
TORONTO – WHOIS Update

Wednesday, October 17, 2012 – 11:00 to 12:30

ICANN - Toronto, Canada

BARBARA ROSEMAN:

My name is Barbara Roseman. I am with the ICANN Policy Department. I'm overseeing the WHOIS studies that were undertaken at the behest of the GNSO and what we have today is an update on a number of different activities that are taking place within ICANN related WHOIS. We will have updates on two of the studies and an update from the IETF on the RESTful WHOIS Protocol that's being developed.

So we have up here with me Steven Pedlow from NORC. They have been working on WHOIS Study 2. Murray Kucherawy, who is the co-chair of the IETF WEIRDS Group. Francisco Arias from ICANN, who will be talking about the roadmap, I believe, and Lyman Chapin from Interisle, who worked on WHOIS Study 4.

So the basic thing here is that there's really a lot going on with WHOIS right now. A lot of different activities and, as we all know, it's a controversial issue within ICANN. The attempt of the studies have been to gain some factual evidence that we can use in planning future policy, and I think that we're getting some very good results from that.

We're going to start today with Lyman Chapin's presentation and then go to the IETF work and then come back again to Steven Pedlow for the other study. Lyman, unfortunately, has to leave us for a few minutes while he runs out to a meeting and then he will be back again for

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.

question and answer. So Lyman, would you like to go ahead and start...do you have the cue thing for the...?

[background conversation]

BARBARA ROSEMAN: Great, excellent, thank you.

LYMAN CHAPIN: Thank you Barbara. The study that Interisle conducted was one of several that were commissioned by the GNSO Council going back as far as April, 2011. And the point of the study that we conducted was to determine whether or not the conditions were available to perform the kind of data collection that would produce genuinely useful data for the council in making policy decisions about, in particular in our case, the use of reveal and relay requests to obtain information in the WHOIS system that was in one way or another located behind a proxy or privacy service that created an additional step in order to find information associated with a registrant.

The reason that we had to conduct a feasibility study at all is because when the GNSO Council originally authorized a study of relay and reveal request processing they did not anticipate the response that they would get, which was essentially, "There are too many unknowns. We..." We being people who might potentially conduct such a study. "We simply don't know who is out there; would they be willing to participate? How



much data might really be available? What would the privacy concerns be?" and so forth.

And that led to the idea that perhaps before we try to design a full study of relay and reveal that we go out and determine whether or not it would be feasible to do so. In other words, if you conducted such a study, would it be likely to succeed? So that was the point of this feasibility survey. It was also, in addition to determining feasibility, it was intended to gather information that would enable us to better design the full study. And also to identify potential participants and give them an opportunity to make themselves known so that if a full study were conducted those people could participate in it could be contacted.

To give a brief high-level summary of our findings and I encourage you to, obviously, read the report for full details, we found a very clear expectation among all the folks that we talked to, whether they were relay/reveal originators or responders, proxy or privacy service providers and others who were knowledgeable about the WHOIS system, that the last thing they wanted to see from any study was simply a data collection exercise that led to no concrete or tangible progress on WHOIS.

And I'm sure all of you are familiar with the difficulty that we as a community have had in making progress on identifying what it is we're trying to do with WHOIS and how we can best do it in a way that respects the rights and obligations of all the different players. So there was, built in to almost all of our exchanges with potential participants, was this expectation that there would be tangible results. And of course that's not something that certainly we could promise. It's not



something that ICANN or the GNSO could necessarily promise, because the process produces the results that it produces.

And therefore, the expectation that there will be tangible results might itself serve as a hindrance, because a lot of people said, "You know, if this is just going to be another data gathering exercise and it isn't going to lead to any improvement then I'm not really interested in participating."

The participation interest would also drop considerably if a full study were perceived as simply an attempt to identify bad actors and track down people who were violating laws or something like that. If it were perceived as primarily a witch hunt or an attempt to assign blame to one part of the community over another, the likelihood that people whose participation would be critical to the success of the study would actually be willing to talk to any study conductors would drop dramatically.

So it would be very important to design the study in such a way that it was not perceived as being intended simply to go on a blame hunting expedition. But we also found that in many cases, in fact in most cases, although people were willing to provide data on the way in which relay and reveal requests were processed at various steps in the chain that those data would in almost all cases be aggregated or anonymized.

That we would not be able to obtain, because the participants would be unable or unwilling to provide, data that could be used, for example, to track the life cycle of a particular relay or reveal request or even a particular class or category of relay and reveal requests. So we would be able to get aggregate data on the way in which these were processed



by the different actors, but would almost certainly not be able to get individually identifiable or specific data.

