ICANN79 | CF – SSAC Work Session (6 of 8)
Tuesday, March 5, 2024 – 3:00 to 4:00 SJU

RAM MOHAN: Good afternoon. Welcome to this joint session between the SSAC and the NCSG. I'm Ram Mohan and Kathy, we are ready.

KATHY SCHNITT: Hello and welcome to the SSAC Work Session 6. Please note that this session is being recorded and is governed by the ICANN Expected Standards of Behavior. To ensure transparency of participation in ICANN's multistakeholder model, we ask that you sign into Zoom sessions using your full name. For example, a first name and last name or surname. You may be removed from Zoom session if you do not sign in using your full name. With that, I'm going to turn the floor back over to Ram Mohan.

RAM MOHAN: Thank you so much and welcome. Just because not everybody knows everybody else, maybe we start off by just quickly going around the table and just a quick name, your affiliation and what you're doing. I'll start. I'm Ram Mohan. I'm the chair of the SSAC. In my day job, I work for a domain name registry company called Identity Digital.

STEVE CROCKER: My name is Steve Crocker. I used to be a pure SSAC type person, but I've recently joined the business constituency. So I feel, in fairness, a full

disclosure and recognition of the inherent tensions that I ought to be forthcoming about that. I'm spending most of my time these days on two subjects, one is the technical aspects of automating the DNSSEC provisioning, but even more time on what I call the WHOIS mess, and happy to talk at great length about that.

JOHAN HELSINGIUS: Hello, everybody. I'm Julf Halsinius. And if you wonder about my slightly strange accent, I'm a Swedish-speaking Finn who has been living in Holland for 25 years just to confuse people. Oh, and married to an American. It really helps pinpoint the accent. I am the chair for the non-commercials, and to just give you a context of stuff, I started with the internet back in the '80s, and I made some fame with some slightly weird stuff back in the day. And my big claim to fame is that I actually got the EFF Pioneer Award two years before the other Swedish-speaking Finn, Linus Torvalds.

MAARTEN AERTSEN: Hello. My name is Maarten Aertsen. I work for a Dutch nonprofit called NLnet Labs, where we make open-source software in the areas of DNS and routing.

JAAP AKKERHUIS: I'm Jaap Akkerhuis and I'm an SSAC member, and also, a long-time internet [inaudible]. That's why I'm now [inaudible] for 25 years to make it complete. And I'm still working part-time for NLnet Labs.

GREG AARON: I'm Greg Aaron, a member of SSAC, and my consultancy is Illumintel Inc.

ONDREJ FILIP: My name is Ondrej Filip. I'm a member of SSAC. My job is the CEO of the CZ.NIC, and I'm also chair of RIPE NCC.

JACQUES LATOUR: Jacques Latour, I'm with CIRA. We run .ca, and I'm the CT [inaudible].

MERIKE KÄO: I'm Merike Käo, a member of SSAC, and I do all kinds of different security consulting.

JUAN MANUEL ROJAS: Hello. My name is Juan Manuel Rojas. I am the NPOC Chair, and I work as a researcher in e-governance, e-government issues.

CALEB OGUNDELE: Hi. My name is Caleb Ogundele. I'm the Membership Chair at NPOC, and I currently serve on the [SSGEC], and I'm just me. Thank you.

FARZANEH BADII: Hi, I'm Farzaneh Badii. I'm a member of non-commercial stakeholder group. I'm a recovering academic. I have founded Digital Medusa to [inaudible] the enemies of the internet, open, secure, global, free.

GAUTAM AKIWATE: Hi, my name is Gautam Akiwate. I'm a researcher at Stanford University. I guess, I'm not a recovering academic. I'm still an academic. Oh, I'm a member of SSAC.

ROD RASMUSSEN: Hi, Rod Rasmussen, SSAC member. I guess, I'm not recovered, but I'm retired as a cybersecurity person. Thanks.

JOE ABLEY: Joe Abley, SSAC, CloudFlare.

JEFFREY BEDSER: Jeff Bedser, SSAC member and CleanDNS.

JAROMIR TALIR: Jaromir Talir, CZ.NIC, [inaudible].

JOHN LEVINE: John Levine, SSAC. I'm wearing a variety of hats, but I think the relevant one is I'm the liaison for the messaging, malware, and something. Basically, it's the messaging, anti-abuse association.

PETER THOMASSEN: Hi, Peter Thomassen, SSAC member, and also running a non-profit that's a DNSSEC hosting platform.

MATTHEW THOMAS: Matt Thomas, SSAC and I'm with Verisign.

BARRY LEIBA: Barry Leib, SSAC. I work for Futurewei.

WARREN KUMARI: I'm Warren Kumari, Google and I'm part of SSAC.

TARA WHALEN: Tara Whalen, Vice-Chair of SSAC, and I work for Cloudflare.

RAM MOHAN: And let me go online and just ask folks online to come off mute and briefly introduce yourself.

ANDREI KOLESNIKOV: I'm Andrei Kolesnikov, SSAC. I work for [inaudible], non-commercial organization.

