
STEVE SHENG: Good morning, good evening. Welcome to today's call on the 9th of April on the WHOIS Review Team Internationalized Registration Data Expert Working Group. On the call we have Sarmad Hussain, Nishit Jain,[inaudible], Jody, Jim and myself, Steve Sheng from staff. So Jim, over to you.

JIM GALVIN: Okay. Thank you, Steve. All of us are here in the Adobe Connect room. The agenda is there and up to date. I apologize for not sending that out ahead of time to folks. Steve's done the roll call. We'll take a step back for a minute and talk about the current status of our interim report here on requirements.

Then I thought it would be appropriate to jump right into talking about a data model. I was going to start that discussion by talking about the SAC 54 report. I think we probably won't use the entire hour, unless folks have questions or comments about that. Then we'll just move forward from there. So any questions about the agenda for today?

Why don't we go over to you, Steve? Do you want to tell us about the interim report, its status?

STEVE SHENG: Yes. Thank you, Jim. The last call was the comment on the interim report [inaudible] last Friday. We have received extensive comments from [Malki], which I went and updated the report. These were not major comments, but editorial clarifications and questions [inaudible]

Note: The following is the output resulting from transcribing an audio file into a word/text document. Although the transcription is largely accurate, in some cases may be incomplete or inaccurate due to inaudible passages and grammatical corrections. It is posted as an aid to the original audio file, but should not be treated as an authoritative record.

document. So, thank you [Malki]. Jim, you also made extensive editorial changes to the document, which does not change the recommendation or the proposed requirement, but brings more clarity to the document so I accepted all that.

We don't have any other comments on the document itself. There was a comment from [inaudible] on the e-mail, which I didn't know how to [fold] that into the document, so I did not [fold] it into the document.

So that's all. The document completes the last call process for comments, and I have sent it to start the ICANN publication process. That is the draft public comment open announcement, and then with the document to be approved by [inaudible] executive. After that it will be posted for public comment.

The public comment date that I currently intend is the comments will open on April 10th. The comments will close on May 1st. The reply period will open on May 2nd and the reply period will close on May 22nd. So from now, from beginning of April, to mid-April to the second half of May, that's where the community will have an opportunity to provide comments and feedback to the report. So that's a quick update. Any questions on that update? Thank you. Over to you, Jim.

JIM GALVIN:

Thank you, Steve. Let's come back to Sarmad's question that he asked on the mailing list. I actually did reply to it. I was replying to it just now while listening to music earlier. I think that Sarmad asks a very valid and important question about whether or not we should include requirements related to translation and transliteration.

What I said on the mailing list in response was I agree that those are important requirements that need to be included. The question that I have is should we be setting those requirements, or perhaps that should be an explicit ask of us to the translation and transliteration PDP group. Whatever decision they make, they need to account for any additional data elements along the lines of original source language, destination language, script considerations.

I'm serious, if anyone has any discussion about that, maybe Sarmad, since it was your question, I'll ask you if you have any thoughts or comments.

I'm sorry, one more point. Part of what I offered was if we think it should go to that other PDP group, it is a message that we could bring to that group. Sarmad and I are on both of these groups. I'm not sure if anyone else is on the two groups with us. But certainly, they meet on Thursday mornings; we meet on Wednesday mornings. It would be straightforward to bring a message to them tomorrow or sometime soon.

Anyway, Sarmad, any reaction? Then of course anyone else?

SARMAD HUSSAIN:

Thank you, Jim. I guess my initial reaction is that the PDP group [inaudible] for a particular element. But at the end of the day, [inaudible] the data model, what needs to be [inaudible], what that PDP requires is something probably [inaudible] within this group.

So I think they are looking at [inaudible] requirements, but also as the data model [inaudible] are concerned, implications of those [inaudible] PDP process [inaudible].

JIM GALVIN:

Samad, one question I have is about timing. If we're going to have the responsibility of making sure that the elements get into the data model – I think that's what I heard you say, that that is within our responsibility as opposed to the other PDP Working Group. That would suggest to me that we would need to wait to not be done until after that particular group was done. Would you agree?