The ability to get any data, useful data, would however depend on the availability of privacy and confidentiality guarantees, and we have a little bit more about that later on. So I have four conclusions from the study to share with you, and we can go in to these in more detail during Q&A later on.

The first conclusion is perhaps the most obvious and the most important, which is that a full study of privacy and proxy reveal and relay would be feasible if it were defined in such a way as to resolve the barriers that we uncovered to participation. And that such a survey, designed with those constraints, would provide some but not all of the data that were anticipated by the GNSO Council going back to April, 2011 when they originally commissioned and authorized this study.

The original intent of the council resolution, going back, was to provide data that would enable you to actually examine the way in which a class or even an individual relay or reveal request was processed, from the originator through the service provider, and in the case of relay, to the next party in line who would be the holder of the actual registrant data. And our results from the feasibility survey showed that it would not in fact be possible to get data with that kind of granularity. In other words, the data that would be available would be anonymized and aggregated.

We did, however, find...and this was somewhat surprising, that pretty much across the board everyone, all different kinds of participants...requestors, processors and so forth, would welcome such



a study. In particular, we found that there was a category of firms and organizations involved in providing proxy and privacy services who were eager to have such a study conducted because they felt that it would highlight the fact that their practices were, in some sense, clean and legal and above board and so forth. In contrast to the evil practices of the few sort of malcreants or whatever...miscreants that were giving everybody else a bad name.

So it's not just the people who are frustrated by their inability to obtain accurate registrant data through the current relay and reveal system who would like to see this study conducted, it's actually the people on the other side of the equation who would like to see the study conducted as well because it would distinguish between people who were, in some sense, doing it right and people who were trying to use it as a way to shield unsavory or illegal activity. That was an interesting conclusion that we hadn't expected.

It was very clear that confidentiality and convenience were both important, so the opportunity to anonymize or otherwise provide confidentiality guarantees with respect to any data that might be provided was important to almost everyone we talked to. So it's clear that any full study would have to enable people to respond, either anonymously or with confidentiality guarantees with respect to the data that they might be providing about their customers.

Convenience actually was an interesting additional requirement. No one that we talked to was interested in doing any engineering or instrumentation work, for example, to collect data other than the data that they collected in the normal course of operations.



So, for instance, we didn't find anyone who was interested in adding to their system a probe or a sensor or any kind of data collection software or even just a process, adding to their process, that would gather information. They'd be happy to share information that they collected in the normal course of their business, but that's as far as they'd be willing to go. They were not motivated to go on, to go beyond that.

And the fourth conclusion, the final conclusion, is that we conclude that a full study conducted according to these constraints and observing the confidentiality requirements and so forth might not satisfy the expectations of either the community or the council with respect to the validity of the data because it would be difficult to establish and confirm the validity of the data. Or independent verifiability, meaning the ability of a third party to look at the data set and be able to confirm that in fact it was accurate, representative, and so forth.

So what we really have to think of is would a study that had to be conducted under those constraints provide data that would be sufficiently useful to the GNSO, that it would be worth conducting the study?

In other words, if you did this study you would have these kinds of data with these kinds of characteristics and maybe these encumbrances. Is that something that the GNSO could effectively use in its policy discussions or would it, in fact, not be terribly useful? Obviously if those data would not be useful with those constraints then it wouldn't be worth doing the study. That was the key point that the council was trying was trying to establish in doing the feasibility study.



We did decide that it would be possible to obtain certain kinds of data from a full study, but of course we were not in a position to make the judgment call as to whether or not kind of data, that type of data set, would be sufficient to meet the expectations of the community or the council.

So I'd be happy to...Barb, do you want me to take questions now or take questions later? That's fine.

BARBARA ROSEMAN: I think it might be easier if we take questions together at the end, because of your having to go and come back.

LYMAN CHAPIN: Okay.

BARBARA ROSEMAN: Thank you. So Murray, would you like to go next?

MURRAY KUCHERAWY: Good morning, my name is Murray and I'm the co-chair of the WEIRDS Working Group at the IETF WHOIS, tacked on the technical side of this...one more slide?

[background conversation]



MURRAY KUCHERAWY: Just keep going?

BARBARA ROSEMAN: Can you put up the next presentation please?