STEPHANIE PERRIN: Stephanie Perrin, I'm an NCSG councilor, and I also have a small consulting company largely in privacy and civil liberties.

GLENN RICART: Glenn Ricart, a US nonprofit, US Ignite.

BENEDICT ADDIS: Benedict Addis, SSAC, and the Chair of the Registrar of Last Resort.

LYMAN CHAPIN: Lyman Chapin, an SSAC member with Interisle Consulting Group.

GEOFF HUSTON: Geoff Huston, SSAC.

RAM MOHAN: Anyone else online who would like to introduce themselves? Okay, I think we've gone through that formality. Let's just look at the agenda. I think we wanted to go to you first, see what topics you may have for us to discuss, and then we have just a couple of things to chat with you about. So we'll hand it back to you.

JOHAN HELSINGIUS: First of all, would you actually mind sort of going through our main points?

UNIDENTIFIED FEMALE: Thank you. So basically, the purpose of this meeting is to tell you who we are, and get to know you. And this is a good occasion to tell you what our concerns are, what we are advocating for at ICANN, how we work, and who we actually work with, and other things. And we are going to bring up an issue on the recent report on urgent requests for access to domain name registrant's data, and we want to also see how we can collaborate and tell you what we do so that you can find out how we can collaborate. So that's it.

JOHAN HELSINGIUS: Yeah. And I think it's important to remind you of our slightly odd structure, which is that we have a stakeholder group that then has two constituencies under it, NCUC and NPOC, who are both represented here. So the policy work and our representation [inaudible] council is all on the top level, and then the sub-constituencies are more outreach. But what we really are about is, of course, civil society in terms of human rights, privacy, freedom of expression, and transparency. So that is kind of really the core values of us.

UNIDENTIFIED FEMALE: So let me tell you a little bit about NCSG and who—so basically, we work on human rights and we work on monitoring and contributing to ICANN policies in a way that there are human rights respecting. Also, they also uphold their rights to freedom of expression, to privacy of domain name registrants, but also, we care a lot about access of the users to the DNS and the internet. And kind of try to facilitate and see what can potentially, what sort of policies could potentially hamper their access.

And it's purely non-commercial, so by non-commercial we mean academics and day-to-day users. Our members are from—they can be not-for-profit technical operators, they can be academics. We have digital rights advocates like EFF and a few others that—we have 500 members, I think. And we also have individuals who, in their individual capacity, bring a wealth of technical expertise, as well as human rights and other expertise. And the mission, as I said, it's about protecting the rights of the users as it comes to the policies that are made at ICANN.

JOHAN HELSINGIUS: And if I can add on a purely practical level, while we participate in all the working groups, and just like everybody else here, where it really comes to policy, we have six councilors on the GNSO Council, so we actually very actively work on the processes there.

RAM MOHAN: May I ask a question? What does success look like for you?

JOHAN HELSINGIUS: An open and free internet, I think, would be the simple answer.

RAM MOHAN: Go ahead, Steve.

STEVE CROCKER: I was expecting some mention of protection of privacy in addition to that.

JOHAN HELSINGIUS: Sorry, that was assumed.

FARZANEH BADII: Farzan is speaking. For us, success does not only lie around data and how much data we have on for example, DNS abuse that is being mitigated, or if there is a decrease or increase. For us, we also care about qualitative indicators. We care about fair processes in place. We care about processes that can facilitate respecting the privacy and freedom

of speech and a host of other human rights for the users and registrants, as well as not hampering their access to that system.

So when it comes, for example, to DNS abuse, for us, success as well as having quantitative measures, is having a due process system in place, having a fair process in place. But that is a really big question for ultimately, we want to see that in our conversations around the corner, in every corner of ICANN, when we consider DNS abuse mitigation or other issues. For example, when we are talking about the public interest commitment at the new gTLDs. We want to hear that there is an emphasis on protecting the rights of the users as well.

RAM MOHAN: Thank you. That's very helpful. Just specifically to DNS abuse, we've been engaged on this topic for quite a long time. We've advocated for several measures. Some of them have been adopted inside of this community. There is some that is going to happen in just another month's time or so. So there's, I think, strong alignment on getting to an ecosystem where there is less abuse, or when abuse happens, mitigation is a rapid-fire thing. Peter?

PETER THOMASSEN: Yeah. So multiple times now, the goals related to end users have been mentioned. And then there's also ALAC, dealing with that. And I guess, you get the question a lot, but I would be interested in how that overlaps or doesn't, and how you collaborate, and yeah?

JOHAN HELSINGIUS: Structurally, of course, the big difference is that ALAC is an advisory organization, whereas we actually are part of a policy-making organization. That is the fundamental change. I think, for ALAC, some of our values are important, but they are just a small part of their total package. They also care a lot more about things like building infrastructure in underserved areas and so on. So while there is some overlap, we are, in a sense much more specialized than they are.

RAM MOHAN: Yeah. So I have you, and then we have a queue online as well. So let's go to you first, Caleb, and then we'll do the rest of the queue.

