BEIJING – SSAC Public Meeting
Thursday, April 11, 2013 – 08:00 to 09:00
ICANN – Beijing, People's Republic of China

JULIE HEDLUND: Hello. This is Julie Hedlund. And this the SSAC public meeting. We'll be starting momentarily, but I know we have someone on the phone. I think I just heard a sound. Who is on the line? Maybe that was just me joining.

Is there someone on the teleconference line? Okay. Well just give us a minute or so, we want to make sure that the slides are showing up in the Adobe connect room properly so that the people joining remotely are seeing them, and so maybe just a minute or two.

And now I see that Adobe connect quit. So it might be more than a minute. Thanks.

[AUDIO BLANK 0:04:00 – 0:04:48]

PATRIK FÄLTSTRÖM: So good morning everyone. Good morning everyone. Okay. The reason why I asked that question twice is just on this stage, like many of the other rooms, I hear absolutely nothing of what's coming out of the speakers. Which might be a good thing because if I pass around the microphone, and you talk into the microphone, then I will not be able to hear what you're saying.

So this morning, in the Security and Stability Advisory Committee, we're sitting in the room with the most conference chairs. It's the morning…

JULIE HEDLUND: I'm sorry for interrupting. Who just joined on the line?

BARBARA: Barbara.

JULIE HEDLUND: Hello Barbara. Good morning thanks.

BARBARA: Hello? Yeah. Good morning.

PATRIK FÄLTSTRÖM: So it's the morning after the excellent gala dinner, thank you very much [? 0:05:48] and all the other sponsors that managed to arrange that brilliant evening. With myself, this morning, waking up sort of having jet lag, and I was thinking, "I have jet lag." That's probably because we're out here in Beijing last Thursday, so maybe it was a good gala.

So now when we're sitting in the room with the most conference chairs, and pretty nice, and we're going to be friends here together this morning, maybe we should just turn down the light and have a little bit more soft than you can sleep. Soft music.

So a little more serious. Let's go into… We have an hour here, and I would like this to be… Excuse me. Who just joined?

MAN: [? 0:06:31]

PATRIK FÄLTSTRÖM: Good morning.

MAN: Good morning Patrick.

PATRIK FÄLTSTRÖM: So what we're going to do here is to go through the, first of all going to present a summary of what we have done in SSAC this week. So we can tell you a little bit about… Just wait a second.

I'm just going to [teach you on my slide 0:07:08] itself…

JULIE HEDLUND: We have a problem with Adobe Connect being sort of oversubscribed so it's kicking me out, like every few minutes. And as we found in the DNS SOC workshop yesterday, it's just really annoying for the people in the room that I have to connect to Adobe Connect every time that happens. But I do need to do that, or the people who aren't in the room won't see the slides. So we're going to share duties here. Sorry about that.

PATRIK FÄLTSTRÖM: So I'm going to try to show the slides so that people here in the room can see at least what we are going to present. Fantastic. There. So what we have done this week in SSAC is that first of all, we have had a

couple of meetings in the following committees and work parties: the administrative committee met Friday and Saturday, almost a week ago.

And the administrative committee consists of myself being chair, Jim Galvin being vice-chair, [? 0:08:37] the liaison to the Board, and staff was Steve Chang and Julie Hedlund that was participating locally. We also had a meeting in the membership committee that Jim Galvin is chairing.

And the work part is to identify abuse metrics. Abuse of the DNS for distributed malware service attacks did meet as well. We also had briefings at a number of public meetings. We gave an overview over the SSAC update on the 2013 work plan. And we have presented about SAC057 and SAC058, and those are the two reports that we are thinking of discussing here today.

In many of the meetings, just because we started with the SAC057, we unfortunately didn't have time to go into SAC058 but as many of you know, the SAC057 report is the one that has created most interest here in Beijing. So because of that, we have the ability of this meeting to give a little bit longer report on SAC057 than what we have done at the various meetings, if it is the case if you are interested in that.