SARMAD HUSSAIN:

Yes and no. It depends on how we want to do it. So another option is [inaudible] for each element if it is going to be translated, what are going to be the [allocation] requirements, if it is transliterated, what are going to be the [allocation] requirements and [leave it there]. [inaudible] that group decides for each element whether they need transliteration or translation. They can pick up the relevant spots and that gets added.

JIM GALVIN:

So you're suggesting we could define the data elements and such and leave it optional, and whether or not they're used would depend on requirements from the other group.

SARMAD HUSSAIN: Well, exactly, because the reason is that even if the PDP Working Group does [inaudible] requirements, then [inaudible].

JIM GALVIN: Okay, thank you. Comments from anyone else?

One additional question that I have is how do folks feel about bringing this message and highlighting this particular question to the TNT PDP? Well, tomorrow, in fact – if not later – if folks want more time to think about it. Do folks think that would be a good idea or should we wait?

STEVE SHENG: I think I agree with Sarmad that we should try to anticipate, regardless what their decision is, which would in some way anticipate the worst or the best, whichever we feel. That is to be able to have some data elements surrounding the language and transliteration standard. I think it's a good idea to raise it with the PDP Working Group.

JIM GALVIN: Thank you, Steve. Anyone else? Please, Sarmad, go ahead.

SARMAD HUSSAIN: So, yes, I think it's definitely good to raise this with the PDP Working Group. That's [inaudible] because I think [inaudible] tell them that we are [inaudible] this kind of input from them, but [inaudible] doing it and tell us exactly what they [inaudible]. But that doesn't mean that we should stop asking them [inaudible] timeline. So if we ask them, it may

be good to ask them if they can tell us when they can respond by so that we can, if possible, [inaudible].

JIM GALVIN:

Okay, thank you. So if no one objects, I will take an opportunity at some point tomorrow during the TNT PDP Working Group meeting to just pass on to them that we're considering this issue, that we're going to respond to this issue of data elements for the purposes of taking note of translation and transliteration and what needs to be recorded with data when that function occurs and let them know that, and also give them an opportunity to comment in any way that they would think is appropriate, and highlight to them of course that when we do have something for review, we'll make a point of calling it out to them so that they can get it careful study to make sure that it's going to accommodate their needs as their work progresses.

STEVE SHENG:

Jim, I have a quick question.

JIM GALVIN:

Yes, please.

STEVE SHENG:

In Singapore, the Expert Working Group gave an update to the PDP Working Group, and there weren't that many people participating. There was less than ten people participating.

My question is during the regular PDP Working Group meetings, are there more people than what we had in Singapore? In other words, do we need to give them an update again via one of their regular phone calls so that more of the working group people know about our work?

JIM GALVIN:

That's a good question. I thought there were more people than that in the room in Singapore than ten people, but I take your point that it was a fairly minimal group. I think that it would certainly be a standing offer to them if they would like an update on the document if they would like that update repeated on one of their phone calls, but I think it's that working group's responsibility. They run themselves the way they need to run themselves.

STEVE SHENG:

Yes, agreed.

JIM GALVIN:

Yeah. If they want, I suppose as part of commenting on this particular issue tomorrow, I could certainly ask them if they would want another update, maybe a more condensed one, during one of their meetings. But that would be up to them – up to Chris and Rudi, I would say.

STEVE SHENG:

Okay.

JIM GALVIN:

Okay. If there's no further discussion on anything else, then I'll jump into talking about data model work. So any other discussion? Okay.

So I've been thinking about our next step in this process, the data model, and what we're going to do and how we're going to do it. I have a particular proposal to put before the committee here, just a suggestion.

What I thought I would do today – the suggestion of course is to go back to the SAC 54 document. The report on domain name registration model, one of the things that we did after we had the WHOIS Taxonomy document, the next thing that SSAC did was to propose a basis for a data model. It was intended expressly as a starting point in a point of discussion as a way to structure registration data. So I take that recommendation to heart and I bring it here to this group as a starting point for our proposal for data model.