CALEB OGUNDELE: Yeah. Just to add to what my chair said, one of the things you will notice about ALAC is that it has a mix of both businesses and business end users, and as well as not-for-profit. So it's a mix of all, everyone in the market, right, that they have. But we have a more narrow focus in NCSG. We are strictly non-commercial, and that's where that name is really from, right? So I'm just trying to differentiate, add some differentiation in addition to what Julf has mentioned, so yeah.

RAM MOHAN: Thank you so much. Stephanie?

STEPHANIE PERRIN: Thanks very much. Stephanie Perrin for the record. I'm kind of back at what success is in our terms, and I just wanted to bring up that the broader concepts of civil liberties are included. Yes, privacy. Yes, free

speech. Yes, due process. But we're also kind of clinging to the vision of openness and development and autonomy and individual rights that were around in the early days of the Internet and sometimes are forgotten these days as we need to move to intermediaries running the Internet for us. And I think that's also a differentiation of the NCSG. Thank you.

RAM MOHAN: Thanks, Stephanie. Julf, I'm wondering whether we—and I want to open the floor to other SSAC members who may have some questions, and then perhaps then we move to the SSAC part of the agenda. SSAC members online and here in the room, any other questions you'd like to have? Greg?

GREG AARON: Hi, Greg Aaron. Not so much a question as a reflection, which is over the last five years, the ICANN community has learned a lot more about privacy. GDPR helped us go there. And SSAC's been trying to keep up with all of that because we want to keep ahead of the topics. And in some of our papers, we've been talking about balance. With GDPR, how do we figure out what legitimate interests are and so forth? And it's also now reflected in our membership. We have people like Matthias and Tara who are privacy experts and have some experience with law enforcement and in the private sector, and they're helping us understand these issues.

JOHAN HELSINGIUS: Thank you, and that's something we definitely appreciate.

RAM MOHAN: Anyone else? All right. Let's shift to the SSAC part of it. The first slide is—Benedict, do you have a question? You do not, okay. Benedict just joined us. He is a member of the SSAC.

So I wanted to share a little bit of the changes, some of the changes that the SSAC itself is going through. First thing is, we've switched now to a default open mode. All our sessions are open unless otherwise marked, and part of it comes from our own realization that a good majority of what we're talking about isn't actually secret stuff. It's stuff that people ought to be able to listen to, especially as we're deliberating on matters.

We had a session yesterday named collision. That was open for folks to listen into. I think those are—so that's one of the things that we're doing. A couple of things that we're doing here at this meeting, on Thursday, we're hosting an informal drop-in session. SSAC members will be there and you can just come directly, connect, have a conversation. So that's one of the things that we're doing.

And tomorrow in the morning, there's a public information session, and we're orienting that session to be not just about, here are all the things that SSAC is doing, and here is all the effort that we're spending on things, but we're trying to orient the session towards, here are our thoughts and our insights on these sets of topics, and then engage some level of interaction. So we're going in that direction. So that's the first thing. On the next slide, I'll pass that to Tara because she is leading something really crucial and important for what we're doing.

TARA WHALEN: Thank you, Ram. So this is Tara Whalen. We hope that because the sessions are now more open, people are seeing interesting work coming out of the SSAC, and they're excited about becoming a part of it. So we have goals for improving the diversity of SSAC, and also, sustaining our membership and keeping it vital and alive. One of the areas of diversity that we are looking at specifically is geographic diversity, our membership. We are represented all over, but we do have a strong bias toward North America and Europe. We'd like to see more representation from Africa, Latin America, Asia Pacific.

We're hoping with, again, some of our connections with people in the other constituencies. We're talking with SOACs and having similar discussions that this will bring more people from different areas around the world because we would really welcome that. And we'd also welcome ideas about places we can do outreach, by the way. We also, despite people saying they're recovering academics, I heard that over there, we are a little bit short of—we're very welcoming of the people who have the research and analysis skills because of the nature of a lot of the work we do, data analysis, we've got measurement, large-scale studies.

We love having people who can bring these skills in and we've had some great new members from these areas. It's also wonderful to have people often who are perhaps graduate students or at that stage of their career where they have finally honed expertise, even if they're not at the very end of their career, we sort of want to make sure that everyone doesn't retire at the same time and disappear out of SSAC. And we're very grateful to have, again, this new blood infused. And you may think you have more to offer. You don't have to wait until you're 55 plus to

have something to offer us. So please step up if you have valuable skills. We'd like to hear from you.

I'd be remiss if I didn't mention gender balance. I would definitely appreciate if we had a few more women up here. So that's another area of diversity that we'd really like to see. If I could get the next slide, please. So we took some words out of our skills survey. We have areas where we ask people about the sort of skills and expertise that they have. And you can see there's quite a breadth of things there. Of course, we're at core a technical group, but this is a pretty broad umbrella of the sort of things that you can contribute to a lot of the documents that we do.

If you look at what we write, you can sort of see the variety of topics and you may see yourself reflected in there or something where you feel your voice would have made the document better. You could have added to it. You could have contributed in a different way. You could have noticed a gap. And if we'd done the document with you, you would have contributed and we would love to know about things like that. And it could be hands-on experience, it could be theoretical experience, it could be any number of these pieces where you could make our documents better. So again, think about how you might want to contribute.

