
COPENHAGEN – ICANN GDD: Thick Whois Implementation Review Team
Wednesday, March 15, 2017 – 09:00 to 10:30 CET
ICANN58 | Copenhagen, Denmark

DENNIS:

So the... This is about the Thick WHOIS policy recommendation that came from the PDP working group, and the benefits there is. And these slides will be uploaded to the schedule on the ICANN website. And you'll be able to download it from there.

Talks about the policy recommendation in three parts. And this will give you some definition about the difference between thin and thick. And we're talking about the, this is the implementation timeline, and of course, February 2017, this is when we had published the policies.

Here are the effective dates. So, 1 August 2017 consistent labeling and display must be implemented. 1 May 2018, thick WHOIS for new registration. And 1 February 2019, the transition of the thin registration data.

Again, this is a [inaudible] form of the CLMD implementation timeline, so we've announced on 1 May and 1 February, and then we're here in March, and we're moving toward very quickly to policy effective date, and we will talk more about it right here.

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.

So, 1 August 2017 is the effective date. So, registry operator must update name of keys in [inaudible] part of the domain name, name server, and registrar objects, and registrars need to provide additional CLMD data elements, such as contact email and a contact phone.

And ICANN updates the WHOIS output, verification two, used by the PDT in compliance. So, let's start with the last one of the task, and the ICANN update. So, I'm going to turn to Francisco here. Francisco, what is the status on the ICANN update, WHOIS output verification two?

FRANCISCO ARIAS:

Hi, this is Francisco Arias. This is done. I'm actually checking right now if we have published, I don't recall if we have. But in terms of development, it is done.

MARK ANDERSON:

So, the PDP... Sorry, this is Mark Anderson from VeriSign. Francisco, I can confirm that. The PDP provider published the updated verification tool, so that's been published and is available on their webpages.

DENNIS: Thank you, Mark. You're right on this, I'm glad to see. So, can you give us the status on the other items that registry operator task here?

MARK ANDERSON: Absolutely. As Dennis pointed out, most of the obligations here with [inaudible] are on the registries. It involves the updating the format of the WHOIS output. Many of the requirements are around changing the labels of the fields. So from a registry perspective, what we're preparing is a notice to registrars, letting them know, basically, the before and after of the WHOIS output.

So, you'll see shortly, we're working on the notification now, but you'll see shortly is a notification to registrars, informing them of the changes, and basically just a sample of the before and after of that WHOIS output.

Once piece though, as you can see from Dennis's slides, there is one piece that does require registrars to do something, I guess. And that's to provide the registrar abuse contact email and phone. And at least in Verisign's case, that's not data we currently have. So, that's something we need all registrars to provide to us.

So, along with the notification, what you'll be seeing is instructions on how to provide that information to us. And you know, we're trying to send that out with as much advanced notice as possible, because the effective date is August 1st, 2017. So, we'll need that information passed to us by then so that we can display it in the WHOIS output.

So, look for that notification to come shortly. And if there are any questions or difficulties, as always, contact our customer support. The information there will be in the email notification.

DENNIS:

By shortly, do you mean like within March or April? To give us an idea.

MARK ANDERSON:

Sure, my hope is within weeks. We have a draft of the notification, and we're reviewing it internally to get it prepared to go out. And this will be specific to the CLND, both the instructions on how to provide the abuse contact information to us, as well as an explanation of what's changing with the WHOIS output.

DENNIS: Thank you. Any questions on this from registrars? Go ahead Roger.

ROGER: This is Roger. Thanks Mark. We're actually in the process of reaching out to all of the registry partners anyway, and providing this data frame. We don't touch, excuse me. Our abuse contacts are already in our WHOIS, so again, we can do it officially through whatever channels. But if you need it right away, there are out in our WHOIS today.

DENNIS: Excellent. Any other questions? Feel free to come up and ask questions from the back here. Yes.

Let's move on. So, this is the timeline for the transition policy implementation. We have used this same slide for about a year now, and we've been able to stay on schedule. Isn't that nice? I didn't have to make any changes. It's really hard to change, by the way.

So, the policy effective date, we know, we talked about, in March. So, we're moving toward a first deliverable that is precisely on 1 May of this year. And we'll talk more about it here.

So, let me go over these things pretty quickly and then we'll come back to it. So, what I have done is laid out the task and deliverables based on timeframe. From 1 February to 1 May of 2017, the registry operator provides the documentation reflecting the system changes necessary to support the registrars to migrate the registration data of the existing domain names.

And then also provides the operating test environment, and you'll hear this [inaudible] term frequently from now, and it's a test environment, of course, which would include an EPP mechanism, and an alternate bulk transfer mechanism. And then we will also, they will also provide a transition [inaudible] metrics, with the IRT, and then ICANN will, of course, continue to provide reports of the progress to the IRT.

It's our intention that this IRT stays together until the completion of the, all of the implementation. So, let me ask Mark how this is going.

MARK ANDERSON:

Thank you, Dennis. Mark Anderson from VeriSign again. A lot of excitement coming up, all right? And you know, I was half expecting we'd have a party here today in celebration the achievements of this IRT. Many of us have been meeting for a

long time, and working towards this milestone. So, maybe we can talk about that at the next meeting.

But that said, you know, we do have some exciting milestones coming up. The first, as Dennis said, is the 1 May 2017 day, and that's the point by which we'll have the [inaudible], or the operational test environments ready for registrars to begin testing the transition from thin to thick.

And you know, what that means, from a practical or technical standpoint, is that that's the point of time in which we'll start excepting context. Really, that's the system change that occurs today. If you try and send us contacts, we won't except it at all. We don't have support or the ability to accept contacts, but starting 1 May 2017, you'll be able to perform all of the contact operations, create, update, delete and manage contact within the VeriSign's systems.

And that also means with the domains, you'll be able to associate the three primary contact types, registrant, admin, and technical contacts with those domain registrations. You know, again, we have an obligation to provide notice in 90 days' in advance, but we realize that this is not an insignificant change, so we're working to give registrars as much notice as possible.

You know, we have not sent out the notice yet, but we have a draft of it put together. And this change does mean SDK updates, so it's probably the first thing that you want to take a look at, what the impacts are to the SDK, and what that will, how that will impact your system. The nice thing about this date, though, is you know, these are backward compatible changes to the system.