MURRAY KUCHERAWY: Oh, I'm sorry. That's it. Great. So there's a bunch of interest, on the technical side, of coming up with a replacement for WHOIS. Some of it is driven by the policy discussions at ICANN; some of it is just driven by community interest.

The main points are that WHOIS, defined in RFC 3912, the conventional WHOIS protocol has never been equipped with the ability to deliver internationalized data, which is becoming obviously very important as we move in that direction. Another important point is that there is no proper data framework to go with the answer.

Two different WHOIS servers you talk to could deliver dates in completely different formats, or not at all. Phone numbers, the same. Things can arrive out of order which makes...it's very easy for humans to find what they want in the output, but for any kind of automated system that wants to access WHOIS and do something interesting with it these points stymie that kind of effort.

There is no current capability to support differentiated access for service providers that wish to do so. For example, if you want to show a different set of fields to different clients the only way you can do that now is based on an IP address. And IP addresses these days can be shared by many different clients. So if you want to do rate limiting, if



you want to show different fields to different clients that simply isn't possible in a reliable fashion right now.

And I try to clarify that when I talk about privileged access to the data, it doesn't necessarily mean I'm going to show PII to these guys or not. It could also mean I'm simply not going to let you do a certain number of queries per hour unless I can positively identify you as someone who should have that access.

So the IETF we've chartered, there's sufficient interest to create a working group called WEIRDS. There's the full acronym spelled out there. We are chartered to standardize a data framework. It's clear that we want to do this, that was one of the major driving points in the beginning and all these other things came along as well.

Deliver objects encapsulated in a framework, a simple framework that makes the software easy to deploy, easy to layer on existing databases, and reuses the public code base as much as possible, which lands us squarely in the field of doing HTTP and RESTful work. Those familiar with the technologies, I'm not going to go in to those here unless we really want that, but you'll know that this makes it...there's a huge amount of code out there on which you services in this way.

We have not forgotten that this is...we're not under the delusion that this is the first run we've taken at replacing WHOIS. There was the CRISP Working Group, which produced a number of RFCs some time ago to produce an attempted replacement that received just almost no adoption at all due to its complexity in other matters.



But the requirements that were spelled out during that work are essentially the same ones we have today. And so we have not forgotten them and we are using them as a reference and we're reminded of them regularly.

One of our primary goals is to produce a simple, easy-to-implement protocol that does support internationalized data. There are several ways to do that, those kinds of encoding, in basic ASCII protocols and we're going to ensure that that happens.

We will support the capability of differential service and address the needs, as much as possible, of both name and number registries in a single protocol. This work started out very clearly divided that we wanted to produce a set of protocols for the names which is what ICANN's interested in, and also for the numbers, the RARs.

There has been a shift in thinking that there's so much in common between the two, there's more in common between the two than there is different. So why not produce one base protocol that can handle both of those constituencies and only differs in a few certain key areas? So we want to try to unify them as much as possible.

So very briefly, what is a RESTful service? It's something based on HTTP, a very simple interface. It is stateless, which makes the back end very easy to implement. Everything is represented by a URI; you can see an example here. The host name tells us where the data is, the middle part is what type of question you're asking, and the last part, the KOSTE-ARIN piece, is the specific piece of data you're requesting.



This is actually a live URI, you can try that and you will hit a RESTful WHOIS server. There are other properties of REST, the REST principles, that mean these services can be deployed and expanded in very interesting, modular, and simple ways.

And we already actually have running code. Five of the six RIRs have, or four of the five...I can't remember, have stood up prototypes of this and, in fact, ARIN's prototype now sees more traffic in this new WHOIS system than in traditional port 43 seas. So the momentum to do this, at least on the number side, is very serious and I'm hoping to capture as much of that as possible for the names work.

So the working group was chartered in April of this year. There are five major documents we're working on right now, each of which will theoretically become RFCs. The basic description of doing this work: using HTTP is the first one. How to form a query using a URI is the second. The reply format, which we've landed on JSON, is going to be the third. We did look at XML and opted against it.

There's a comprehensive security considerations document that talks about how you would go about doing certain data privacy, client distinction, that sort of thing. And a fifth one about how to do redirects the way ARIN does. For example, if you ask ARIN for information about an IP they don't own they'll simply redirect the query to the new location.

The current set of milestones has this last one completed by December of 2013. They don't all have to go in one big omnibus package for approval; we might get the first several done while the last one is still being worked on. So that is not a full deadline. The approval policy, the



process through the IETF, is usually a couple months beyond that. So we're looking at probably first quarter of 2014 before it's all finished, but some of them may be done well ahead of that.

