KOBE – SSAC Private Meeting (5 of 5) [C]
Tuesday, March 12, 2019 – 17:00 to 18:30 JST
ICANN64 | Kobe, Japan

ROD RASMUSSEN: All right, you sat down? John, ready to go? All right, quiet down, folks. We got lightning talks, the fun part.

[JOHN LEVINE]: Shazam, this is a lightning talk. Okay, so I always qualify my audience. Who heard my similarly titled talk yesterday? Okay. Well, this is going to be like if you ever used to subscribe to Highlights for children, they would have two pictures that were almost the same and your job is to figure out where they're different.

All right. So as you may have heard, it is almost possible to get copies of all of the contracted zone files if you are diligent and so forth. And so there are 1,200-some zone files available and the total number of names in all those zone files is 193 million – I counted them – of which roughly half are in dot-com and the rest were in other places. And while I was counting them, I figured it's pretty easy to [grip] four lines and start with XN, dash, dash, and it turns there's slightly under 2 million IDNs in the 200 million names, so it's about 1% of the names. Only 40% of the contracted TLDs have any IDNs at all. But once I realized

there's only 2 million names, so in a computer, that's a totally tractable number.

So here's an impossible to read graph about the sizes of the zones. The blue lines are the sizes of the zones. The dark lines are the numbers of IDNs. And mostly, what this tells you is that other than all the way at the left where the largest two zones have the most IDNs, which are, of course, com and net, beyond that is totally random. There are some zones that are almost entirely IDNs and there's dot-voting that has two.

And then we have … And the percentage of IDNs in the zones and you may notice an artifact in the middle of the graph there. There is a large number of zones that are exactly 50% IDN. Okay? Not 49%, not 51%. No, I checked. It turns out the reason is there's a whole bunch of IDN TLDs that are in the root but haven't actually done anything yet. So they have a total of two names. Yeah. There's the TLD, which is an IDN. And there's NIC.TLD which is not an IDN. Okay? So that's that artifact.

So anyway, what did I do is I wrote a little Python script that runs on my server at home and it can go through all of the zone files in an hour and a half, which is nice because it means every time I find a bug in the script, I can just re-run it. And it collects statistics on the IDNs and I also attempt … I have attempted to find out for each IDN that was interesting, of which there are

only a few thousand, I attempted to do WHOIS lookups to find out how old they are and I discovered that even without the … And the GDPR redactions here are totally irrelevant because nobody is redacting the original registration date. It's still nearly impossible because for a large number of, particularly the new TLDs, the WHOIS is rate limited to a point where they might as well just not bother.

Oh yeah, and finally, yeah, this is a couple of Python scripts. If you want to play along at home and you have the zone files, just ask me and I can send them to you. It's basically one script that does all the analysis. There's another script that does the WHOIS stuff and then it puts it all into a MySQL database where you can do whatever you want to do.

So what did I check? First, is the name valid under the old IDNA rules or are they valid under the new IDNA rules? It turns out the new IDNA, even though the new IDNA rules are written in, they attempted to write them in a very precise way, the Python IDN library and the widely-used GNU Libidn 2, interpret them slightly differently. Libidn 2 complained a lot more.

And the last thing I did is every … I believe yeah, at this point, every contracted TLD that accepts IDNs has published label generation rules, like these are the scripts we accept and for each script, these are the characters that are valid for labels in

that script. So basically, I tried to match up every name with a script valid for that TLD.

So yeah, in theory, the TLD is all published at label generation. They all publish their script rule. But the contract specifically lists what languages a TLD can accept and the TLD, according to my reading of the contracts, is supposed to send the file to IANA in this format and then they're supposed to follow the rules.

What actually happens is a lot of TLDs kind of forgot, and particularly, there's a lot of TLDs that have names in Chinese but their contracts do not allow them to register names in Chinese. My understanding is the contract modifications to add languages are totally trivial. This is not because it's hard. It's just because they're lazy.

The other thing is everybody in the IETF knows nobody actually reads the spec and the tables … Yeah. It's like the file is supposed to have the hex code and then either "end of line" or semicolon. How hard is that? It turns out phenomenally hard and people got it wrong in all sorts of ways. The wrongest way they did it was they took these text files and turned them into HTML. Just who the hell knows why? So I had to de-HTML-ify them.

A lot of TLDs have not sent in the script files for all of the languages. Well, a lot of them haven't sent them in for all the

languages they're allowed to do, although there's a fair number of cases where they are allowed to accept languages but they actually don't. And finally, yesterday, somebody said, "Oh, there's this new XML syntax which is being used in the global LGR rule for the root." And he claimed that they're all going to send in nice, new XML files but I haven't seen them. If they do, that would be nice because they're more likely to get them right since they're XML syntax validators.

So what I did is I wrote a parser and then I kept adding regular expressions and stuff so I could try and parse the stuff out properly. And what I did for each file is I simply, I parsed out all the character codes and I turned them into a Python set so it's really easy. In one line of Python, I can say, "Are all the characters in this string in this set, yes or no?" And then I took all the sets for each language because for dot-com, there's dozens of scripts. I merged all the scripts for the TLDs so I can make a pre-pass and say, "Is this string valid in any possible script for this TLD?"

