LONDON – Expert Working Group on gTLD Directory Services (EWG) Final Report Overview
Monday, June 23, 2014 – 15:15 to 16:15
ICANN – London, England

JEAN-FRANCOIS BARIL: So if we can take our seats, we would like to start in a few seconds.

So I believe that we should get started. So good afternoon and a very warm welcome to the expert working group public session.

My name is Jean-Francois Baril and I'm the EWG facilitator.

I'm very pleased, of course, to be here today along with my colleagues from EWG for an overview of our final report on next-generation gTLD registration directory services, so-called RDS, which was posted on June 6.

So for the next hour or so, we plan to cover the following.

After a short recap on EWG that I will make, I will pass the microphone to my colleagues on the panel to discuss more important elements of our report. Then we'll have the next step. And then we have prepared, later on, two Q&A sessions.

The first one, we have had the opportunity to pull forward for practicality, because I think a lot of people will be busy at 6:30, so we are proposing from 4:30 to 6:30 in this room, and that will be two hours of hopefully fruitful and more in-depth dialogue on the substance of the report itself.

Then on Wednesday morning, from 8:00 to 10:00 a.m., we will finalize the discussion and clarification on report elements, and if necessary, also discuss some of the elements of the EWG process.

So coming back to the origin of our EWG.

As a reminder, this group was created under Fadi Chehade's initiative and ICANN's board request in February 2013, and this is to overcome a decade-long deadlock on WHOIS, with a mission to, number one, re-examine and define the purpose of collecting and maintaining gTLD registration data; then to consider how to safeguard those data; then of course propose a next-generation solution that will be better to serve the needs of the overall global Internet community. And I repeat: Overall global Internet community.

Then this work should serve as a foundation to help the ICANN community through the GNSO process to create a new policy for gTLD directory services.

As you very well know, working with a consensus-based and bottom-up multistakeholder model is a very, very challenging and difficult task to create a convergence of view. At the same time, it can be very rewarding and can generate very solid conclusions when people are genuinely aiming to positively contribute to the benefit of the overall.

But it requires very direct and frank discussions, and we had a lot of those, with participation from everyone as individuals and not as an advocate. Important also.

But also, the courtesy, the cooperation, and compromise by all parties.

It's very important to underline the last comments, the last elements of "compromise of all parties." Otherwise, we stay as a status quo as it is today and we'll never, never find a solution.

Of course with that, it comes with an extraordinary amount of work and commitment, but this is probably the price that we have to pay if we want to contribute to a more trusted Internet. This is probably a very good opportunity to once again present this magnificent team of volunteers. Everyone here is a volunteer working out of their normal 24-by-7 already work, and they are, in fact, the artists behind this final report. I'm not the expert, but these people have been very, very impressive to do the work.

So quantity of work is one thing, but what I am even much more excited about, and to report here, is the quality aspect and the great spirit that have transpired from the long -- from this long journey, and this is a very strong and great complement to the team.

In addition to the 15 EWG members -- Pekka, Lanre, Chris, Steve, Scott, Jin, Susan, Nora, Michele, Michael, Stephanie, Rod, Carlton, Faisal and Fabricio -- it's my great pleasure to praise the exceptional efficiency and effectiveness from the ICANN staff; namely, Denise, Margie, and Lisa.

So just to make sure that we're -- you recognize -- and I'm sure you know all these people, but I would like Lisa, Margie, and Denise to stand up for everyone to recognize these people for the exceptional amount of work.

[Applause]

JEAN-FRANCOIS BARIL: And very honestly, without them, this report content would not have been possible because of this incredible complexity to integrate the result of our vibrant discussion -- and these people are very, very creative -- and not so easy to assemble all of the dots in a correct way, which makes sense, and a consistent way. So that's why I wanted very special appreciation for this team.

So at the origin, to form the team each member was very carefully considered and selected over more than 70 highly qualified experts across the overall Internet ecosystem and all geographies.

Decision-making factor was probably the main element based on capability of good balance between operational and down-to-earth experience and also the necessary soft skills to make the impossible possible.

In this case, it means strong attitude to build consensus looking at the big picture, rather -- to serve the better -- to serve the global Internet community, rather than promoting own direct interests, attitude to speak freely and frankly and not be mentally blocked by position relative to a particular constituency. That means thinking in our own name. But also the capability to innovate with a positive attitude to change and capability to think out of the box, to find the best solution, and to avoid status quo.

Very honestly, I can proudly report and testify of the utmost observance of the demonstrated skills that I just mentioned. And this is probably the major part of the -- if any, of the magic which has created the 180-plus principles and recommendations of this report.

The approach we used within EWG is reflected in our final report.

First of all -- and once again, very sorry for the 166 pages, but despite of all of our best will to make it shorter and simpler, the extreme complexity of this longtime blocked situation to drastically improve the current WHOIS took us much more time and effort than I ever anticipated.

I was probably very naive, but it is -- may be one of the reasons I was able to embark the team with the spirit to make the impossible possible.