I'm going to skip these two unless we need to talk about them. But to get involved, if you want to monitor this work or participate, the next full meeting of the IETF is November 4th to 9th in Atlanta.

[break in audio]

FRANCISCO ARIAS:

...a WHOIS protocol, which is happening in the ITF. But in parallel to that there is a need to work in the ICANN circle to work on the adoption of the protocol, the future adoption of the protocol. So the genesis of that is a report by SSAC, the SAC 051 report which is a report on the WHOIS terminology and structure. And it has some recommendations that I'm going to talk about in a few slides.

Before that, there is just some background on the current environment where this report happened. There is a GNSO/WHOIS Service Requirement Report that lists various limitations of the WHOIS protocol and calls for some of the things that would be nice to have on the WHOIS protocol.

The results of the joint GNSO/SSAC Working Group on internationalized registration data resulted in requirements that should be considered for including registration data in WHOIS that is beyond asking this to support other writing systems like Chinese, Japanese, Arabic, etcetera.



And the results of the WHOIS Review Team final report also has some related recommendations.

So before SSAC produced SAC 051 they had produced in the past several advisories related to the limitations on the WHOIS protocol. But in SAC 051 they clearly are requesting the replacement of the protocol to address these limitations. And the board took that report and requested the producers of the roadmap to implement those recommendations.

This is a timeline of how things happened. Last year is when the board produced the resolution. There is the text, and we produced this draft roadmap that was subject to public comment in February, during February to April I believe. And then during that public comment period we had a very good consensus of what the patch should be and so we produced an updated roadmap that was produced on 4 June 2012.

In terms of the context of the SSAC report, what it does is clarify the taxonomy, makes some observations and offers some recommendations. The observations, there is no uniform data model for the domain registration data. What does this mean? Every registry and registrar, well not every, but many registries and registrars have different forms to represent the data and the elements of the lists.

Of course I already mentioned efficiency of not being able to handle internationalized data, and potentially also the lack of support for differentiated access to data that is perhaps some access to those parties that have some authentication. This is something that already exists in the ccTLD and gTLD world.



So the recommendations from the SAC 051 report are to adopt the terminology, a terminology that was listed on the report. And recommendations two and three can be summarized in the replacement of the protocol to address the issues listed.

So now talking about the roadmap that was produced by the request of the board in response to SAC 051 document, with regards to the terminology one of the things that were revised in the updated version of the roadmap in directly adopting the terminology is to have a small implementation assistance group to develop a proposal on the terminology. The issue was some people found the proposed terminology by SSAC to be a little bit too long. So this is something that is still to be done.

And this is the rest of the steps, basically to prepare a summary so that everyone in the staff and the community know the proposed terminology. And then have a slow transition with the reports for the future. Just to give you an idea of the kind of things that this terminology relates to, perhaps the tool, the best example that I can think of right now is the differentiation between the data itself, the registration data. For example, the domain name, the name servers, the contacts associated with a domain name. That would be the registration data.

On the other side you could have the protocol that allows you to access that data, for example the old port 43 WHOIS. That's one way you can access that data, but you can also access that through web interface or through one of the protocols that had been produced in the past like Ares.



In terms of the recommendations to replace the WHOIS protocol, here are some of the highlights of those recommendations. It's to promote the participation ccTLD and gTLD registries and registrars in the development of the protocol. In Murray's presentation you saw the link that gives you access to the pointers that you need to know if you're interested in participating there.

On the ICANN side, there is a recommendation to have eventually a PDP to replace the WHOIS protocol. What's the timing of this? That's still unclear. I will say whether we should wait for that protocol to advance a little bit more before this should be started or should this be started immediately so it runs in parallel with the forum in the ITF. This is still something undecided.

In the meantime, things that can be done is negotiating the inclusion of provisions in registries and registrars' contracts, as appropriate to regarding this replacement of WHOIS. In this I can say there are two examples of what has been done. VeriSign, for ".com" and ".name", in the recent renewal cycles of the agreements they agreed to include a provision regarding the future adoption of the replacement of the WHOIS protocol, and also showing their commitment to developing this protocol in the ITF.

There is also another thing that came up as part of the public comments, this is the second recommendation shown there, is to have an open-source project that will be available for anyone in the community to use. Registrants, registries, etcetera, so they can easily and cheaply implement this RESTful WHOIS protocol once it's standardized. In that I can report now that we have been talking...we



did a [ERP] and we selected a provider and we are in the later stages of finalizing that agreement so soon there will be an announcement. So we already have this provider selected.

