

**Translation and Transliteration of Contact Information PDP Charter DT Meeting
TRANSCRIPTION
Thursday 10 July at 1300 UTC**

Note: The following is the output of transcribing from an audio recording of Translation and transliteration of Contact Information DT on the Thursday 10 July 2014 at 1300 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-transliteration-contact-20140710-en.mp3>

Attendees:

Jennifer Chung - RySG
Rudi Vansnick – NPOC
Peter Dernbach – IPC
Jim Galvin – SSAC
Howard Li – SOI
Klaus Stoll – NPOC
Wolf Ulrich Knobon – ISPC
Wen Zhai – NTAG

Apologies:

Chris Dillon – NCSG
Petter Rindforth - IPC

ICANN staff:

Julie Hedlund
Amy Bivins
Glen de Saint Gery
Lars Hoffmann
Terri Agnew

Coordinator: Okay recordings have started.

Terri Agnew: Thank you (Tara), good morning, good afternoon and good evening. This is the translation and transliteration of contact information PDP working group on Thursday the 10th of July 2014.

On the call today we have we have Rudi Vansnick, Howard Li, Jim Galvin, Klaus Stoll, Jennifer Chung, Wolf Ulrich-Knobon, Peter Dernbach and Wen Zhai.

We have apologies from Chris Dillon. From staff we have Lars Hoffman, Julie Hedlund and myself Terri Agnew. I'd like to remind all participants to please state your name before speaking for transcription purposes. Thank you very much and back over to you Rudi.

Rudi Vansnick: Thank you Terri. So there - with this we have executed point 2 of the agenda and as usual we will check the statement of interest. If there are any changes please speak up.

Hearing none and not seeing any comments in the chat room we can then enter into what (Chris) has been extensively working on and proposed for this meeting call to discuss the final report from the expert working group on gTLD directory.

Did I hear somebody speaking up? No, okay well I don't know if we can have - well we have the list of pages that (Chris) has mentioned and that are of interest for our working group and especially for the place where we have to start producing some reports, written reports on what we got back from the different stakeholders being the SO's and AC's.

As we know we still are waiting for others - for other reports to enable us to have a full overview. However (Chris) has also been working on another document that we will probably try to share with you in the - for the next call or the call after next week.

But I think that indeed it's good if we could go through the pages that he indicated being of importance for us and I think that on some of the pages if we have some questions we can call on him to give additional comments on the pages.

I don't know if we can have the Page 14 of that report in the Adobe Connector. Okay, thank you Lars.

Julie Hedlund: Hi this is Julie actually I'm manipulating it. I took the document off synch just now so that everybody can go ahead and page through it as they like. And as noted in the suggestions that I had put up there we're looking at Page 14 to start with.

Rudi Vansnick: Thank you Julie. While we have the diagram presentation of the flow of information into the proposed new mechanism being the RDS. And what is of importance for us is especially as far as I have recognized is the (unintelligible) is mentioned validators.

The normal process is that data is going to the registrars and from the registrars goes into the Whois database, which in this case is the gTLD registry data storage where we will have the doc collected and then the data will be synchronized in the registry directory services platform.

I think the validators is the one that could be of concern for us if validators are eventually translators or official or bodies that are willing to validate the data itself. I think it's important that we recognize the - this party in the process.

I don't know if Jim you have any additional information on the definition of the validators in this case?

Jim Galvin: Thank you Rudi this is Jim. No I mean I think an important question is whether or not translation and transliteration would be occurring there, you know, then we'll need to consider more of the requirements or responsibilities on that particular component.

But I don't think we've really committed to where we think that function might occur if it's going to. So a little hard to - difficult to comment to strictly, thank you.

Rudi Vansnick: Thank you Jim for the additional information. Indeed it's an important question that we will probably have to put on the table in the next week for clarification if that's what would be considered under the label validators.

But we will have other topics where we will probably have similar questions and doubts and we will probably need to address the ICANN board to have more clarification if that's part of our task yet or not.

I would then move to Page 20, I think printing them out was a bit easier to follow in case I couldn't have Adobe available on my screen. So on Page 20 you will see also a diagram and that the Figure 1 that is called points analyzers.

