GNSO/SSAC
International Registration Data Working Group
TRANSCRIPTION
Monday 26 April 2010 at 19:00 UTC

Note: The following is the output of transcribing from an audio recording of the GNSO/SSAC International Registration Data Workign Group on 26 April 2010 at 19:00 UTC. 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. The audio is also available at:

http://audio.icann.org/gnso/gnso-ird-20100426.mp3

All recordings and transcriptions are posted on the GNSO calendar page:
http://gnso.icann.org/calendar/

Present for the teleconference:
Edmon Chung – Group Leader
Bob Hutchinson, GNSO Commercial Stakeholder Group
Owen Smigelski

ICANN Staff
Steve Sheng
Gisella Gruber-White
Dave Piscitello
Francisco Arias
Julie Hedlund

Absent apologies:
Jeremy Hitchcock - DYN-DNS
Rafik Dammak -- GNSO Non-Commercial Users Stakeholder Group

Coordinator: This is the operator. I need to inform all participants that today's conference is being recorded. If you have any objections you may disconnect at this time.

And I'd like to introduce your host for today's call. We have Ms. Gruber-White. Ma'am you may begin.

Gisella Gruber-White: Thank you (Lori). Good morning, good afternoon, good evening to everyone on today's IRD call on Monday the 26th of April.

We have Edmon Chung, Owen Smigelski, Robert Hutchinson. From staff we have Steve Shang, Julie Hedlund, Francisco Arias, Dave Piscitello and my self Gisella Gruber-White. Jeremy Hitchcock may not be able to join today. If he does join, it will be later on in the call.

And if I can please remind everyone to state their name when speaking, thank you, over to you Edmon.

Julie Hedlund: And please also (speak loud). And Edmon I do have a suggestion if you would like on how we might proceed with this call. I had sent around a revised matrix from the last call just to remind those who might not have been on the last call we had a discussion of the matrix that Steve Shang has provided that has in it a description of three different models.

And we discussed that extensively on the last few calls and I made some revisions based on the discussion on the last call.

And Edmon I was wondering if you think it might be useful to maybe look at the three different models and the matrix and consider how this might help us to perhaps arrive at some recommendations.

Edmon Chung: Thanks. Yes, and I think that's a good idea please (go forward).

Julie Hedlund: Right. Well I was wondering. I know you didn't have a chance to be on the call last time since you had a conflict.

Have you had a chance to look at the matrix? I didn't know if you had any thoughts to sort of get us started. In particular, you know, as staff we're really not trying to make any determinations between the different models. And we're really hoping to be, you know, guided by the work team.

Some of the things we discussed fairly extensively on the last couple of calls were, you know, to take out some of the requirements relating to translation

and change those to transliteration and sort of to further define the three models.

Do you have any thoughts on this remodel to get us going or should we just go and open this up to discussion?

Edmon Chung: Honestly I'm looking at it right now. And please jump right into it and will provide my input as we go along.

Julie Hedlund: Great. Thank you, Edmon. Well perhaps to get us started. Anybody who's on the call here, if you'd like to speak up, you know, we have three different models. We've looked at those models based on the impact to registrars to registries and registrants.

And then we'd also add in a new section on the impact to users of Whois and the technical impact to applications.

Just as a reminder and I did send around a matrix, Model 1 is requiring registrants to provide a must be present language to Whois and in addition give the option to provide data in local languages.

Model 2 has registrants providing their registration data in a script that can be accepted by the registrar and the registrar to provide a point of contact for transliteration and abuse issues based on request.

And Model 3 has registrants providing their registration data and a script that can be accepted by the registrar. And registrar provides tools to assist the registrant and providing for transliteration services and for publishing it in a must be present language in Whois.

What is the sense of this working group with respect to whether or not Model 1, 2 or 3 might be a useful starting off point for a set of recommendations coming out of this working group?

Don't all speak at once? Okay.

Robert Hutchinson: Yes Julie.

Julie Hedlund: Yes.

Robert Hutchinson: This is Bob Hutchinson.

Julie Hedlund: Yes.

Robert Hutchinson: I guess from my perspective I was more interested in the technical implementation of what we were proposing and it seems that the three models if you will for language, transliteration and who's doing it and so on and so forth, that forms the core of what this matrix is about.

But all three of those result in the same technical implementation. Is that what we're essentially saying is you put the internationalized characters that are in the Whois data either in UTF form directly in - in an ASCII UTF form in the Port 43 data or in the case of U labels and A labels punycode I guess?

So is and that is the proposal that people have in front of us for the technical implementation. Just to be clear so they're down here where it says impact to applications, none of these make any differences as to they're all the same in terms of their impact to applications.

Is that correct the way I'm reading this matrix? And maybe Steve or somebody can comment as to whether I'm reading it correctly or not.

Steve Shang: Hi this is Steve.