And so this gets pretty close. There's a few context-dependent rules like the Japanese middle dot which we will discuss later, which is only valid in certain contexts.

So anyway, I did the 2003 check and the 2008 check, and then I checked to see if they're valid in a merged set and if they are

valid in the merged set, then I actually ran through every script for the TLD to see whether it's good. That's not quite right because it [inaudible] into the context dependent rules, but it turns out that that was not all that important. The other thing is a lot of TLDs have updated their script files and since I couldn't reliably get WHOIS dates and I couldn't reliably get the old files, it's hypothetically possible that when something was registered, it was valid under the old LGR rules but not under the new LGR rules. Although, that doesn't seem to have happened much, and as far as I can tell, when they update the rules, they tend to increase rather than decrease the number of valid characters.

So then I find … The first thing I found is that the vast majority of the 2 million names are fine. I only ran into a few thousand that had any issues at all. They are valid. They're valid under 2003 and 2008. They match the label generation rules. They're valid under one of the scripts.

So yeah. So I found 509 names that are bad under the old rules and depending on which library you use, either 1,000 or 4,800 names that are bad under the new rules.

Okay. All the names that are invalid under the old rules appear to be recently registered names since ICANN told people to switch to 2008 and the vast majority of them either have the German [S set] which was not valid under the old rules but are

under the new ones, like in dot-berlin and dot-hamburg, there's lots of names with [S sets] because that's how you write German.

Also, 2008 allows you to embed numbers inside Arabic strings whereas the old ones didn't because numbers go left to right and Arabic goes right to left, and they didn't know that that was actually a normal thing to do in Arabic.

So as far as I can tell, the ones that are invalid under 2003 are basically showing that 2003 has been superseded so we're cool there. What's more interesting are the ones that are invalid under the new rules. 3,800 of those are arguably length errors, which I'll get to in a slide or two, and there are 962 other invalid names under IDN 2008, of which nearly all are names registered a long time ago like … Well anyway.

Here's the issue with long labels. We all know that IDNs, there's the A label, which is the ASCII version, XN, [dash, dash] whatever, and there's the U label, which is the Unicode version. There's a one-to-one mapping between the two.

The Punycode is a variable length mapping and it turns out that, and in particular, if you have the same character repeated in your Unicode, it compresses it really well. So there is … That's an actual name in dot-Tokyo and I think it's about, the label is about 60 characters and there's the string in Japanese, which is,

I'm sure you'll really recognize means "We could have nice shoes without killing animals – dot-tokyo." Okay, but if you expand that out into UTFA, it's 69 characters.

Yeah, okay. And so, well, Patrik. Where's Patrik? He bugged out. Okay.

UNIDENTIFIED MALE: He was here a while ago.

[JOHN LEVINE]: Yeah, okay. Well anyway, I talked to Patrik and to John Clemson about this. And in fact, this is … It's a known X in IDN 2008, which is, an X is either bug, feature or thing we couldn't resolve.

And if you are sufficiently perverse, you can have about a four-to-one expansion if it's the same long character repeated over and over again. So that you can construct a DNS name, which the A labels are less than 256, whereas the expanded U labels are 1,000 characters. And although from the point of view of the DNS, this is not a problem. From the point of view of applications that are allocating buffers into which people will type domain names, it probably is a problem. So this is something I'm going to bring up with my Universal Acceptance buddies and I think it would probably be a good idea, just as a best practice, discourage people from registering names that

expand to more than 63 bytes of UTFA because even though it's legal, it's probably not going to work very well.

For the old stuff, there's a lot of names like Section Sex or Euro Bank or AThousandDegrees.com. And again, when we're talking about a few, like 400 names, all the ones I spot-checked, none of them were in an active use. They were all parked or for sale or like that.

The best one is there's two registered in 2014, which is Google.com, and if you look very carefully, there's a little … See that little [doo-zit] under the first G? Okay yeah, that is an underscore something, something. This is valid under IDN 2008 and apparently, it was valid under the label rules that dot-come was using in 2014. And so unfortunately, as far as I can tell, somebody just did this to prove that he could. But there is still room to leak through evilness when you're registered IDNs even if you're following the script rules.

Okay, what I have found though is that particularly in the new TLDs, the sloppiness just is overwhelming. Dot-club and Dot-art have all these Chinese names even though they're not allowed to register Chinese.

The third most popular IDN domain is this one, which I gather means dot-web in Chineses. But yeah, the largest number of IDNs, there's about 900-some thousand in dot-com. There's

slightly over 200,000 in dot-net and then third one is this Chinese thing which has about 200,000 names total. But it also has 200,000 ASCII names. You know? If you look at the contract, it doesn't say ASCII. They're only allowed to register Chinese. So I don't know what's going on there.