So, the changes will go into effect in the [inaudible] environment in 1 May, and subsequent to that, in 1 August 2017. These do not require any changes for registrars. Everything you're doing today will continue to work after those dates. Everything we're changing is new functionality, mainly the supportive contact information.

So, you know, this is not a date that is requiring changes to the registrar. But the two key milestones there is that you'll be able to start back-filling data for existing registrations, and for new registrations you'll be able to start creating them with contact information. So, pause there and see if there are any questions.

DENNIS: Questions? Joyce.

JOYCE: Yeah. Mark, maybe the last item ICANN report progress led us to the IRT, how does that work? You report ICANN, or...?

DENNIS: That's really me. This is the ICANN ORG reporting to the IRT in continued manner. You will receive emails from me, at least once a month, but we will also be having monthly meetings as we have been, to make sure that we are all on the same page as we move forward.

This is a long implementation project, it's two years long. And we want to make sure that in the middle, you know, things don't change. It's exactly as we had designed it, and we will deliver as we have promised.

That's the concept. Okay? Go ahead, you have a comment?

MARK ANDERSON: Yeah. Mark Anderson again, for the transcript. I would like to highlight two other things before we move on. In addition to the ICANN reports, we look at one bullet up, there is transition progress metrics. Those will start on 1 August 2017. And there is sort of two flavors to that there is a report that VeriSign will provide to ICANN, and there is a report that Verisign will provide to the registrars.

So, that's part of the required policy. And so, that's a monthly report that basically provides metrics around the percentage of domains that have been, have contacts associated with them, have been backfilled or have been transitioned to thick.

For registrars, you'll see that delivered using the typical VeriSign delivery mechanism. So, if you're familiar with it, we have a SFTP place where you can download it, as well as our UI portal, our VeriSign customer counsel, so that report will be available at both of those locations for registrars to download.

Also, if you're... I'm looking at Roger. You were asking me about this a couple of meetings ago, I think. We'll be updating some of our existing reports, and there will be more information in this, but the purpose of that is to provide more information about which domains have contacts associated with them.

So, in the notice we send out, we'll specifically detail which of the existing reports are being updated. And again, the real purpose of that will be to make it easy for registrars to know which domains have contacts and which don't, to facilitate the back fill for registrars.

Dennis, one other thing, there is a bullet point on there on the alternate bulk transfer mechanism. So, in addition to the EPP mechanism, we'll be providing information about how to provide the bulk transfer of contact information. And they'll be

sort of two flavors to this. There will be one mechanism for the uploading, or the creation of contacts with us, and there will be another mechanism to take those contacts and associate them with domain names.

And again, that information, more detail will be coming in notification itself, but you know, two mechanisms, one EPP and for registrars that don't want to use EPP, we'll also be providing information about the bulk mechanism. Thank you, Dennis.

DENNIS:

Thank you. You're quite busy. So, we'll talk about the other deliverables, subsequent to 1 May, which is an important milestone for us. And from 1 May to 1 August, this is where the registrars have to ask, we have work to do. A registrar obviously will be reviewing all of the notification and documentation that are available to them, from Mark here.

And you're going to be testing. And also, from 1 August 2017, this is when Mark, as you heard, will be deploying the system officially, in a production mode. I have one question that, if the registrars are testing the system, and using real data to transfer before 1 August 2017, do they have to do it again? Or, is the data already transferred space?

MARK ANDERSON: They have to do it again. The [inaudible] environment is a completely separate environment. It has, you know, it has to mediate in there. And one of the things about the [inaudible] environment, registrars actually have two accounts, so they can test transfers. So, it's, you know, there is no collation between what they do in test and what they do in production.

So, the short answer is everything they'll do in the test environment, they'll have to do again in production.

DENNIS: Yes, thanks a lot. I just wanted to make sure that's clear to all of the registrars. Go ahead, Joyce.

JOYCE: Hi Mark. When you do the EPP or the bulk transfer, do you have any sum criteria for the registrar? Either by the number of the domain names, or any special requirement, or it's completed up to the registrar's choice to choose?

MARK ANDERSON: It's completely up to the registrar's choice, which they want to use. There is no criteria, both mechanisms are available to all registrars. Registrars' choice.

DENNIS: Thank you for providing both mechanisms. I know that creates more work for you, but it really helps the registrars, and we thank you and Mark, and VeriSign.

So, from 1 August to 1 February 2019, this is when the transition, or migration of the data happens from the registrars to registries. So, from 1 August, you can start right away, and you can be done in August, if you like, but you do have until 1 February 2019 to complete this task. And we're envisioning that registrar, some registrars will jump on this and get it all done, and then the other registrar will have to probably do some outreaching, and make them aware.

And we have that planned as well. And more so, we just talked about it. We will probably do a, conduct a webinar, as an outreach to the registrar. And Theo, Roger, I mean, new registrars will be invited to that, but Mark has graciously agreed to support that webinar when we do that.

And the... Go ahead, Joe.

JOE: Just to comment on webinars. Typically, when we do these types of webinars with registrars, we do them at multiple times of the day, just to make sure that we've got a reasonable coverage period so that, you know, the stats from the

registrars... A lot of the time, you have technical people, and we want to make sure that we give everybody an opportunity to participate when they're on shift.

DENNIS: Do you do three? For example, to cover three time zones live, and do you record them?

JOE: I'm not sure if we record them or not, but it's a capability.

DENNIS: So, we typically record our sessions, and have it, provide it, but we'll talk about whether we're going to do one, recordings available, or do three live, and it depends on your schedule, mostly Mark. Thank you for supporting us so well.

So, as far as the transition progress matrix is concerned, and that is a matrix or report that VeriSign will be sending to the registrar, this monthly, as Mark mentioned, and I'm envisioning the very first matrix report will be 1 September 2017, right? Because it's one month after the month of the reporting period. Did I get this right?

UNKNOWN SPEAKER: For those who can't see video, I'm nodding my head.

DENNIS: Thank you. Okay. Let's move on. So, next steps. Next steps are really continuing work on the implementation as we have been doing. The IRT meeting will be meeting again on Tuesday 18 April 2017, that's already scheduled. And as we go, we will see if we really need the meeting, if we take care of things online, and if there is a webinar on top of that, we may switch, we may not have the meeting, or we may actually use that particular IRT session to have the webinar.

And we'll see, and we'll work together.