And what I thought I would do, having not introduced this before to this group, is I would take some time today – maybe 10 or 15 minutes – to give an overview of this document, its structure, why it is the way it is, and then give people a chance to go back and study it, think about it, decide if you have other ideas for how you think we should do this. Or just what would be our next steps if this was an approach to the data model that we wanted to take.

So that is an introduction. Let me pause for a moment and just ask if anyone has any questions or comments. Okay. So let's jump in here then.

The document is set up here on the Adobe Connect so that you can scroll through it yourself and see what's here. What I wanted to say is in much the same way that when we approached talking about requirements for data elements, we took a step back and broke them up into categories rather than trying to example a complete list of individual elements. It seemed appropriate to group them together and consider them as group development.

This allows for easy updates, easy changes. You can add and remove elements from an appropriate group, and then the requirements are appropriately inherited. It's obvious if you create a new set of elements that you need to take a step back and define some new requirements for them. So if they don't fit naturally into a category that's there, in the same way, what SSAC proposed here was grouping data elements, again, into categories or collections and putting them together.

This, at the time, seemed like a natural thing to do, because in much the same way that we did it, you can have all the elements associated with a domain name grouped together, all the elements associated with contact information grouped together, and also elements related to the state of a domain name.

So expanding on that a bit, the state of a domain name, one of the things that was important in defining the data model to SSAC at the time was how do you know what elements need to be present and what don't need to be present? What are the core required elements versus what might be optional or only for a specific use case?

The Data Services Expert Working Group is focused on a model of proposing data elements based on use cases. It's slightly different than this model that SSAC proposed. It's important for us to be aware of that distinction and to consider how that distinction affects the data model that we want to propose here from this group.

SSAC chose as a basis for the data model the lifecycle of a domain name. A domain name comes into existence, it goes through a series of events, there's data that you need to record or store relative to those events, and then the domain name expires and goes out into the pool that's available for the next person.

During that lifecycle of creation to deletion, you need to record certain things about the domain name in order for the system as a whole to function. That was the way that we had approached looking at the data elements that have to be collected, and we created groups of data elements based on that. So there's an operational set of data elements in addition to contact information data elements, the name server data set, that kind of thing. So I just want to point that out for you to be aware of.

There was another reason why the grouping of the data elements seemed to make sense, and that was because as we've already seen ourselves, gTLDs are different than ccTLDs. There's certainly a fair amount of consistency among gTLDs even today, although there are some differences in today's world because most of this is driven by registry agreements or registrar agreements and different registries putting different things in their agreements. There is a core, which is the same.

ccTLDs, of course, they're somewhat more varied – rather more varied in terms of the set of data elements that they have, what they require registration and what they don't.

One of the other reasons for this data model being proposed this way, and even going forward with gTLDs, as you have a gTLD which is specific to a community or a particular service or function, there may be additional data elements that are necessary only for that. They're not generally necessary.

So the model here is that you collect these differences and you make them their own group and they become an optional set of elements that may or may not be present, depending on whether they're needed.

By dividing up the data elements into groups, you can make each group – there are probably some groups that are required, but there are other groups then that would be optional and would only be present if they're needed. And if you use XML as a language in which to describe your data model, all of this is relatively convenient to document and specify and deal with, so you can have collections that are optional, and if they're present, do you know what to do with them? And if they're not, then you don't. The idea being that any registry that needed its own collection of elements would provide also the [DTB] that would describe those elements and all the strengths that go with them.

So I think that is my overview of this document and why it is the way it is and what's here. Let me pause there and ask if we have any questions. Actually, Sarmad, you and I were both part of the

development of this. Maybe I should ask first if you have anything that you'd like to add or call out with respect to this document.

Sarmad, if you're speaking, we can't hear you.

So to anyone, any questions or comments, concerns?

STEVE SHENG: I'd like to add.

JIM GALVIN: Please, go ahead, Steve.

STEVE SHENG: One of these documents in Appendix A, very quickly scroll over there. It consists of various tables. For example, on the element, the label, the format, minimum/maximum length, [inaudible], language, script tag and then at the end of that, there is also an XML schema which is very much similar to the [EBT] schema. That will be a useful reference for the working group to consider, to look at, as well when you review the doc. Thanks.