And we see several icons popping up that are of importance for us. Again it's the block at the or the round oval at the right under - right corner, validation and accuracy requirements, which is again the question of what kind of validation should be done and what are the accuracy requirements.

And as far as I've been in the report I have not seen yet any clear definition also about what is considered being in - under that topic is that also taking in consideration translation and or transliteration? So far I didn't see any.

But it's important that we have - we can point back to the report and highlight the fact that it is mentioned that a certain validation could occur. It's again in the context of and the initial question that we have is it desirable to have translation and or transliteration that's the initial question.

And having these items being mentioned validation and - yes sorry Jim we are on Page 20 at this moment on Figure 1 (unintelligible) analyzers. I will try to not forget that I mentioned the page each time I'm changing the topic.

I don't think that there is much more to talk about on this slide on this presentation and this figure. It's good that we know that it is taken into

consideration and that we can eventually point back to the fact that if we need to respond to the question is it desirable we can probably point to that point in the report.

I don't know if there is any other remarks or comments on this or if you just raise your hand I can follow up on the Adobe screen and let you come in the discussions.

So the next page is Page 21, the next page where we will go to the bottom of the page. It's about the users it's on - you have the top the Figure 2 the users are in fact here the users all - the database of the registration system.

We have all -in fact all the stakeholders mentioned in that last in that figure but there is a detailed description that follows below, which is the table just at the end of the page, Page 21.

Where there is a clarification on - with considered being registrant. Yes I'm still on Page 21 and the content is on Page 22 where we have examples of data that is going to be addressed and used or provided by the registrant and further on we have this definition of protected registrant.

Where it is considered that the data, the contact data would be under the topic of privacy and should be hidden for public access. And we have the - I'm on Page 22 at this moment in the table, we have the technical issue resolution where for technical reasons there is needs to have access to the contact data.

I presume that most of us have already touched upon these different items. There are certification authorities mentioned in that table too and I think certification authority means again that we most probably will have a kind of validation done by that authority.

And the question is would they consider translation and or transliteration of the data yes or no. It's maybe a stakeholder group that we need to address if we want to go a bit more in depth on the responses to our questions.

And we have the individual Internet users still on Page 22 where it is everybody in the world that wants to know who is the owner, who is the registrant of that domain, that's the general public I would say.

And at that point the stakeholder we have been talking about also but then it's the individual who will be enabled to translate and use translation or transliteration to - it's available to him or her in order to be able to translate.

In that case I would say the response to is it desirable would probably be no because it is available for any user in the world to have a translation of data that is provided on the screen at least that's my perception.

I see Jim's hand raised yes Jim you have the floor.

Jim Galvin: Yes, Rudi thank you this is Jim. So I want to say that I agree with you and add something. I think that as we consider here whether or not there will be translation and transliteration and we consider where it's going to be.

I think there's a valuable matrix if you will or cross relationship that we can show between the purposes of data and the users of that data because we could define - I'm imaging that we could define the purpose and place of translation and transliteration based on how we might align users with purposes.

We might decide that a particular user needs translation and transliteration, which will drive its existence in a particular purpose or we could take it from the other direction.

And I think that, you know, we could draw from this report and these diagrams quite nicely if we were to do something like that, thank you.

Rudi Vansnick: Thank you very much Jim for this additional information. It can be something that we have been talking about a few months ago to try to provide a kind of matrix that would enable us to see the purposes of the translation or transliteration.

I said the question is it desirable should have probably more than one condition. It can be desirable for the purpose of yes it's good for people to have access to data that they are not able to read due to the language differences.

But it's also as you mentioned the purpose of it who is using the translated information or the transformed information that states - I think that's a quite interesting discussion that we could have later on.

I think that at this point it's quite clear that there is a need for an additional work on the purposes of the questioning. I'm just scrolling through the other table I'm Page 23 and 24.

You will see that we have the other stakeholders and there is a list of stakeholders mentioned on Page 32, it's in fact a list that indicates already what are the purposes of each stakeholder and that can be very helpful for us to - with regards to the matrix we want to produce.