This 166 pages of recommendations are, in fact, the culmination of an intense 15 months of work with thousands of hours of in-depth research, digesting 2600-plus pages of public comments, responses of different surveys, research of all kinds, 19 public community consultations, 35 days of face-to-face EWG meetings -- 35 days -- 42 EWG calls, formal ones, and 200 plus calls from the sub-team that we've created adjacent to our work.

And of course countless interactions with outside experts and community members.

Even so, our consensus-building process was not perfect and we were, in fact, expecting, at the end of May, from one of our EWG members -- Stephanie Perrin -- to write a dissent to be published along with the EWG final report. And this is normal that we cannot agree offering. This is fine.

Unfortunately, discussion could not happen before the publication, but we are currently working between us on this EWG and we are pretty sure that we'll be able to publish this dissent soon in our EWG blog.

So overall, all this tremendous amount of work to answer what? To answer a simple question: Is there an alternative to today's WHOIS to better serve the global Internet community?

How to answer this interesting question with respect of difficult privacy issues, data quality and accuracy, data protection, and striking a workable balance between access and equitability is not an easy task but this is the purpose of our report within the EWG.

Sure, this final report, even if it comprises an enormous amount of analysis, collecting all the brains and hearts of most of the experts in and outside of the community, it's probably not 100% perfect.

Once again, it's a result of many, many compromises from everyone, and nobody can say "I compromised more," "You compromised more," whatever. I think we have all got our share of incredible compromises to balance very diverse and divergent needs, but always very impressive intellectual honesty.

Intellectual honesty was very much one of the core values that has guided us through this 15 months of this process.

So along with everyone within the EWG, I'm very, very confident that this report provides a solid foundation to support the overall ICANN community with the respect of the WHOIS replacement and the real breakthrough of the Internet looking forward.

And we are all very confident that this report also fulfills the ICANN directives, ICANN board's directives and will be the beginning -- that's not the end; that's just unfortunately the beginning -- of a constructive dialogue and a fully success GNSO PDP.

With that, I would like to invite Fabricio Vayra to explain why we have decided the paradigm shift to replace WHOIS.

FABRICIO VAYRA: Thank you, Jean-Francois.

So the question was, is there an alternative to today's WHOIS, and our group decided that the answer was a resounding yes, there is. There's actually an alternative to anonymous access to what is usually inaccurate data.

So as you see from our little friend here, the garbage can, we heard plenty from the community that WHOIS was full of garbage.

So it is time to upgrade and our recommendation hopefully goes to that end.

So what does this 166 pages involve? What does it entail? What does it contain?

It contains over 180 recommendations for a new registration directory service, the RDS. What we tried to do was really strike a balance between accuracy, access, and accountability, and one of the mechanisms that we did propose for accountability is purpose-based access, meaning that today you can get access to data -- it may be

garbage, but you can get it -- anytime without any accountability in the system.

What we're proposing is something where it levels the playing field and there's accountability for those searching for data as well, and it has to be tied to specific purposes, and they're permissible purposes.

So in our gated data model, what's left outside the gate?

Well, it's minimal data that's left outside the gate.

With regard to personally identifiable information, it's really just an email. Everything else that you see outside the gate is actually selected by the registrant to be put there or is what we would call system data or metadata, meaning who's your registrar, who's your registry, when was the domain created, what's your DNS, et cetera.

If you want information about the registrant, you actually have to get authenticated. You have to get behind the gate. And so getting behind the gate, you have to, as I mentioned, identify your purpose.

And the reason for this is we really thought that accountability was something that was missing across the board. You hear a lot of people talking about accountability for those who put their data into a database, but there's also accountability for those who access it, there's accountability for those who store it, there's accountability for those who disclose it. And so having a purpose-driven and authenticated database was our answer to that.

And of course one thing that will be obvious to you when you read the report -- and I understand if many of you haven't yet; it is pretty thick --

there are two new parties in the ecosystem now -- those who validate and obviously those who run the RDS -- and those would be contracted parties.

So that's what you can look forward to at a high level within our report, and hopefully you think, as we do, that this is an improvement and a step -- a very, very large step forward and a necessary model but something that needs a lot of correcting.

So with that, I'll pass to Susan.

SUSAN KAWAGUCHI: So why create a new RDS?

Right now, the WHOIS provides a one-size-fits-all public access to anonymized -- to anonymous users. That doesn't work. We all know it doesn't work.

Little accountability or abuse remedies, limited individual privacy protection or ability to conform to differing laws, limited ability to ensure data integrity, lack of security and auditing capabilities, cumbersome contact management, and inefficient communication.

So we tried to address all of those, and we've hopefully come up with something that may fix all of those.

So our solution is the purpose-driven access.

Some registration data would remain public, to promote the Internet's stability and meet basic DNS needs. This public data would still be accessible by anyone for any permissible purpose without

authentication. So, yes, you'll get some information. You will not get what you see today in the record.