And the last slide will kind of tell you why you might want to do that. Of course, we're just a lovely group of people to work with, but mostly, what you would get from this is, of course, if you're interested in deep problems around security and stability and you want to make the Internet better, then this feels like the kind of group where you're very

eager to improve the Internet. This will allow you to apply and sort of deepen your expertise in these areas and definitely good for career enhancing for people who are looking for that sort of thing in terms of sort of visibility, maturing of your roles, give you some opportunity to advise the board and the DNS community and just making things better and moving things toward a positive direction and building a safer open Internet.

So if you want to do that, we now have an online process for doing that, thanks to our staff. And we have a link there to show you how to apply. And again, very happy to talk to you as well. You don't just have to do it online. You can actually talk to our friendly faces while we're here and we are happy to connect with you. So we hope to see some people in your communities applying for SSAC, so thank you and we'll take questions as well.

RAM MOHAN: It doesn't look like I can click your website on the PowerPoint here. So what's really the website?

TARA WHALEN: So the application is in the SSAC component of the ICANN's website. The area that we're on in there. So thank you for noting that.

JOHAN HELSINGIUS: So after hearing that presentation, darn. If I wasn't above 55, white male from Europe, I would apply.

TARA WHALEN: You're still allowed to apply even if you fill that demographic. You're still welcome. It's balance we're after. And online, we have Tomslin. Oh, behind me. Oh, I see. Please come to the mic.

TOMSLIN SAMME-NLAR: Yeah, Tomslin. I just want to—if you could explain a bit more about what you mean by security, because I'm just trying to understand what do you intend to serve as technical security expertise or even someone with just policy research expertise in security matters?

TARA WHALEN: I think it will very much vary on the context of what's being worked on because our own expertise is not as deeply policy driven as some people. It does tend to be very much more, I think, about—again, it could be just because you've learned issues of operational issues at scale, for example. Again, it could be, but mostly it's that the core of it is something that involves—I know I realize that the technical policy split can be really difficult to, I guess, define as a person who keeps a foot on both sides of that division. Even I sometimes say at what point does the stack split between technical and policy.

I guess it means you don't necessarily have to be a computer scientist to have some ability to engage with technical material. It's just that it will have its roots somewhere, I would say, in technical aspects. So the maturity level should be that you should be able to, I guess, comprehend and engage and understand the implications of where the policy and the technical meet. And then the depth of that, again, the

requirements of that may vary and you have technical experts around you. But there should at least be some overlap between the two, which can be a little hard to do a priori. And we have someone else who wants to step in over me. Thank you.

BARRY LEIBA: Yeah, this is Barry Leiba. I'll point out that there are—that one aspect of security is trust models and how people trust things and the perception you have of how things are working, whether it's a safe and secure internet. The safety part is a lot of that too. And so, there are aspects of that that are non-technical, that we need the views of people on how the system works and how it's perceived by end users. So yes, there are certainly non-technical aspects to security as well.

MATTHIAS HUDOBNIK: Tara, can I just say something? Matthias Hudobnik, speaking. I just wanted to say also for the people which are interested like risk management, information system management skills are, I think, also very important per se, not purely technical, let's say. So if you have expertise in this field, I think it also could be very interesting.

UNIDENTIFIED FEMALE: I know we are kind of over time, but I see this word safer. And so, are we working like we have with this word safe has a lot of connotation. And also, law enforcement uses it in one way. Trust and safety field uses it in another way. It kind of moves away from that objective security and operational aspect and moves into more social and legal field that could have many implications and like doing content moderation that

we don't want ICANN to do or stuff like that. So I just wanted to know if you have a good definition for this safe.

RAM MOHAN: To be honest, it's something that we use interchangeably, but thank you for bringing the deeper seriousness of using those words. We will go back and consider it. Our charter or our remit is about security and stability, not about safety.

ROD RASMUSSEN: Yeah. Let me just put a point on that as well. I think in the context we look at, it's that this is more of a trust thing, right? That what you are going to try to connect to is what you want to connect to in that perspective. But there is an element as well when you start thinking about things like DNS abuse that brings in more of that safety aspect. But I believe that the genesis of this was much more around the reliability and integrity side of the equation.

JUAN MANUEL ROJAS: Sorry. This is Juan for the record. Just a following question. It's about, okay, you are talking about security, right? But how is security and safety is not related? That's my question in this case.

RAM MOHAN: Just paraphrasing it, you are asking how is security and safety interrelated? Was that the question?

JUAN MANUEL ROJAS: My question is, why are you saying that it could be not related.

RAM MOHAN: I didn't say they're not related. I said that our remit, our charter, if you will, is about security and stability. Our remit is not about safety, right? So I'm just being quite precise about it. And as Farzaneh said, safety has sometimes larger connotations and we are focused on a fairly clear and narrow path. And most of our advice is on security and stability matters. Some of that advice, if implemented, could actually build a safer open internet, right? But our focus is on providing clear advice on security and stability topics. Jacques?