So Page 34 I'm moving over to that page now. It's a long list it's Page 32 and 33 up until Page 34 if I'm not wrong as the fullest of all the stakeholders and that are in fact - that have been addressed in this report and it's a very extensive list very interesting.

As it gives a view on who is involved in or would be involved in the new RDS platform. And then on Page 34 I made myself some remarks, Page 34 and

talking about the second paragraph that starts at the same time registrants of domain names may desire returning on the local jurisdiction to privacy.

They may not want their contact details made public. That's indeed difficult the privacy of contact information that would not allow to have a public translation and or transliteration but could only be done by a specific group of stakeholders.

I think it's important that we keep this distinction that there is as far as I am concerned a difference between what is publicly available and what is not publicly available due to privacy rules.

And I think that that could be one of the conditional elements in our recommendation that if it's publicly not available while I would say it not desirable to translate if it's not available for the public and it's for a very specific group that the translation or transliteration would be made available.

And in the third paragraph it isn't clearly mentioned, halfway the paragraph the sentence and is at least however in the domain name must have one or more publicly - published accurate and reachable contact that may respond to (Excel) queries and provide a point of reference for permissible purposes of external actors who are affected by the existence of operations of a domain name.

It's a quite important sentence where in fact I feel a little bit of contradiction with the fact that if it's - if I want to keep my privacy I would not allow to make data public.

And I don't know if we - the expert working group have tried to make a distinction between what is officially allowed to be hidden and whether it is officially not allowed to be hidden.

Maybe Jim you've got something that the workgroup has been discussing the separation between what is allowed to be hidden and what is not allowed to be hidden?

Jim Galvin: So this is Jim I'm not sure I understand the question, which workgroup are you referring to when you ask about a discussion?

Rudi Vansnick: I'm talking about this expert working group that has been working on this report. I don't know if you have any additional information.

Jim Galvin: No I don't have any special information regarding this directory services report I was not part of this working group. If you're talking about the registration data working group no we haven't talked about the privacy considerations in particular.

That's not really a question that's in scope we're really talking about the data and the requirements on data that exists. We're not talking about the scope of data as it's used.

Rudi Vansnick: Thank you Jim, which makes clear that something that (Chris), Julie Hedlund and myself have tried to figure out already and created discussions and we had a call yesterday to prepare this call also.

That there is a lot going on in fact the same kernel that we have to work on but that there are - there is no definition or global definition across the working groups and studies on what is being considered in - for instance as I explained what is considered being allowed to be hidden under the umbrella of privacy and what's not allowed to be hidden.

That's the question that I'm trying to figure out for myself for a while already. Yes Jim you have the floor.

Jim Galvin: Yes thank you Rudi. I guess I want to add to my comment before and maybe clarify a little bit with respect to the registration data working group that I'm involved with.

I said before that we're not worried about the use of the data, that's not really a completely accurate statement. Our charter is about requirements for the data in order to support mission and display.

So, you know, display obviously is use of the data but we're not - when I said that we're not thinking about the use, you know, are concerns are not the same ones that this directory services group was concerned with.

I mean we're just concerned that the data itself technically can be displayed whether or not it should be displayed and who should have access to that display are the use questions that would be out of scope for the group that I'm involved with.

This directory services group is attempting to answer those questions or at least speak to those issues. So I don't know if that helps this discussion but I realized that for the purposes of the archive here I wanted to clarify my comment from earlier, thank you.

Rudi Vansnick: Thank you Jim, indeed it helps the more clarity we can have on where the limits are and what the purpose or the goals and objectives of the charter of the working group are the better we will understand if there is - additional working group yes or no. I see also Peter Dengate you are speaking you have the floor.

Peter Dengate: Yes Rudi I was just going to say that I had a slightly different reading of that portion that you were quoting from Page 34. At the very least however any domain must have one (MR) publicly published accurate and reachable contact that may respond to external queries and provide a plan of reference for permissible purposes of external access.

I understood that to be that if the individual registrant went through a privacy or protection or a proxy service that the registration data had to show published, accurate, reachable contact information for that privacy or proxy service and that they would be in a position to respond to external queries.