And so in comparing what we're proposing to the existing WHOIS record, in the top diagram is the current WHOIS. It's entirely public data. If you put your -- if you put correct information, accurate information, in the WHOIS record, anyone can see that, and it's entirely anonymous access. The registrants cannot provide contemporary or alternate data.

There's a lot of new ways that have come up recently to contact people and the current WHOIS doesn't provide for that.

Also, contacts cannot prevent inaccurate or fraudulent use. My neighbor's -- someone could use my neighbor's information or they could use my company's information. There's nothing to prevent that.

So we've looked at all of those issues and we've -- with the purpose-driven RDS, we think we can prevent some of those misuses.

So you have the existing domain name data, and you see all the little crosshatches there? These are all gates. Gated data. But don't look at a gate as one big gate. You don't unlock the keys to the kingdom and walk through. You go through into a hallway with a lot of doors. I would have preferred maybe to call them doors. Or a lot of different gates. And it depends on your purpose and what you've told us about you to access the data.

So you have to -- you know, when you go to the front door, you don't get to just walk in. You have to tell us who are you and what are you

using this for, which I think is very important. If you want to know somebody's personal data, you have the duty to provide your own.

So most of the data is going to be gated by default. The contact data is validated.

IDs. Everything provided is linked to an ID, so this is an eight-digit number in our proposed --

>> (off microphone.)

SUSAN KAWAGUCHI: Or, yeah, could be. Whatever works for the community.

And that is what you get first is that contact ID. The contact ID does not mean you get the number, the address, the phone number. None of that is minimum public data.

And these are all -- then there's also additional roles we've put in place, or proposed, that you have to -- there's different purposes for contacting a domain registrant, and so you can categorize that, if you want to.

You can act -- a registrant can act as their own purpose -- PBC, but Rod will explain that a little bit more. So all of this, there's a very minimum data set that's outside of the gate.

So when you look at the full -- the record, the minimum public data, which I'll show you in a second, but there's three categories of this.

So there's a large category there on the left, domain name data supplied by the registrar and registry.

That's creation date, renewal dates, you know, updated dates, statuses, things that you need -- that others may need to use and view that's supplied by the registry and the registrar.

And then the top right is the registrant contact ID. So there is an ID associated with that -- number, again, associated with the registrant, but there's only minimum public data available.

And then also in that minimum public data is the purpose-based contact IDs. Once again, a number. Nothing else.

So if you are not telling us who you are or telling the RDS who you are and why you want the information, you're not walking through the door to get more information.

So this is an example of that same -- basically the green chart I just showed you, so you'll see to the left is the registrar and registry data. All of that is just important for everyone to see and there's no public -- or personal information on the left-hand side.

The right-hand side will tell you the domain name, the name server, the registrant type, which you can leave as undeclared. We've provided some options. If businesses would like to say, "Yes, I am a business," they can declare themselves in that category. But you don't have to. It is a choice. You receive a registrant contact I.D. There is the number again. And then also we thought it was important to know the last time the information for the registrant was validated. Is this done once a

year? Is it done every six months? Is it done every ten years? That's important to know when it was done last.

And then the registrant email address, everyone needs that email address for basic technical management of your domain name. Consumers need to be able to contact you. And then registrant country. In the presentation, we'll talk about the rules engine, being able to associate the domain registration with the applicable law within that registrant's country. And that's why we've put that element -- that data element out to the public.

And then we get into the purpose-based context, the PBCs. Once again, all you receive is the numbers. That's the only things you will see, so you can then, if you need to contact someone, beyond the email address for the registrant, if you want to talk to their technical contact, then you're going to have to go through those doors or gates. But you have to choose the right door to get the right information, and you have to provide information about yourself to get that.

So, if you want to just get the minimum public data, then you would be -- you would go to the system and basically search for the domain name. But you would only return -- be returned the public data. You have to declare a purpose. If there is no purpose declared, you're not going into that gated access. And we've lined out some of the purposes. Have we figured out every purpose you might need information from a domain name registrant? I doubt it. We tried really hard. But I doubt it. So you can see that, once again, the return-only public data is the minimum public data.

And then I'm going to hand this to Rod for more detail.

ROD RASMUSSEN: Thank you, Susan.

So, let's talk about that hallway of doors. I like that analogy you used.

The gate is not a single gate. The concept is really around the ability to protect data at the appropriate level for the purpose it is being used for.

If you take a look at the overall qualities of what we're describing here, it is really from an engineering perspective, it is a role-based access control system. Something that we're very familiar with in, for instance, the corporate world where I can have a single sign-on to something. And depending upon my privileges and the things I'm trying to do, I can get access to the types of data I have the privileges to get within that system. It is really pretty straightforward from that perspective when you think about it that way.

And for our system that we've proposed here, it's driven by a combination of who you are, how you've authenticated, the purpose you've declared, and the rules of the system itself for providing you the data.