JIM GALVIN: Yes, thanks, Steve. Appendix A in that document is an explicit example of what was proposed in the SSAC document. You might want to give people back the ability to scroll in that document, Steve.

STEVE SHENG: Just did.

JIM GALVIN: Refresh, okay. Any other questions or comments from anyone? Please, Jody, go ahead.

JODY KOLKER: Just a general question. We're using this document as a proposal for how we would like the data to be laid out or as a starting point?

JIM GALVIN: Right. I'm offering it as a starting point for our discussions. It at least is a model. It's a starting point. A critical thing is recognizing that the Directory Services Expert Working Group has a different starting point. SSAC here started, again, with domain name lifecycle as the basis for how it did what it did.

It's possible that what SSAC did is still applicable, but it is important to understand that its starting basis is different than what the Directory Services Expert Working Group is doing and we're going to need to pay attention to that difference in our group here going forward.

JODY KOLKER: Okay. I didn't realize that they actually had a language tag on every element. That's something that we've been discussing for a while it seems like, whether we should have a language tag on each element or overall, etc.

JIM GALVIN: Yeah. In some sense, that's kind of an XML thing. When you sit down and you actually want to propose a specific XML schema, you do have to think about where to put the language and script tags and what they need to be a part of. There are different ways to do that.

I agree with you in general. You could probably have one at the top that says, "By default, this is the language and script for everything in here unless otherwise specified." You probably need a mechanism that says you can optionally specify language in [inaudible] anywhere else that you want inside the schema. I would imagine creating the template technically in that way.

JODY KOLKER: Yes, I agree. The group that produced this document, what was their intent of the data element I guess?

JIM GALVIN: Well, like I said, Sarmad and I were at least two of the people that were directly involved in this and I actually don't recall that we had a specific intent beyond what I've said. Insofar as it's just the starting point, I guess the historical context is interesting, but certainly not required at this point.

JODY KOLKER: Okay. Curious, are registries required to use this data element? When I see min length and max length or a contact name and e-mail address I

believe is on here too, I'm pretty sure I know for a fact that these are not followed by some registries, by some gTLDs.

JIM GALVIN:

Good question. You're asking a compliance question and that's a detail I don't know for certain. Whether or not you're mandated to do it versus whether or not you do it is an interesting point. It certainly would be easy enough to find out. I guess we should. Are you suggesting we should know that?

JODY KOLKER:

No, it's just a general question. I'm new to all the ICANN formalities and I'm curious what this document actually does or what it's for.

STEVE SHENG:

Jody, I can answer that.

JODY KOLKER:

Okay, that would be great.

STEVE SHENG:

The SSAC is an advisory committee to the Board and the community. It issues its document in [inaudible] form. So it's up for the community to discuss and the Board to take that up. This is not a binding document that registrars have to do.

JODY KOLKER: Okay. So the group that produced this, this is our advice to the Board. Then it's up to the Board to say whether they want to require that from registrars or registries.

JIM GALVIN: Well, I would characterize it...in principle, you're right, Jody. But let me suggest the following characterization in working groups. Keep in mind the distinction between our working group and the TNT – the Translation and Transliteration PDP Working Group. PDP working groups are working groups that have special standings, because their work product becomes binding on their communities. That's the very definition of policy development process.

So any decision made by that other translation and transliteration working group is something that will ultimately find its way into contracts and become a compliance issue for the communities in which it's relevant.

This working group, however, is functioning largely like a design team. We've got a few technical experts or experts in various areas and we're creating a set of recommendations. In theory, we're going to look at the entire space and make some recommendations.

Our recommendations would ultimately have to go through some kind of process in order to be accepted as a binding commitment on the communities to which they're relevant.

The same applies to the Directory Services Expert Working Group. It's going to make some recommendations. Presumably our

recommendations and theirs carry some amount of weight in terms of credibility. Presumably, you've got the right set of people on a working group, so the working groups, their work products tend to stand on their own merits. Or in principle, they'll stand on their own merits in an effective way.