So that the privacy would still be protected by another level there but that the users who would either have public access or gated access to information would be given access to accurate information that would identify one party that could be a reachable contact that could respond.

Rudi Vansnick: Thank you Peter, yes that's a good point it all depends on how you read the sentences and that makes it interesting to have as many eyes as possible on this report in order to be able to get clarity on what is meant by the wording.

Indeed it could be that the consideration here is about proxy services for instance and in that case indeed there is information that should be available if the proxy service is great.

But it needs again a clear definition that's one of the things that I'm really trying to wonder how it would be possible to have clear definitions on what is meant by certain information that is provided.

At the end I think a glossary on what's discussing the different reports would help a lot in order to make it easy for anybody to understand what the definition or the goal of it is. Yes Peter you have the floor.

Peter Dengate: This is Peter Dengate again for the transcript. I completely agree with you on that point but I think that a glossary or something where we can get some agreed upon terms would be very helpful because we're looking at this report, the report of the earlier study and there are a number of different materials that we're using that are using slightly different terms for similar concepts.

So I think that would be very helpful. I was - I just thought that when you were saying that I just didn't want to be confused. I think that that portion was saying that even if it's protected through a privacy proxy that there should be publicly available information so that the external actors who do want to have a contact do have an accurate - accurate information and a reachable contact.

But that that contact may well not be the registrant but may be a privacy or proxy service, that was my understanding. But again we can - a number of issues that we could get more clarification on.

Rudi Vansnick: I fully agree with you Peter on this point and as you said there should be in the case of a proxy for instance the data should be available of the proxy service itself and that's the point where eventually as we know that proxies have some specific objectives that they could say well we put the information in a language that nobody or that only a certain group in the world can read.

Then the question is it desirable to translate? I would say yes indeed in that case but that's again the point of what's the purpose of the translation. It's an interesting point and I think we will have some discussions later on in the next weeks and calls to try to go a bit more in depth on the definitions to a lower us to be clear in our report what we mean with what we are bringing as a recommendation forward.

I think we have quite extensive discussion on this part of the report already. Going to Page 43 that is - again it's a - the list of items in a table and we are talking about Page 43 - and it's about Points 30 in the table. It's all about the contact information.

And again it's touching upon the issues of the privacy as well as the proxy services. And there is a sample of data that would be provided. In that case it looks to me that that could be indeed very good qualified data that should be made available.

I don't know if there are any questions or comments on this item. I think it's good that we - eventually out of this table what is written there in order to allow us to point back in our reports to what other working groups have brought forward.

And then we have - I thought there was another point that - well if we go on Page 44, that's correct? Yeah, 44, we see a list of in Number 32. There is - there are data elements to be defined to be mandatory. And one of the information is mentioning specifically - it's Point G Registration Agreement language so there is a definition and a clarification on what is the language of registration in that case.

It's one of the first times that in the data model, there is a reference to language. However I'm not sure that that is - that the inspiration came from the fact that there is another working group thinking about translation and transliteration but you never know.

I think that that is also important to pick that up for our future reporting. And then in - on Page 45 in Point 34 we have (unintelligible) is informing us on validators that may collect, store or disclose additional data elements for internal use that is never shared with the (IPS).

That could be, in our case indeed the translated data that could be eventually stored (unintelligible) the (RDS), that's at least what I understood from that report. This is a way that the validator could be a translation office or a body that is validating the information through translation or transliteration. That could be a case for our discussion I think so.

Then going to Page 58 we have the data access principles. (Unintelligible) that page. I don't know if anybody else has seen anything interesting in the previous pages. Don't hesitate to put your hand up and bring it to the table. I'm moving into Page 58.

No one - my screen jumped all of a sudden - and - another page? Back.
About - sorry, I was on the wrong page. It has a list of the data of registrant that is mentioned on top of the page. This is again a list of all the items that should be used as being contact information for the registrant and that is initially what concerns us for our working group.

I think it's just a question of taking note of all the information that is provided on that page regards the registrant. I think that there is a more important information on Page 60 where it's about the Items 53 and 54 where we have the (unintelligible) Point 53 to be truly global the RDS must accommodate the display of registration data in notable languages, scripts and character sets including internationalized domain names.