We also regarding DNS SAC, we were running the DNS SAC for everybody this morning. We had implementers, DNS SAC implementers gathering on Monday evening, thanks very much to the ones that helped sponsor that. And then we also had the normal workshop DNS SAC workshop yesterday.

We would also like to very explicitly thank everyone that participated as speakers. And we talked a little bit this morning about the format and we could probably have run that for two days instead of half a day, because there were so many things to talk about. And we'll see how we're going to make that more efficient next time.

But obviously that is something that is very, very interesting for people, and we would also like to thank them for influencing aside as well for trying to do this coordination collaboration between the ISOC 360 program and what we are doing.

I think that's yet another sign of good collaboration. And then we had a full meeting for all of SSAC late Tuesday, where we discussed all the 2013 work plan. On the other hand, we of course were discussing the aftermath of SAC057 quite a lot and name space overlap. A part of going through quite a large number of administrative things that we, of course, always have to talk about.

We did go through, for example, processes used regarding membership, we also went through how to handle our bios and keep both bios and statement of interest information up to date, and make that information by ISOC members more clear on the ICANN website.

We have had various meeting with the community. We have met with the At-Large advisory committee. ccNSO tech day, we presented there. We met with GNSO. We also had a meeting with the counsel registered stakeholder group and ISPC in separate meetings. We also had meetings Monday this week with law enforcement representatives.

We both had SSAC individuals that participated in large portions or even the whole day of the Monday. But then we also had a specific meeting between SSAC and law enforcement, when we were talking both how we could do more things together and exchange information and try to basically help each other to be more efficient and create something in the future.

We'll see what's going on. I also had a meeting with ICANN Fellows yesterday morning, which was once again, a great success. So I'm really happy with the corporation that we have with the people running the fellowship program. And at the previous… Around the previous ICANN meeting, we actually one of the ICANN Fellows was accepted as a SSAC member.

So some people come from Fellows directly to SSAC, which I think is a very good thing. So while I switch the slides, is there anyone that has any questions on that? So now over to the sort of the normal program. So the Security and Stability Advisory Committee, we were initiated in 2001 and began operations in 2002.

And we are 38 people at the moment, and the charter is to advice the ICANN community and Board on matters relating to the security and integrity of the internet's naming and address allocation systems. I've already gone through this.

So the publications that we have released during 2012 and 2013, divided into categories, are the following: regarding DNS security and abuse, we produced SAC053, 56, 57, and 58 where, as I showed earlier, 57 and 58 are the ones that we are going to go through more in detail later on today.

Regarding internationalized domain names, we released SAC052 on Delegation of Single-Character Internationalized Domain Name Top-Level Domains. Regarding WHOIS, we sent in some comments on the WHOIS review team final report, and also released a report on the domain name registration data model.

And now a question to you. We have two alternatives here, either I go through sort of an overview of SAC057 and we can have a discussion, or I let Warren up on stage and he can give show you more precisely what the implications of SAC057 actually is. Do you want… How many people would like to see what Warren has actually found with 057? Okay. Good. I'm not going to ask the other question.

[AUDIO BLANK 0:15:21 – 0:15:54]

MAN: Good enough.

[AUDIO BLANK 0:15:57 – 0:16:02]

WARREN: Hey everyone. So you probably, or you might have already seen it, what was it? Monday? Tuesday? I'm going to go through it again, hopefully a little slower this time. And if any of this is unclear, please stop me, ask questions, etcetera.

So background. When you connect over to an HTTPS site, basically with something using SSL or TLS to secure the connection, your browser actually uses a public key to encrypt the data. And it gets this public key in a certificate, and the reason that it trusts the certificate is that it is signed by a certification authority.

And this signature binds the public key to an identity. So basically them signing it, binds the public key that you get to something like www.example.com. Or something like that. And then when you actually use the certificate, your browser makes sure that the site it is connected to matches the identity that's in the certificate.

So for the CA can actually issue the certificate, they need to make sure that they're giving it to the right person. If they don't do this, somebody has showed up, if they don't do this then you can perform a man in the middle attack. Somebody can get a certificate for a name, that person can then present it.