This access control allows you to protect data. It also allows you to hold people accountable for their usage of the system and looking at what the system looks like conceptually, it is fairly straightforward where you have somebody who has been accredited to some level, has an access credential. They access the RDS system which may pull data from registries, registrars, validators where it's been stored, and present that back to the requester based on the system -- or the request they have made and the privileges they have.

There can be multiple methods. It is really important to understand that. It is not just necessarily a person sitting in front of a computer as is depicted in the picture there. This allows for tying into other authentication systems which we will talk about as far as the validation section later on. But that's really important to realize that this is not a single look up a domain name at a single time with a person sitting at a user interface. That is obviously one of the methods; not the only one.

This should look familiar to, I think, most of you who've been following along as we've been going for the last year plus as registration data permissible purposes.

I would point out that we have added in the last final report concepts around DNS transparency because that's important to ICANN and the overall community as being able to provide transparency into how the system works.

And we also have domain name research defined in the report and suggest taking a look at what we mean by that. But it's really being able to do the kinds of studies that ICANN has done to provide information around quality of information, how domain names are being used, et cetera. So those are kind of the new ones that you may not have seen in the prior reports. Everything else is pretty much as we described it in our earlier versions.

So every permissible purpose has data needs. They all may be different, however. Depends on the domain names that are involved. You can get data that is public. You can get registrant -- and that's already described, the data that surrounds the actual provisioning of the domain name itself, the registrant data, public or gated -- gated data.

And then there are also provisions in here for what's called WHOWAS, which is historical information, reverse queries which allow you to determine domain names that may be related to a particular contact.

Some of these purposes would be widely used and have a fairly minimum set of standards for being able to get access to that information. Others are going to require fairly formal accreditation processes where that data is more sensitive and would require you to go through a stronger -- you know, more hoops, have a certain responsibility around controlling your credentials and be subject to further scrutiny for anti-abuse.

So let's take a look at the purpose-based contacts. The idea here -- there's many ideas here that are encapsulated in the purpose-based contacts.

They allow for a lot better control of personal information by a subject matter, whether it is a person or a company, of their contact information. It also allows for much more efficient portability amongst domain names of contact data. So you have -- the role of the registrant is pretty obvious, is the controller of the domain name. But you also have these various purpose-based contacts which encapsulate very common usages of domains that people would want to be able to get -- external parties would want to make contact with that registrant or the person or entity responsible for that aspect of a domain name.

So, for example, a technical contact, you would want to be able to provide answers to technical issues. So a logical kind of person or entity for technical contact would be an ISP, Web hosting firm, registrar, somebody like that, who could answer technical questions. That isn't

necessarily the person who has registered the domain name. They may not know anything about running the Web site or the like. They farm that out to somebody else to do. This allows you to get directly to a person or company that can take care of technical issues.

Same thing for the various other types of contacts here where you may have abuse issues to deal with, legal issues, administrative control of the domain name. The two on the -- I guess it is the right side from your angle -- actually from my angle, too, now that I think of it -- they are optional. Depending on the type of registrant, if you are using a privacy or proxy service, that contact information would need to be included there so that you can go a hold of the appropriate privacy proxy service in order to fulfill some sort of a reveal process or get a hold of the registrar or what have you behind that.

Also, if you are a business, you can self-declare yourself as a business and provide some information that makes it easier for consumers or your customers to be able to get a hold of you to resolve some sort of issues or what have you. So those are the different purpose-based contacts.

Now, this contact data itself that's involved here, as already mentioned, is largely gated. But as we said before, there's different doors you can go through to get to different levels of access. And that's what I'm going to run through right now.

Let's see if you can make that out. It is fairly tiny on this screen there. My apologies for that.

Oh, okay, so they are up on the session page, the slides are, if that makes it easier for you to follow along and see what the tiny print is.

So for this example, somebody has come in and wants to get in touch with somebody involved with a domain name particularly for legal actions. This has been one of the ones that's been discussed fairly -- at fairly great length.

So how does this process work from a process control flow? Well, first of all, have you been authenticated? If no, you are not going to get any of the data other than the public data that was pointed out before. Are you authorized for this purpose to come in and be able to request this?

Let's say you have gotten into the system but you have been authenticated maybe for just technical reasons. You can't go and then get the legal contact data based on that role, right? That's the concept here.

So let's say you are authorized for that purpose for legal -- getting legal information. So are you requesting the sensitive or gated information around the purpose-based contact or not? If no, then your request -- the data returned from the request would be that information that was listed under that contact for the more public viewing of that data. So that would be, in this case, the legal contact information which would include, I believe, the name -- I can't even read it myself -- in the contact address so that you can serve legal process for that particular one. That's what typically a legal contact is for, is to serve legal process of some sort.

If you have requested gated data and you are approved for that -- if you are not approved, again, all you get is that more public side. If you are approved, you get into all of the information around the legal contact. Then it would return that full set of gated data.

Now, which -- what are the approval processes and all that gets determined by policy. And that's for our policy development process to hammer out.