So it is clear that, from the work group - the Expert Working Group in this case, that there should be space in the data model to allow registration in multiple languages. And that's in fact something that points back to the work group - Expert Working Group where Jim has been involved in trying to define the data model for the RDS.

And I think that that has been also mentioned in that report too that data could be stored in - for instance as a sample in XML tag format so that you can identify which language it is.

The data model itself, as far as I remember, did not specify for instance a street in a native language and a field street in translated language. But again be sure and asking confirmation by Jim, is that correct that there is no duplication of fields in the contact data that allows registration of two or more languages?

Jim Galvin:

So this is Jim. Yes, actually. We - the current data model - the data model itself, which is under discussion and draft we only had principles that we had done in the interim report for the registration data - Internationalized

Registration Data Expert Working Group. But certainly one of the topics that is under discussion is dependency on the results of this working group.

You know, if a decision is made that there must be a single language or script and then that translation and transliteration, you know, could occur we will probably - the discussion topic is recognizing that you will probably need to store two versions of data. You'll need the original - you'll always want to keep track of exactly what you've got from the registrant.

And then also taking on recommendations that came from the commissioned study that was recently published on solutions for translation and transliteration, transformations more generally is the term that they adopted using.

You'll need to keep track of the original language and script and then if you transform it in any way you'll need to store that separately and indicate that it's a transformation and also keep track of when that transformation occurred. And then you'll have issues of keeping those two things in sync and, you know, whether or not you actually did validate that they match and how you validated them.

So the presence of transformation brings additional constraints and requirements on registration data and all of the players here in the system that we're talking about. Thank you.

Rudi Vansnick: Thank you very much, Jim, for the clarification very helpful. As I'm - I have a technical background I'm too often jumping to the technical details. And it is clear and I thank Klaus and Julie to help me staying aligned on what the goal and the objectives are of the - our working group. Our working group has not to define the technical solutions of how to do the translation and/or transliteration.

But indeed I think it helps that we can point to solutions that could be used and in fact Point 53 and 54 in this report gives a clear view on - and even 55 - gives a clear view on what is important and gives some directives to other groups on what to consider if this was - this would happen.

I think that we can move forward now to - we still have 40 minutes left. We can move to Page 62 where we have the Point C, it's at the end of the page almost. It's - and type of RDS user accreditation principles.

As noted in Section 3 the purposes - and some purposes require access to all data elements or approved subset of gated data elements. The principle on the Number 64, any purpose requiring access to gated data requires use accreditation. That's a quite important sentence.

It looks like the user who has to do the accreditation in this case and again the gated data is still, for me, not that clear except that when we look into the graph that is just above - it's at the top of the page the data that is, in fact credit is called the gated data here except if I'm misunderstanding the (unintelligible) of that information.

And then it's mentioned - all access must be purpose-based returning only data elements permitted for the stated purpose. And it goes back to what Jim has said earlier also that all depends on what is the purpose of the data that we want to address or query and what would we do with the data that we are gathering.

But I didn't see any specific items that would help us in clarifying the answers to our questions. It's clear that there are many stakeholders involved at the stage of collecting data or gathering data from queries. But at that point it's rather, for me, the use of the data and not the storage of the data. And use of data can always be handled by anybody to any tool that is not engaging in this case ICANN in the responsibility of the data that is made available.

I don't know if there are any other additional comments or questions on this page? If not (Chris) mentioned Page 69 as being the next one. I have another call in 10 minutes.

Going to Page 69. It's mentioning the validation of data, the validation principles. Trying to jump - the page in front of me. What is meant here as far as I could understand from the report it's a kind of prevalidation of the registrant or contact information of the registrant.

And when reading through the text, increase the occurancy of contact information by (unintelligible) prevalidation to check data prior to use for a new domain name and to promote consistent data across all registrations which would reduce error and fraud of course.

That indeed if - if the language is - you're not able - or the validator is not able to read or understand it could make this assumption difficult. And next bullet is avoid the need to validate registrant or other (PBC) contract data each time the registrant registers a new domain name by performing validation once and then reusing their block of contact data for several domain registrations. It would simplify the process and these work requirements.

