean, the proposed timeline. One thing that I noticed that it's, I mean, it's good to have a more inclusive phase approach to have everybody to comment on the repertoire itself or the LGR itself.

But one thing that actually Han Chuan and myself and others like Edmon in the room we've talked about is that many community members may come in late and then to start the process, you know, for some reason one thing that we found out in the Chinese LGR process and the communities nearly put together, you know, the finalized proposal but it seems that we also need to incorporate the comments from the Japanese and the Koreans and potentially - I'm not sure if the Vietnamese also need to be incorporated okay so that's no.

So we're still kind of waiting for the comments from the Japanese and, I mean, also from the Koreans which they may or may not have the right solution or right thinking towards the proposal. I'm just trying to get a sense what is ICANN, you know, standing at this point and then (unintelligible).

Han Chuan Lee: So with respect to Chinese - or the (unintelligible) script for the P 2.2 as you know the Chinese has written their proposal and we have been in talks with the - or rather the Japanese community has actually come out and said that they are in the midst of forming the generation panel and as such the Koreans have also come up to say that they are actually forming a generation panel and they have been in communication with ICANN staff on that.

In fact during this meeting there are coordination meeting being set up for this (unintelligible) to sort of work. And even before this meeting, last month - or earlier this month in Shanghai the Chinese had started engaging the Japanese and Koreans to look at the coordination work necessary for the Han script and the integration panel was also present.

And I think everybody has a very good understanding of like how coordination work that's required between these three community. So in terms of progress for the Han script I think it is moving quite well and well ahead of other scripts as well.

Sarmad Hussain: Thanks, Han Chuan.

((Crosstalk))

Man: I have a question. Are you limited to work only on the TLD, not on secondary level - second level?

Sarmad Hussain: Yes, that's correct.

Yoav Keren: Okay so I must say that I'm quite surprised because we had several discussions about this work I think also in the Council. And I remember I personally, and some other people, I think also from the Council, recommended that the work will not be limited to a TLD.

I don't know if everyone noticed the new gTLDs have already been - started implemented and chosen and submitted for so we currently don't have any, it looks like no variant issues being - happening.

But the things - the problems are going to be really, really, really bad on the second level if we don't take care of that very close very soon. The issue is, you know, we're going to wake up in two, three years once IDN TLDs are implemented and having so many phishing and pharming attacks using variants if there is no solution. So this is taking too long and I don't understand why is it limited to the TLDs? I don't know who needs to change that but that has to be changed.

David Cake: Yeah, thanks. Yeah.

Edmon Chung: Edmon Chung here. So I guess building on what Ching and Yoav were saying, I have three particular comments. First of all just directly on the presentation this - it's kind of called the IDN variant TLD program but as I understand it the LGR is required for non-variant languages and scripts as well.

So I think, you know, this is one of the things that is very important as you - I think you pointed out that only the two scripts are now being worked on,

Arabic and Chinese. Obviously they are more variant-oriented. I want to make sure that the community understands that this project is not restricted to IDN variants; it's all the scripts and languages need - we need volunteers from all of them.

So I don't know - I understand that the program is called IDN variant TLD program and I understand that that's that the team is called. But you've got to find a way to let people know that, you know, other scripts and other languages need to come. So this is number one.

The second one I'm going to, you know, I've said this many times in previous meetings: where's Project 7? Without Project 7 we can't move forward. We need more information than Project 7.

By the way, what I mean by Project 7 is with the IDN variant TLD program right now the Project 2.2 that the team mentioned about is the variant table, you know, how the mechanism works in selecting variants.

But then Project 7 in the project, did I get the number right? Project 7 is supposed to be the process for which it's delegated. So the actual ICANN and eventually IANA processes to actually get it delegated. Without that even if we have all the tables in the world we cannot actually delegate the TLD. So that is as important right now as I think Project 2.2. It's - I think that's an important point I think.

The third item that I want to bring up is building on what Yoav said there is a mechanism on - from ICANN on second levels and that's called the ICANN IDN Guidelines. And I think I would encourage the team to re - it's about time we re-look at that, the ICANN IDN Guidelines with all the case studies and all the experience that has been learned over the last few years it's time to revisit the ICANN IDN Guidelines and also to relate that to the universal acceptance work that is being done right now.

And also, you know, especially with Sarmad here, I think that should also relate into the user experience study that was done before. Universal acceptance, the ICANN IDN Guidelines in terms of the - and also in terms of the user experience. So I think that is as important. You know, those three things are pretty important I think in terms of the overall program. Thank you.

David Cake: Thank you for some really helpful comments, Edmon. Do we have any more questions? Yeah, I'm not...