And finally, even in dot-Tokyo, that middle dot is legal but only in a context where it is next to a Japanese character because that dot is used to separate, Typically, when you have katakana, transliterations of foreign names. Like if it's Bill Gates dot … Bill Gates in Japanese would be Billu-dot-Gates. But anyway, Taylor.Swift.tokyo is an actual registered name that resolves to a website somewhere whereas I'm not too worried about Taylor Swift getting phished, but if it ere PayPal or Middle.com, Dot-Tokyo or dot-something, that would not be great.

So again, they shouldn't have registered that. It's not valid under IDN 2008. They're not doing the validation they're supposed to. We also have things like this. Yes, that Japanese. You'll notice there the Punycode ends with As because this got supercompressed and that character means top or above, so this name basically means "Top, top, top, top, top." Yeah.

However, if you do a WHOIS on it, it says it's reserved and then if you do a host lookup, it's on the DNS. And so my understanding is that if a name is reserved, it's not supposed to be delegated,

right? So anyway, so again, we have a certain amount of our new TLDs not following the rules.

So here's my summary, the vast majority of IDNs are fine. There's a long tail of old junk which as far as I can tell is harmless and the new TLDs are not following the rules. Or let me say some new TLDs are not following the rules. And pressure does work because I found five names in dot-Asia that were completely invalid under all of the IDN rules and I mentioned it to a few people I know there and they said, "Oh yeah, those names. They're gone." Apparently they were tests from ten years ago and they forgot to delete them.

And so nothing that I've done is particularly difficult. I took the names and I gripped them out and I just ran them under the standard mechanical validators that the domains should have been using all along, but some of them haven't. So anyway, that's what I did and if you would like to play with this, the scripts I use, both to analyze it and to download the script files, are yours for the taking. Just ask.

Questions? Yes.

UNIDENTIFIED MALE: So do you think it would be useful to give this to ICANN Compliance and say you guys ought to clean this up?

[JOHN LEVINE]: Well, I've already given it to OCTO so they can play with it. But my guess is if we gave it to compliance, it's like I would rather help ICANN establish a process so that they can audit every IDN every day but at trivial cost. I'd rather you help them build a process rather than give them a hit list of what's bad today. Yeah?

[ANDREI KOLESNIKOV]: I have two questions. First, do you have access to [fast] track ccTLDs zone files to run the scripts? Or they don't provide it. Nobody provides it.

[JOHN LEVINE]: The only ccTLDS I have access to are dot-US that has no IDNs and dot-SC and dot-NU, which since they are run by Swedes are perfectly clean.

[ANDREI KOLESNIKOV]: Okay. The second question is this is because we launched this dot-RF a long time ago and the EPP entrance, we had the IDN table which every string you want to register goes through the normalization and checked against the table. How come these

domains can be registered and they don't have a check? That's a simple question.

[JOHN LEVINE]: I don't know. If I were writing their code, I wouldn't have that problem.

ANDREI KOLESNIKOV: You physically cannot register inappropriate domains if you run through the checks.

[JOHN LEVINE]: One thing I haven't done is to try to correlate who's got funky names versus who their back end is. My assumption is that some back ends check and some back ends don't. And it's also possible, in dot-Tokyo, there's four bad dot-things, so I wonder whether somebody knew somebody who somehow let them evade the rules. We're not talking about large numbers here. I'm talking about four.

Rod?

ROD RASMUSSEN: So on the ccTLD question, you reached out to Farsight or somebody with passive DNS where you could actually get a quasi zone, so to speak, to take a look at that?

[JOHN LEVINE]: I have. Paul's given me a password to Farsight's thing. I certainly could ask him to see. Could we flip it around and take a look? My guess is that most of the ones that are invalid were registered to be cute and they are very rarely resolved so I wouldn't count on Farsight actually having … These scripts are not very complicated. If I can get the data, I can run stuff through them and see what we find. It's a reasonable thought. It's a reasonable thought keeping in mind I'm doing this for free, so this is like a three Sunday afternoon project. Let's not just see the same hands. Anybody else? Okay.

[ANDREI KOLESNIKOV]: Regarding the passive DNS, after the last ICANN meeting, we actually did some research and found out that all this passive DNS strange names are the results of typos and switching between, for example, Chinese and English or especially Russian and English. So you end up in the passive DNS, with really weird names, but it's not intended. It's just switching of the keyword layout. It's crazy.

[JOHN LEVINE]: So just to respond to that, if the passive DNS is being monitored for answers, you shouldn't get that unless they're wildcarding.

[ANDREI KOLESNIKOV]: No answers. This domain doesn't exist. You just log all the DNS information, DNS queries.

[JOHN LEVINE]: Yeah. Oh, that's different. Anything else?

JAAP AKKERHUIS: This language table, do you use the IANA language table for this?

[JOHN LEVINE]: Yeah. I scraped these all off the IANA website.

JAAP AKKERHUIS: Yeah, a bit my fault. Just a bit. But IANA, it's kind of weird that IANA actually now states that these are not really binding to anything and for information only. And that was [inaudible] IDN. When IDN popped up and Polish registries started to cough at the whole world, the Polish guys are asking me, how can we tell people how to do that, what we support and not? And this is a bit of history. They said, "Well, you know, maybe IANA can help you out here."