JACQUES LATOUR: Well, I was going to mention that at the DNSSEC workshop, for example, where we have a new topic, which is digital trust, and in there is building a new fabric, a new layer on the internet to build trust between the different entities. And that should bring a safer internet eventually. But we're part of the solution. It's bigger than just a SSAC.

RAM MOHAN: Thank you, Julf. I think we are out of time. But this has been a great start to this interaction. I hope we have more of this. I think our respective staffs will be able to get together. We welcome this kind of interaction and we look for you to come and your members to come attend our other sessions. Come, talk to us on Thursday. And for those of you who have a keen interest in kind of being further involved in the matters that we look at, please talk to our folks and, you know, apply.

JOHAN HELSINGIUS: Thank you. We will definitely take you up on that and we really appreciate being invited here. Thank you.

RAM MOHAN: Thank you. That wraps this part of the conversation. You're welcome to stay if you like. You're welcome to leave. It's your call. We have a second part of our conversation of this meeting, which, Joe, I think you're on first. And after you comes John. Over to you, Joe.

JOE ABLEY: Very good. Thank you. A couple of weeks ago, a very particularly rainy morning in Amsterdam, I went and had breakfast with Jim Cowie, who is a guy who founded Renesys a long time ago, which was curator of an awful lot of active measurement data and passive routing data across the internet. He built an entire company out of it. The company was subsequently sold to Dyn and then it was sold to Oracle and then the products that related to it were all turned off and then the data was deleted.

And this made Jim, who had spent quite a long time collecting and curating this data, quite sad. And apart from Jim's personal sadness, which, of course, is not nothing, I think Jim thinks there's an opportunity here to consider what happens to these things. How do we record? What is the fossil record of the internet? How do you look at measurements that happened in the past and preserve them in a way that they can still be trusted and they can still be interpreted and still teach us something?

And this got me thinking. That he was in Amsterdam because he was off talking to the RIPE NCC because they have a lot of routing data and he was interested in talking to them about how they can preserve it. But since this meeting was coming up, I thought, we also have data sets. So what we used to call the IANA and various parts of ICANN org do maintain data sets. That's the core function for a lot of what ICANN does. The root zone is an easy example that we can recognize because it intersects with an awful lot of policy with the IETF.

Uses the IANA to record all kinds of protocol assignments. We have the numbers. Let's not forget the numbers. We have the gTLDs, the zone contents of the TLDs are stored. They are made available for a certain degree of accessibility and usefulness through CZDS. And there are other operational metrics to do with things like L-Root and other type of metrics that are produced. And sometimes these things are shown with historical context. Sometimes they're not. and other type metrics that are produced and sometimes these things are shown with historical context. Sometimes they're not. Sometimes they're easy to cite. Often, they're not. So, if we skip to the next slide.

Often these things are really just a point in time. CZDS gives you excellent record to the last zone that was collected by CZDS but if you want a history of zones, if you want to see what's happened in the Comm zone over the course of the past five years. You probably ought to have been collecting that data every day for the last five years. And if you weren't, you had to look or you have to wait five years before you can answer that question with any kind of data and maybe that's not ideal.

So I know lots of people, including some here, who collected CZDS data, just to give an example and I dread to think how much of glacier cold storage is containing duplicate compressed versions of the Comm zone but I think it's a lot of it. Yeah, yeah. Several hundred dollars, which otherwise could be spent on something like more vegan options in the break or something. But I think the thing here is not simply that the data is hard to collate, because often it takes real time to collect data. That the history is not available. It's also a question of, how do you judge the authenticity of this data?

If you're going to cite this data and try to determine trends over all of this data, how do you know that the data you're using is authentic? Because you have no authoritative copy of it. So this sounds like an interesting topic. And this is not completely obvious. This is not a proposal of any kind. This is me hand waving and throwing observations at the wall and hoping that some of them stick. So next one.

So, I was wondering, I don't think it's as strong as thinking anything, but I think there is actually a stability and security angle to all of this. So I think understanding history is part of how you know whether things are getting better or worse. And I think, if you don't have a reliable data source about what happened in the past, it's very difficult to say which direction you're heading. So I think maybe there's an angle here where SSAC could say something, but like I said, this is not a proposal. Next one.

So one of the things that Jim is working on is he has the idea that if you take really valuable data sets, part of the historical record of the

internet in a technical sense, like that's routing data or [inaudible] data, it's trace routes, latency data, active measurements across the internet, something, and he wanted that to be preserved, there is a body of expertise in how you preserve things, but a lot of that expertise is based around objects that are literally objects.

If they exist in a museum or a particular first edition of a book exists in a library or something and this is a slightly different problem because we can make unlimited copies of these things, but then we have a different kind of problem because it's not that we have one authoritative original copy, not one particular Grecian urn or something, but there's value in replicating this stuff. But that implies the problem of knowing when it's authentic. So knowing how you can trust it.

And also, being able to measure how much this data is replicated, so how safe it is? How do you know when one piece of data is about to disappear from the world and needs to be copied by other people versus others that might already have lots of copies? This has other interesting parallels with things like peer-to-peer networks, which I think are interesting. And there was a Wired article I remember years and years ago about the lost years of Usenet, where Google was building Google Groups and they were missing 10 years of history, and it turns out to have been on a tape in the zoology department at the University of Toronto because that person happened to have backed it up because they were enthusiast and various people flew to Canada in order to collect the tapes [inaudible] business because apparently the internet was not, I don't know, reliable or something. I don't know why they flew there. I think Wired sometimes just likes to just be dramatic.