Again, the larger point here is that there are PDP working groups and there's everything else. In that everything else category, there are advisory committees like the ALAC and the SSAC, and each of those committees gets to make whatever comments they want to make about whatever they want to make.

In the case of SSAC, SSAC enjoys a particular standing in the community that has developed over years, because of the quality of the advice that it tends to give. Its advice stands on its own merit. SSAC actually makes advice to its – SSAC serves at the pleasure of the Board. It was created by the Board and it exists because the Board wants it to exist, but in fact, SSAC comments to the community at large, to ICANN community, to ICANN to the Board. It makes recommendations.

What ultimately happens is we submit our reports to the Board and what the Board would do to act on them is they may direct staff or other bodies in ICANN to take notice of SSAC recommendations and do something about them, and to cause working groups to come into existence and that kind of thing. Or to create a PDP to evaluate a particular recommendation.

That's expanding a little bit. Does that help overall in terms of the process and what it means?

JODY KOLKER: Thanks, Jim and Steve. Yeah. I appreciate it.

JIM GALVIN: So SSAC made this recommendation. We were considering the issue of a data model. We've actually said a lot about WHOIS over the years. This is one in a series of documents. It proposed a data model for a discussion. It was born – when the WHOIS Review Team was running, SSAC had published a couple of documents – one about taxonomy and WHOIS taxonomy and this one. There were some prior documents going on in parallel when the WHOIS Review Team was doing its work.

I'm simply bringing this forward and saying here's at least one starting point for a data model, and in fact a recommendation in here was that a data model should be created. The WHOIS Review Team also talked about the need for a data model, so everyone is operating fairly consistently in that respect. So this is just something for us to think about, and at least is a starting point.

Any other questions or comments?

SARMAD HUSSAIN: I just posted something on the Adobe Connect. I guess looking at this data model, we've been talking about [inaudible]. But when we talk about the display side of things, the example that you had in the documents, the label of the content is also [inaudible]. So that's also [inaudible].

So is there anything – when we're doing this model, are we only doing things with internationalized content of this [inaudible] as well?

JIM GALVIN:

So from a technical point of view, I think that in the XML, there's no need to internationalize the label. You can use XML in its defined format from the standard. However, for display purposes, just as we showed in our interim document, we had kind of a mixed display of labels for data. I absolutely agree and believe that if I get an XML-labeled thing and I now know what the data is, I should be able to put the label in a localized format, in a localized language and script, and then dump out the data that goes with it.

The fact that the XML label here and the XML [DTD] itself might be essentially ASCII based is all hidden. It's just behind the scenes in the system.

SARMAD HUSSAIN:

So, yes, I understand that the labels themselves [are a fit], but what I was I guess suggesting was that [inaudible], let's say that this is the content information and this is the label information.

JIM GALVIN:

Maybe I'm not understanding the question. There's a label that's displayed when you do a directory services query. When you do a WHOIS query, that label itself is not part of the data. And then inside the XML, there's something which says [address one] and it has content. That's not really part of the data either.

SARMAD HUSSAIN: Right. So what I was suggesting was if we're talking about [inaudible] content that we may also [inaudible] address label [inaudible] content itself. Then [inaudible] and then the address itself. So the label ends up being [inaudible] explain this.

JIM GALVIN: I think you're asking if the label in the XML needs to be internationalized. I think that's what you're asking.

SARMAD HUSSAIN: No. I guess what I'm saying is when we do our WHOIS display, normally we would say "name: John Doe" or something. Then we want to do the same thing [inaudible]. We don't want to say "name:" and then the name [inaudible].

JIM GALVIN: Right. I agree that on display the name string could be name in U.S. ASCII in English and then the content. Or if it's being localized, that name string could be replaced by an appropriate label in the local language and script. That appropriate label, though, I would think would be decided by the client. The data model doesn't have to say anything about that.

STEVE SHENG: I would like to comment.

JIM GALVIN: Please, Steve, go ahead.

STEVE SHENG: Sarmad, are you talking about localization rather than internationalization? Because I would think if it looked like this and, for example, the label, it should soon be left to the client who implements this to determine whether they wish to do that.