And when [inaudible] came, the head of IANA, I happened to talk to him. He said, "Use this ID," and that was the first thing he did

by IANA by doing these tables without any thought. And so that's how this started.

And it started out as ASCII only in any format you could do, and only very recently Kim Davies put in the XML specs. But since it's not required or doubly checked anyway if it is required, we made such a mess. It's the mess it is.

[JOHN LEVINE]: Yeah. It's true. It is. Actually, I have one more. Yes, while I'm pointing fingers here, I was looking through dot-mobi which is … Mobi is an old domain, but their contract was updated in 2017 and the 27 contract, to me, appears to give this list of languages. They have one language table file for Chinese, which is fine. And I went through and took a look. And during … And earlier this year, somebody registered a name which is [Wah Wei] in Chinese, digit 5, ASCII letter G, mobile phone in Chinese. Not a valid domain name because you can combine Chinese with digits, but you can't combine Chinese with ASCII letters. But for some reason, affiliates let it through anyway.

UNIDENTIFIED MALE: And probably it should because that's …

[JOHN LEVINE]: No. It's semantically reasonable, but again, they have this character table on file and if they want to register something. I looked in the character table. It's got all this Chinese and it's got digits but it doesn't have letters. If they want to allow letters, they should update the table to say that they're going to allow letters. This is not super-hard.

UNIDENTIFIED MALE: Okay, so you're talking about a missing table update. Fine. We'll take care of that.

[JOHN LEVINE]: Yeah. Well, I can't tell whether it's a missing table update or it's a weak validation.

UNIDENTIFIED MALE: But that's the problem with these somewhat arbitrary rules is there isn't a way to say 5G in Chinese. There isn't. So you have to put 5G.

[JOHN LEVINE]: I believe we have a longstanding consensus here that not every string that you might like to register as a domain name should be registered as a domain name.

UNIDENTIFIED MALE: We could argue about it for weeks. Never mind.

[JOHN LEVINE]: I'm just saying that … I'm sort of being … I'm being totally [fanatic]. I said it's like you said here are the characters you're going to allow and here are the names where you use something else, so at least one of those is wrong. Warren?

WARREN KUMARI: I think what I was going to say is yes, not every string that people want to register should be allowed. However, every string that people want to register equals a dollar, and therefore, seems to be allowed. Right? This is not technically correct, but people get paid and so it works out for them.

[JOHN LEVINE]: Yeah. Well, I understand that and also different. Registries have different priorities and dot-moby is trying not to shrink too much. I don't think it's inherently a stupid name. It just happens to violate the rules that they published. Yeah.

Yeah?

UNIDENTIFIED MALE: When Patrik and I were doing the emoji research, I actually went through all the gTLD contracts. I think other than dot-asia who

there was no contractual language for 2008 where others have it. So I was wondering in your research for dot-asia, did you see that's a problem? The names are valid in 2003 but not 2008?

[JOHN LEVINE]: I saw some, but I believe most of the 2003-only names are pretty old and that the new names tend to be valid under the new rules. Also, there is an incentive to use the new rules because it allows that middle dot, which apparently, people think is really cool. Yeah.

UNIDENTIFIED MALE: And regarding the IDN tables, I think the 7940, the draft that Kim wrote, that's made it into the IDN Guidelines.

[JOHN LEVINE]: Okay.

UNIDENTIFIED MALE: So that's Op Version 4.0. I think the registries are going to implement that soon.

[JOHN LEVINE]: Certainly. If they published nice, tidy XML that I could parse mechanically, I would think that was great. Yeah.

My half-hour is up. All done.

UNIDENTIFIED MALE: Thank you, [John].

ROD RASMUSSEN: Do you have a guest, Andrei?

ANDREI KOLESNIKOV: He might show up.

ROD RASMUSSEN: Okay, so Andrei has invited one guest. So if he shows up, just keep in mind there would be a non-SSAC person in the room.

ANDREI KOLESNIKOV: I'll do a short presentation of the MPA 77. This is actually the assigned, already assigned to Russia, prefix, global prefix. And the guys who is in charge of national domains, dot-RU and dot-RF, they also taking technical function for the handle infrastructure.

Administratively, Rostelecom, which is the largest incumbent operator in Russia is the primary administrator for the 77 and don't assign this global prefix to Rostelecom in September last year. As I said, the technical function, however, will be managed

by the same guys who are managing the Russian domain names, which is kind of cool I think.

Also, what they do, I'm pretty sure some of you, probably most of you know how the handle system operates, how the global registries connect to each other and Russian changes in the global infrastructure, that they have a contract with a donor foundation and they exchange the global data with each other.

And there are currently a few MPAs up and running, but should I say that there is not enough operational information to make a conclusion about how good this architecture works and what are the numbers, how many millions of the handles have been registered. And according to the donor requirements, the contract actually, you can see it online, MPAs must support the multistakeholder model as well as ICANN does, right?