So anyway, I think maybe there is a role for ICANN here, partly because some of these data sets that ICANN maintains are valuable, in the public interest sort of valuable, not valuable commercially, but also because it's an opportunity to set a good example because ICANN is far from the only source of these kinds of data. And I think lots of people could benefit from some homework being done in a place that can be shared openly and transparently so that other people have a model if they have to do this.

So I think it's interesting. Like Jacques had mentioned before, I think on the weekly call, Steve you brought up this idea of cryptographically proving transformations of data sets or even images to take a use case which is to do with deepfakes and things like that. I think this is the kind of information that could usefully be combined. Jacques is working on how to trust these kinds of [inaudible] projects using DNSSEC, which is another more direct sort of tie-in to what ICANN does.

There's an opportunity I think here to make some stuff available, not just in the technical sense of working out the mechanisms, but also working out why these things should be preserved, the general principles about why these things are worth thinking about, and also, for ICANN specifically, the idea of when does a data set stop becoming a commercially sensitive thing that needs to be protected by access agreements, and when does it become public data because it's part of the historical record.

I don't think anybody's asked that question. I think it might be nice at some point if we could say actually there's a statute of limitations on protecting access to this data and past a certain point perhaps, we

could all agree that this stuff is just available because it should be. I think there might be one more slide or maybe two. The slides I think have been distributed, I think. Could I send them in? And obviously, Kathy's presented them so she has copy. So that's a blog post from Jim Carrey that got me thinking about all of this stuff.

I think I've mainly waved my hands around everything else there already. Is this the last slide or is there one more? Oh, yeah. So, I think it's possible that SSAC could look at this if there's some interest and some energy to write something down. I think we have some people from academia who know about data sets. I think we have a reasonable perspective on this. I think the work that Jack is doing is interesting. I find that interesting anyway. So I don't know whether this is a work party or whether it's something else, but I'm interested in what other people think.

I think there's one of the things that I think is interesting is that final policy question of when does data stop being commercially sensitive and when does it start becoming just part of history, and I think that would—I'm getting a good group understanding of those kinds of things would require a lot of coordination with other groups here. Like the business constituency is going to have a different opinion from our non-commercial friends that we just heard from. And I think, we might be in a good position to be able to have those conversations. And if there is some broad policy possible, then maybe by talking we could actually get there. So, I'm interested to see what anybody else thinks.

RAM MOHAN: Warren is in the queue and Steve and then John.

WARREN KUMARI: I'm assuming that you looked at opendata.icann.org, yes? Which has, like icann.org makes registry reports available, blah, blah, blah and the open data license.

JOE ABLEY: So it does. That's still a single point of data being produced. It's not answering the question of how you preserve this stuff by encouraging it to be replicated. It's not talking about authenticity of the data and mostly it's summary data. So I think it's related, but I don't think it's the whole answer. Steve?

STEVE CROCKER: Quick question, and then I have several points. On this last one, were you pointing out storing things in multiple places so that there were a higher chance of being around when you want them? Okay, so that's part of the sort of the craftsmanship of once you've decided what the data is, how do you make sure that it's going to be there over a long period of time? I have not heard of this effort, and we've had no interaction. So for the benefit of everybody, everything I'm going to say is off the top of my head response to what you're saying.

I like it a lot. I see several different issues. There are security-relevant aspects of wanting to preserve the data, but there are many other aspects. So, one thing that occurs to me is that if SSAC is going to get involved, which we might well want to look for partners or allies who are interested in preserving data for their various reasons, and when it

comes to trying to explain why we're interested, we have a much broader case to make. That's one thing.

Second is a small anecdote. As you alluded, keeping copies of the root is a cottage industry, which there are some untold number of things. We had one in our office. I had it printed out or organized by TLD, and it came to pass that I found myself in a conversation with a senior official in the US government in the State Department who was countering a rumor that Syria had been taken out of the net, I think.

And I said, well, no, because here's the data, and it's been recorded every day, and you can check this data six ways from Sunday. And he thanked me very much. I have no idea whether that was helpful, but that particular rumor had been circulating, and I had encountered it more than once. So there's a specific instance in which it was helpful to have that data.

RAM MOHAN: Steve, I'm going to ask you to have the rest of your comments offline because there's a queue behind you.

STEVE CROCKER: No, no. Okay, I'll just say one quick thing. In addition to everything you want, what about the registration data and all the [inaudible]?

RAM MOHAN: Thanks, Steve. John?

JOHN LEVINE: Well, as we alluded, I'm one of the people keeping daily snapshots of the zone files because you never know when they might turn out to be useful. And I have sort of a thought and a half. One is, I am always extremely sensitive to mission creep. I mean, I do not think that ICANN should do this. I do think it might make sense for ICANN to enable this and look for existing organizations that do—maybe the internet archive, maybe the Computer History Museum, who I've had, back when I was doing the ITF stuff, did some archiving stuff.