That is the process flow here. That is the intent of how you walk through -- walk through a door, a gate, what have you, to get to various data elements that you would need to get to in order to satisfy your purpose.

So I think I'm passing this over to -- no, I got one more slide, I think. Right? Yeah, that's mine.

So -- these are some thoughts on what could be in purpose-based contact data. This is important to realize it is not necessarily the data that is directly to that entity or person. It can also be privacy proxy information that is inserted in there as well. So you can still proxy through and use that as a purpose-based contact.

If you have not made any choices on this -- and many people don't have those kinds of contacts and don't want to -- they may have their own information there. That's perfectly fine. There's two or three different choices you can make as to what kind of data you want to have in that purpose-based contact.

And then each contact holder themselves may decide whether to gate data or not. Let's say I'm an intellectual property attorney and I am

representing several people who hold large portfolios of domain names and I would like them to be able to contact me if they have an issue with my customers' domains, I would publish and want public all that information so it would make it easy for people to get a hold of me if they want to talk about a UDRP issue or something else that's of import.

I, as that attorney, can use that contact and I can use it across all the domain names of all my customers, and I only have to -- if I change my phone number or address or what have you, I only change it once and it is reflected throughout the entire system. That's the idea there.

Okay. I'm going to pass it over to Lanre.

LANRE AJAYI: Thank you, Rod. The RDS introduces new measures, new concepts to improve better quality. And very significant concept is getting access that has been extensively discussed by Susan and Rod.

But a point to note here is that when registrants are cautious of the fact that their sensitive data is put behind the gates, there's that tendency not to supply incorrect data intentionally. So by getting sensitive data, we tend to achieve accurate -- we tend to achieve better quality improvements.

Another concept is the separation of contacts that are actually from data of the domain name directory. They are not separate data entries. The individuals and companies cannot continue to maintain their data within the RDS.

More importantly is the introduction of a new actor called the validator. So if the registrant is not motivated enough to get access concepts to provide accurate data, then the validator is there to do its job, to ensure that data that's inside the RDS are properly validated. It is a new concept. And it is a very important one to ensure data quality.

And there is the new -- the identity theft issue is reduced by the concept of identity validation. There will be more on that later.

The reusable contacts also improves consistency, especially when you have a large domain. You are able to update your contacts once through the validator. So that also helps to improve data quality.

And letting contacts use any validator is in compliance with local data protection laws. The validators can be a local one or a global one. If you use a local validator, then you have a chance of complying with local data laws.

So these are the new measures that were introduced into the RDS to improve data quality.

But now I will just talk through the process of validation. In this illustration, there is a man standing there with the name Z. He has an email which is z@isp. The guy submits data elements; in this particular case, the email address to the validator. And this are the processes the validation will take.

The system will first check if the data is technically valid. In other words, is the format okay? In this particular case, it is email. Does he have the @ sign? So if the system is -- if the format is okay, it goes through the next validation which is operational validation. The

operational validation is to check if they make a dissent to that particular address.

Then if the registrants are desires to have identity validation, it is passed through that process. And if the identity validation is correct, is valid, then is assigned a contact I.D. and credential.

If all of these are not correct, the registrant gets some kind of error message. So that's the flow of validation process. We should note here that there are three kind of validation: System validation, which is mandated for every data element; the operational validation, which is mandatory for some of the data elements, not all; and the identity validation, which is completely optional. It is just whoever desires it, asks for it. But it is not optional -- it is not mandatory for you to do identity validation.

So on this slide, we just want to focus on the importance of RDS contact directory because it's a new concept and it is something that we believe will enhance the data quality.

The RDS contact directory is completely separate from the domain name directory. Registrants and their public -- purpose-based contacts will create and maintain their own contact data using the validators, as mentioned, before proceeding to domain name registration.

But by separating the contact validation from the domain name registration, we tend to achieve the following: One, we intend to take away the difficult task of validating contacts away from the registries and registrars. But that is not to say that the registries and registrars

cannot act as validator (indiscernible), so the burden of validation is essentially taken away from the registries and the registrars.

Registrants can choose to have the local validators, which is very important when we are talking about local languages. I mean, the IDN and other stuff. And it's also important when we are talking about applicable local laws. And I guess that is --

Okay. Before I pass the mic to Faisal, there's a little illustration of the -- just to emphasize the importance of the contact ID principle.

The first box is the domain name for a registrant with the registrant contact ID Number 1, and the second box is a domain name registration for a contact ID Number 2, but both IDs are acting as their own admin contact. What is common to them is a tech contact. If you look at the boxes, they both have a tech contact, ID 3.

The important message here is that when the contact ID 3, the tech contact, updates his own contact, that update is reflected on all the domain registrations across the board. It doesn't -- the technical contact does not have to go and update his contact on every domain. He only has to do it once, and that is the beauty of the contact ID concept.

So at this moment, I pass the mic to Faisal and Scott to discuss the various models.