So what's going on right now? So they have a global handle registry that's now up and running. It's already there. Anybody can check it. If you're interested, I'll send the link to where. Prefix already delegated. The current work is to launch the production instance for the GHR and slowly, not fast, of course, move it into the high load cluster depending on, basically, market demand on how many handles will be handled. It's really cool. Handles handled. So what they also do is now developing the MPA policies. There are some interesting discoveries in the

prefix allocation because there are two models of prefix allocations. It can be fully autonomous, so the local registry can delegate the prefixes, the handles, or not delegate them. So there are a bunch of interesting ways. And they will start the local handle registry accreditation.

If to compare with the domain name system, it's basically the registry and the registrars. That's how it works within a certain prefix. That [is] currently planned. I'll speak about it a little bit. And the [test] access to the [dual] IP will be provided. [DOAIP] is actually the protocol which lies behind all the system to make it up and running.

In April, basically next month, the global handle registry will be in production mode. This month, guys is doing with the [PKI] infrastructure because in order to run it in a production mode, or have it publicly available, provider needs the [PKI] infrastructure in order to connect all these things together. There will be a replication test with other MPAs. I think they already have an agreement with a couple of MPAs to have a [predication] test because they all share the same copy.

And the [funny] thing will begin this month and will go through March until May. We planned a couple of interesting tests. First of all, listed the food life cycle test was a production factory and retail, so there will be a number of handles. Of course, there will

be a local handle resolver and a couple of handles printed on, basically, sausages in the store with a QR. You can scan it. But the actual resolver will go through the production handle resolver and direct the requester to their front end as a web server because a lot of companies, I don't know, maybe not only in Russia, they're trying to differentiate the products by providing some proprietary indexes basically on their products so you can scan and see what it is. All of them use a different kind of identification but this will be pure handle-based.

Who do we talk to? We talked to transportation, the infrastructure with [Wacos] Railroads, spare parts, all kind of logistics related to the infrastructure might have an acceptance of the handle as a basic system for the identification.

We actually look at this in China. It was an interesting meeting in early December last year. DUA has a big global meeting and most of the presentations was from China because I think that's the only real handle cases we can see and touch and see how it works because we couldn't find any other products available.

So [direct] [inaudible] is interesting area of application. Utility meter life cycle support, every meter everywhere on the street and on the house for the life cycle control must have some kind of identification so we decided why don't we just try with metering devices?

And of course, traditional ways of using the handle is similar to DOI, which is digital objects, articles, media, whatever. But it's all preliminary ideas. Not significant work has been put into this area.

And of course, the question is the handle for IOT and IIOT as a unique identifier. There is interesting models already tested in the lab, not in production of course, but it also … I don't know if you know that there is a whole ID initiative in parallel, how to manage the Internet of [six] identifiers so the handle and OIDs kind of do the same job and we don't know how it will end up and who will win. Maybe nobody will win. But I said that the main problem is that we don't know the case, except DOI, Digital Object Identification, for the books and articles, with the production systems handling millions of queries per second. We just don't know. We don't have this experience. We don't have access to the data.

As I said, there are a few cases, mostly in China. They do interesting stuff, but they use handles for the person's identification in the governmental services. They use it in transportation. They use it at the factories. It's interesting that the concept, what they use is that this identification helps them to connect informational systems. So it's a kind of tool to connect different sources of the information into the simple request from the identification.

And the major problem, of course, is that we don't have approved cases with monetary value. I will not say that we're going to make a lot of money out of handle. We don't because there are no cases we can show and demonstrate to the people that, "Look how cool it is. You can make money." And this is a major problem for the technical community that, unlike [Bind] for DNS, there is no community work on the technical part of it. Basically, once you sign up with the donor foundation, you receive a package of the software, mostly Java, and you install it, you play with it, but there is no community doing cool things based on this technology. And that, I think, is a very serious problem. If the guys want to move this technology to the outer world, make it big. They do need the community to work with these applications or they need to have a lot of money like, I don't know, Microsoft or Google to focus on the development of the core application.

That's basically it. A short presentation and why I'm giving this page? Because the guy is asking me to look at the administrative part of it, accreditation, different kinds of resolvers, how to deal with local registrars, how to deal with industries, what to ask. This is all new. There are no answers for that, so kind of helping them with administrative stuff. That's it.

UNIDENTIFIED MALE: Please [inaudible]. Not Alexei. Go [inaudible].

PATRIK FÄLTSTRÖM: Thank you. I thought I heard you said Patrik, but anyway. So some of you might have heard that there is currently quite a lot of noise within the ITUT regarding [inaudible] and it was kind of interesting at the plena-[inaudible] where some member states of the ITU. specifically the ones that are not members of GAC, were pushing quite hard that ITU would make a statement that DOA, that this would be used instead of DNS. It was kind of interesting discussions.