Robert Hutchinson: Hi.

Steve Shang: I think Model 1 and Model 3 both require a must be present language. And also Model 1 would require essentially a language, a local language, a script in the Whois.

So in terms of the technical impact Model 1 and Model 3 are very similar to the applications.

Model 2, it's very different. I think it stretches the limit further because, you know, in Model 2 we don't have a limit or any requirement what language user would put into. And then so that has I think greater impact from a technical standpoint to both registrars and registries. That's my understanding.

I hope others on the call can clarify and help on this issue. Thanks.

Francisco Arias: It is Francisco. May I speak?

Julie Hedlund: Please Francisco go ahead.

Francisco Arias: So my reading of the options I think all three models will have the same impact in the - the same technical impact since they all allow for localized (play stations) that would be (foolish) (unintelligible).

So doesn't which one we choose. There will be some technical impact. That's my reading of this matrix.

Steve Shang: Yes, I agree Francisco. Yes.

Julie Hedlund: This is Julie. Bob is that a helpful explanation with respect to your question on the impact among the different models, technical impact I should say?

Robert Hutchinson: Yes. I guess I'm still confused by Steve's answer because I'm looking for the impact to applications that are reading Port 43 data off of a Whois server primarily as that impact.

And maybe that's not what he's focusing on in his response.

Steve Shang: So Bob the question is the impact to Whois client running Port 43 for each of these?

Robert Hutchinson: Yes.

Steve Shang: Okay. Let me think. So the - let's see, so for Model 3 the registrar will provide the transliteration service and then put the transliteration of the Whois into the - of the registration data into the Whois record.

Robert Hutchinson: Right.

Steve Shang: So for that one then it's not too much different from today, what we are seeing today, right.

So for Model 2, you know, registrants provide, you know, can provide any kind of script to a registrar. And registrar will pass that along to the registry.

So what we are seeing is the kind of from a Port 43 perspective will be a full version of, you know, internationalized registration data. So like, you know, we could have, you know, everything internationalized. All the local contact information will be internationalized, right.

So for Model 1 I think we're seeing somewhere in the middle because registrars would still provide - registrars would still require a must be present script.

So we have some continuity in here because some of those are still, you know, incurring US-ASCII format.

But we have an option to add localized script. So in terms of impact (fine grain), I think Model 2, you know, is fully internationalized and you - and Model 1 is internationalized but it depends on what the user wants to do or the operation of the registrar and what they require.

Is that - does that clarify things a bit?

Julie Hedlund: Bob.

Robert Hutchinson: Yes but...

Man: (Unintelligible).

Woman: (Unintelligible).

Man: (Unintelligible).

Woman: Okay.

Julie Hedlund: So Steve this is Julie.

Woman: (Unintelligible).

Julie Hedlund: A question just, you know, for my own edification. I mean so you were saying Model 2 then is fully internationalized.

((Crosstalk))

Woman: (Unintelligible).

Julie Hedlund: And just remind me again Model 3 would be most similar to what we have today and 1 is somewhere in between?

Steve Shang: Yes. I think so. I mean others can jump in and correct me if I'm wrong because...

((Crosstalk))

Woman: (Unintelligible).

Steve Shang: ...what we are saying in Model 3 is the registrant can provide...

((Crosstalk))

Woman: (Unintelligible).

Steve Shang: ...that the information in local script. But the registrar would transliterate them.

((Crosstalk))

Woman: (Unintelligible).

Steve Shang: And publish that and also provide that to the registry. So that's essentially the same as what we are having today.

But then the burden of the transliteration is being put on the registrant's shoulder.

Julie Hedlund: I see.

Woman: Much.

Julie Hedlund: I hear someone speaking in the background. I don't know if someone thinks they're on mute but they're not. Anyway there's some discussion on the line. I just want to let someone know if that's someone on the call.

So Bob as far as technical impact and the Port 43, applications on Port 43, does that - did Steve address your question as far as that impact?

Robert Hutchinson: I guess I see from the Port 43 client applications that either one of these models is going to wind up with a text file, an ASCII text file being spit at the Port 43 client that contains UTF-8 characters in some cases that represent registrant data in local scripts of various sundry types.

In no case are we proposing at least in any one of these models whether or not that script defines the language or in other words it's just UTF-8 characters. It doesn't go beyond that, okay.

And it's a representation of what the registrant data or the - well it's the registrant data primarily in their local script, okay.

Now that's all there is from a technical standpoint at this point. I guess what I'm looking for from a technical standpoint also perhaps is an example or a couple of examples of what we think the Port 43 text content for a few applicants might look like given these three scenarios, okay.

And I don't know if somebody could go and do that exercise and provide us an example, okay. It will be helpful I think for the group.

Steve Shang: Okay. Yes.