The way they do the validation, at least for DB certificates, is they send an email to the specific address at the domain you're trying to go a cert for. So if you're trying to get a cert for example dot com, they'll send the email to host master at example dot com, or webmaster, or even the email that is in WHOIS.

And then whoever receives this, just replies back and that proves that they own or control the domain. There is another type of certificate, which is called an Internal Name Certificate, and these are designed for internal only type things. Microsoft recommended for a long time that you use these, they don't recommend them anymore, but a lot of

people have deployed them for Exchange, Active Directory, similar sorts of things.

Lots of corporation use the internally to secure stuff. And the identity in these certificates is of the form www dot corp, or www dot accounting, or mail dot test. And the only thing that makes an Internal Server Name Certificate different to a regular certificate is that it doesn't end in a TLD.

That's the only difference. Because it doesn't end in a TLD, the CA can't actually send this validation email, because it isn't a real domain to send it to. So they kind of skip that step. They are supposed to perform other validation, this doesn't always happen. So the fun part is when one of these internal server names certificate becomes a real certificate, basically if the last part of the identity gets delegated.

And when this happens, potentially really bad things can happen. So to demonstrate that, I applied for a certificate for www dot site, site is actually an applied for TLD. And because I knew that the CA wasn't going to send me a validation email, I put in an interesting sort of organization name[? 0:19:15] forge. And not enough people laughed last time, and I'm really happy with my pun so.

Okay that will do it. Yeah that's go enough. So the CA actually popped up a little box warning me that this wouldn't actually work on the internet, and mainly just to protect myself. And so I said, "Okay, yeah, I realize that, I'd like to cert anyway."

Two or three hours later, they emailed it to me and here is my newly minted certificate, and you can see the identity there is www dot site. It

also has another identity, which is just site, which I think was nice of them, but anyway. And then to demonstrate the issue with this, I setup a fake instance of the root, I delegated dot site to myself.

So basically I made it look as though the site TLD now existed. And then I setup a webserver that was serving the cert. And this works fine in Safari. Basically I have now performed a man in the middle attack against what would be site if it existed in the real world. I mean, it's all fair that it works in the browser. The browser is doing what it is supposed to do, it matches the identity to what is in the cert, the cert is validly signed, and so the browser is doing the right thing.

Same thing happens in Chrome. I didn't bother putting it in the slides, but also IE, and Firefox, and Opera, and all the other browsers. I mean, this isn't unexpected, this is correct behavior. But what it does means, is that an attacker can get the list for applied for TLDs, and then they can go to CAs and get a certificate for things that they are fairly sure will be delegated eventually.

So well-known names that they will expect will show up in this TLD. They then wait until the TLD actually gets delegated, and then they perform a man in the middle attack. Basically they go up to a Starbucks, and they setup a fake SSID, or they hang out on a hotel network, or potentially they run a small country, or something like that.

Cache poisoning, HTTP poisoning, etcetera. And when a browser goes to one of the sites that the attacker has a cert for, they just present the cert, the user has no way of knowing this isn't correct behavior, and they fill in their banking credentials. The user has a lock icon, they have no reason to suspect that anything odd is happening.

And then of course the attacker runs off with all of your money. So are we going to talk about the timeline? Or just what we suggested. Okay. So we wrote up this advisory. We did have a little bit of trouble because our advisories are normally public, and so we had to figure out a new way to publish this so that it had limited distribution.

We got legal involved, etcetera. And we ask the security team to take over much of the litigation to reach out to the CA/B forum, who is the sort of industry group that represents a large number of CAs. Not all of them, but a reasonable number. And we asked that they start treating applied for gTLDs as TLDs as soon as possible.

The security team also developed a vulnerability disclosure policy, largely based on this I believe. A communication plan to inform the affected parties, and a contingency plan in case this vulnerability leaked too early. And I think that's the end of my slides. Do you want to do the… What happened? Any questions before we, or questions after.

PATRIK FÄLTSTRÖM: I can go through what actually happened first.