And I also think that some of the commercial—when your stuff become commercially sensitive, is not that hard. I mean, ICANN already has a precedent that the registry reports are embargoed for three months, and then you can download them without asking. So that we might open. But I think this is worth looking at, but I also think that before we get too far, we should find out, are there organizations that are interested in talking to us? Because simply saying someone should do this. We've already done that.

RAM MOHAN: Thanks, John. John, an unrelated separate scheduling question for you. We have only 11 minutes left in the session. There's a long queue here, so we may end up sacrificing your session if we continue with the queue, or I could say we draw the line here and switch to yours. Okay.

JOHN LEVINE: Let's do mine and then people can talk about whatever they want to after that.

RAM MOHAN: Okay. So let's go to you now. Let's make sure that content gets through, and then if we have extra time, we'll come back for you, Joe. Back to you, John.

JOHN LEVINE: I think Kathy has my slides. Well, here they are. So I presume everybody here has heard about KeyTrap, which is the DNS disaster of the month. So if we can look at the next slide, please. So in case you don't know what KeyTrap is, this is an extremely oversimplified version of how you check a DNSSEC signed set of records. I mean, there's a DS record that points probably indirectly to a DNS key record, which is in your zone, and the DNS key record has a bunch of parameters.

And then for whatever name you're looking up, you're looking up mydomainexample.com, spelled wrong, which is a real name and then there's an RRSIG signature record, so you can check that the signature in the signature record matches the key. And if it does, then the signature record has vouched for those A records. To make things easier, you can have multiple DNSKEYs, so that checking the RRSIGs faster, there's a 16-bit key, which is 18-1-1-1 in this particular one and the key ID is actually a 16-bit check [inaudible] of the DNSKEY. So that before you do any checking, you figure out what the ID is for each DNSKEY are, and you only check the ones where the ID matches.

However, the interesting fact that nobody noticed is that it's actually a checksum, not a hash. It's a pretty lame checksum. It's a 16-bit one's complement check sum, so it's easy to force it to particular values. So next slide, please. So the problem with KeyTrap is that it allows you to—if you were a malicious person, you could put in 100 DNSKEYs all with

the same ID, and then you could have 100 signatures all with the same ID, so that in order to see whether the signature was valid, you needed to check all 100 signatures against all 100 keys, which is 10,000 checks, which turns out to be pretty slow.

Even though it had been possible for a decade, nobody had actually done it until a few months ago, and they said, whoa, that will hang, bind, and stuff. So everybody went to fix it. So there were some extremely complicated solutions to this, and I started wondering, how common actually are these duplicate IDs? So since I have all these zone files that we just talked about, I looked at every zone file with more than a million entries. And I also looked at .se and .nu because they're available and they're fairly heavily signed.

So I looked at 200 million domain names, and I just went and counted. About 8.5 million are signed. So I fetched all the DNSKEYs and I fetched all the IDs and I figured, how many zones actually have two keys with the same ID? The answer is, next slide, 107 out of 8.5 million, yeah. And so, of those 107 collisions, 86 of them are a key signing key and a zone signing key that just happened by accident to have the same ID, which I'm pretty sure is an accident. And I actually found about 20 pairs of zone signing keys where there were two zone signing keys with the same ID. And I found one pair of KS keys.

And if we flip to the next slide, here they are, and they're all sort of like random things, and I'm sure they're accidents. The only place I found a pair of KSKs was in .nu, which I presume they were in the process of rolling it over. And then the other key thing I found out is I counted how many zones there are with three colliding key IDs, and the answer was

none. So the way you fix this problem is, first it hardly ever occurs, and if you simply stop after two or three, I mean, you will validate all the real stuff and you won't lose anything real.

So I think this is the last. I think that's what the next slide says, yeah. There's a handful of collisions. There's none of three or more. You don't need any fancy mitigation. So you simply add this to the list of things that you check before you do a whole lot of work. The end. There was no easy way to count the RR sigs, but I have no reason to believe the answer would be any different.

RAM MOHAN: Warren has his hand raised.

WARREN KUMARI: So I mean, the bit that maybe you glossed over is that much of this was you could intentionally make a zone which causes resolvers to fall over with one query per second.

JOHN LEVINE: Right.

WARREN KUMARI: Yeah. You don't need fancy mitigation. Just stop when you see a certain number.

JOHN LEVINE: Three.

WARREN KUMARI: A certain number.

JOHN LEVINE: Yeah, yeah.

WARREN KUMARI: That's what people largely ended up implementing, but the spec said you must try all. So this is—if you implement according to the spec, bad stuff happens. Once people realized this was a problem, they fixed it by capping the workload to [inaudible].

JOHN LEVINE: Yeah. Although I think you can argue that the spec has always said—I mean, we know lots of ways where if you don't cap the workload, bad things will happen. Like people cap the number of CNAMEs they will chase. So I think that on the one hand, the spec says you have to check all this stuff, but the spec also says, you're allowed to stop when things get stupid. And I think this—well, go back. I think it's in 1034. In practice, I mean, ISC had a very nice page that—here's all these places where we put in work limits. We simply, okay, we had eight of them, and now we have nine.