Robert Hutchinson: The reason I say that it will be helpful is that it depends a lot on how you think of the impact to applications.

So if I'm building a Whois viewer of some sort, either a JavaScript for a web browser or something like that, if I don't - if I just present to the user a series of UTF encoded ASCII characters it will just look like, you know, kind of a hex dump of what the information is, okay. That's an unchanged standard Port 43 application today.

But I would assume that eventually people will modify those Port 43 applications and they will represent or present to the user not only the UTF-8 representation in the ASCII characters but they would represent the actual internationalized character string itself, okay, as, you know, as it's intended to be displayed.

And the impact of the application in these cases is that you have to change from a standard flat ASCII file display to something that's farced and presented kind of in internationalized characters, okay.

And that's I think primarily - it's the same - that implication or the change necessary to update the Whois displayers is the same for every one of these models, okay.

See I'm looking at this primarily from a technical standpoint as to what we're asking people to do and what there will be available coming up in Port 43 which is I guess the holy grail of Whois data representation, okay.

Steve Shang: Right. So I think there are - if you characterize the applications that use Whois, you know...

Robert Hutchinson: Right.

Steve Shang: ...you have the Port 43 client, right, the standard UNIX client. You have the web Whois, right, you know, using web browsers and to view Whois. And we have these customized applications.

So for web browser - for the Whois to display properly on web browsers it depends on both the operating system support and the browser support.

Robert Hutchinson: Yes, exactly.

Steve Shang: So if I have support for UTF-8 and a browser can correctly parse it then, you know, it should be no problem, you know.

You know some operating systems like Windows, maybe Windows '98 or, you know, IE or maybe IE5 may not be the support internationalized IDNs either. They don't have the full support for that.

Robert Hutchinson: Right.

Steve Shang: And that I mean it's probably, you know, the simple thing to do is to upgrade to a newer browser which is very little cost and it's widely available, personal opinion.

So I think from a web browser perspective, you know, the display wouldn't be a problem. I see the bigger problem of display comes for these customized applications and also Port 43.

And in terms of the impact for all of those three models I agree with you. Those have very - those have similar impact, very similar impact, you know, to both the Port 43 and to the customized application.

And we haven't discussed those yet. But I think those are independent of these three models.

Robert Hutchinson: Okay. That's what I wanted to hear is yes, you agree then that the impact on applications is the same no matter which model of the three we're picking...

Steve Shang: Right.

Robert Hutchinson: ...because the technical implementation at the ASCII Whois file level is the same for every one of these. You know the data will change, okay.

Steve Shang: Right.

Robert Hutchinson: But the technology that we are using to store and display the data in its correct form isn't changing across the three models that you're talking about here.

Steve Shang: Yes. And we - and I think, you know, that's still the elephant in the room we need to address, you know, too but, you know, those are independent of these models so.

Robert Hutchinson: Right.

Steve Shang: So kind of tackle kind of the easy ones first.

Robert Hutchinson: Well I think actually the three, picking between the three models going back to Julie's opening question, I would be okay with, you know, floating a proposal that - are you looking for something to be floated out to Brussels, you know, meeting Julie?

Julie Hedlund: Yes. This is Julie. So what we're trying to see if we can determine here is whether or not there is something that we'd be prepared to say at Brussels along the lines of some, you know, preliminary recommendations.

I think in either case we'd like to be able to have, you know, something to report as far as our discussions and the progress thereof. And, you know, I wouldn't know for sure whether or not we'd have recommendations ready.

But it would be I think helpful if we could sort of move in that direction. Sorry, if that's kind of (a long) response.

It's really more what the working group would like to do and how, you know, you proceed with, you know, with what - with this discussion and how we as staff can help.

Robert Hutchinson: Right. I guess the piece that maybe was some part of the elephant that Steve was referring to is there's currently no formatting standard at all in how you present the Whois data. There's just a requirement that you present it in, you know, in a human readable form.

And the format's come back, you know, virtually impossible to write an automated parser to parse the data because every registrant or every registrar uses slightly different formats for say their dot-com and Whois data and so on and so forth.

So is that part of what this protocol recommendation is also going to do or are we essentially taunted on that and said we're just going to leave it the way it is and the only thing we're going to deal with was these transliteration scenarios for the internationalized data but we won't deal with trying to make the Whois data form uniform in any form or fashion across the, you know, various Whois databases?

Is that a correct assumption or...?

Dave Piscitello: This is Dave. Can I answer that?

Robert Hutchinson: Sure. Go ahead Dave.

Dave Piscitello: We - in the parallel activity you may be aware that we're looking at future Whois service requirements. And Steve Shang had written a fairly, you know,

fairly lengthy paper about that essentially gathering an inventory of service requirements.

And as part of the effort for that activity one of the things that we're looking at are the data that currently comprise a registration record and whether or not there would be value in establishing some, you know, normalization or some form of data structure that we could reliably use as, you know, as part of these who developed services or who developed client applications.