[AUDIO BLANK 0:22:56 – 0:23:16]

PATRIK FÄLTSTRÖM: So this report we finished in early January. And just like Warren said, we came to the conclusion that this was pretty serious. And we cannot make this public immediately because then, even though it is sort of known that these kind of things happens, if it is the case, that we're

rocking on the nose of some people, it might be the case that there is a big risk that people would run away and request these internal certificates before we actually had had communication with CA/B forum.

So we decided, as Warren said, to hand this over to the ICANN security team under Jeff Moss. And what they did was that they alerted the browser forum chairperson on January 23, brief the CA/B forum at the annual meeting on the 5th of February, which was sort of kind of pure luck but also good thing. It was very convenient that they actually had a yearly meeting already planned.

They wrote together, Ballot 96, on new TLDs was brought forward and passed by the CA/B forum on the 20th of February. And the ballot says that CAs will stop issuing certificates that end in an applied for gTLD string within 30 days of ICANN signing the contract, and they will revoke any existing certificates within 120 days of ICANN signing the contract with the registry operator.

At this point in time, we together with the ICANN security team when we worked through like what are the actual impact of this, we decided and we are now in early March, we decided to make the whole report public. So the document SAC057 consists of our report, and then the timeline and list of events after the report was found, which is what you see on the outcome here, but in a little more detail. You see that in Appendix A of the report itself.

So now any questions? We should have at least one microphone. Okay.

MIKEY O'CONNOR: This is Mikey O'Connor. What's the proportion of certificate authorities you think that are covered by this? And maybe the converse question which is, how many of them are left? And what do we think the vulnerability that that is?

PATRIK FÄLTSTRÖM: We happen to have a person from CA/B forum here that I think is the right person to respond to that.

[AUDIO BLANK 0:25:46 – 0:26:03]

MAN: Okay. Whoa. Yeah. I think your question was how many CAs actually issue these type of certs? And how many certs there are?

MIKEY O'CONNOR: The real question is, how many of them are not going to be abiding by this outcome?

MAN: All of the major CAs are going to be abiding by it. And in fact, as an update, yesterday Mozilla started the process of adopting it as a mandatory part of their root requirement program. So every CA will be… If it passes through the Mozilla process, which it will [laughter], every CA will be abiding by this new rule.

MIKEY O'CONNOR: Great news. Thanks.

PATRIK FÄLTSTRÖM: Clarification. I assume that was a means registration authority, because eventually they chain up, right?

MAN: Yes. The CA/B forum rules are always applied to the delegated third parties.

PATRIK FÄLTSTRÖM: [Rom? 0:26:56]

MAN: Thank you this is [? 0:26:57]. Just a clarification question. Up there it says that the CA will stop issuing certificates, and then there is a 120 day time period where CAs will revoke any existing certificates. Can we clarify for the group, what ICANN intends to do?

In other words, is ICANN going to wait on… If they sign a contract with the new TLD registry, does this mean that ICANN intends to delay the delegation of that TLD for that 120 days?

PATRIK FÄLTSTRÖM: I guess I'll try this. So I believe the wording is that ICANN will stop delegating 30 days after contract signing. I think that…

MAN: You said stop delegating. You mean stop issuing.

PATRIK FÄLTSTRÖM: Sorry. Stop issuing. Yes. Thank you. Pardon? What did I say? I take back what I said. The CAs will stop issuing these 30 days after contract signing. I think that it is likely to take more than 30 days between signing of contract and delegation anyway, but I don't actually know.

I'll let somebody else do it because…

MAN: Do we have anyone from ICANN here that can respond to that question? John, yeah.

JOHN: I can respond to that question, but I think that's a question best put to the Board and the executives, so maybe to the public forum. Get somebody to put it there. I don't think anybody in this room is going to be able to answer that.

PATRIK FÄLTSTRÖM: Yeah. Thank you. So let me clarify a little bit, because it's a little bit confusing this week on the response to that and similar questions. And the reason why I don't answer, is that there is a difference between what are the SSAC work ICANNs, which are the actually toss that we are working with, and the issues that are discussed sort of in SSAC and in ICANN community.

