

Transcript
IRD Outreach Webinar
Wednesday 16 February 2011 at 2000 UTC

Coordinator: This is the Operator; I do need to remind all parties that the call is being recorded at this time. If you have any objections you may disconnect. And thank you, you may begin.

Steve Sheng: Hello, good morning, good afternoon and good day everyone. Thank you for joining today's call on the Internationalized Registration Data Working Group. My name is Steve Sheng. I'm the ICANN policy staff supporting the working group. The working group has asked me to give an update of the interim report to the community.

There are two goals for this presentation. First we want to brief the community of the working group's work. And second, more importantly, the working group has just issued the interim report that is in public comment.

And we will very much like to hear the community feedback on some of the proposed recommendations. In this presentation I will introduce the background of our work and some of the recommendations and at the end open up for questions.

So as a introduction what we're talking about, IRD, the internationalized domain name, the IDN, the labels, there are guidelines today already exist for those - how they composed and submitted. There's the IDNA guideline for use of IDNs in applications.

However as of today there are no standards exist for the submission and display - I'm sorry, there's a question there, oh - there's no standard exists for the submission and display of the domain registration data in directory

services. So the domain registration data in the ICANN context is commonly referred to the Whois data. And the directory services most frequently refers as the Whois services.

So realizing these problems the IRD was formed to study the feasibility and suitability of introducing submission and displaced specifications to deal with the internationalization of registration data. So there are two questions here, there is the suitability question.

So traditionally all the domain registration data is in US ASCII. So going into the IDN era do we want to keep that in US ASCII? And if we allow multiple scripts in the Whois data, you know, how do we kind of avoid a Tower of Babel effect where we can think of, you know, many, many scripts in Whois but registrants cannot understand and contact each other. So that's the suitability question.

The other part is the feasibility question is - is the Whois protocol able to support the submission and display of the internationalized registration data? And if not what kind of recommendations, you know, we should do - we should do about that. So those are the two questions.

So just some background, domain registration data are the data that registrants provide at time when registering a domain. It includes, you know, the following items on the slide. And this information is usually displayed using the Whois protocol at least for gTLD registries and registrars. A subset of the data is also displayed, you know, used in ccTLDs registries as well.

So the working group deliberated about, you know, how to internationalize the data. And one consensus was various data elements could be separately internationalized. So for some of the elements there are already internationalized in standards for that.

So for example, you know, there's telephone, fax, you know, there's the ITU standard, the email, there's a new RFC, you know, study in these internationalization. So the working group just used these standards and - that I've described here. We would also like to seek the community's feedback on, you know, the community's thought on these approaches.

The key question the working group would particularly like to seek feedback from the community is internationalizing the contact data. So the contact data, you know, we have the - the domain owner contact, the technical contact and the administrative contact; this includes names, organizations, postal addresses.

We have discussed four models to internationalize the data. Before going to detail of these models I want to highlight some of the principals that guide our discussion.

So the first principal is we want to accommodate, you know, different scenarios where user or registrant may want to or wish to submit data in his or her own languages. But on the other hand we also want to avoid the Tower of Babel effect for registration data.

The second principal is we also want to balance between cost and usability of the registration data. So for example, you know, one of the models discussed, you know, is translation and - versus transliteration.

You know, although maybe translation is more accurate there's also cost implications so there's a tradeoff between cost and usability. These are great policy questions.

And the third guiding principal is to consider the needs of the human users which are regular Whois users as well as legitimate automation. That means applications that parse and analyze the registration data.

So balancing, you know, with these guiding principals in mind the working group has come up with four models. So the first model is more or less the status quo today is registrants provide domain contact data in a must-be present script.

You know, we have a Russian example and Chinese example here. It's pretty much the status quo today. Model 2 is registrants provide data in any registrar accepted script. And the registrars provide a point of contact when there's issues, you know, when registrants need to contact a particular domain name holder.

So in this example we see a registrar - a fictitious registrar, you know, with a URL, (neat).ru, you know, put his contact information, phone number, fax and email number. But the actual contact information for the domain name is in Russian; it's the same for the Chinese example. So that's the second model.

The third model is registrants provide data in any registrar accepted script and it's up to the registrars - responsibility and - to provide transliteration tool and publish in a must be present script. So it is of somewhat similar to Model 1 but, you know, kind of going one step beyond. Registrars provide a transliteration.

Model 4 is very similar to Model 3 but here instead of transliteration is registrars provide translation tools to publish in a must be present language; in this case is English. So let me compare - we've been talking about translation and transliteration. Let me compare the two for you.

We have a convenient example in Model 1. So in Model 1, for example, in the Russian - in the Chinese example the country is, you know, the People's Republic of China; that's the correct translation of the country name. But transliteration of the country name is different it's, you know, is a phonetic transcription. You would say Zhonghua as, you know, how you mean the country of China.