Inside the ITU, the current state – and this is also an area where I think you could help quite a lot – where more help is needed is that in Study Group 20, Question 6, there is a document at the moment which is pushed by the Study Group Chair which is requiring this for Internet of Things, that every Internet of Things object [inaudible] to you need to implement this. Sweden and the UK, specifically the UK and Sweden and a few other states, unfortunately, quite a few actually because there are not many people that are following question 6 in Study Group 20, objected to that document because it mixes two different things: one, what requirements do we need for a naming mechanism for Internet of Things object which is very much what kind of requirements that could come out of work like the one that

Kristen is doing here in SSAC. It mixes that with an explanation, how DOA and this solution is resolving these issues.

So the Swedish objection, and it's also pretty complicated to talk about how a certain one out of many directory services might solve the solutions in, for example, the context of GDPR. What the Study Group Chair is currently doing, which makes Sweden very uncomfortable, is that they are trying to push that document through what is called the alternative approval process in ITU. Yes. You know what AP is? Yes. Okay, and what is currently happening is that there is a meeting in April, a two-week long meeting just about this document and they are even more crazy than what we are in ICANN.

So what Sweden and the UK and other member states are doing, and will bring into that meeting is a request to, first of all, move the document from the alternative approval process to TAP, which is the traditional approval process so that it ends up being visible for all member states and then we'll see what's happening. But where engineers need to help here is to try to iron out and split that document – this is my personal view – split the document in two, one which really, seriously talks about the requirements on naming systems for Internet of Things objects which Study Group 20 is working on, and the second one is to really try to explain in a more formal way how this naming system that we just heard is working and what kind

of solution, so we know what it can be used for and not used for. Thank you.

ANDREI KOLESNIKOV: Interesting. It was not a question because I don't have all this data you have.

PATRIK FÄLTSTRÖM: No, this was information to the others as well and where, I think, you can help as well.

ANDREI KOLESNIKOV: I've looked particularly for the Internet of Things because it is my primary job. I looked at the [OID] system because that's alternative path which goes in the ITU, by the way, also. So in my previous experience with the Internet, it's not necessary that something accepted in the ITU level will become standard for the Internet. I am sorry to say that.

BRITISH MALE: A couple of things. If I'm right, DOI doesn't use DOA resolution. Does it?

UNIDENTIFIED MALE: No.

BRITISH MALE: No. So the Chinese, the experience you talked about, is the only production DOA resolution going.

UNIDENTIFIED MALE: They use a DOIP [inaudible].

BRITISH MALE: Right. Okay. Do you know whether they have fixed the protocol and have they pushed those changes back into the protocol?

UNIDENTIFIED MALE: It is expected.

ANDREI KOLESNIKOV: I believe it is expected that the new release will be available in April/May this year.

BRITISH MALE: Right. So this is going to be …

ANDREI KOLESNIKOV: Yes. They track the requests and inbox.

BRITISH MALE: Right. So this will be used in the new release that, presumably, the Chinese have fixed at that point.

ANDREI KOLESNIKOV: Yeah.

BRITISH MALE: Is there any security assessment of the protocol that you're aware of or anything?

ANDREI KOLESNIKOV: Unfortunately not.

BRITISH MALE: Excellent. I think it would be really useful – this is a separate thing – I think it would be really useful to get Jim Reed to come and talk about the protocol and Study Group 20 and things. And the things about this. So for those of you who don't know, this is a very interesting geopolitical piece of software. This is effectively the alternative DNS for authoritarian countries.

ANDREI KOLESNIKOV: No, it's not replacement for the DNS. It has different functionality for God's sake.

BRITISH MALE: It's not far off been attempted to be used at that in different ways. I think Patrik wants to say …

ANDREI KOLESNIKOV: I have no idea how to use it as a DNS, to be honest.

PATRIK FÄLTSTRÖM: From my perspective, I have always viewed this as a directory of services protocol on a layer above the DNS in the stack. That said, we all know that you can implement abstraction layers in naming systems with the DNS protocol, which means that you can implement higher abstraction layers with the DNS protocol. So one of the problems in this discussion, first of all, as I said that in Study Group 20, they mix up requirements with solutions, which is never good. And by the way, that is something we are also very good at doing in the Internet world. But the second thing is, of course, that we have problems as well that the word DNS or the domain names is also, like it can mean multiple things: the database, the data structure, the protocol, whatever it is.

Regarding Jim Reed, just to clarify, Jim Reed is doing quite a lot of the heavy lifting together with [inaudible] for the UK government and I do the similar for the Swedish government, so we are peers for disclosure.

BRITISH MALE: I'm not suggesting it doesn't have a potential application in IOT. The DOI application has been around for many years and works well in those things. It's just that within the ITU, it's particularly being pushed by a particular group of regimes, shall we say, in certain circumstances that seems a political push for it and that's the unusual thing. But this is just a technology thing that you're showing us here, so I don't mean to offend anyone by saying that.

ANDREI KOLESNIKOV: Well, you cannot offend anyone by saying that. For me, I still don't understand the political part of it but what I really know that people who may be doing some lobbying and political stuff, they have no idea about the technology. That's, I think, the core of the problem.

ROD RASMUSSEN: Okay, we're at the end of time. Thank you, Andrei.

Jaap, I believe you're up next.