When in SSAC very, very explicitly start to work on a specific issue that is based on a question or a trigger that is issued from someone, by someone, including SSAC members that can initiate an issue, a discussion topic themselves. Like in this case, it was Warren that brought this to SSAC, that notified SSAC about it.

SSAC then does work and then release a report, which we just did. When we release the report, that is when we stop working on that specific issue formally. That doesn't mean that we stop thinking about those kind of topic related items. But the SSAC formal work has, at the moment, stopped.

That doesn't mean that we can restart it, that we are not thinking about or discussing these kinds of issues. So that's where we are, and that's why it's not really – it's not up to SSAC to decide and answer the question instead, it's something to be brought up the Board, requesting the Board because we have now released our report and now it's up to sort of ICANN to act.

MAN: Thank you Patrik. Just as a quick follow up. I've heard some folks come up to me, I mean I'm their SSAC liaison to the Board, and say, "I've heard that the SSAC recommendation is 120 day delay is necessary." So even though I recognize that it's not something that we have recommended to ICANN, or something that we even necessarily thought of, there is some level of conflation between what we're seeing in our recommendation, versus what some folks in the community are hearing about ICANN taking the 120 days, and saying, "Until all the

revocations are done, no delegation is done, because that is the right security and stability approach."

So that's a reason why I think there is some conflation going on in people's minds.

PATRIK FÄLTSTRÖM: Absolutely. And personally speaking, as personally, of course, as Chair of SSAC I also drew specifically the last two days as something I got back question quite a large number of times. Of course, that kind of discussion is exactly what my trigger instructionally, actually do something or act, but there is no – I don't see any sort of resolution to that discussion.

Did you want to say something?

MAN: I actually have heard a couple of times here, and people have said this, I'm sorry [Rom] but even you just said it here too. We just need to be careful to understand that it's the CAs that have that 30 day window, or 120 day window. It's not ICANN.

So it's the CAs that are going to take that action. And I guess the concern here is your suggesting that ICANN negotiated for the 30 or 120 days, and that is what the CAs are going to do.

PATRIK FÄLTSTRÖM: It's early in the morning, and we should not dig this hole in the ground deeper. We have our report, we are reporting back that this is the

resolution from the CA/B forum, okay? Now the question is, does that imply that ICANN should have some extra policy or something to delegation?

That is up in the air. We have issued our report, you can read it yourself, that's where the SSAC formal work stopped. But once again, that doesn't mean that we as individuals, you sitting there, regardless of whether you're a SSAC member or not, have stopped thinking about the issue. Okay. I think that's where we are.

MAN: Patrik [? 0:33:05]. I have a question for the CA/B forum. I know that ICANN released the full [purpose 0:33:14] of the certificates. But couldn't they measure and aggregate per [? 0:33:21] TLD?

Say they asked their members to say, "Oh how many dot com certificates were issued?" And then they report, in aggregate to the community. Not as the CA has searched, they will know that, but they will report that was – they will just report this TLD has that many certificates, this TLD has that many certificates.

So we would have a list to base risk analysis.

PATRIK FÄLTSTRÖM: So do we have a response there? Yeah.

MAN: I think if we leave the microphones on, I think it would be better.

MAN: Yeah. Okay. That work is actually being done through a group called the [CASC 0:34:06], which is the CA Security Counsel, and I'm hoping that we will publish it shortly. It's a lot though because this was a recommended practice up until 2011 by a lot of the server software. In fact, it would auto-configure your network that way.

So you're seeing a lot of small and medium businesses that actually have it pre-configured.

PATRIK FÄLTSTRÖM: Next question.

GREG AARON: Greg Aaron. [Rom] I think one of the questions that should be asked is, once ICANN signs a contract, will the staff have a process for immediately reaching out to CA/B and letting them know?

PATRIK FÄLTSTRÖM: Next?

STEVE YORK: Steve York. I want to switch a little bit on a topic, and maybe it's just that it's early in the morning and I've only had three hours of sleep, but Warren, when you talk about this, could you explain a bit about the actual attack service, I guess I would say? Because if I think about it, in your example where you have www dot site, if that's something that is out there, in order for this attack to really impact me, I've got to be on a

network where somebody else in the middle there has a www dot, and in effect be going to a www dot bank dot site, first of all.