And so one of the things that we - I'm speculating, one of the things that this committee might do is weigh-in on that topic and say part of what might make this work for IOD would be to have, you know, have a more formal data structure and/or more formal separators or semantics and syntax for existing data.

Robert Hutchinson: Excuse me. Where I connect the dots here is without a frame for, you know, what this data is if you present it to me as an English or an ASCII, this romanced character reader, if you don't even give me the context of what these internationalized characters are supposed to be expressing like a name or an address or street and so on and so forth, I'm going to have a hard - really, really hard time, you know, I mean after looking at - I can look at it in English street name and address and so on and so forth.

But if you do it in completely in Japanese or in Chinese or some of the other - I don't know what they refer to those (clips), high density (clip) languages. I'm not going to be able to even tell what I'm looking at, okay, let alone try to be able to translate it.

So anyway that's to me that's where it sort of connects. I - you know if you have transliteration between a ASCII, you know, standard ASCII character set and an UTF-8 label, without knowing that and if you just scramble it all over a Whois form it's going to be really difficult for somebody to figure out, you know, who's not a native reader of that foreign script.

So that's my point about that.

Does that make sense Dave?

Dave Piscitello: Yes. It makes perfectly good sense. I'm particularly, you know, trying to think of the right answer for you. And I don't - I want to try to avoid having duplicate and worse, you know, different opinions in terms of, you know, what would need to be done.

I agree that, you know, if we continue along the path of arbitrarily tagging or not tagging data elements and we incorporate, you know, (blue) space languages alongside, you know, Roman characters and other numbers it's going to be very difficult for humans to read.

And I think automation may have a, you know, may have a less of a problem because of the code pointing (neatness). But even there I can see, you know, complexity in writing parser so.

Steve Shang: Yes, point well taken. Do others have some thoughts in terms of different models and merits? No.

Edmon Chung: This is Edmon. Just curious got a - I have to admit that I'm trying to catch up to the conversation.

But in terms of the format we're talking about is this - how should I say? What's the - what has been the discussion? I'm trying to flip around. I'm trying to figure out the last couple of meetings, again I apologize for not being there, but where are we with just generally the - we talked about format. But I'm not sure what -- how should I say -- what led us to this discussion?

And I'm a little bit worried about trying to set rules that somehow have implications to the existing Whois and just want to make sure that. I

understand that with the internationalized data we'll have - it's more likely that we will need some sort of format that what has led us to that discussion? Just very briefly, and I apologize for asking this.

Steve Shang: Edmon this is Steve. Let me try to answer. And others feel free.

I think Bob's point was saying he - if the Whois completely internationalized or isn't internationalized substantially a non-native speaker of these languages may need some context to be able to identify, you know, what those fields about. Especially labels to the fields.

So, you know, to understand this column is a name of the registrant. This column is the address. This column, you know, is the technical contact.

You know I think that's the - Bob is that - does that characterize your intent correctly?

Robert Hutchinson: Yes, you did. Thanks.

Steve Shang: Yes.

Edmon Chung: So this is for the in staff's life data or in general given the whole thing isn't just staff life?

Robert Hutchinson: Well for example let me - if you look at some Whois data coming out of Port 43 for some registrants the Whois data is buried in a sea of text that has to do with, I don't know, registrar contact information and what I would consider to be almost quasi advertising by them and so on and so forth.

And if you have them internationalize that stuff and put that in the Whois, okay, it becomes for a non-native reader of that script. It becomes very difficult to parse out, you know, what pieces are - you know what is this awful

lot of (wheat in it) and, you know, (wheat and chat) kind of situation for a non-native reader of that script.

So that's the other part that I think, you know, it would help us if we had some examples of what people think this internationalized Whois Port 43 data could and should look like or could and would look like given what today's data looks like, okay.

Dave Piscitello: Well I'll offer - this is Dave Piscitello. I'll offer, you know, some speculative examples.

Yes, so let's imagine that today you're looking at registration information from .org, .org has a specific format where...

Robert Hutchinson: Yes.

Dave Piscitello: ...they have a tag with a colon and then a value.

So in the example where or in the model where the registration data would be in both a mandatory ASCII 7 and an international character set or yes, you could envision a record that is larger than what we currently have today where one piece is exactly what you get from .org as it reads today.

The second piece has the labels, the tags, registrar colon and then whatever the language is that the registrar actually collected the data in so that's one example of what I might imagine in the Whois.

And I'm - you know another example might be that everybody says we can't do this without XML. And we go out and, you know, all agreed that here's a data structure that the IDTF blesses and that the GNSO blesses and it all happens yesterday. And so the data would all have XML tags for both the ASCII 7 version and the international version.

So those are two examples of what you might see coming down the pike on a Port 43.