It's the same for, you know, you know, (Sinjin) Guangdong Province, that means a province within China. But, you know, that's in translation but in transliteration you would (Sinjin) Guangdong (shep). So that's kind of a - kind of very high level difference between translation and transliteration.

So those are the four models. To summarize we can think of these models in a two by two table. You know, for the column we have transliteration versus translation. For the rows we have, you know, registrant's responsibility or the registrar's responsibility. And Model 2 is not in the two by two table it's - because of the registrar provide merely a point of contact.

So the working group has issued a set of preliminary recommendations. The first recommendation sets up a basic service requirement for the Whois protocol in the IDN - in the IRD world that it must be able to support - accept query of either U-label or A-label format and must be able to display results in both U and A-label for the domain names. And there must be - something must be done to handle IDN variance in the responses.

The second recommendation touches on the various data elements to be internationalized. For the interest of time I won't go into it. These are in the previous slides. I would like to come back and really see (unintelligible) or discussion on these three questions.

Which of the four models for internationalizing registration contact data is most appropriate? And are there any other models that the working group should consider? And so that's the first question, the four models.

Second which of the preliminary recommendations if any are feasible? This refers to both setting a service requirement for the Whois protocol as well as, you know, internationalizing those specific data elements. Are there any related recommendations the IRD working group should consider?

The third question is issues of language tag. The question is there a need for the contact information to be in multiple languages and scripts? Let me qualify that statement, during the working group discussion, you know, we were discussing a point to see is there a need - a use case in particular for example for contact name for the domain holder contact information in one language but administrative contact or technical contact information in another language?

Or even further are there any use cases where for a given contact (unintelligible) the name and the address those could be in different languages and scripts? So that's - that has some implications - technical implications on how the data should be tagged.

So those are the three key questions the working group would really like to seek out (community) discussion. I think with that I'll conclude the presentation and you're welcome to - questions.

Before I do that, Jim, do you have anything to add? Jim is our co chair for the working group.

Jim Galvin: No, thank you Steve, nothing to add.

Steve Sheng: Okay thanks. Okay any questions? Let me see. I saw some discussion on the chat room. Could we have the PowerPoint presentation? Absolutely. Oh, Paul raised his hand. Paul go ahead.

Paul Diaz: Hi Steve, it's Paul Diaz from Network Solutions. I'm waiting for a response and Owen has given me something here. From a registrar's perspective I think looking at this group - let me take a step back - it's largely populated by registry, operator reps or others in the technical side of the community. There's only one registrar rep that I'm aware of.

And my colleagues are very concerned that this - any of these proposals - the four models that have been suggested sort of have a unfunded mandate feel to them in that, you know, one way or another it's registrars who are going to be bearing the cost of providing transliteration or translation services.

We just came back saying well why aren't registry operators in the mix here? You know, I saw Chuck's note saying that registries don't have a direct relationship with the registrant. While that may be true I guess my concern is that look ultimately registry operators want to have as clean or perfect information in their database as possible.

It strikes me that if you - if we leave the transliteration - translation in any of the four models proposed you run a high risk of garbage in garbage out because there's going to be any number of different perspectives on how to translate or transliterate the information.

I'm just - I'm sorry, as a registrar hearing that, you know, that this group is recommending that registrars or the registrants directly do all the heavy lifting on this seems to me just wrong-headed. And I'm really wondering isn't it in the registry operator's interests to provide a standard tool to translate/transliterate data that comes in so that ultimately, you know, the information that appears in their database will be accurate or as accurate as it can be?

Steve Sheng: Thank you Paul. I've taken note of your questions. I think these are very important. With regards to accuracy the - I was giving a previous presentation to the At Large community and they raised this questions as well.

My sense is at least we don't want the Whois data to be more inaccurate because of the, you know, internationalized registration data. So I think on that point the incentives is aligned.

I mean, your other points - some other points the incentives may not be so well aligned so I really welcome you to submit a comment through the public comment and the working group will take up that and consider that carefully.

Chuck, go ahead.

Chuck Gomes: Thanks Steve. Paul, you're absolutely right that registries are concerned about the validity of the data that's in there. So - and I think that's something that we as registries and registrars would want to work together on.

One of the things I would suspect will happen though whoever does the translation or the transliteration is that there's going to have to be corrections made, adjustments made because we all know that there are translation errors; I'm sure there would be transliteration errors too.

And so there may have to be communication between - with the registrant. So that would put registries - if they have that primary responsibility at a disadvantage in fact an awkward situation where we would have to go to the registrar and then the registrar would have to communicate and so forth.

These are all really good issues. I'm not suggesting that registries would not want to work with registrars and the community in terms of the best overall solution but that's where the - that's at least one area where the fact that the registry has no contractual relationship or doesn't even know for sure who they should be contacting without using registrar Whois information would be a challenge that we need to work on.