I've got to go to that first. And then as I go there, somebody has to be in the middle, somewhere, in order to be hijacking this. So what's the realistic risk here in terms of people and exposure that's out here? I totally agree, this is a bad thing. So don't get me wrong about that.

I'm just more looking to try and understand the risk. Thanks.

WARREN: So I mean the, yes. If these get issued, it's not going to be happening everywhere. You would need to be on a network or a location where somebody is able to present the certificate. We have seen instances of this happening in organizations. Some place where an employer would like to know what their employees are doing, that's potentially reasonable.

There were also sorts of things where in certain countries, people have had a large amount of the country going through something like this. I mean if you look at, like the [? 0:36:31] event, it could be used somewhere like that. But also if it's an organized group, they could have a number of people for example, standing up fake SSIDs in airports, or coffee shops.

So this isn't the end of the world.

STEVE YORK: But I guess my question, what's that? I mean, seeing the tactics that you mentioned there with [? 0:36:54], with SSI… Any of that type of

thing could be done against any SSL certificate. I mean, it's a generic attack that could be done against anything that's out there.

WARREN: Yeah. The thing is, you need a signed certificate for the…

STEVE YORK: Oh, okay.

WARREN: And what this does is that it gives you a validly signed certificate. So a difference between…

MAN: Okay.

WARREN: …that you actually have a valid certificate. Basically what this means is any place where somebody can perform a man in the middle attack, they can now potentially provide you with a real certificate that your browser would believe.

When I am at Starbucks, and I go to my bank, because I have the lock icon currently I believe it's secure. I would not necessarily be able to after this sort of issue.

MAN: Okay. Thanks.

PATRIK FÄLTSTRÖM: Steve.

STEVE: Greg, regarding your question about notification….

MAN: Can you hold the microphone closer to your mouth please.

STEVE: Sure. ICANN setup a notification service through the CAs and browsers, and so there are some CAs not on the CA browser forum. What we did is, we reach out to Mozilla and Mozilla asked all of their CAs that used their root list, to sign up for this service.

So even if your CA not on the CA browser forum, because of this Mozilla requirement, they had to sign up for this service. So really got a lot of subscriptions. In terms of service itself, we will push out first all the list of applied for TLD strings, all of their priority numbers.

So that will give the CAs a sufficient warning ahead of time. After that, whenever each contract is signed with ICANN, we will push out that notification, for the CAs. And we're also working with Microsoft for them to get their CAs, signed up for their service. Thanks.

PATRIK FÄLTSTRÖM: Warren.

WARREN: And just a quick follow up to what Steve said. It's not just Mozilla who uses the Mozilla list. A large number of people sort of inherent the Mozilla list of which CAs are in the group. So.

PATRIK FÄLTSTRÖM: David.

DAVID: I'm curious if ICANN staff or anyone else has actually looked at sort of the non-browser software side of this. Because, for example, I'm sure ISO has deeply embedded in its developer's kits mechanisms which go, and fetch, and validate URLs through [? 0:39:30]… I'm just wondering if there's been sort of an attempt to broaden the scope beyond just the browser world.

PATRIK FÄLTSTRÖM: So yeah. I think I can respond to that. There have been a number of suggestions that one of the places to mitigate this is in the browser. So the browser would look to see whether their certificate was issued after the TLD went live, and things like that.

To me, that seems like the wrong approach because there are a huge number of protocols that rely on these certs. But the great thing is because this CA/B forum members, and it seems like most of the LCAs are going to stop issuing these, that problem sort of solves itself.

If it's not a validly signed certificate, or if it's expired, it doesn't really effect anyone. But great question.

PATRIK FÄLTSTRÖM: And I think that is a valid question for all definitions of [? 0:40:23]. Hooray, people are laughing a little bit. Hey they laugh for my pen. So any other questions, or should we move on to the next topic? So let's do the next topic. Jim, over to you, and I'll skip this slide so let me know when you want to swap.