Robert Hutchinson: Right.

Dave Piscitello: Right. Now let's take a look at what happens when it gets at the end of the pike.

So there are two possibilities at the end of the pike. One is that somebody is going to transform what just came down into a web page or render a web page...

Robert Hutchinson: Yes.

Dave Piscitello: ...out of that data, okay. Then it's a matter of as Steve said is this browser capable of displaying both - not only ASCII 7 but characters in the international characters set, you know, that is in the registration data?

Similarly if it's a client application on a Windows machine or a Linux box then the client application has to be able to display, you know, in the appropriate character set. Of course those have tendencies on the libraries on the client to actually support not only the, you know, not only the Whois application but the rendering of the characters and fonts and (lists) and the like.

Was that helpful?

Robert Hutchinson: Yes.

Edmon Chung: Yes.

Robert Hutchinson: That's...

Edmon Chung: That's great. It's good. I had a question originally was - I understand that part. It's - or actually I should restate. Yes, I understand it correctly.

So what we're right now discussing about is that there should be a sort of a part where, you know, there's the existing Whois output and then there would be an additional part where in essence that's the internationalized part.

That part would be more structured. Is that - and there needs to be some sort of a structure to indicate that the - this is the internationalized part?

Is that - the question was really - is that what we're talking about?

Dave Piscitello: Well I think we sort of, you know, arrived at that. It wasn't - you know the original question is I believe Bob posed was, you know, how do I distinguish these things when I'm the client and when I'm the - you know when I'm the user.

And what will help us as users or application writers to be able to discriminate when, you know, when something comes down - something is delivered over Port 43 to a client that is not in ASCII 7.

So, you know, there are some conventions that exist today that basically have a lot of clients that ask for, you know, only ASCII 7 or for, you know, some server dependent character set for example in .hp. You can add something to a Whois command line that says I want the Japanese characters, you know, for whatever you happen to have.

But we don't have, you know, we don't have anything that formal yet. So I was simply speculating that, you know, we could go further and say, you know, and at least talk about, you know, if not recommend some of the consequences that Bob is looking at to try to, you know, understand what the user experience is going to be as registrants or as consumer of Whois information.

Edmon Chung: And generally I think I think that's a good idea. I think I'm asking the question is whether the discussion -- I apologize again that I feel lost -- is that the question is this is something new. This is something to be added for internationalized data, right.

We're discussing about a suggestion that would be added for internationalized data rather than as a general Whois function.

Dave Piscitello: Okay so let me - so I see where the disconnect is. When Bob had asked the question I had mentioned that in - the GNSO Council asked ICANN staff to put together an inventory of service requirements and an inventory of known or speculative deficiencies in Whois.

And Steve Shang primarily did the work on that.

And one of the recommendations or one of the steps of the technical deficiencies that we identified was that, you know, that there is no normalized representation of registration data. Each registrar actually has, you know, and registry has a different formulation of what that looks like.

So if you go to .org as I said there's a tag for the registrars, it says registrar colon. If you go elsewhere there may not be a colon. If you go somewhere else the contact information is not uniquely identified as name, address, phone, fax but it's just one opaque string, you know, that contains all that information. And it's all under contact.

So one, you know, one suggestion that Steve's paper offers is that it will be very valuable to have some normalization or some convention that people can rely on so that automation in particular and now in the case of IOD representation in multiple languages would be facilitated.

Edmon Chung: Okay. So the reason why I brought up the discussion is really I'm not sure. And for the internationalized data part I think that's in a sense safe and, you know, I think that's within the scope of our discussion.

I'm just wondering whether we're venturing into sort of defining the format of Whois in general. That's sort of the source of my question and...

Dave Piscitello: Yes, and this is Dave again. That's not our intent. I was simply making the observation that there is a parallel effort and that we should all pay attention to the fact...

Edmon Chung: Okay, cool.

Dave Piscitello: ...we're looking at that.

Edmon Chung: Yes, that's fine.

Dave Piscitello: Yes. I mean I don't see that it - personally that I think that what we say here and what, you know, what the GNSO Council deliberates not only in this - in the IOD issue but in the service or the inventory of service requirements for Whois, is all, you know, above our (page rate) so to speak.

Steve Shang: I like that Dave. Yes.

So my personal observation is in terms of these three models that we present, the technical implications are not different for each of these three models.

So I think that this is more of a policy discussion and deliberation and ultimately a policy decision coming out of this working group.

You know my sense is the way that we present these three models is we start with the list of the things to see how we can separately - a list of the

registration data and see how we can separately internationalize them. So we talk about email address, phone numbers.

And those, you know, we - there are already (extenders) exist for internationalization and there's not an issue.

When it comes to internationalized registration like user contacts and we had kind of different opinions on what scripts and what languages that we should require and I think the goal of coming up with a matrix is trying to move forward that discussion.