Man: Okay (unintelligible) Ching's question, so the first question is the fact that other languages or other scripts needs to become, I think we're well off and I think in the (unintelligible) mentioned actually we are doing IDN outreach to the other scripts as well.

And which is why for MSR 1 we are focused on the scripts that were right (unintelligible) applications and we are actually in touch with the variants committee to make sure that, you know, delegation panels are being formed, that the case of those 17 scripts that we identified. MSR 2 would then take care of the other scripts that - that were not taken care of in MSR 1. So we are working together with our (GSE) folks as well to do IDN outreach to get all (unintelligible) from as well.

David Cake: Thank you. Ching. Yeah.

Ching Chiao: Okay, follow up own that Edmon just said to - so are you in the position - so in the position to tell us what would be the timeline for Project Number 7 since it's really related to the policy work that the Council potentially need to work through that the potential amendment of the (AGB) if the next round is going to adopt the varying parts.

This is - we are seeing, I mean, this, I mean, definitely estimated of good potential number of works that needs to be done, I mean, the delegation part, the cost part and, you know, and also I understand that you - I mean, you don't feel like going there but still we need to resolve the confusingly similar

extreme (unintelligible) the variant can, you know, really actually go beyond the principle confusingly similarity string.

So, I mean, going back to my, you know, the question is that are you in the position telling us, you know, what would be the estimated timeline for Project Number 7?

Sarmad Hussain: So at this time we don't have - we do not have a timeline for P7. And the reason for that is that all the focus and energy has been put on getting the generation panels together and getting the LGR out which is eventually a prerequisite to P7 project - the implementation project.

There is some internal work from what I understand. I'm just joining the team. I'm - so there - from what I understand there is some internal work. And obviously the community is going to be consulted as that work proceeds.

But at this time, again as I said, we do not have a timeline. Our main focus is to get the LGRs. I will take this opportunity to also add that we have two sessions on Wednesday focused on this program. And we are actually seeking community feedback on how to organize the releases of - and (unintelligible) of MSR and the label generation LGR 1 release.

And that is, you know, we need to understand and talk with the community on when they can put in the proposals for LGR so that we can plan an LGR release accordingly. And we'll use that feedback also to eventually in a way plan our P7 implementation work.

David Cake: Yeah. Thank you. Sorry, I can't see you, you're silhouetted in (unintelligible).

Man: We've got two questions in the chat and they are - (Jason) is asking, "Who's funding the panels and their work?" Who's funding them? And also, "Which members of the Unicode Consortium are involved?"

Han Chuan Lee: Han Chuan here. Allow me to answer those two question. So the first question is who is funding the generation panel? The generation panel is a community-driven work made up of volunteers in the community so ICANN is not funding anything. But we will provide support to the generation panel in terms of, you know, if they need to have a mailing list we will set out a mailing list for them.

And to facilitate any kind of meetings that they want in terms of like (unintelligible) assist in that manner. So we provide support but not funding. So for the who are the Unicode experts on the integration panel we have two members who are (unintelligible) one is (Michel).

So we do have Unicode experts on the integration panel and as well for the generation panel we also encourage them to make sure that they also had their own language experts who are also familiar with Unicode - Unicode experts work on (unintelligible) were part of the panel. And in both the Chinese generation panel and the Arabic generation panel (unintelligible) there are such experts in the panel to take care of the Unicode points.

David Cake: Thank you. Thomas, you had a question?

Thomas Rickert: Thanks. I guess my question is a little bit broader but listening to Yoav's question earlier about the second level and listening to your comment, Ching, about similarity issues, with the various initiatives going on in parallel do we know for sure that all aspects of other scripts, regardless of what level and what is the variant question or the similarity question are being addressed somewhere and that there's sufficient awareness between these initiatives of what the other groups are doing so that we don't find out in a couple months or years' time that we've failed to look at certain aspects of this whole complex.

Han Chuan Lee: Quick reply is that as we've talked about the IDN Guideline has been there for a while for the second level and is - the LGR, I'm sorry, the VIP project

talk - the primary to address the top level. You probably recall that in the Singapore meeting that we closed the JIG, the JIG Working Group.

But I did have the proposal to have somebody in the community potentially, I mean, maybe Edmon or someone else could serve as a liaison for - to closely watch the development of either the VIP and also the potential necessity for the revision of the IDN - I mean, the Guidelines. We haven't take any action on that in fact.

Edmon Chung: I'll just in I guess. Edmon here. So actually on what Thomas raised and Ching mentioned I think last time when we - in Singapore when we had the wrap up meeting, Jonathan might remember, there was a dialogue with Ram about what the situation is.