FAISAL SHAH: Thanks, Lanre. So the EWG looked at a number of different models. You'll see the models that we analyzed on this slide. And each of these models differ in how the data is being queried or copied into the RDS.

And specifically, we looked at -- I mean, we looked at the current WHOIS and really analyzed elements there. We then looked at a federated model, which was a distributed model that pulls data directly from the registries and the validators.

Then we looked at a synchronized model.

And here's where the data is pushed to a common hub and then distributed to multiple datacenters.

And you'll see that we have, on -- on the synchronized model, an asterisk, and so we had changed the wording from "aggregate" to "synchronized," and it was deliberate and I'll let Scott maybe -- if you want to talk a little bit about that.

SCOTT HOLLENBECK: Sure.

One of the reasons that we wanted to change the name is because we'd received some comments after the Singapore meeting asking questions, "Well, what does 'aggregated' mean?" And as we got together as a group and tried to come up with a good definition, we realized we didn't really like the word either.

And we thought that, you know, "synchronized" better described the movement and management of the data and the function in a controlled, consistent, and secure manner.

So rather than try to justify "aggregated," we thought it better to rename the model to "synchronized."

FAISAL SHAH: Yeah. And to add to that, I think it was somewhat of a misnomer because there was going to be distribution of data in multiple databases and datacenters, so we wanted to get that across in the name as well.

So the other thing we looked -- the other models we looked at were the regional model, a model where the data is copied to several regional hubs, the opt-out model, a combination of the synchronized and federated models, in essence really allowing the registries to -- a choice of either one of them.

Then the bypass model, where the data is passed from the registrars directly to the RDS, skipping the registries.

And you should note that, you know, a number of these actually came from the community and that's why we included them and analyzed them.

And the key for the EWG in our analysis was that across all these models, the RDS remained the authoritative source, providing authenticated gating and logging access, and at the end of the day, whatever model that we recommend, it should remain distributed in architecture with data located in multiple databases across multiple datacenters, like I discussed before.

So the -- so we took all these models. The EWG analyzed each of these models against a set of criteria and it was a rigorous analysis that

included a number of factors, including, for example, you know, what were the security implications for each of these models.

What were the jurisdictional and privacy concerns that we had to -- had to address.

We looked at implementation requirements. Would one particular system be better equipped over another to handle the requirements that we were actually putting forth in our final report.

We looked at accreditation requirements. Was one model better or worse than another.

And we looked at some of the operational issues.

After that analysis and a lot of time, we then came up with two models that stood out from all the rest, and those are the ones we wanted to focus in on, and that was the federated and the synchronized model.

So then we went back again and took a hard look at those two specific models and then asked a third party, which was IBM, to analyze the costs attendant to these two models, and IBM released its implementation model cost analysis report that it did in March -- I think February/March, which was posted in -- on June 6th, this month.

And based on this study and based on our original analysis and the specific requirements that were laid out in the report and what we were -- we were trying to achieve, the EWG's recommendation was to adopt the synchronized model.

So on the next slide, we put up the synchronized model, and, you know, it -- there's probably some complexities to this, but you can see it's well

laid out here, and, you know, just kind of following the flow, the domain name data is collected by the registrars and/or the validators. A validator can be a registrar or a third party.

The validator stores the contact data and then pushes that to the RDS. And then the registrars would then simultaneously push the domain name data and the contact IDs to the registries, which is then pushed to the RDS that -- you know, we talked a little bit about that. Susan talked about that, as well as Rod.

In the synchronized model, the RDS is responsible for the storage of copies of the validated data. It handles all the queries, public and authenticated, authorizes access, applies these gating policies, returns allowed data, audits, and then provides additional services and those were some of the services that we talked about, which was the reverse query -- reverse queries and WHOIS services.

There's some -- obviously some differences between the federated and synchronized and I'll let Scott maybe talk a little bit about his thoughts on that.

SCOTT HOLLENBECK: Yes. I actually wish -- I wish we had a copy of the federated model just for comparison purposes because the biggest difference is that model has arrows going in all kinds of different directions that you just don't see here.

This is conceptually a simpler model, and when we say that it's synchronized, if you look at the movement of the data, following the arrows here, what you're saying is data collection coming from

registrants and contacts ends up in a registrar or validator place, who then synchronize that data with the registries or the RDS, who then synchronize that data at the next point in the path.

So balancing everything we had in terms of principles, requirements, and whatnot, we thought that the synchronized RDS was the simpler model to describe and the simpler model to implement.

FAISAL SHAH: And by the way, I tend to -- well, you know, as just a -- as an aside, you know, you know a big nerd when I -- I think it's really cool to sit next to Scott, who actually wrote the EPP, so, you know, he knows his -- his -- his tech stuff. So if you want to talk about the next slide.

SCOTT HOLLENBECK: Sure.

All right. So the next slide. We haven't actually said anything about protocols yet, and I think that one of the points I want you all to walk away from, when you look at this slide, is that a lot of the protocol work that would need to be done to make this model work either already exists or is well -- well underway.