And these are, I think are really policy discussions because the underlying technical implications are essentially similar.

And so hope - you know maybe we can deliver it on these options and after that, you know, we dig more into the technical implications more.

That's my thoughts.

Edmon Chung: It's Edmon again. I again apologize for missing a couple meetings. But in terms of the models we talked about sort of transliteration or translation.

Has there been a discussion about what that really entails especially on the concept of "Transliteration?"

There's a general sort of definition of transliteration that in the Whois context is there - are we talking about just presenting it in a certain format that's relatively readable or are we talking about in some sort of accurate transliteration based on certain standards that for which the transliteration could be used for other purposes.

Has that been discussed? Where are we on that?

Julie Hedlund: Edmon this is Julie. We did discuss transliteration a bit on the last call.

And a couple of people on the call mentioned that for some languages there are standards for transliteration and, you know, for other languages there aren't.

But they - I think the sense was that, you know, we - I mean where standards apply, you know, local standards apply say in Japan or, you know, in Russia those would be used or elsewhere.

But we didn't really get that far into what exactly the transliteration would entail. You know exactly how would that work, you know, whether or not we - I think the sense was we wouldn't try to suggest standards but that, you know, the whole standards if there were any would apply.

And I think that your question is a good one perhaps. We need to get into a little bit more detail as far as what that would really mean.

Dave Piscitello: Yes, and this is Dave.

Edmon Chung: The reason why - sorry.

Dave Piscitello: I just wanted to point out, this is one of the hard things that, you know, from staff perspective is it really would be great to get you, (Jay), (Andre) and the gentlemen from (CNNIC) all on the same call because I think that the four of you represent the most experience with, you know, with dealing with character sets that are not easily transliterated.

And so if we could carry this conversation over to the email list perhaps, maybe we'll get a little bit more traction. You know we've sort of been, you know, been working on, you know, on parser strength so far as an observation. If we could get everyone together on one call it might actually help us as staff catalog the many issues around this particular point.

Edmon Chung: Yes. I think that's a good point. And I agree with it.

But I guess the question originates from this relatively (loaded word) about transliteration and the question is whether we really mean transliteration or do we mean a - some sort of format that is somewhat readable.

And let's say US-ASCII. Are we talking about just representing some set of data in a format that is US-ASCII or are we really talking about transliteration, is really the crux of my question because it's quite different?

Steve Shang: Edmon this is Steve. When I first wrote out the draft matrix, you know, I wasn't sure, you know, to what percentage we need translation and what percentage we need transliteration. You know for the Russian language it's probably easier to transliterate than Chinese, you know.

So yes, it's like I really - as a staff I don't have too much expertise on that. I really appreciate.

Edmon Chung: No. I guess yes and no. But Chinese is relatively - it's not - it's - let me put it this way. It's not impossible to transliterate it.

But my question is really when we talk about these three models we talk about transliteration, what do we really mean? And do we - should we reestablish some sort of a vocabulary that we just say a kind of representation in US-ASCII?

Is that what we're really talking about especially in Model 3? Because overall I think Model 3 seems to be a direction that somehow makes sense.

But the question is what are we really talking about because if we talk about translation or transliteration that's quite prohibitive. But if we talk about a -

that there needs to be some sort of a representation in US-ASCII maybe that's the much better way to describe it and a way that's much more doable.

That's sort of reason why I brought this whole discussion up.

Steve Shang: Right. So my understanding is to have it fully US-ASCII there are different levels of accuracy associated with it.

So take Chinese for example, if we have a Chinese address, you know, we can just put the (P) in. You know that's like Roman character US-ASCII and put that into Whois, right.

But more accurate perhaps is to put parts of the street in (P) and, you know, translate part of that, you know, like (DNO), translate that into actually street or avenue or apartment, you know. That's my initial thought.

I think that there are different...

Edmon Chung: Right.

Steve Shang: ...parts of ASCII associated with it.

Edmon Chung: Exactly. That's sort of where I'm coming from is that if we specify transliteration or translation then exactly, the level of accuracy and, you know, how operable we need the data to be. If we just say, you know, this is a kind of data that, I guess sorry for jumping ahead, but this is data for machine to grab and then can do some analysis with such as issues with trademark or phishing, right.

And then the requirement is quite different than if we say we need a translation or transliteration where it's completely operable and if we put it on an envelope it would reach the person.

Steve Shang: Right.

Edmon Chung: Right. So again sorry for jumping ahead but overall if we're talking about former than I think personally I think it's fine and it works well.

But if we're latter than I think it would be quite a significant job for registrars. And that's, you know, they're in lies the sort of the question.

Steve Shang: That is more of a policy question.

Edmon Chung: Right.

Steve Shang: It depends on...