In terms of the overall program I agree very much with Thomas what you raised is that there needs to - there should be some way of looking at making sure that there are people in those groups that can relate to all the other policy stuff that is happening at ICANN.

And in the design of the LGR actually one of the requirements supposedly was that the integration panel would have advisors that include ICANN policy expertise.

So far at this point - so far from what can be seen in the makeup of the LGR team it's very focused on linguistic and technical and Unicode aspects. And that is the missing piece right now which we raised in Singapore and perhaps we should follow up on.

And in terms of the dialogue with Ram, yes, I did follow up with Ram. I apologize, I dropped off my table and I didn't respond directly back to the Council. But, yes, I think Ram understands where that missing point is now and, you know, I think we should pick up on it and follow up on it.

Jonathan Robinson: Edmon, just so we don't lose track of that can we just capture what we're expecting, you know, what we're expecting any particular person or...

Edmon Chung: So very concisely right now the LGR, which is the Label Generation Rule Set Team is divided into two main groups. One is what is called the Integration Panel which integrated different scripts and each particular script have a generation panel.

The design that was created in terms of the document that create the LGR team supposed to have in the integration panel to have advisors that have policy expertise. So far we have not seen that yet. So what I guess we should follow up on is remind - the Council should remind the staff team that is one of the expected elements of the integration panel. And if they need help identifying people or volunteers willing to serve in that role we can follow up with that.

David Cake: Thank you, Edmon. Thomas, you had a question and I think we're out of time for this item so unless anyone's - this if the final call to be on the queue if anyone has a question after Thomas.

Thomas Rickert: Yeah, maybe this is something for the Council to discuss with the Board and maybe additional resources need to be set aside for this. But I think that it should be in everybody's interest to keep a close eye on this. I mean, this is closely related to the universal acceptance issue, at least in my view.

And when this program is going to be evaluated the questions will be has the user base at the global level benefited from the program? And if there are issues such as that unresolved or even if they're unresolved, if they didn't get sufficient attention I think that's going to be bad in the evaluation process.

So I think we should make this a priority action item and potentially discuss what the Board or at least re-table it at the wrap up session and maybe Ching and Edmon can help by, you know, sending a small paragraph on what their asks are to the Council list so that we can take it further and operationalize.

David Cake: Yeah, thanks for that comment. I think that's a good point. And this is clearly an issue - clearly an area where a lot of quite specialized knowledge is needed often and it's not always coming up through our standard volunteer process I think so that's well worth commenting.

Daniel.

Daniel Reed: Yeah, beyond the technical issues of implementation, which are absolutely crucial for the reasons just mentioned I think this is an issue where investing appropriately is really crucial because the underlying issues are tied up with the broader context of Internet governance and multi-stakeholder participation.

And the more that we can push this forward in an inclusive way the more we bring everyone's voice to the table. And I think that's a really crucial maybe understated but essential aspect to why this matters so much.

David Cake: Yes.

Han Chuan Lee: Han Chuan here again. Just address Edmon's question. I think (unintelligible) Project 2.2 is indeed a technical and linguistic implementation. There is no policy to be decided in P2.2 is basically what is P2.2 basically the scripts - the (unintelligible) come together and look at all the core points for the (unintelligible) that could be used as TLD, all right. And then they decide if for a particular point is there a variant? And if there's a variant should this variant be locatable or should this variant be blocked?

So there's no policy consideration in P2.2 at all; it's basically the linguistic and Unicode is still coming together and looking at, you know, what should be a disposition, all right, in other words block or locatable for a particular critical point.

So that's all P2.2 is; it isn't purely an implementation (unintelligible) and P7 basically takes whatever the label generation rule, which is like the same thing, you apply for a string in the top level (unintelligible) say okay for this particular is there a variant?

If it's a variant is the variant locatable or blocked, if it is blocked fine; it is blocked then it's locatable then we look at whether it should be - or how should the (unintelligible) process be (unintelligible). And in terms of string the confusingly similar strings, I think that's already been addressed in the (unintelligible) unless the process changes or there's a change in policy we should be following whatever insight the (unintelligible).

So again the outcome for P2.2 and P7 has nothing to do with policy, it's purely an implementation project.

David Cake: Thank you. Do we have any - anyone else wants to speak on this? In which case I think I said we're over time for this one and I thank everyone for one of the more helpful useful discussions of this long ongoing topic. I think it's been really productive.

And with the conclusion of that topic I will hand back the chair to Jonathan.

Jonathan Robinson: Thanks, David. We can stop the recording then on that and we'll pause for a moment while we get set up for the next session.