JAAP AKKERHUIS: Okay. This is supposed to be a lightning talk and it came about because so many remarks were made about the system on the

mailing list. So I decided to give some background of where it comes from, what it is supposed to do and whatnot. And actually, I am not working on this at all but a lot of [inaudible] people are and the [back office] is implemented by two colleagues of mine, so that's how it comes. And note the new logo.

Anyway, the crew of people behind this is [inaudible], [SNGM], the Ministry of Economic Affairs, SURFnet, ISOC, and one we should not forget is the Forum for [standard] [inaudible]. And that happy place was kind of the start of all of this stuff.

And this forum of standardization, we're trying to define what people should do with standards to follow, and it's not really a [stick of] carrots, although it helps when you make a request for proposal for products. I mean, this has a couple standards you need to do or you should explain why you don't do it. There might be good the reasons why you're not following everything.

The Internet dot-NL domain name is owned by ISOC I think. But there would be a way how people could attest whether or not an easy way for whether they apply to standards in general and it's not exclusively security standards. There's also other stuff. And it also tries to create awareness that these things exist.

So there are three categories which the website looks at, and that's the way, how you connect, how your web server looks like,

and e-mail still. This is kind of the most popular places there are. And so basically, it's pretty easy. You [inaudible] the domain name, do a click and then you get some members. The idea is that the numbers are relative and there's a waiting factor of which part works, which not, and you can't get 100% scores or everybody is very happy, but while there's not a [inaudible] comply or complain, explain, it's not really required for everything, but it gives you a [inaudible] ID and also [inaudible] user [inaudible] ID how the ISP is doing or how the bank is doing. You can dive in relatively [inaudible]. There are explanations try to be put in a layer, in end user terms, why things fail, though you [inaudible] into the technical part as well. And [as that], there might be [inaudible] reasons why some tests fail.

But it is not a debugging aid or a learning tool. A lot of people are asking, "Why can't you do this test and these details as well?" And it's not meant for a debugging aid. If you start to debug your own stuff, you probably should know what you're doing in the first place. You cannot do everything.

It's also designed by a committee that researched all the scoring [inaudible]. You could talk endlessly about what you should include and not and how many points it is, and there are some people doing that, but not us. It's not [inaudible], although we [inaudible] in the discussion. But this is really the collective, what we saw earlier so it starts to get pretty popular, and so now

there is also … People have asked where we could suggest APIs for doing last test sets. I think [inaudible] wants it. So people can do a complete block of domain names put in it and get results. And that's the idea for testing all the government domains on a regular base. They want it as well.

And this API might be [inaudible] for a large audience, but it's not really meant to be. But the website can be used by anybody and this was getting popular and what you see is that you see that there is some adoption to the [inaudible] where we style it and this is for the government, the [inaudible] disparity, the [inaudible]. That's kind of a government layer which only deals with water and [ice]. And so that's where it started, and we see that slowly things are getting adapted.

So without putting any pressure, like putting out rules or [inaudible], things are improving slowly but continuously, which is actually much better than trying to force people to do this or else you get a fine and it's way more effective. That's the idea behind it.

And you see some of the standards below there, how the [inaudible], quite [inaudible] actually are not there because they were [entering the test]. So that's why you see the empty spots there.

So anyway, next one I've got … Yeah, this was planned but it's actually [just done], so that's why. And there's a [inaudible] really [signed] so it can be used because a lot of people are asking to do … We have to translate it to Polish but then this is a really bad idea to have that on something which is not designed for that because the maintenance is a big problem then when things change. And so we've just got to redesign, also for the API and to help people plug in their own language parts and they also plug in how the whole scoring stuff is now more dynamic depending on the needs of various countries and the regions, stuff like that.

And it's [outsourced]. It's just released last week, pre-release actually, so anybody who wants it can take it and adapt it to its [inaudible]. It's out of 5K [inaudible]. You can put all of your complaints about which test fails and why, and start a fine discussion with the committee who is designing all of this stuff. Anyway, these are the details.

[Test] side, [inaudible] have fun. More information about the details behind it and this, if you can [inaudible] help desk somewhere from the [inaudible] commission which might be answering questions. So this is all I have to say, I think. Yes.

UNIDENTIFIED MALE: So yeah, I tried these tests and for the websites, it was pretty easy to get all of my websites to 100%. For mail, I got 96% and then one of the tests was wrong so I figured I wouldn't. I think by and large, it's a good idea. It's just I think it's a committee so there's a certain amount of "This looks cool. That looks cool, too. Oh, maybe people should do that." And I get the impression they don't have enough help from people who actually have deep expertise in the protocols they're attempting to secure.

JAAP AKKERHUIS: Yeah, well it's partly you should complain to the committee [inaudible], but [inaudible].

UNIDENTIFIED MALE: I filed bug reports but nobody does anything with them.

JAAP AKKERHUIS: Okay. Well, that's another thing to say. Just [a basic] way to doing the redesigns so they might be improved. With that, I am just the messenger here. But this is also the locator that you can [fill in] the issues that people are actually looking to that, so I don't know how they do the priorities about the complaints they get.