And lastly, is promote adoption of this replacement protocol within ccTLDs. As some of you may know, in the ccNSO the subject of replacing the WHOIS protocol is outside the scope and the policy development process so there is no option to liken the gTLD space to do something in that regard. So it has to be voluntary adoption by the ccTLDs.

This is just a graphic of what I just said, and so I think I'm going to skip it. Thank you.

BARBARA ROSEMAN:

And can we have the next presentation please?

STEVEN PEDLOW:

Thank you. My name is Steven Pedlow. I'm a statistician with NORC at the University of Chicago, and along with my co-workers Michael Jugovich, who is the IT Data Leader, and Ed Mulrow, who is the Project Director, I'm pleased to present some highlights of analysis for the WHOIS Registrant Identification Study. I would like to thank Lisa Pfeiffer and Barbara Roseman, who helped a great deal...

[break in audio]



STEVEN PEDLOW: ...potentially commercial activities were present or not. And if any of them were detected then we classified the domain as having potentially commercial activity. And this pie chart shows that we detected potentially commercial activity for about 56% of the domains. A full breakdown by the facets will be in a full report that will be published soon.

So that's the background on the coding, and I've allowed some time for some preliminary draft results of our analysis. These are still under review, of course. Going back, thinking back to...

[break in audio]

STEVEN PEDLOW: ...domain users who are apparently natural persons in our sample. So this is a large difference, but not large enough based on a small sample size.

This slide shows all of the other statistically significant difference by domain user type in the facets of potentially commercial activity. And so you can see clearly here that domain users who are apparently legal persons show higher rates of e-commerce, offline membership content, offline promotional content, and pay-per-click ads.

Moving on to our second focus questions, we're actually comparing three different types of registrants, apparently legal persons, apparently natural persons, and privacy/proxy registrants. This set of pie charts show that the...



[break in audio]

STEVEN PEDLOW:

...commercial activity is highest for the privacy/proxy registrants, and it is significantly...I'm sorry, statistically significantly higher than for registrants who are apparently natural persons, which is 55%. Coming in right in between, at 60%, are the registrants who are apparently legal persons. And that's not significantly different from either of the other two groups, so right in the middle.

This slide shows other statistically significant differences in the facets of potentially commercial activities, and you can see here that privacy/proxy registrants have a lower online promotional content but are higher in host banner ads and pay-per-click ads. And none of the differences you see here between apparently legal persons and apparently natural persons are statistically significant.

I just have one slide on the third focus question, which is comparing those domains with potentially commercial activity versus those that don't. And here you can see that those with potentially commercial activity have a statistically significantly higher rate of having privacy/proxy registrants than domains with no potentially commercial activity. And that's 23% versus 17%, which is highly significant.

So, in this talk, I just showed some highlights of our study. NORC will complete and publish a draft report for public comment soon. The community will be invited to post comments on ICANN's public comment forum. Webinars may be offered...



[break in audio]

BERTRAND DE LA CHAPELLE: Yes, thank you. My name is Bertrand de la Chapelle; I am on the ICANN Board. Two points, just to highlight the importance of the aspect of differentiated technical modalities of access to the WHOIS data. And I'm extremely pleased to see that this topic is bubbling up in the discussion. It was long in waiting.

The other thing is just a very, very concrete question. I was intrigued, or maybe I misunderstood something on the slide eleven of your presentation. There is a pie chart regarding the natural registrants and it seemed to imply that 100% of them are for profit, and maybe I misunderstood something.

STEVEN PEDLOW: No, 100% are non-businesses for the natural persons.

BERTRAND DE LA CHAPELLE: But that's not what the pie shows, if I'm...

STEVEN PEDLOW: You may be right. There may be an error in the color scheme. Thank you.



JUDITH VAZQUEZ:

My name is Judith Vazquez; I sit on the ICANN Board. Wonderful work. Thank you so much. Wonderful. We have a lot ahead of us. The first question of the community is what should the right name be? Not WHOIS, it's WEIRDS. At any rate, from Murray's presentation, particular to the practical logic of WHOIS and the similarities between...because we do have a difference, there is a WHOIS for names and a WHOIS for the numbers and it makes sense that what will be WEIRDS, or WHOIS Version 10, be both names and numbers.