BEN ANDERSON: Ben Anderson, [inaudible]. Just a point of clarification, on your providing an [inaudible] environment for the EPP method. What kind of testing will there be, or platform there will be to test the bulk method?

MARK ANDERSON: The, you'll be able to use the bulk mechanism in [inaudible] as well. You know, everything we do in production, we do first in [inaudible], so the same exact functionality will be available for both.

DENNIS: And as Mark said, there will be a notification coming up in weeks. And we'll learn a lot more, and we'll have subsequent conversations about that. And in terms of the timing of the webinar, maybe the webinar should be better suited after the, all of this comes out with instruction, right? That will be more fruitful.

So yeah, let's plan on that.

Just one quick announcement. We are calling for volunteers for the RDS review team, registry data, right? Services. So, the extension for... I mean, the deadline has been extended. So, those of you who are interested, please submit your expression of interest. Any questions?

BOB: Yes, Bob [inaudible]. So, once this transition occurs, you'll be maintaining a port 43? Is that correct? So, I take it, the question is, around the WHOIS output where we continue to have port 43, as well as the web version. And at least, for the scope of this IRT, yes, those both continue to be a requirement.

UNKNOWN SPEAKER: Okay, so my follow-up question to that is, what mechanisms have you decided to put in place for port 43 throttling, or

protection of the WHOIS data from abuse and scraping? That type of thing.

BOB: Actually, [inaudible], do you want to tackle that? You just spoke on that, I believe.

JOE WALDREN: Yeah, thanks, this is Joe Waldren from VeriSign. So, we already run thick registries with other TLDs. We obviously have a lot of experience in operating WHOIS. We have pretty well established processes for rate limiting.

I think that the systems that we have in place right now, you know, we've looked at those, and I think we're prepared for the changes in the display.

DENNIS: Any more questions? So, look at the time. It's 9:29, we finished with one minute. Okay, go ahead.

MARK ANDERSON: Mark Anderson, again. I do want to make one more note. I think everybody is aware of this, but I'll just state it again. VeriSign is the registry operator for com and net, but there is another TLD, dot jobs, that's transitioning along with it.

We're not the registry operator for it, but we do provide backend registry services. So, for registrars, expect to see separate notices for that. But [inaudible] is the technical provider of that. You know, we're trying to make it as similar and seamless as possible for registrars, but just to be aware, there are three TLDs transitioning and one of them is dot jobs.

THEO:

Thanks. And this is Theo for the record. One final question. We are still moving the data to the US servers, right? In the USA? I mean, it's actually a serious question, because you also have these gTLDs in a lot of geographical location. I just want to make sure.

UNKNOWN SPEAKER:

I'll take that one. So yeah, we're moving the data to the same data centers where we currently have data, and that is the same data centers that we operate com, net, as well as our new gTLDs that we either operate for ourselves, or as backend registry. So, there is really no change in what we've been doing from that perspective.

DENNIS:

Anything else? So, we will be concluding. This is the end of the thick WHOIS IRT session, and we're moving right into the

translation and transliteration IRT session right now. And I know that some of you, we have the same IRT members, please remain. And those of you who are only on the [inaudible], please feel free to come up and join us at the table.

Just make yourself comfortable. This is a working session, and that's how we're conducted. Go ahead. I'm going to turn it over to Bryan here. Go ahead, Bryan.

BRYAN:

Yeah, thanks everyone. It's a very exciting IRT. It involves computational linguistics. You'll love it.

But give us just a minute while we make a switch here.

Everyone ready? Great. Okay. Thanks everyone and welcome to the translation and transliteration of contact information IRT session. As Dennis has mentioned, this is a working session, but I'll give you a quick update on the project's status, and then we'll move into some very fun, as I mentioned, computational linguistic type stuff.

And you'll see just how deep the rabbit hole goes on some of this. So, right now, we're working through a decision tree to address complexities related to implementation of language and script tags. We're discussing some final issues surrounding data provisioning for script data.

We still have yet to discuss language data provisioning, and registration data transformations. So, there will be some significant discussions around that. Our current timeline is somewhat pending. You'll see we do have an implementation plan posted, indicating that we're looking to have a public comment on policy language sometime in Q2.

But underlying the tentative there. Some of this is going quite, getting quite nuanced. So, you'll see when we get into the decision tree. So...

So, you'll see, just kind of... I know you can't read any of that, but you'll see how we sort of worked through this. And I'll just point out to sort of the colors. [SPEAKER OFF MICROPHONE]

What we're doing is, we're asking some questions. We're working off the recommendations that were given to us from the PDP process. Working through this tree. Red indicates that support so far has not been indicated by our IRT members for a given provision. We've spent a lot of time, if you'll look at box one at the top, just discussing what the slash meant in the and or or in one of the recommendations.

So, I think that took about a session and a half to get through that. So, very nuanced. The second question we asked was, how are scripts to be identified? And that's where we'll move on in this session, and if you'll scroll down just a few sort of inches

there, the sort of big discussion yet to be had is, how are languages to be identified?

You'll see that scripts are quote/unquote easy to identify. There are automated means to do that, using Unicode script properties, that kind of thing. Languages are kind of a different animal. There is no automated means to detect a language. The example we refer to a lot of the time is, take a city name like Los Angeles. Is it Spanish? Is it English?

So, there is a lot of nuance to that. Okay. I think with that, are there any questions on just sort of the overall sort of plan or way we're proceeding before we get into this path to here?

Anything? Great. All right. Let's jump right in. So Amanda, if you'll move down just to the right, sorry, so we're on script. In fact, can I take over the mouse for a sec?

[SPEAKER OFF MICROPHONE]

Bear with me with one moment. So just so I can walk you through how we're... Okay. So, we're on point two here. How are scripts to be identified? And I'm going to show you where we are. So, you'll look at these three parts of the tree. Scrolling over, we discussed having just a simple Unicode script property detection, which would be completely automated, versus the

option two, which would be providing, provisioning a script sub-tag per RFC 56-46.

When we have that discussion, we talked about a kind of hybrid option of having a Unicode script property or automated detection, plus an optional provision for those who wish to populate script data for a RFC sub-tag, but it still brings you back to the question, how do you get that script data?

Moving over, we can either have it provided by the registrant, which brings up a number of problems. Or, this option one, highlighted in green, you can infer it, or have the registry or registrar infer what script is being used, using an automated detection method. Moving to the right. How do you infer that?