JIM GALVIN: Okay. This is SAC058, our report on registration data validation taxonomy. In the same way that we had approached the WHOIS term, sometime ago, and published a taxonomy for that, we took a look at the discussion about validation, and have done the same thing here.

And we simply wanted to realize and separate for the community that there are several different topics for discussion here in validation. And hopefully, contribute and facilitate to a better discussion about validation in the community, validation registration data.

So next slide. I think it's already self-evident that the quality of domain name registration data is important and useful. There are a variety of reasons, and different stakeholders have different purposes for that data.

So we took a look at not only the taxonomy, but we also looked at some of the techniques that are used for validation. Also in an attempt to separate the discussion and get people to realize that different things are more effective than other things. So next slide.

Okay. So taking a look at the specific findings that we had. The first one, again should be fairly self-evident that the quality of data is really relative to the registrants and their purposes. There are different

stakeholders for different elements of the data that is collected, and they all have different reasons for wanting it, and therefore they also have different quality that they might be looking for.

The reason why this is important is because there are different mechanisms that you can use for validating the data, and if you need very high quality data then you want to imply a more expensive or time intensive validation technique.

And if you have data which is – which you don't need to be – which you understand could be volatile, might change frequently, you might not want to use a very expensive technique for validation, you just need a mechanism that confirms that it's effective.

The second finding, of course, is that some validation technique can be automated and some cannot. And this is something that we really wanted to focus on a bit. In our experience in watching some of these conversations about validation, you find that different stakeholders, because they have different things in mind, different purposes for validation, different mechanisms in mind, they're often talking past each other.

They don't appreciate that things that can be validated, things that could be automated, I'm sorry, would typically be less expensive, less costly to implement. But some things cannot be done in an automated way. So we want to try to look at those kinds of things that can be automated and separate them from the mechanisms that cannot be.

So that we can be clear about what we are talking about. And then the third finding is obviously the different data elements have different cost

structures. Again, the obvious example is looking at an email address and whether or not it's effective. It's typically something that you can partially automate.

If you look at a postal address, or contact information, it really would be impossible, nearly impossible to automate that and determine whether or not it's a valid address. And so it's just important to call out those different cost structures. Next slide.

So we have three recommendations for the community. The first one is the taxonomy itself. So we have three phrases, we've separated the taxonomy validation into three categories and the first thing to emphasize about these three choices is that they are peers to each other. They are not intended to be ordered or suggest that one is any better than another.

Different validation techniques and mechanisms will be appropriate for different circumstances, and that's what's essential here. So syntactic validation is all about whether or not the data actually looks like what it is supposed to represent. Easy examples are, is an email address really an email address? Does it look like an email address?

Is it a left hand side an at and a right hand side? A phone number. Is the phone number actually look like a phone number? Does the phone number actually have… Is it being input in a syntax that it looks like it should for the region of the contact information that's been submitted? Okay.

These kinds of things are also expected to be automated. The expectation about syntactic validation is all of these kinds of tests could

be done in line with the registration process. Because it should be possible to automate this process and thus make that decision very quickly.

And ideally, that would be something that, in fact, would be done by everyone. That could be a baseline or a minimum, and an expectation that one could have of everyone. Operational validation is the assessment that the data that you are looking at will actually work in the way in which it's intended.

So again, in the case of an email address, can I actually deliver an email message to that email address? Will it take a message from me? In a case of a phone number, can I make a call? Can I attempt to connect a call and will something answer on the other end? So is it a connected line.

And note explicitly that this is distinct from identity validation which is giving an email address, is it an email address that actually represents the real world identity behind it? Is it the same real world identity that's the registrant that you're looking for? And the same thing with a phone number.

They might enter a valid phone number, but is it actually their phone number? Or a phone number that is a way to get at them? In the case of operational validation, the expectation is that some of this, though not all, could in fact also be automated. The email address is sort of an easy example because email servers in today's world actually do a variety of things.

There are some that are simply – they won't answer a question of whether or not an email address is valid. You actually have to try to deliver, or not deliver. So I mean there is actually a command that allows you to query email servers and ask them if an address is valid. But in today's world, there are many servers that won't simply respond to that query anymore.