But I did notice in the WHOIS roadmap that the PDP referred to does not include the NRO and the ASO, with two representatives on the ICANN Board. So I think this is something that is very urgent to look at, is my comment, and should be brought up to the ICANN Board as well. So this is just a comment, no need to answer here. We can take this offline, but this is something I observed. Thank you.

FRANCISCO ARIAS:

So I think I have an answer related to the other supporting organizations. What happened, there is that a PDP will not be scope in order to require the support for the protocol. However, as Murray said, four of the five RIRs already have either a pilot or a production service. For example, ARIN, who is basically the leader here, they already have a production service and it's already receiving more queries than the port 43. So even in total they may not be covered by the policy aspect here in ICANN, they are very active in this so it's probably not a concern.

[background conversation]



FRANCISCO ARIAS: Sure, sure.

STEVE METALITZ: Hi, I'm Steve Metalitz from the Intellectual Property Constituency. I had two questions. One, I see we've had another presentation now on SAC 051, which is a report that came out a year ago and has been the subject of at least two presentations at prior ICANN meetings, but no presentation on SAC 056, which is the most recent statement from the SSAC on WHOIS. So I would be interested if there's someone on the panel who can discuss that and perhaps respond to questions about that.

And my second question really is to try to understand better the dividing...and this may be very intuitive to Murray and others on the panel, but I'm having trouble figuring out the dividing line between what IETF is doing in the WEIRDS group and what ICANN will be doing.

Let me just pick up on Bertrand's example of the capability for differentiated access. I think that's important, but there are other capabilities that many in the community are important and don't seem to be listed in the WEIRDS mission statement that you put up there. For example, the ability to obtain historical data on domain name registration records through this domain name registration data service, which sometimes is called WHOWAS.

That's one example, and the other example is the ability to determine from the data that you're returned whether or not the data has been verified or validated, and if so, when and by what method? Are those



topics that will be addressed within the WEIRDS group and within IETF, or are those topics that are probably more appropriate to address within ICANN?

MURRAY KUCHERAWY:

The first question was the distinction between what parts of this is ICANN doing and what parts of this is the IETF doing? Okay, I was asked that during the ISP Constituency yesterday and my answer was basically that the IETF is fairly policy-allergic. What we want to do is make sure that the protocol can support whatever policy you might want to implement, but we will not place any policy requirements in the protocol.

So we are strictly interested in how you ask the question and how you form the answer. What you put in the answer is kind of up to you, but just the syntactic way you prepare and deliver the answer is the part that we're interested in. So in a layering model, you're talking about layers eight and nine and we're talking about the lower ones, I think.

Second question was about requirements. It sounds like there's a few that we don't know about yet, and so my slide about "please come participate". I would reiterate that one. The two examples you gave are...one I hadn't heard of at all, the WHOWAS thing. If that's something that we should be tackling then we'd love to hear about that sooner rather than later.

The other one was verifiable data. I believe we've...I'm interested in hearing about all of these surveys about how WHOIS is being used now, because we have some people in the working group that are running



their own to see...sampling a large number of different WHOIS services. What fields do they return? One of them might be "this was verified by...or in this manner" and therefore, if a lot of places are doing that then we think that there's an interest in doing it and we would include it in the model.

That's kind of the path we're on right now. So we're doing some of that research on our own, and I'm excited to hear that that's happening here. We should trade information.

STEVE METALITZ:

Thank you, that's a very useful answer. But I would suggest that ideally we should have a system that doesn't just depend on what people are doing now. I thought one of the goals of having a successor to the WHOIS protocol is to do better and to be able to be more useful. We have brought some of these issues to the IETF and we will continue to do so if that's the right place to make sure that they have this capability. I totally get what you say about policy allergy, but I think capability is really the important question.

MURRAY KUCHERAWY:

Exactly what we're interested in.

STEVE METALITZ:

Okay, thank you.



BARBARA ROSEMAN: I think one thing to keep in mind is that the WEIRDS protocol is intended to be extensible, which means that new fields can be added to it once the basic protocol has been established. The WHOWAS might be a little more complicated than just a field return, so I think that might be one that needs to be integrated in to the basic protocol. But the idea of creating a field for validated data would be something that would be addable after the initial work is done.

STEVE METALITZ: But why would we wait until after the initial work is done to add it? I understand that it's extensible, but let's roll out a good version first.

BARBARA ROSEMAN: Yeah. The problem is that, as Murray said, they did do a survey of the current data fields that are being returned and that would not have been one of them. And so it just was invisible to them as a requirement, that's all.

STEVE METALITZ: I understand, but I think our goal is to improve, not to duplicate the...