Well, the method of inference could be subject to additional policy discussion. We didn't want to go around that route. We think we have what we need to figure it out here within the IRT, so the method of inference should be left to the discretion of the registrar or the registry, with guidelines in the implementation, in the policy implementation notes.

So, and to the IRT who has been working on this, this is the kind of new stuff based on our last discussion, and you'll see, I'll just walk you through the language here, we were talking about using Unicode, and we discussed having a minimum version of Unicode to recommend for script detection.

We have been in touch with an expert on Unicode, unfortunately, he couldn't make it to this session, but we still have plans to have him hopefully on our next call to sort of provide his input on that. The next one, second yellow box down. A script tag, if it's provided in RDS systems, must be valid for all Unicode script points detected through automated script detection, and registries must validate that the script tags match with the Unicode script points in the data entry.

And I'm sure that you'll have some comments on that, but I'll just read through all of it before going back. And there was some discussion about validation of the script tag by inference, saying, so basing it off their locale. So, if someone was in Russia, and we didn't know they were using... What's a better example? I think Han, Han script is the one we usually go to.

And that can be used in Korean, Japanese, and Chinese, so you can sort of infer script based on where they are, or where they say they are. So, we'll discuss that. And then we had a talk around private use language tags. Script is included within a language tag, so I know that might be kind of weird that language is used there, but in quotes, I suppose. So, the provision we came up with for discussion is that general private use sub-tags, as described in RFC 56, shall not be allowed in generating script tags.

However, use of private use sub-tag under lang tag in RFC 56-46 shall be permitted, and we also have a language tag expert that we'd like to get onboard, that I actually haven't heard from yet, but we'll hopefully get him in our next meeting or one after.

So, just to go into that one a little bit, that's, I don't know how many of you have read through the RFC 56-46. It's fairly detailed, but there are two sort of different kinds of private use tags. There is a sort of a general high level private use tag that private parties can setup. If I'm describing it correctly. Then within the general language tag, there is sort of another sort of private use sub-tag that could be, we're thinking we can permit in certain cases if need be.

So, that's the kind of new stuff for the IRT to talk about, and I guess we'll walk through all of the boxes and see what we thing, and hopefully get them green or red, so we know which way to go with the policy document. So, go ahead.

I see Jim is raising his hand. Go ahead.

JIM GALVIN:

So, Jim Galvin for the record. Why break from tradition? I'll be the first one to talk. Actually, I should step back for one second. Did you really want to step to the boxes one at a time, or could I just jump right...?

say, gee, this particular script is more clickable in this locale, so let me make this one the priority one.

So, just need that specificity in there, not the notion that I would just go look at the locale and decide what script tag to use. You still have the predecessor moment there of knowing that it's a valid script tag for the code points in use. Makes sense?

BRYAN:

Yeah. That does make sense. So, it's sort of the backup. If you can't figure it out... If you can't tell that it's Han, Korean Han, Japanese, or Han Chinese, you can theoretically look at the locale, and that, and populate the data that way, correct?

Great, so that's helpful. I think we can... So, we'll... That kind of eliminates that far right yellow box, or not eliminates it, but we'll update it. Does anyone disagree with that characterization by Jim? Or sort of disagree with the next [inaudible]... Francisco, go ahead.

FRANCISCO ARIAS:

Francisco Arias. ICANN organization. I don't disagree. I was just going to clarify a little bit. I think what I'll do I shan't speak for others, but I don't consider myself an expert, but it doesn't... That still doesn't seem completely right what we have here. I think we should, I mean, we should advance as we just said,

regarding the inference based on locale, but we should check that and the other boxes where the preliminary decisions we have here, with experts to ensure that we have the proper guidance in the policy.

BRYAN: Fair enough, yeah. We'll definitely cross-check all of this. Any other comments on that pending provision?

Okay. So, I think we'll look at the first box. Do we agree or disagree that a minimum version of Unicode should be recommended for script detection? At least preliminarily that we discuss it. Any objections to that?

Not hearing... Jim, go ahead.

JIM GALVIN: Could you just clarify what's meant by minimum version of Unicode? Just say that out loud once.

BRYAN: Only that the Unicode is, based on my knowledge, is an evolving code point system, that different versions come out every so often. I'm not the Unicode expert, but that we would recommend a minimum version that would have to be used, and as far as I understand based on, I think it was [inaudible] who

said it in the last meeting, that each subsequent version supports the previous one, which I think we were kind of wondering if that was, in fact, the case, but... Does that answer your question?

JIM GALVIN:

Yes. And Jim Galvin again, I guess, for the record. So, for the... The interesting thing that comes up here is, the phrase just sort of struck me, I assume that that was what it meant, that we need to specify what version of Unicode all of this stuff applies to, when we're doing this. And there are a variety of reasons why that's important, and one of which comes to mind is, to some extent, when you're just talking about, you know, contact information, there is no transformation to anything that goes on with it, and we don't use it.

But, on the other hand, you need to know which code points are valid and allowed to be used at any given point in time. Since code points get added with new versions, that kind of thing. And other kinds of rules might change in there. But I also worry a little bit about, it does not apply in this particular case, because we're only talking about contact information, right?

So, it doesn't apply to domain names, but the other thing that matters is the IDNA stuff, you know, fails to work as new versions of Unicode come out, and that becomes very important when

you're talking about domain names, and which code points are valid in domain names. But I don't think that applies... Domain names is not... These rules don't apply at all to domain names in any way, right?

FRANCISCO ARIAS: This is Francisco Arias, ICANN. Correct, I don't think this is about domain names. This is only about the rights of the registration data.

UNKNOWN SPEAKER: So, I don't know if minimal version is the right phrase, but you know, a version of Unicode is sort of what we're after here. But, someone can tell us what they're really after here. [Inaudible] issue since we had this discussion the last time.

BRYAN: Okay. We'll run that by him, and the Unicode expert, which sounds like they're buddies, so we'll hopefully have them both on the next call. Okay. I think we're reaching resolution on that point. So, this one I think may be a bit more complex, in terms of what sort of obligations registries would have. Excuse me.

A script tag, if provided in a RDS system, must be valid for all of the Unicode script points detected through automated script

detection. Registries must validate that the script tags match with the Unicode script points in the data entered. How do we feel about that one?

Thumbs up. I'm getting thumbs up. Jim? Nobody?

Roger, please.