WARREN KUMARI: Yeah.

JOHN LEVINE: Yeah. But my point here is that it's turned out that this is a totally contrived thing. This is not like, oh, we have to be careful because we might break something that's—

WARREN KUMARI: Well, I mean, it's contrived in that—if you make a contrived zone file, one query per second makes a huge machine disappear.

JOHN LEVINE: Oh, absolutely. But I'm saying, with CNAMEs, originally CNAMES were just supposed to be a temporary shortcut. So you would never follow more than one or two, but now we have CDNs and stuff. So in fact, it makes sense to follow four or five. The limit has to be more than one. But I'm saying in this case, if the limit were four, I think it is unlikely that you would ever go past that.

RAM MOHAN: Thanks, John. Did you have your hand for this, Maarten? No, okay. So any other questions for John? All right. We have about five minutes. Shall we go back to Joe's piece? And if I review the queue, I think, John, you had just gone. So Maarten is next.

MAARTEN AERTSEN: Yeah. So real quick comment. It occurred to me while you were talking that you were not the only one talking about longitudinal data sets, but that has been coming up a lot in NCAP too. So given that there's a reason to do longitudinal data sets or additional longitudinal data sets in that work, this may be nice timing for the pitch you're doing.

RAM MOHAN: Thanks. Farzaneh? Farzaneh left, okay. Greg? Oh, okay. Got it. Next, Jaap? You put your comments in the chat and just to read out what Jaap wrote, I wanted to point out that there is a wonderful book about libraries and how they play in shaping in the world. It's Papyrus, by Irene Vallejo, and I suggest to Google. So for those who are not on the chat. Warren?

WARREN KUMARI: Yeah. I mean, I guess, I'll just mention again, ICANN's opendata.icann.org says, let us know what other stuff we should archive. They archive stuff you ask. They do it longitudinally. They give you an API to suck it all down. It's under the open data agreement. So if anybody wants it, they can suck it down and stick it in a repository. They don't currently have the root zone, as far as I know. But they could just add that.

There's currently 36 or 37 data sets in there already. So I think this is more just, ask ICANN to add another one and then ask somebody to make a copy. I think Google might actually even be already copying this into the open data repository, whatever they call it. Or maybe not because they don't care.

JOHN LEVINE: I'll check it out. Maybe I didn't look closely enough. I think there are elements there that are not quite as broad as I was imagining.

WARREN KUMARI: We should definitely improve it. It's also, if you try and use it before logging in, it looks completely empty, which is an awful UI. So you have to log on and then suddenly they're like, oh, now there's actually data. Yeah, sorry.

JOHN LEVINE: Yeah, but I think a technical part of this is the decentralization of the archive function and trying to make sure that you have a measure of how decentralized particular data sets are as well as a confidence in their authenticity. So, I think, yeah, I will check it again. Maybe that's already answered.

WARREN KUMARI: And it is under the open data agreement, which lets anybody make copies of it as well and as long as you don't call it a derivative work or [inaudible].

RAM MOHAN: Thanks, Warren. Gautam, you have the last comment.

GAUTAM AKIWATE: So I had two comments. The first was John Kristoff during Hamburg had a presentation. I think the ICANN org, like the staff, compiled a list of all of the data sets and the different things. So that might be also a resource that we want to consider. And I think there was some interest by John to continue looking at it. The second was, I think, to your point about how does ICANN facilitate? One of the things we've noticed, because we also archive the zone files daily, and we have been doing

that for the last 10 years and that's part of where the register name management work party would happen, because like we could see 10 years of data.

The reasons that it was possible, and we were able to archive 10 years of data without killing ourselves, was dropping information that we thought was not necessary, and sort of summarizing it. So maybe one aspect that we can think about is how do we best archive it instead of just having a daily [inaudible] that just saves and you have 10 years' worth of data, daily snapshots. What information is not necessary so that we can summarize the information and that way queries are faster, the storage footprint is smaller, and maybe that is something that we can think about is how do we facilitate better storage of this data?

JOHN LEVINE: I mean, I think that is definitely part of our answer. I think there are lots of examples though, where because of summarization assumptions that have been made, there are questions that can't be answered. You can't always predict the future. So I hear what you're saying. I think there is some benefit to looking at raw source data in as raw source format as possible. But even if you do transformations, then having the transformations be verified as being authentic is also important.

RAM MOHAN: All right. Thank you very much. We are almost out of time. There's a minute, and I'll use that to say thank you for this session. Kathy, what's the rest of the schedule for today?

KATHY SCHNITT: So we have Warren up next for the lightning talk and then the registrar name management work party after that.

RAM MOHAN: And we have a 15-minute break in between. So we come back at 4:15 local time to listen to Warren. I will not be in that session, but Tara, you'll be here, right?

TARA WHALEN: Yes. Of course, I'll be here to listen to Warren's session.

WARREN KUMARI: It's not a very good one, I should warn you.

RAM MOHAN: Okay. Thank you, folks. We are adjourned.

[END OF TRANSCRIPTION]