Steve Sheng: Thank you Chuck.

Jim Galvin: So this is Jim Galvin from Afilias. And speaking in part as a co-chair of the SSAC side of the co-chair of this working group one thing I want to be careful about is to point out that the quality of the data itself in terms of its accuracy is actually not in scope for this working group. It is true that the quality of Whois

data is quite generally a very serious issue but that's not an issue that this working group is speaking about.

The second thing I want to point out is the scope of this working group is about setting requirements for internationalized registration data. And that does not speak to precisely how this would be implemented. So Paul raises an important point which is how would this be implemented and where would most of the burden be borne in the implementation of these suggested requirements?

And while those are important - that is an important question and it deserves due consideration that question in and of itself is also not in scope for this working group. The purpose of this group is simply to specify what are the requirements that are needed so do we need a translation service? Do we need a transliteration service? Des the data need to be presented in a (double) - (two) quorum?

(Unintelligible) that and it would offer up one of the four models as the most likely candidate for how this could be done. So I would encourage you in making comments into the public forum to draw that distinction between your support for internationalized registration data which presumably we are all in support of.

We certainly would like registrants to be able to work in their own languages and draw a distinction between that and how you would like to see it implemented. If you don't like these four models because you think that they drive you towards a particular implementation ideally it would be most helpful to propose a different model and at the very minimum be very specific about what it is about a model that concerns you in terms of its expected implementation. Thank you.

Steve Sheng: Thank you Jim. I saw Paul want to respond. Paul.

Paul Diaz: Yes just a follow-up. Thank you Jim, excellent points. The question I had for you is should the community consider this - the output of this working group the first of basically several iterations. In other words you guys are probably taking the first then a subsequent group will take the findings and focus more closely on implementation issues or whatever.

Is that the way this should be read that this is not some ICANN rush to get IDNs out there because some particular interest may be, you know, pushing very hard for it - is it akin - or more akin to, you know, almost never-ending - let's hope it doesn't go there - but the never-ending Whois where there's just slice after slice as we try and work our way through all the various issues

Jim Galvin: Well to some extent, Paul, I agree with you. This is, you know, we're sort of biting off a little piece of something here and hoping that we can come to consensus as a step in the right direction.

My expectation is that this group does have, you know, fairly limited scope in that sense. So yes I would expect another iteration with some other group. Perhaps this group will be reconstituted in a different form with a different scope and different requirements in order to begin to address implementation issues or maybe that'll be handled differently all together.

This group does have a very specific mandate and the ideal goal is to simply meet that mandate, specify those requirements and then move on and allow this to move forward, you know, according to the natural order of things.

Steve Sheng: Thanks Jim. Chuck.

Chuck Gomes: Yes and these are really good issues that are being brought up. If I remember - if I remember correctly this group is clearly not a policy development body. So there's probably another step that will need to occur or steps in between the work of this group and whatever recommendations come out of it and the

implementation and that's policy development which would be handled by the SOs as appropriate.

Steve Sheng: Thank you Chuck.

Jim Galvin: Yes and this is Jim again. Thank you, Chuck, that's a very good point and you're exactly right.

Steve Sheng: Any other questions? (Unintelligible). No? Oh.

Paul Diaz: I don't mean to ask all the questions today but I'll just follow up one that I had in the chat that I haven't heard any of the experts answer. Can anybody help - are there existing translation/transliteration tools that already exist that experts feel do a very good job that they might even recommend using and one of the four models?

I mean, I guess it'd be a lot easier for people to get their head around what will be involved if there's something we can look to, start testing or tinkering with today as opposed to just this big bogey out there that oh we're going to have to do this at some point when it gets operational concerns.

Steve Sheng: So the tools that - let me think. There are two tools that I can think of. The first two - and both are open source I think. The Google translation has an API for both translation and transliteration so that it is kind of a web based Ajax tool.

The other tool I have to find out for you. There is also an open source. If you could just bear me with a second I'll find out for you. There is a tool called - there's an institute called ICU - they have a transliteration tool freely available. Let me put that into chat room.

So those are the ones I know we can - that question, you know, I'll take back to the working group and we can find and get you an answer for that.

Jim Galvin: Well Steve...

Steve Sheng: Go ahead.

Jim Galvin: ...this is Jim again. Yes, let me just comment that we have in our discussions, you know, talked about the existence of tools and, you know, and talked about how they work and what it means to have them and that kind of thing.

Again, you know, it's not within the purview of this group to recommend particular tools. I mean, it's clear that automated tools are never going to be 100% and I think it's fair to say that there was consensus on that point. You know, obviously these automated tools have varying levels of success which then goes to the issue of quality of Whois data and that of course is a whole separate discussion which is not in scope for this group either.