The other problem that you have is there are also servers that will simply take all messages in, and they decide later what they are going to do with it. So just the fact that you can give it a message doesn't mean anything to you, you still have to be able to process any failed message that you might get, or message notification that you might get after the fact.

So and that's true for a lot of things. So some things can be automated as far as operational validation is concerned, but many things cannot be. Identity validation is the other end of the extreme. The expectation there is that's probably largely manual, and most things would be, and it would require human intervention.

So anything that falls into that category like whether or not a postal address is a deliverable address or not, is definitely something that could not be done in line with a registration process. You would have to do that after the fact using some other manual mechanism.

Next slide. So the second recommendation to try and lay out the questions that need to be asked in deciding how to apply these three types of validation. And how do you make a choice as to which one you want to use? So what we've proposed here is at least these four questions that would need to be considered and answered as a way of

deciding which one of these validation categories, and which technique you want to look for, which one you would want to use.

So obviously in the first case, again getting back to this point, we collect a fair amount of data. Different people have different purposes for that data. And for that reason, they have different qualities or different levels of validation that they might require for that data. And so it's important really to look at the data and to think about what you're doing with it, and who is going to use it.

To a large extent, this is a question which is going to be answered by the expert working group that we have on directory services. So those things are happening in tandem, but our expectation at this point is that question would be mostly answered by that group, ideally.

In the second case get to the question of whether or not validation should be done in line or not. And we have tried in our separation, the validation techniques to speak directly to this question. Anything that falls into the category of syntactic validation, again, the expectation is that you could do that in line and it could be automated, and therefore it could be something which perhaps could be declared for everyone to do, and become a minimum part of the registration process.

And other things such as identity validation, there would have to be different processes set in place to handle that, but you couldn't do it in line because it would simply take too long, too great a delay in the registration process.

In the case of the third question, obviously there is different cost factors associated with all the validation techniques. And you really do have to

ask yourself what value you're getting out of applying a high cost validation technique to data which is low cost. And an economical example here is email addresses versus postal information.

Email addresses are quite volatile and would change quite frequently. And in that sense, you wouldn't typically want to apply a very high cost validation technique to testing an email address, since you would have to be applying it quite frequently.

Postal addresses on the other hand, tend to be less volatile. So they don't change quite as often, and therefore applying a higher cost technique is probably more appropriate. And then of course, the last question is the natural tension between accurate data and the privacy wishes of the registrants in particular who are contributing the data.

This is also a question which we expect to get some guidance from the expert working group on directory services which is now engaged. So hopefully there will be a significant contribution to that question for future consideration by people applying validation.

And the last recommendation is explicitly to call out that we suggest the community seek to identify validation techniques that can be automated. We have identified just a couple as examples in this document. But clearly some effort should be put into identifying those things that can be automated, so that we can have this baseline, and thus in fact improve overall the quality of registration data.

There certainly is a great deal of value in doing that. We do have to answer some other questions, it's not a singular question. For example the privacy considerations, but we should seek to find those techniques

that could be automated, and then incent their development and deployment so that everybody has them and we can all do them, and improve the system in general. And that's it, thank you.

PATRIK FÄLTSTRÖM: Is there any questions?

MILTON MULLER: Hello. This is Milton Muller from Syracuse University. I just did want to flag, when you talk about identity verification, you certainly are talking about a policy issue of enormous consequence. If you're talking about global linkage of the registrant to real name identity cards, on a global basis that's a pretty, let's say interesting change in global policy.

Maybe you're not talking about that. Maybe you're talking about these more automated telephone number validation techniques and so on. But my question about that is, okay, what is the net gain in the actual security and stability of the internet by automating that? Have you actually been able to get a handle on that in a quantitative sense?

In other words, are you constantly validating it? Or are you, at the initial point of registration, validating it? Yeah. When you actually do have an accurate telephone number, are you trying to stop people from making mistakes when they enter the data? Or are you really trying to stop people who are trying to….

[END OF AUDIO]