ROGER:

I'm definitely not opposed to this, because I don't have to do any of the work. So, but the one thing I want to, I suppose, clarify a little bit is, if a code point does not match, is a contact, create, fail, contact, update, fail, what's the process when that happens?

JIM GALVIN:

So, Jim Galvin. I would assume that that means whatever transaction is in progress, is rejected. And there will probably need to be some kind of error message that explains why and what's going on. I mean, if the obligation is going to be to check, then it seems that the only natural conclusion here is that you reject it because it didn't match.

So, although it's interesting you say you don't have any of the work to do, you might actually want to do the checkup once anyway.

BRYAN: Okay. So, it might be useful to include some kind of, should I say, guidance on what, and in case there is sort of that error or some things don't match, what the process would be, or is that something that should be handled as the sort of local implementation issue? That kind of thing, which I know is a phrase we've used.

Or, should we look to provide guidance on that [inaudible] implementation notes?

JIM GALVIN: So, Jim Galvin again. You know, from the point of view of standards and consistency, I think that if we agree things should be rejected, then it should say something about that. And then the other side of that of it is, there should be some kind of standard error too, so that clients can figure out what's going on.

I guess I'm not prepared to make suggestions of what that might be, but I can certainly ask our team and get some advice that they can bring to this discussion about what might be appropriate there. And you know, then we'll just see if anyone else has any comments about it.

BRYAN:

Okay. Yeah, Jim, that would be appreciated. Okay, so it sounds like, on this one actually, we may need to provide some guidance on what an error would look like, and what kind of path we would need to pursue under those conditions, but otherwise, more or less, this is reaching resolution as well, I think.

We might move into the language discussion today, that will be fun. Anything else on this one? Okay. Let's talk about the private use tags. I'll read it again. General private use language tags, as described in RFC 56-46, shall not be allowed in generating script tags. However, use of the private use sub-tag under lang tag in RFC 56-46 shall be permitted.

We looked into this a couple of weeks ago. How do we feel about this wording?

UNKNOWN SPEAKER:

Okay, so, what I want to say here, just in the larger community for folks here, is that this reads really badly. But for those who know what it is, I understand it, and it's right, I mean, I'm good with it. Obviously, the issue that we have here is 56-46 describes a way to construct a tag. And unfortunately, it overloads the term language inside it.

So, this makes sense if you understand 56-46, and if you don't, you're probably completely lost, and that's completely understandable. I mean, I'm hopeful and I would expect, and I think that you'll do this, when you ultimately are trying to craft the real language for this, you're going to sort of explain that fact, and what's really going on, you know?

So, that this becomes obvious what it is. Because we're using, what we're doing is using the 56-46 construction of tags to create the script tag. And that's kind of what's going on here. So, we just need to be clear about all of that, and anyway. Just to comment about being careful when you get around to writing this and maybe trying to explain where the, why this is complicated and convoluted.

BRYAN: Sure. Roger, go ahead.

ROGER: Yeah, hi, this is Roger. I'm just sitting here thinking that I have no idea, so this is just a thought. Is it possible that somebody would send multiple language tags on a contact for different reasons? And the only reason I bring it up is, we're saying they couldn't use a private... We don't want them to use a private for

the information that we're sharing, but could they use a private if they're sharing with somebody else?

FRANCISCO ARIAS:

Francisco Arias, ICANN. Clarification question. So, you're talking about having a language stack, sorry, more than one language stack per element? So, for example, for the name of the contact? So, [inaudible] is knowing two language stacks. And now, to your question, if this is a public system, right? So, everything that is put there will be visible to everyone, if they are adding one language stack that is private use language stack, what will they share it in a public system if it's on the internet to be used by certain party?

Should they be communicating privately?

ROGER:

It's Roger. Again, I just that the idea... I don't know that anybody would actually, I don't even know if RDAP would allow it. I don't know if you could pass two language tags for an element or not. But I just thought of the idea, because we're specifically saying they can't use it.

FRANCISCO ARIAS: Francisco again. So, in any case, I think we should check all of this four boxes with the respective experts to make sure that what we are putting here is, revise to be correct, and we put the correct language in terms of the different specifications that are being used. So, in that sense, we could check, I cannot remember if [inaudible] for more than one language tag.

So, that's the first thing that we need to check for example, in that case.

UNKNOWN SPEAKER: And I'm completely okay saying this, I just didn't know if anybody else knew that this happens or not. And again, can't even happen. But again, just sitting here, it just popped in my head.

BRYAN: Okay. Yeah, I think it would be good to have a language tag expert to come and give his input on it. It's a him. We know. Oh, okay. Thanks Jordan. Jim?

JIM GALVIN: Yeah. Roger just triggered something in me. Yeah, I think we have some consistency checking to do across here to make sure this works the way we want. It's not just what Francisco was

saying about, you know, RDAP. I'm less concerned about RDAP and whether or not you can add multiple language tags, I'm more concerned about EPP and what I can specify there in the transaction.

And what's allowed. And I actually don't know offhand, so we need to go check that. But it sounds like Francisco knows, if you want to... Do you want to say?

FRANCISCO ARIAS: Yes, Francisco from ICANN. So, I can tell you right now, that is not in there. If you want language text, you have to extend EPP.

UNKNOWN SPEAKER: Yeah, that's right. Now, we're down to the low key stuff and what's [inaudible] what's not, okay. Yeah, so there is no language tag information there. The other thing that occurs to me is, I suddenly realized that as we're talking about having to validate the script tag, and that it has to be valid for all of the code points, there is... You wouldn't have a...

No, the private use thing... Yeah. So there is an awkwardness in this wording that's here. There is an awkwardness in the wording, in this last box here. So, see if you can track this with me here, Francisco. So, you do actually have to specify a language tag and a script tag, and those have to be a valid thing.

The private use part is an extra label on the end. You know, think about a DNS name being a sequence of labels, right? A dotted sequence of labels, if you will. The tag business, that 56-46 defines, is a hyphenated list of labels. The first of which is the language. The second of which is the script, and then you get a reach, and then you get some optional stuff, and one of the optional things is private use thing.

So, you get to extend what the script label is. It's important to keep that distinction here. So, that last box, it's... You're allowed to use the private use sub-tag, but that does not interfere with the script tag, and the script tag checking, that has to be a valid known script tag only known for the code point. So, the other stuff can be there, doesn't have to be checked. It's just carried along and it goes with the data.