Whereas internationalization, we just need to make sure that it has been standardized in a way that can be localized.

SARMAD HUSSAIN: So I think that we [might have] something in [inaudible] on Adobe Connect. Because if I'm seeing [inaudible]. But we need to develop [inaudible] localizing to view the content so that we [inaudible].

JIM GALVIN: Sarmad, I have to admit, that particular time I really had a hard time understanding what you were saying. You always sound muffled. I'm not quite sure why. Maybe you're too close to the mic or something. I confess I really did not hear most of what you just said. I just could not quite understand it. Steve, did you get it? Can you state back what he said or do we ask him to try and speak again?

STEVE SHENG: Yeah, I had a difficult time understanding that too. Sarmad, would it be possible to type it into the chat so that it's easier for us to capture and also for note purposes?

JIM GALVIN: Okay. Sarmad is typing. Any other questions or comments from anyone? Let me ask explicitly while Sarmad is typing – not to get ahead of him here, but are folks okay with using this as a starting point? I want to take some time to take a look at this carefully. Anyone object to doing that?

JODY KOLKER: I don't.

JIM GALVIN: So I think what Sarmad is typing here is he's suggesting that we should be specifying in the XML what the label is. And I think that you would use an XML tag. I don't think I would say contact. Yeah, that isn't the way that I would use XML, Sarmad, and I'm not aware that that's common to be honest with you.

You have your XML tags, so you would have something called content name in [angle] brackets, and then you'd have the John Doe, and then the slash contact name in [angle] brackets. The client would then see the contact name in [angle] brackets and it would know what to display in terms of a label.

So it could put “name:” or it would put the appropriate localization of “name:” into the actual output. Then of course with the content of John Doe.

Right. I agree with your conclusion, Sarmad, with what you’re trying to achieve. It’s just it has not been my experience with XML that one would do it in the way in which you’re proposing. Normally you would just say, like I said, contact name as a one-word element in [angle] brackets, and then it would be up to the client to localize that and display it as either “name:” or something else in Arabic if that was appropriate. And it would simply do that. But that particular translation of that label would not be part of the content, would not be part of the data, would not be part of what the registry maintains.

While Sarmad is typing, I’m actually on vacation next week, so it’s kind of up to the group as to whether the group would like to meet without me. I’m sure that Steve or if someone else wants to step up to drive the meeting, that would be fine, too, to conduct the meeting. The alternative would be to not meet next week and meet in two weeks.

While Sarmad finishes his thought here, anyone have any preferences, want to express an opinion about whether we cancel next week or not?

JODY KOLKER:

I’m fine with canceling next week and meeting in two weeks.

JIM GALVIN:

So Sarmad, when you say this can be done by the registry, I want to be more specific there. I think you mean the directory services agent,

which would presumably be hosted by the registry. You're saying yes. Sarmad, did you have anything else that you wanted to say? Have we covered your issue for now? Okay, thank you. Anything else from anyone else?

Let me summarize for a moment here. We won't meet next week. We're going to accept this document as a starting point for a data model for us and folks will have two weeks to review this and we can come back in two weeks and discuss in more detail about whether or not this will work for us or not, considering of course the requirements that we've laid out already in our interim report.

In particular, I want to call out the fact that the starting points at the SSAC document is different than what the Directory Services Expert Working Group did. We need to at least take note of that difference and make sure that the data model we propose is going to cover what's going to come out of the Directory Services Expert Working Group.

Any comments on that? Steve, you've captured that action and make sure it gets distributed with notes to the group here.

STEVE SHENG: Yeah.

JIM GALVIN: Okay. Any other business from anyone? Okay. Then I would say that we're adjourned. Back over to you, Steve, to close the meeting.

STEVE SHENG:

Thank you. As an action item, we will not meet next week, but meet the following Wednesday, which is on the 23rd of April and we will start to discuss the data models using the current SAC 54 as a starting point.

With that, thank you so much. Good morning, good evening, and good afternoon wherever you are.

[END OF TRANSCRIPT]