Edmon Chung: Right. It is. And that's sort of the reason I brought it up whether there was already extensive discussion on this topic or this is some of - this is one of the things that we should really talk about in terms of the three models.

I think overall three models gives a good structure for discussion. Right now I'm really bringing up the issue with transliteration because of I guess personally the tendency towards Model 3 of if that is being resolved. Well no, I shouldn't say resolved but if that is being better described in the way that, lack of better word, easier for registrars to implement and registries in some sort.

Julie Hedlund: So Edmon this is Julie. I would suggest if you'd like that we could try to get the discussion going on this particular question on list because I think it's a very important one. And it does, you know, have an impact on how we consider each of the models.

And that is what, you know, the question being what do we mean by transliteration?

And I think you've mentioned, you know, a couple of different levels and, you know, resulting in the different levels of accuracy and perhaps we at staff can try to frame that question to get some discussion going, you know, particularly amongst those working group members like (Shawn Kahn) and (Andre) who and (Ron) as well who have, you know, particular expertise in this area.

Would that be a good way forward?

Edmon Chung: Yes. But just one thing is that rather than trying to sort of "Define" transliteration, I would suggest going forward and just saying that transliteration or some sort of representation in US-ASCII is whether that works because if we go down the path of trying to define transliteration, because transliteration is in a sense it's defined in linguistic terms.

I'm just wondering if we can - in a sense the three models are being proposed at this point. But I'm just wondering whether we should change the wording a little bit and saying that, you know, transliteration or some form described in US-ASCII because that would somehow allow us a little bit more flexibility as we go down the path. Because if we talk about transliteration it does bring about quite different conversations so that's really the thing.

And the thing is that if we go down the path of getting the mailing list - going to the mailing list and getting everybody to try to figure out what the "Definition" I guess of transliteration would be for this purpose it would still create the same issue.

And so I agree that, you know, the general approach but with one slight adjustment is that we not talk about the definition of transliteration but we really talk about whether transliteration or some representation in US-ASCII is acceptable in a way.

Julie Hedlund: Right. And this is Julie. Edmon thank you. That helps clarify it for me. We're not asking for a definition of transliteration. That already exists out there as you say linguistically.

But whether transliteration or representation in ASCII, those two options, you know, those are the two options we're considering for each of those models.

Did I get that correct?

Edmon Chung: Yes. That's really what I'm suggesting is transliteration is to me it's going to be tough because you need to follow a lot of rules.

But if we talk about a model where some form of US-ASCII representation then it could become much easier for the registrar. And it could be a model where it allows certain let's say mechanical, some sort of spider to come across and analyze the Whois data and without, I mean without a lot of enhancements to that mechanical process. And would be able to analyze the data meaningfully but then it doesn't put a heavy burden on registrars to actually produce - to be quite direct. It doesn't require the registrant to produce data where for example the contact information, the address, could be used as actually put on an envelope and send to the person is that really the big sort of issue there that I think.

Julie Hedlund: Right. Edmon thank you. This is Julie. I was wondering if you might be willing to put this, you know, these thoughts of yours, these questions into an email to the working group. I mean I think I understand what you're saying.

But and I know it's a difficult hour for you right now so just sometime in this week perhaps. Would you be willing to do that?

Edmon Chung: Sure, yes. I'll try to describe what I'm trying to drive at over email.

Julie Hedlund: Thank you Edmon. I really appreciate that. I just I think (unintelligible)...

((Crosstalk))

Steve Shang: Julie this is Steve. I just have one point before I know we're past the hour.

So we - with this working group every time we kind of have different people on the call. And, you know, I think people have different opinions on which model they prefer.

What will be a process that the working group think that we go ahead and kind of to reach us a consensus of, you know, knowing that not everybody's on the call at the same time and...

Edmon Chung: Yes.

Steve Shang: So that is my question for the working group.

Julie Hedlund: Steve I could comment and Edmon of course please chime in.

But I, you know, I'm on several of these types of working groups and it - because of time zones it is often difficult to get everyone on the same call which is why of course we're having rotating times.

I think suggestions like Edmon's to put a question out to the list is very helpful to get responses and to get people to begin to coalesce around a particular set of recommendations.

And I think that as we, you know, sort of distill that discussion we can pull that into perhaps, you know, one or two or three sets of recommendations and ask each working group member to respond to those on the list, you know, and of course encourage people to attend the calls as well.

Edmon do you have any thoughts on that?

Edmon Chung: Sure. I think in general given the situation it's best to sort of vet any decisions we make over mailing list, sort of treated kind of ITS approach, if you will for lack of a better description, to get the consensus on the mailing list.

I'm interested to put forward a question is that mentioned a - different opinions and stuff. I wonder on this specific topic, sorry I haven't reviewed all the Minutes from before. On this particular topic what are some of the - what are the views, if you don't mind but...?