BARBARA ROSEMAN: Yeah, so Steve, as to your other issue of SAC 056, I can speak to that a little bit. The response that SSAC gave identified a couple of different issues. One being this question of how current WHOIS policy is defined and collated as information, and it's difficult to locate and difficult to pin down, as the WHOIS RT Group recognized as well.



The other is that there may be real value to splitting the discussion on WHOIS in to two separate elements. The registration data element, the collection of the data and storage of the data, and then a separate piece, which is what we currently gloss as port 43, which is the display of the data. What gets displayed and at what level of detail and all of that.

And if you look at the terminology document, the taxonomy document, you'll see that this is sort of the direction they're moving in. And it allows you to focus on the requirements that are needed to really associate the registration data for the life cycle of that domain name with that domain name.

So that means it could be associated with it through a transfer, through a registry or a registrar failure, that the key data would stay with that domain name for the life cycle of that domain name. So if the domain name doesn't get renewed, that data doesn't need to stay with it any longer. It's useful to think of it this way because that helps you go a long way towards defining the purpose of why we're collecting registrant data and why we're holding on to it, why we think it's important to have as ongoing escrowable data. You know?

The question of display is obviously where we get in to the differentiated access issues, and I think is a much more complex one in some ways. And it has a lot to do with how we want to have access to that data, but it doesn't need to define the data itself. The data itself is what is necessary to keep a registration associated with a domain name for its life cycle.



There are a lot of different groups that are looking at this now, and SSAC is trying very hard to spread this terminology more widely in the ICANN community. To talk about registration data and display of data, and really keep those two discussions a little bit separate. So SSAC identified as the key issue, moving forward for ICANN, to really focus on what is the purpose of registration data collection and make that the priority before even taking on many of the other recommendations that the WHOIS RT made because they believe that by defining the purpose will have gone a long way towards answering some of those questions, or at least opening up the space in which to have those discussions.

STEVE METALITZ:

Do you know if the SSAC, in reaching that conclusion, looked at all the work that was done in the GNSO on the question of what is the purpose of WHOIS data?

BARBARA ROSEMAN:

I'm sorry; could you ask your question again?

STEVE METALITZ:

Yeah, did the SSAC...I didn't see any reference in the SSAC report, but the GNSO conducted a lot of work on the question of what is the purpose of collection and access to WHOIS data. And I recall, this is a while back, but I do recall a lot of debate about that and even to opposing formulations that were put to a vote, and was all that looked at by the SSAC?



BARBARA ROSEMAN: I don't know how much of it was examined because I wasn't doing SSAC support at the time, so I'm just unaware of how much they looked at it. I do know that they have looked at documents produced by the rest of the organization. They're not trying to ignore those. And I don't think they were trying to say that there hasn't been a discussion about the purpose. They're saying that it hasn't been a defined, accepted "here's what our purpose is" for this data.

STEVE METALITZ: Thank you.

BARBARA ROSEMAN: Thanks for your questions, Steve. Bertrand?

BERTRAND DE LA CHAPELLE: Hi, this is Bertrand de la Chapelle again. I want to piggyback on this discussion, because WHOIS has been an acronym floating for ages with bad omen. It's almost a Batman. We're all trying to find the analogies. I promise a great success in promoting the WEIRDS acronym, which is going to have some traction. One point...

MURRAY KUCHERAWY: It's a well-earned name.

BERTRAND DE LA CHAPELLE: What's that?



MURRAY KUCHERAWY: It deserves the name it got.

BERTRAND DE LA CHAPELLE: Yeah, well I think...and so a lot of the problems that we encounter in this topic, and I witnessed it first-hand when was in the GAC and before I joined the board, I was participating or at least observing one of the WHOIS working groups. And the debates were endless. And at the time introducing differentiated technical modalities of access didn't fly. I tried, it didn't go.

But what is striking is that we still are talking about everything in the singular. We're talking about "the" WHOIS data, "the" WHOIS tool, and "the" purpose. And all the discussions show that the challenge is that actually we should be talking about the "purposes" of keeping some types of data regarding registrants of different sorts to be used by different types of actors. And the moment you diversify this you begin to ask the community, "What are the needs?"

And the needs are different. There are needs for trademark lawyers, there are needs for law enforcement people, there are needs for consumer protection, there are needs for whatever you may think...academics. I mean this is a trove of data. If it were in an open data type of repository you could do wonderful searches. But if you want to, for instance, get the bulk of the whole historical data completely anonymized this would be extremely interesting, including for statistical purposes.