I don't really know what the final answer is going to be about translation and transliteration. And certainly if Steve wants to put together some URLs and make them generally available in some way as just sort of observations that we've made over time that are related to this work I think that would be helpful and a good thing to do.

But I want to be careful just to take note that it's certainly not a, you know, an official position of the working group to sort of point out tools and in any way try to be comprehensive about that or speak to their quality. Thank you.

Steve Sheng: Thank you Jim. Any other questions? Okay the public comment - the document is in public comment and the public comment will close on March 14. So I will strongly encourage you if you have comments to submit through the public comment and then at the end, you know, the working group will consider - oh, Chuck has another question.

Chuck Gomes: Well my question - sorry, I didn't - you could have finished your thought there. My question has to do with feasibility. Does the interim report talk about feasibility at all or just on what we've seen today?

Steve Sheng: Yes the interim report does talk about feasibility. Jim, could you - would you want to say a few words on that?

Jim Galvin: No, I can't think of anything to say thank you.

Steve Sheng: Oh...

Chuck Gomes: Well I plan on reading the report so that's okay but I just thought maybe there was another part of the presentation on that. But that's fine. I have not had time to read the interim report yet but I will do that shortly.

Steve Sheng: Okay. There's a section in the report that talks about that existing Whois protocol is not able to support IDN consistently. So that's I think one finding of the working group's work.

Chuck Gomes: And that of course is no surprise to most of us. And I guess that's where I was, you know, why I was asking the question. So you kind of left the impression though, Steve, that the existing protocol handles - does a better job I think than it really does with regard to IDNs. Was that intentional?

Steve Sheng: No it's not, I mean, we want to - it was not intentional at all. We just want to - these four models are currently the ones - open issues that the working group is discussing.

And the working group really want to hear from the community what they think about these models. That's why we put them in the forefront of presentation. It does not imply at all that, you know, the Whois protocol is able to handle them.

Chuck Gomes: Yes, thanks.

Steve Sheng: Thanks.

Jim Galvin: Yes, Chuck, I guess - so this is Jim again. I'll just add I guess I wasn't really sure what the context of your question was. I mean, I think we're all in agreement that the Whois protocol, as defined today, is insufficient to meet the needs of internationalized registration data.

And with that in mind, you know, we're out here specifying the requirements for something to replace that. If that's what you meant by feasibility, I mean, I hope that that answers your question. The purpose of these four models is, you know, giving some thought to what it means to internationalize registration data and the ways in which that might be achieved.

And so we're very interested in what people think about those things. I don't know if that adds to your feasibility question or makes it worse. Does that partly answer your question?

Chuck Gomes: Really, my question I think has been answered okay. I was really going back to the beginning of the presentation where there were two parts and one of them was feasibility and then - and maybe that was just the report and not this presentation so that's probably where I read it wrong. It's fine. The - my question was answered sufficiently.

Jim Galvin: Okay, thank you Chuck.

Steve Sheng: Okay so will make the slides available. I'm trying to think what's the appropriate mechanism to distribute the slides? So maybe we can put it on ICANN Website and give a URL to it.

Jim Galvin: Right, the working group - the working group has a place, I mean, the same place where you're going to post the recordings of these webinars just put the slides alongside it right.

Steve Sheng: Right okay will do that. And so will do that and if people want to request the slides you can also send a email - Liz, you want to say something? Say - yes...

Liz Gasster: So I just wanted to say that I thought Chuck's question was - or, you know, comment was actually a really good one because although this is an interim report and the purpose of this is to do outreach to, you know, so that the broad community understands what's being considered with these models ultimately going back to the charter for this group the board asked the working group to study the feasibility and suitability of introducing displaced specifications to deal with the internationalization of registration data.

So I think the question about feasibility specifically is quite appropriate for, you know, the group ultimately to be opining on to the degree that you have, you know, the expertise - the technical expertise to do that.

Steve Sheng: Thank you Liz.

Liz Gasster: And that might be a future step or, you know, etcetera. Understand that's not the focus here.

Steve Sheng: Okay. Yes we will take these comments in and if there's no more questions - are there any additional questions? If no more than we can end it kind of few minutes early. Okay I will put my email address in the chat rooms if people want to request slides you can email me.

And we also post the slides in the IRD working group's - now is the confluence wiki page. Okay? Well I want to thank you for your participation and hope you have a good day.

Paul Diaz: Thank you Steve.

Liz Gasster: Thanks Steve.

Steve Sheng: Bye-bye.

Chuck Gomes: Thanks.

((Crosstalk))

Gisella Gruber-White: Thank you (Holly). Enjoy the rest of your day.

Coordinator: Thank you, you as well.

Gisella Gruber-White: Thank you. Bye-bye.

Coordinator: Bye.

END