But again, in the context of translation or transliteration it makes it not that easy. If you do the validation starting from the translated data I think that we would end up in maybe more errors than we want, yes.

And the third bullet is avoid delay in the processing of a domain registration since validation has to take place at the time of registration. And that's indeed an issue for us too in the sense that translation and/or transliteration could require rereading if the automatic system would not allow to have a correct translation done. It means there is validation process of that translation and that could take some time.

So in that case I think we have to keep that in mind that it (unintelligible) block the registration of a domain name and then the whole question is if that blocks the registration of that domain would the same domain name registered by somebody else in the native - normal language, for instance, script language, have priority on the one that is in validation process?

That's still a question that I think will be needed to be responded to in a later stage. But it's important that there is a kind of point back in this report for our questions. I don't know if there is any other comment on this text?

If not let's move to Page 71 where this - the prevalidation process explained. And it's almost like Point (D) took some notes earlier. It's about the Point C where there is again pointed to the validators. And a sampling in this case was the (unintelligible) the ccTLD managers, the telephone companies that would be able.

And we have been talking earlier already about (UPU) and that is using certain standards for validation of contact data. And I think that's one of the points that they are trying here to - it could be helpful for our report later on to pick up on this page too I think.

Then we have costs and impacts which is a very interesting one. And that's on Page 117. We make a big jump now. (Unintelligible). Okay (unintelligible). And the costs and impacts (unintelligible) page.

When we read through that - those two paragraphs and I'm reading especially the second one where it is mentioned the cost of operating rules are unknown today. That include costs to the entire ecosystem or just to the registries and registrars who offer the Whois services.

We have this question also in who would burden the cost of translation and/or transliteration. And we have here already a sentence that is mentioned that

the costs - it's actually unknown so far for the management of the Whois data
- operating the Whois data.

That means that it is probably not easy for us to have - give an indication on
this one. They are also referring to the use of the trademark clearinghouse
where the holder of a trademark will pay for the service that the brand
protection companies or commercial Whois services are going to deliver.

It could be quite similar to our approach that the costs and who pays for it
could be indicated through a same approach saying, okay, this can be
covered by those who are asking for the translation and/or transliteration.

So being close to the end of the hour Page 119 is about the international - the
international registration data. And it's also tracking on to the RAA. But I
would presume that we can just jump that page and go to Page 121.

Where is the discussion on the use of EPP. Going to Page 121, as I said. It is
clear if we need - if we want to get the data stored in - or delivered by the
registration process in different languages that the EPP will be - there will be
need for a review of the EPP process as actually there is in the existing RFC
there is no information on - or there is no definition of several - use of several
languages in the contact data. So that would probably also have an impact on
the whole implementation process.

But as mentioned earlier in this report that the - if there is need for registration
in notable languages there should be a modification done in the data model
and that will have impact on the EPP anyway and impact EPP means
registry/registrar that could be huge costs if that has to be done for the actual
rollout of the new gTLDs. I think that that is almost unfeasible to date
because the RFC should be first rewritten, there should be a new version
before we can start having new EPP developments.

So that's what our colleague (Chris) has been bringing forward for this call. And I think there is a lot of usable - useful information and data that we need to keep track on and maybe have bring them in our document in order to be able to provide and first initial reporting on - based on the report that - the answers that we got back.

I don't know if there is any other information or remarks or comments that we want to talk about now? Is there anything under the all other business that we need to talk about?

We will - based on the Doodle it is clear that the time slot that we have been chosen from the beginning seems to be the most desirable one so we will keep on doing the calls at the same time, the same day which leads me to the next call next week where I think that (Chris) is also not available for that call.

And so I will prepare that one together with Lars and Julie in next few days also. But I propose that based on what we discussed today that you go through the report eventually also and try to highlight things that you think are of concerns.

So not seeing any additional comments or remarks, no hands up I would like to thank you all for this call and see you all next week. Have a nice weekend.

Coordinator: Thank you for joining...

((Crosstalk))

Man: Thanks, Rudi. Bye-bye.

Coordinator: Thank you for joining. This concludes the conference call. All participants may disconnect at this time.

END