And so my question is, following what Steve was saying regarding the articulation between the IETF and the work within ICANN, of course



code is law and the framework that is going to be adopted as a protocol has implications. But in the reverse, just like privacy by design in web applications can also be an approach, the development of the protocol without IETF getting in to the policy-making will be, I hope, very much fed in to by the discussion on what are the users of the types of data we want and the types of access.

And I think there has been significant progress, as indicated in the very early discussions, about separating how data is entered and how data can be used by the different actors. Just like in the Trademark Clearinghouse discussion, actually. And so what I am very keen on is to make sure that the two discussions in the IETF and ICANN work hand-in-hand, and that you get all the input that you need, both by the active participation of people, but also by just exchanging.

And that vice-versa people who have been discussing for ages about it should be A or it should be B understand that there's actual work done on this differentiated technical modalities of access. And so I don't know how it will go on, but I hope that in Beijing we have a good opportunity to make this presentation with all the interfaces between the different actors.

And I support Steve's comment; we need to think about the users in a broad and very straightforward approach. It's not an all-purpose rifle that is supposed to serve all needs in the same manner. The one-size-fits-all doesn't work here, and I like very much the way it's going.

BARBARA ROSEMAN:

Thank you.



MURRAY KUCHERAWY: I suppose if there is anything for me to take back from that it's I wonder if there should be some kind of more formalized channel between whoever on the ICANN side...I've got to say too, ICANN is about as bad as we are at acronyms, so half of what some of you are saying is really confusing. And I could turn it around and confuse you too, but I won't. If there's some kind of relationship that should be established there to make sure what you're talking about is happening and what the other gentleman's talking about is happening, I'd love to talk about it.

BARBARA ROSEMAN: I will say that there are a number of staff who are following the WEIRDS work quite closely and who are very familiar with the discussions within the ICANN community about the various issues around WHOIS and particularly about the display of data and access to the data. I think that there are probably requirements that haven't been well articulated to the WEIRDS group. But I also think that everyone is really interested in seeing a follow-on to port 43 come in to existence.

And so there's an attempt to not burden it with too much detail at the moment. I do think the WHOWAS thing is probably worth discussing. But again, I would emphasize that ICANN has established in contracts now that they have the ability to identify a follow-on to port 43. And we all call this thing WHOIS, but in the contracts it's always been referred to as port 43 access. And that is a method of displaying the data.



And so I think that you may, over the next few months, see the conversation grow in to this distinction between talking about the registration data that needs to be acquired and held and the registration data that needs to be displayed, and that that might get us further along in the WHOIS discussion than we've been able to get for some time. Any other questions? Yes.

[CHANDRA]:

Hi, my name is Chandra, and I found the presentation on the study that analyzed the natural persons and legal persons in commercial activity very interesting. And I was just wondering is it connected...like what's going to happen with this data? Is it connected with any ongoing work within ICANN and also is this something that would be connected with the WEIRDS group? Is there some discussion about the protocol, whether or not it will display whether someone has sort of self-selected as a natural person or a legal person? So I was just wondering sort of what the output of that data would be and how it would influence these other areas.

BARBARA ROSEMAN:

The data is being collected under a study that was commissioned by the GNSO, basically, for aiding the development of future policy work. And so I think that there's no question that it's going to be used to further additional work in the WHOIS area. As for making it a required field, I think that's a lot more challenging to think about.

But again, the value of the data is whether it tells us is there something here that we need to be looking at more closely or was this a sufficient



cut on it to get us further along in the discussion. The process that NORC used in identifying, or classifying I should say, classifying the various sites that they looked at was pretty extensive and that will become part of the material that's published shortly. In their appendices they talk about their process quite a bit.

So we intend to publish this probably within the next month or so, depending on the finalization of the draft, and NORC is working on it, not ICANN. But it should be available soon, and then there will be a comment period and I think that would be the right time to ask those kinds of questions.

STEVEN PEDLOW:

I'll just add that some of the variables were much easier to code than other variables, and that will be described in the report that's published.

MURRAY KUCHERAWY:

And since you asked about whether that information will be in the protocol or not, it's probably far too early to tell. But the E in WEIRDS is Extensible, so it probably could be added and clients that know what to do with that value could do something with it and clients that don't would simply ignore it.

BARBARA ROSEMAN:

Do any of you four up here want to add anything or say anything? If not, I think we'll call this session closed and thank you very much for attending.



[End of Transcript]