Of course, the other side of it is, as Roger was pointing out, EPP doesn't support any of this anyway. So, this is really only about the transformations that may or may not be happening, either before or after a RDAP transaction. So, Francisco, did you track all of that? Did I say that right?

FRANCISCO ARIAS:

I cannot really tell if it's all right, but one thing to remember here, that this is not about transformation, we haven't even really get to that. This other language text that, according to the

policy recommendation, will go with every single registration from the moment this becomes an effective policy.

UNKNOWN SPEAKER: Sorry. So, say that last part again. The language tag will go with every single registration. What does that mean?

FRANCISCO ARIAS: Yeah. Francisco again. We are not... We haven't even talked about transformation yet. This is a recommendation about adding a way to identify the language slash script, in the registration data. So, this would apply to every single registration, starting whenever this becomes an effective policy.

This is not about transformation yet.

UNKNOWN SPEAKER: You trying to suggest that this tag is going to have to be stored with registration data? And if we don't get it, if a registry doesn't get it from a registrar, we're going to have to vent it and stick it in the database?

FRANCISCO ARIAS: Francisco here. No, I didn't say you have to invent it. And you have, given the language here, I think that would not be allowed. But potentially this starting effective date, I guess we will want

the policy to say that every new registration, the language stack for each element of the registration data, would have to be added and perhaps...

Of course, this is subject to discussion that there would have to be some migration of the existing registrations, so that you add the language stacks or maybe not. Grandfathering whatever. Some position would have to be made, but at least for the new registrations, every single registration would have to carry language stack for each of the applicable registration that the data limits.

That's what, that would be the conclusion from the policy recommendation.

UNKNOWN SPEAKER:

Okay. I don't want to derail the rest of this conversation. I think we're going to have to have a different conversation and start this, because I actually was thinking that all of this stuff only applies if the transformation is about to be in progress, and it doesn't otherwise apply. So, we need to have a different discussion, and we don't have to have that here right now and derail where you're headed here. Thanks.

UNKNOWN SPEAKER: I'm going to add on to the not derailment part, but again, I voiced my concerns several times on this. Recommendation one says that this is not mandatory. Recommendation seven says, when a system is available. So, again, I'll say it once more is, and you'll hear it from me again, most likely. I think that we're solving a problem if somebody chooses to do translation, not forcing everybody to do translation.

FRANCISCO ARIAS: I'm surprised that you guys are surprised. I thought we had discussed this point already. If you read recommendation one, it is correct. It says that transformation of data is an optional element. I absolutely agree with that. But if you read recommendation two, that's the identification of language, sorry, language slash script. That has nothing to do with transformation of data.

And that part, there is no optional element there.

ROGER: This is Roger. And this again, because we've disagreed on this multiple times now, that is not the way I interpreted the PDP decisions. It was a whole, not an individualistic, so. Again, we don't need to run down this, we can... Bryan has purposefully

taking us down the path of solving the problem, and then coming back to that.

UNKNOWN SPEAKER: Yeah, and I think it would be good to have members... The full members of the... As many members of the PDP working group [inaudible] who couldn't be here today, on and I think we'll step back out and have that discussion again, because I know we've had some back and forth on that, and we'll want to definitely clarify that, how we interpret those recommendations, because they are...

When you really start reading into them, they're kind of vague, and as we can see can be interpreted in different ways. So, we'll want to go back and look at the original intention behind it. And Theo, you were on the PDP working group, am I correct?

THEO: No. I wasn't. But I had been following it all of the time. And from my observation, that was never the requirement, the working group envisioned it should not be mandatory. Right.

ROGER: This is Roger. Jim and I were on that, and Jim does have different recollections than I have on some of the pieces, but we were both part of that PDP. So.

UNKNOWN SPEAKER: Definitely, we'll come back to that. That's probably one of the bigger discussions we'll have to have. [Inaudible] has a comment. I think it's referring to, let's see, box two is it? Yes, box two, the middle box. I don't think registries need to validate that the script tags passed by registrars are accurate.

This would be analogous to asking registries to validate that a city or postal code belongs to a country or territory. This validation operation belongs to registrars, in my opinion. Registries will simply store it. Responses?

THEO: Okay, this is Theo for the record. As a registrar, talking about the bit where registrars should do a validation over multiple fields, like the cross-field validation, that discussion is still ongoing outside of this group. Thanks.

JIM GALVIN: So, Jim Galvin. My expectation actually in this process is that nobody is going to specify the script tag. It will, in fact, be

unusual for the script tag in the language tag [inaudible], then we haven't had the language discussion yet, but we're going to get there. So, that will be an exceptional circumstance, not a typical one.

And in that context, nobody is going to validate anything, because it gets script tag for free. You know? Why say what it is, unless it happens to matter to you for some reason. And on average, it's not really going to matter, so you're not going to say, because anybody who wants to know is going to look it up and find it for themselves.

So, with that in mind, if it does come across, I would think that registrars should validate, they should know that they're giving you valid script tag data, because everybody can check it. You know, you look it up in the code points, and you've got to pick the right one.

And if the registrars have to specify something that's valid, it seems reasonable to know that it's there. I mean, this is, I don't know. I'm speaking as a registry. You know, we can go look to, and we'll check and make sure it's the right one, because everything fails. If you don't get this data right, everything fails downstream.

You know? I mean nothing works. So, the registrars should check that it's given you the right script tag, and the registries

should check that it got the right one, if you specify it. But my expectation is that on average, no one is going to set it, because you don't have to.

ROGER:

It's Roger. I think that probably the tripping point here is validate. And that registrars aren't validating it, they're looking it up and providing it. And it's a difference where we're basically generating the data, so we're not validating the data.

So, maybe just a terminology thing.

UNKNOWN SPEAKER:

Yeah. You need to provide a valid script tag, and you get a valid script tag by seeing what script tags are allowed for the code points in use, and selecting an appropriate one from that short list. And sticking it there. But you only have to do that if it's important to you to say it, and on average, I would expect nobody is really going to do it.

UNKNOWN SPEAKER:

Okay, I think that makes sense to me. I'm glad this is being recorded though. All right. Any more discussion around that? Around Dennis's comment?

JORDAN: Doesn't sound like anyone actually.... So, Jim and [inaudible], are you guys agreeing that a registry does not have to validate in that scenario? If a registrar is obliged to generate a valid tag, does it therefore mean a registry doesn't need to validate it? Because it can be assumed that the registrar did the right thing.