EPP has been used as the provisioning standard for domain names, you know, for -- for like 13 to 14 years now. All of the gTLD registries know how to implement it. The registrars know how to implement it. And while we did identify the need for some extensions to the protocol, it doesn't need to be scrapped and it can do a lot of things we need it to do right away.

[Laughter]

SCOTT HOLLENBECK: Yay!

FAISAL SHAH: We actually brought the confetti for this particular purpose.

[Laughter]

FAISAL SHAH: This is a celebration.

[Laughter]

SCOTT HOLLENBECK: All right. Now, the other side of this, now EPP is the provisioning side.

We haven't talked about the query side, right?

And the query side with WHOIS is the WHOIS protocol. You know, a very, very simple protocol that is specified in such a way that it leaves an awful lot open to interpretation.

So for the better part of, oh, two years or so, the IETF has been working on another protocol in a working group called WEIRDS, right? Web Extensible Internet Registration Data Service. The name of the protocol is RDAP, or the registration data access protocol. If you haven't been -- or haven't had a chance to see the -- what's going on in the WEIRDS working group or if you're not familiar with RDAP, I would encourage

you to take a look. The working group is probably within a few months from starting to push documents through IETF last call, which means that very, very soon we will hopefully have proposed standard mechanisms for making this model work. All right?

So now let's jump to the -- a more specific example of how EPP and RDAP can be used in these models, building on some of the examples that Rod and Susan talked about earlier.

So imagine that you've got all this data that has been pushed into these repositories using EPP. Right? The data's there, it's been validated, and all the people who are looking to make queries have been validated and have credentials.

So User X sends a query to the RDS using RDAP, and this query will contain identification information so that the RDS knows who is asking -- right? -- and it knows what principle -- I'm sorry, not what principle -- what purpose they are claiming.

So in this case, the person is asking and they are claiming that their purpose is, "I want to resolve a technical issue." So for example, "This Web site is not resolvable and I'd like to talk to somebody about that."

When the RDS receives the query, it's going to authenticate the person making the request. Right? And assuming that the authentication passes, the next step is to compare the stated purpose to the purposes for which this person is entitled to issue queries. Okay?

Assuming that passes, the RDS will then, in the back end, look up the appropriate data for a particular domain name that has been authorized for release for this purpose of technical issue resolution.

Right?

The system will then look up the appropriate gated contact information, to include the registrant contact information and the technical contact information, aggregate it in a structured way, and return that information using RDAP.

Okay?

Carlton?

FAISAL SHAH: Just one point. On this particular slide, you'll notice, you know, it's looking up the authorized data for the domain name, the registrant contact ID and technical contact ID, and the -- if this -- in the federated model, it would have to go to several parties to get that information versus the -- the synchronized having it in one place, so...

CARLTON SAMUELS: So you've heard so far that quite apart from a very small set of data elements that are available for anonymous access, the rest of the data elements are for purpose-driven disclosure based on permissible access.

And you might think, therefore, that -- I hope you notice that this is grounded in protection. We know who is accessing the data, for what purposes, so we know if there are any problems, we can take mitigating action.

So data protection principles are at the heart of this discussion and we spent a lot of time talking about it and you will see from what we have produced fealty to the principles and we worked diligently.

There's another reason, and a big one that you should keep in mind.

The compliance effort -- if you've been around ICANN for a little while, with 25 gTLDs the compliance effort has been very, very, very difficult, and with the growth in the namespace, you would understand that they would be challenged. And one of the things we wanted to do was to make sure that with the growth in the namespace, WHOIS as it is would not be appropriate or even up to the task.

So we have to do something to take care of what's happening with the new gTLDs.

So mechanisms have to be adopted -- and this is now organic -- to facilitate routine legally compliant data collection and transfer between the RDS ecosystem actors, all of the persons involved in handling personal data.

And we have looked at it and decided that you could do -- get this, you could achieve this through a couple of ways.

The first one is by way of standard contract clauses that are harmonized with the privacy and data protection laws, codified in policy -- notice "policy"-- and enforced through contracts. Pretty much the framework that exists today.

And then you have a rules engine that is engaged to apply data protection laws by jurisdiction.

And of course RDS storage localization so that a high level of data protection is implemented.

By that, we mean where the RDS is localized, you have the best data protection available under law.

Pass it on to Stephanie.

STEPHANIE PERRIN: I would just note that that's not going to be a trivial challenge, building that system.

So in addition to compliance with data protection, we also must accommodate the accredited privacy proxy service -- I believe we've talked a little bit about that already and I have the honor, as does Carlton, as does Michele, of serving on that working group that's working on that -- and an accredited secure protected credentials service.

Now, this is a new service and the next slide talks about it so I'll wait a minute and...

No, I might as well move on to the next slide, I guess. Very good. Thank you.