Steve Shang: So I think that a call two weeks ago or I think four weeks ago, there were some discussions thinking, you know, Model 1 may be better.

And some in the mailing list for example (Jay) think Model 1 and Model 2 are both okay.

So they're just - you know I'm trying to anticipate, you know, we're already heard different opinions which of these models people think benefit most.

And there's no agreement yet.

Edmon Chung: Yes, right. I understand. Was there any discussion or were there any discussion on transliteration just (unintelligible)?

((Crosstalk))

Julie Hedlund: Edmon this is Julie. We did have discussion on the last call, the last two calls. Two calls previous, the sense was that we would move away from recommendations on translation and then have further discussion on transliteration versus translation.

And on the last call the working group members reiterated that it shouldn't be a recommendation for translation but for transliteration. That as I mentioned

there wasn't detailed discussion on what would be required with respect to transliteration.

And I think that the particular aspect that it will be helpful to perhaps you pose a couple of questions on to try to get a little more clarification from the working group members in that specific area.

Edmon Chung: Okay, that's good. Thanks. Because I don't want to - I apologize for using this time here. But I just want to get a quick background on whether we have a discussion and what people thought about stuff, yes, that's useful. Thanks.

Julie Hedlund: Good. Thank you. Well since we are a little over time if it's okay with you Edmon, perhaps we'll close up the call.

I'll take a first stab at writing up some notes and have Dave and Steve look over those.

And in the meantime if you wanted to send, you know, around some thoughts when you had a chance with respect to transliteration to get the discussion going there, that would be great as well.

Edmon Chung: Sure. Yes, I'll do that. Just before we close, are we - what's the plan going forward and especially driving towards Brussels and whether we would have some activity there?

Julie Hedlund: So Edmon this is Julie. I think we definitely would like to have a presentation at Brussels. If we don't have recommendations ready at that point I think that we would like at least like to be able to say what we've been discussing and progress of those discussions.

And this would be I think probably a public forum, you know, type of event as opposed to - at the last meeting we just had the presentation at the (FA

Open) Meeting. I think we'd probably try to have a separate, you know, public event on this.

And so I think that we're trying to see if we can move forward to something a little more concrete in the area of recommendations. And if not it would be more along the lines of a progress report.

Edmon Chung: Okay. All right, I'd like to sort of suggest that we put out something even if we know that we don't have somewhat of consensus over it as a straw man thing because that would help. I think that would help participation in Brussels and, you know, help the whole thing move forward.

That's my opinion because if we just have a very general oh let's talk about this then we might not get the kind of participation that we want to actually drive towards a recommendation eventually.

Julie Hedlund: Thank you Edmon. This is Julie. I think that's an excellent suggestion. And I think, you know, as staff, we'll try to help see if we can, you know, coalesce, you know, and distill the discussion among the working group members into something, you know, fairly, you know, like framework as you say or something along the lines of recommendations if we can.

And I think that I should note too that I'll look back at the charter. But I think that for consensus we don't require agreement from every single working group member. And I think we also have the option for, you know, minority reports and so on. Those aren't, you know, ideal but what I mean to say is that we'd like to try to, you know, get as close to consensus as possible.

But failing that I think, you know, if we have a majority we could still, you know, have something that we could say.

Edmon Chung: Sure. And I totally agree. And I guess my suggestion is really just put up something and let people, you know, shoot it down in a way. That way we

might get more people to participate in discussion, because it seems to me, again I apologize for my own absence, but it seems to me that that might help us overall in driving towards a - you know the final recommendation.

Julie Hedlund: Thank you Edmon. Well we'll - staff will take on the task of trying to put together something that might be, you know, a straw man at least to have people react to.

And I think that, you know, I think we're getting - we're definitely moving in that direction. The matrix is helping with that.

And I think we can, you know, try to see if we can move further that way and help the working group move further that way as well.

Edmon Chung: Right. That's very useful. And again I think staff has been - the work has been phenomenal so far. And I would encourage to, you know, keep moving forward.

And take a look at the time, we probably have a - two or three more meetings to go before Brussels. And thinking about how to lay out that time would be useful as well.

Julie Hedlund: Thank you Edmon. That's an excellent suggestion. We'll try to come up with maybe something more concrete to have people look at before the next meeting and maybe be able to focus the discussion a little bit more on some specific recommendations.

Edmon Chung: Sounds good.

Julie Hedlund: Great. Well thank you everyone. And thank you in particular Edmon for taking all of this time. I think it's been an extremely helpful discussion.

And we'll meet then at the alternate time in two weeks but we'll expect to make some progress prior to that on the list as well.

Steve Shang: Thank you Julie.

Robert Hutchinson: Thanks Julie.

Julie Hedlund: All right, thanks everyone and have a good morning, evening, afternoon. And we'll talk to you next time. Thanks.

Robert Hutchinson: Thanks. Bye.

END