UNKNOWN SPEAKER: No. What I'm saying is, if a registrar decides to provide one, then the presumption is that they're going to provide a valid one, and the registry is going to confirm that they did that. But in general, I would not expect them to provide one, and therefore there is nothing for the registry to check.

UNKNOWN SPEAKER: Right. So, right now there is text up there that says registries must validate. And I think... [CROSSTALK] was registries shouldn't have to do that, so it seems like you guys talk through why registrars might be... The fact that it might not happen very often, and the fact that registrars should generate a valid tag, but we didn't actually get consensus on whether or not you actually the registries, whether or not to keep that sentence or remove that sentence.

UNKNOWN SPEAKER: I think what we agreed to is, the sentence needs to be revised. I mean, we're clear about what's going on. We sort of get the principle that's being suggested there, but the words taken in isolation don't match what we're really thinking.

UNKNOWN SPEAKER: And these yellow boxes mean this is all sort of evolving language. So, usually what we do is go back, listen to this, and update it.

DENNIS CHANG: This is Dennis Chang. ICANN org. So, just to be very clear, when we redo the language, we will break it up into what registrar must do, and what registries must do with those requirements language to be used. Registrars must select from a valid list, that's what they have to do and provide it to a registry.

Registries must validate that they have, in fact, chosen [inaudible] to keep the records clean, and that's what we were intending to do. So, there is an obligation on both parties.

UNKNOWN SPEAKER: Right. And the only thing to say in front of that is the registrar does not have to put a script tag in there, but if they do, this is the way you do it.

UNKNOWN SPEAKER: Right. I guess like, I really don't care that much, it just feels like no one has actually provided a rationale to... So, someone has said, I don't feel like registries should have the obligation to do the validation, and we've had a bunch of conversations in this room and literally, not one person has said why a registry should have to do validation.

So, I don't think... I'm not advocating... I'm not joining that point of view that they shouldn't have to do it, but it feels like there should be a... If you're going to impose a requirement on a registry, someone should be able to articulate why that requirement exists.

UNKNOWN SPEAKER: You can't lie about the language and script in use. So, I mean, you're trying to make... With the push towards accurate, you know, registration data, okay? You need to have accurate data. I mean, the downstream consequences of getting it wrong, means you can't do translation and transliteration if the label it goes with it is wrong.

UNKNOWN SPEAKER: [Inaudible] you just put it in an affirmative obligation on the registrars to provide a valid... Like, if the registrar... If you

move the obligation to the registrar, than the registry doesn't have to do anything, and if it's screwed up, it's the registrar's fault, not the registry's fault.

UNKNOWN SPEAKER: Okay. Well, speaking as a registry, I'm very happy to put the obligation on the registrar.

UNKNOWN SPEAKER: [Inaudible]. And again, and we're not taking any more responsibility than we already are, so I mean, the first part of that is making us responsible. I think that Jordan's right. I think that's what Dennis is saying, he doesn't feel that registries need to do anything besides accept it. And again, to me, that wording doesn't matter to me, because I'm not a registry and I don't have to do any of that.

We still have to do the first part, which is our obligation of providing a valid one. So.

UNKNOWN SPEAKER: So, I would suggest you remove that sentence. It sounds like everyone agrees that the registrar would do it anyways, and there is no need for the registry to also do the validation.

And if it's screwed up, compliance can get mad at the registrar.

ROGER: This is Roger. I'm not suggesting to remove it. Again, that's not my choice.

UNKNOWN SPEAKER: I'll agree with Jordan to not require the registry to validate. And you know, you're right, just put an affirmative action on the registrar to do the right thing, if they're going to do it.

UNKNOWN SPEAKER: That's helpful. We'll revise. Dennis?

DENNIS: Yeah, so in terms of requirements, we have multiple languages to use, must validate versus should. So, you might consider changing the language to should, because I'm concerned about the accuracy of data going in, and I'm not saying that registrars will do something wrong, but in the transmission of data, anything can happen, and some garbage may get in there.

It would just be very nice for registries to automate their checking process to ensure that what they have in their system is not corrupted. That's all I'm saying.

UNKNOWN SPEAKER: The word should seem like an excellent choice. If you put it in capital letters. So that compliance understands that it's in the IETF form, not in the...

DENNIS: Yeah. So, what we are doing with the policy implementation at ICANN is adopting the IETF language for the requirements. So, we are going to make every policy come out using those terms, and must, and should. And we're actually thinking about just remaining, or anyone, that's besides the point.

So, yeah. It's all caps. And that's what you are looking at if you looked at our latest thick WHOIS policy language. We preface all of the language document with the paragraph.

UNKNOWN SPEAKER: Right. And so, this is... We were going through iterations of a policy language document, Jordan, just to sort of bring you up to speed. We realized that it sort of, you hit the sort of roadblocks, and we created this decision tree to kind of help us work through that.

And what we're doing is, we're giving that some of the recommendations can be interpreted in different ways, we're kind of working through this tree to see how far we can take it before we hit these kind of roadblocks. Sounds like we're

getting there already. So, but this discussion does get integrated into each subsequent policy language document, which we should have in a few weeks, maybe a month.

So, yeah. I mean, RFC 56-46 specifically says, only use as much of the language tag as you need to use. So, it doesn't even recommend using script tag, it just says, use as much as needed to get your result.

UNKNOWN SPEAKER: That's an important point, yeah.

UNKNOWN SPEAKER: I think we've had a good script tag discussion. I think we have enough to go off of to clean up the language in the third box to make more sense. And change... What did we say about validation? I'll have to go back and listen to this, but we're not validating, we're confirming. Was that the...?

We're not doing anything.

UNKNOWN SPEAKER: Registrars, if they're going to specify, they must specify a valid one. And as is consistent with the rest of registration data coming to a registry, the registries just take it and keep it.

UNKNOWN SPEAKER: That makes sense. Francisco?

FRANCISCO ARIAS: Francisco Arias, ICANN. So quickly on another regard. I posted the text of recommendation four. It says the working group recommends that regardless of the language, less script, use it is assured that the data fields are consistent to standards in the registration accreditation agreement. [Inaudible] consensus policy, original WHOIS information policy, and any other applicable policies.

Enter contact information data, are validated in accordance with the aforementioned policies and agreements. And the language slash script used must be easily identifiable.