This service would be available to people who are at risk who can come to some kind of a process or tribunal and indicate that they are at risk or that their speech rights are going to be prejudiced and apply for basically what amounts to anonymous credentials for a domain name registration.

Now, I know there will be folks in the audience who recognize that if you're seriously under threat, the domain name registration system is the very least of your worries, and that's true. Nevertheless, it's still one tiny cog in protection that ICANN is responsible for. So ICANN has a responsibility to make sure that people who are at risk -- at least has a responsibility and I rather say that rather glibly. In my view, they have a responsibility.

And so we are proposing this structure that would allow people to get a secure credential, which are on the market -- there's a couple of providers -- and apply for a domain name probably through some kind of proxy, and then they would get the secure credential and take it to a privacy proxy service provider. That's sort of one additional layer. And if you're at risk you're going to have several layers.

So you will see in the report we've identified five different categories of people. Life is easy if you are a New York Times reporter and you're in a hostile country. Your paper is probably going to go and get your credential to run your blog. It's more difficult if you're a women's rights group in a country where you're under threat. That becomes more difficult.

But we are reaching out to the civil society community to help figure out how this would be done and who would accredit these folks.

So that's it for me.

CARLTON SAMUELS: So we pass it back to Michele.

MICHELE NEYLON: Thank you, Carlton, Stephanie, and everybody else. Hope you all are still semi-awake. I know this room is rather toasty. We'll go with "toasty."

Okay. The -- I'm seriously of getting a T-shirt which says, you know, "Keep Calm and Read the Bloody Report."

It is 166 pages long. It's very, very dense. It's complicated. We're not going to lie to you. So there's a bunch of other topics --

If somebody could move forward the next slide, please, because -- so that they know what the hell I'm talking about, I'd appreciate it.

Thank you.

There is a load of other topics that we have covered in the report. Some of the principles around user accreditation, how we're going to handle things like law enforcement access, how we're going to handle compliance, the privacy and proxy services. Cost, of course, is another important thing.

There's a lot of information in there. It's all covered, well, in gory detail, I suppose is the best way of saying this.

And of course another one which we put in there is, you know, benefits comparing what we're proposing up against the 2013 RAA and the WHOIS in that.

Also, as was mentioned earlier, the slides from this -- from this presentation are on the schedule page, so if you're having difficulty

deciphering some of our beautiful graphics, please download the PDF and do whatever you need to do on your own computer.

So the obvious thing is, you know, what's going to happen next. Where do we go with this. What do we do. Do we turn the report into a doorstop or do we do something much more interesting with it like, you know, actually implementing it.

So over the next while, there's going to be more sessions like this. And after this session today where we presented, there is a Q&A session where you get to ask us polite questions. Please be polite about it. Please don't attack me.

And, also, we'll be running Webinars and other options for community to engage with us so we can go backwards and forwards with you to go through as much detail as you want or need.

As this report -- as this work was mandated by the ICANN board and CEO, it is pretty much up to them how they handle this moving forward.

So I believe that they will be looking at -- at the report in detail over the next couple of months, and then they will push it onto the GNSO or not. I don't know.

I mean, Chris is making muttering noises here beside me. So we will just ignore that.

Obviously, the question that a lot of people are asking is: Is what we have proposed better than the current WHOIS? It is an open question.

You know, please have a look at the report and decide. And, of course, you can disagree with us.

And I know that Kathy will.

I'm sorry, Kathy. You are sitting straight in front of me. I couldn't resist.

If what he are proposing isn't a good answer or doesn't contain elements that you consider to be a good answer, then is the current WHOIS suitable? These are the questions we were asking. This is what we spent the last 16 months -- my look of what the future is, is getting back to my original life. Thank you very much. I mean, I love them all to death. They're lovely. But it would be nice to actually see my own staff and family.

So we're holding the discussion sessions today. There's another one. If you can't make the afternoon one today, there's another one early in the morning on Wednesday. And I know that the concept of discussing WHOIS at 8:00 in the morning is cruel and unusual punishment. But maybe for people in different time zones, it is more suitable; I don't know. For those of you from other time zones, maybe it will fit in nicely with your jet lag.

Most of the EWG members are here with us on stage. We're more than happy to discuss these -- the elements the report with people as well. And Jean-Francois wants to say a couple of words.

JEAN-FRANCOIS BARIL: No. Thank you very much, Michele, for finishing this first part of this public session. Before I hand over that microphone to Chris, who has proudly volunteered to be our moderator for the next two hours, I will suggest that everyone of us stand up in this room, stretch our legs, and

then we sit down in respect of getting maximum momentum and time for the benefit of everyone.

CHRIS DISSPAIN: Thank you. I think that's a great idea. And it gives Nancy a chance to close the session and reopen it. So over to the voice of God.

>> We are now going to go into the expert working group final report discussion session. Thank you.

CHRIS DISSPAIN: So I noticed that people are taking -- what do the Americans call it? A comfort break. So we will start in two minutes. Two minutes' time.

[END OF TRANSCRIPTION]