UNKNOWN SPEAKER: Does that...? Is our discussion consistent with this?

UNKNOWN SPEAKER: I think so. I'm not quite sure what you're driving at, Francisco.

FRANCISCO ARIAS: I was simply pointing out that we have specific recommendation in this regard, and I will up to you what it means.

UNKNOWN SPEAKER: Oh, are we back to the question of whether or not language and script tags have to be visible with all registration data and stuff?

FRANCISCO ARIAS: Not at all. This is about validation.

ROGER: It's Roger. And I don't think that we're inconsistent with this. One of the keys in this is the data that's provided to registrars, we have to validate. Script tag, language tag does not provide it to us, we're generating that.

So, we don't have to validate that. We're providing valid data. So, I think we're consistent with that recommendation.

UNKNOWN SPEAKER: Yeah, it says entered contact information data are validated. You know, if you specify a script tag, and they will affirmatively specify a valid one, and you know, that's sufficient. Yes, we've covered it. But if they don't enter it, there is nothing to do, right?

UNKNOWN SPEAKER: Do we agree?

UNKNOWN SPEAKER: We're going to have to come back around to this anyway after we have a language discussion. There is going to be nuances in here we already know. We have to go back and revisit all of the recommendations once we've gotten through the technical details of how this stuff works.

UNKNOWN SPEAKER: Right.

UNKNOWN SPEAKER: What if the contact was entered incorrectly? Registrar. And the registrar doesn't validate that?

UNKNOWN SPEAKER: The registrant doesn't give you this information, the tag value, anyway. The registrant will never give you the tag value. Well, I don't know. I shouldn't say that. I suppose a registrar could set up an interface that lets a registrant tell them what they want to tell them. Honestly, I just don't see that as practical. Certainly, I can't expect anyone to do that at scale, but you know, who knows what individual people are going to do.

I think the general rule applies that Jordan had said earlier, which is, you know, garbage in, garbage out. The registries are just going to take it, and somebody, somewhere along the way is

going to notice it's wrong, and they're going to complain, and then all of the right, you know, complaint processes come to bear.

UNKNOWN SPEAKER: Jordan, did you want to come in?

JORDAN: I was just saying, there is an affirmative obligation to provide valid data here, it's just sitting on the registrar. And so, that's who compliance goes to if they're unhappy with... You know, if there is an instance in which the tags are not valid, then compliance goes and says, registrar, why are you providing invalid tags?

UNKNOWN SPEAKER: Makes sense. Well, we have four minutes. I don't think we're going to move into anything about the language discussion, which is good. We'll save that for our next call. Jim?

UNKNOWN SPEAKER: Could you move it down to where the first question is anyway?

UNKNOWN SPEAKER: Sure. That's a good idea. Just a preview where we're headed.

UNKNOWN SPEAKER: Yeah, stay to the left. I just want to see the beginning part. And yeah, leave that up there. Go ahead.

UNKNOWN SPEAKER: So, we have this question, just for you to think about on your airplane home. How are languages to be identified? We have a similar set of options. We have this sub-tag we've been talking about. We can populate data for that. Question remains how you obtain language data on that sort of top yellow box 3A, I think that is.

And then there is option two, which language inference would be left to the discretion of the transforming party? Francisco?

FRANCISCO ARIAS: Francisco for ICANN. Are these really two different options? RFC 56-46, this should say language [inaudible], oh no, no, no. Sorry. That's fine. But so, the language [inaudible] is the [inaudible] is what? It's one [inaudible] to... Is a protocol to provide information. Option two, which I don't think is really an option. Two is about how you get that information.

So, I don't think this is correct.

UNKNOWN SPEAKER: All right. So, let's not jump into that now. But you can see the kind of discussion we're going to be having more or less preview, and we'll work on the nuance of that. But Roger, go ahead.

ROGER: This is Roger. Now that we're starting to move into this, I wonder if this is something that we're going to have to take back to the GNSO, through the discussions I've heard several members, I don't know, mention that the practicality of this doesn't meet what they thought during the PDP.

And I wonder if this is a IRT responsibility to go back to the GNSO that we can't, I guess, resolve this issue without further possible further policy development, or the GNSO provide guidance on it.

UNKNOWN SPEAKER: Yeah, this has been sort of hinted out as we move through this and seeing what it actually starts looking like. I think [inaudible] mentioned that, a meeting or two ago, so I think that's something we'll start having to look into fairly soon. So, Jim.

JIM GALVIN: Yeah, so even to agree with Roger, and maybe speak more directly to the questions and comments that Francisco has with respect to the recommendations, I think I've stated pretty much

unequally as we got down to this thing, my thinking on all of this has evolved since this working group took place, and I think that the recommendations overstate what's possible.

Roger, much more tactfully, phrased it as the practicalities of doing this are a little bit different than we thought, and would have expected. So frankly, I'm going into this, we're doing this discussion right now of what's possible to do, on a very specific technical level.

Then we're going to have to go back and map that onto the recommendations, and I fully expect that we're going to find that the recommendations are not going to be able to apply, and that we're going to have to use the process to push back on the council, with appropriate justification and everything, and all of the explanation to look for some modification in some of those recommendations.

I don't think it's going to come out the way the recommendations are. We're going to have to push back, and push back in a positive way. I don't mean to make that a negative thing. This is a good thing. We're all going down this path of doing this. It's just the practical issue of exactly what it takes to do some of this, and what's possible.

And you know, we'll sort all of that out once the language discussion, thanks.

UNKNOWN SPEAKER: No, that's well put Jim, and thanks for that. Yeah, we want to get it right. We want to make sure the community agrees with it and can do it. And it's not unduly onerous. So, I think we're at time now. This has been a great discussion, as usual.

I think this group makes this kind of stuff fun. So, appreciate that. And we'll go back, revise the script discussion, get back to you in a couple of weeks and have our next call, and hopefully get the Unicode and language tag experts on board for that one. Jim?

JIM GALVIN: Just as a logistical thing, meeting in a couple of weeks. Yeah, the IETF is in two weeks, so we're not going to meet next week, maybe we meet in three weeks?

UNKNOWN SPEAKER: Okay, I'll make a note of that in my scheduling. So, thanks.

UNKNOWN SPEAKER: Anything else? We have about 10 minutes left. Oh, I see... This comment, one second.

[END OF TRANSCRIPTION]