UNIDENTIFIED MALE: Yeah. I've been using this. It happened to give me a score of 81%, but I didn't really care what the score was. What I looked through was the list of errors that it thought I should improve and I thought that was a really useful thing to do. It's sort of like some of the other things, like the SSL labs test, which just tests SSL stuff and it pointed out some things like, yeah, I should have [inaudible] addresses for name servers. I had known that, but I had kind of forgotten about it so I thought it was a really useful thing, kind of like the score didn't mean much but the list of tests I thought was useful and comprehensive and reminded me to go fix some things.

ROD RASMUSSEN: Absolutely. My websites didn't all start out at 100%. Like, oh, my [dane] records are screwed up. Fixed.

JAAP AKKERHUIS: That's a [good] problem. Who cares how much percent because if you do a lot of mail or connections to Australia and Asia, things are IPv6 routing is a mess there and so the mail doesn't end there. We had the same problem trying to get for another project, getting data from Geoff, which [inaudible] perfectly of IPv4 but not of IPv6. Geoff, are you listening? Anyway.

UNIDENTIFIED MALE: I can explain what the issue is, that I have a secondary [MX], which is IPv4 only and that's what I have to have to make sure that mail actually arrives. But that gives a big red mark because not all of my [MX] records do have IPv6 support. So in my environment, of course, that is a sensible configuration. But in general, I completely agree. It should give a warning and this kind of thing. So doing tests is hard, and specifically, scoring.

But on the other hand, personally, I think that if it is the case that people immediately dive in to complain about the scoring, that for me is an indication that the tool itself is good.

KATHY SCHNITT: Jaap, Geoff is online and he said, "I am listening. The problem is ECMP and long delay."

UNIDENTIFIED MALE: Has he considered moving Australia?

UNIDENTIFIED MALE: What Jaap was trying to do was to move a few hundred MGs between Australia and Europe continuously and what happens with TCP when you get ECMP is that out of order packets causes the TCP session to simply collapse, and unless you're using something like BDR, now it just never works anymore.

Interestingly, America to Australia doesn't have the problem because the delay is slightly lower. But once you get up above 300 milliseconds and you're moving large data files on the public Internet, either use DVR or use the postal system. Nothing else works.

UNIDENTIFIED MALE: I was going to observe. I know the OTA had the online honor roll, their honor roll where they did something similar for this but they did it manually and they're part of ISOC now. And I think they're about to spin it up again. So I don't know if they have a tool to do that, so this could be a really good opportunity to solve that problem. Are you involved with that already? Oh, okay. Did I miss that? Did you say that?

JAAP AKKERHUIS: Yeah, I mean the [inaudible] files we talk about [inaudible].

ROD RASMUSSEN: Well, that looks like it worked out perfectly because it's now time for Tim's presentation. Thank you, Jaap.

TIMOTHY APRIL: I'll start with the slides and then we can move into the live demo. And sorry to anyone online.

And presenting won't work?

So this topic came out of John Levine in San Francisco a couple of weeks ago pointed out that he saw some fishy data in the WHOIS logs, the data that registries are required to submit to ICANN showing how many WHOIS queries they've received over the last month. So if you look at the agreement that the registries had to sign when they wanted to apply for a gTLD, there's this section on page 54 or something like that, that says a registry operator shall provide one set of monthly reports per gTLD, blah, blah, blah, essentially a whole bunch of different data fields that they have to provide in two different files once a month to ICANN.

So there's this webpage on ICANN's site that has all of the gTLDs and you can click on it and you can see all of the historical reports that they've ever submitted except that's usually six months behind. They just released November's data yesterday. Greg, was it or today or something like that?

GREG AARON: Yeah, it's released contractually about three or so months after.

UNIDENTIFIED MALE: That's what it says but it's more.

GREG AARON: It ends up being more and this has been a requirement since 2001. It's the basic metrics about registrations and renewals and all those kinds of things.

TIMOTHY APRIL: So John and I both have scripts now that will go and pull all of the data, so there's this big CSV file that has two rows in it every month that has all of these large things of data. So far, I've only looked at the WHOIS queries. At some point, I'll modify my script to actually look at more interesting things or other interesting things in the file. So this is just an example from dot-com. And then when I plot them, so this is the data that [New Star] has submitted for the month of October because I hadn't pulled the data for November yet. So over at the very left side of the screen, I think that's dot-biz is that peak and then it slowly trails down and there's a very long tail of WHOIS queries. If I get rid of most of those peaks, you can see that there's a fairly consistent slope down on the long tail except for those two down at the end. I think those were delegated that month.

So after talking with, I think Danny mentioned that it may be monitoring queries that make up most of that data. So New Star so far has from the registry operators I've looked at, has the most likely data that I've seen. So for each registry or registry

back end, I compute this table where I show the gTLD, the number of queries they receive that month. I compute nominal queries per second that they're receiving and then number of standard deviations away from the mean and the difference from the mean. But it seems kind of weird that some of these TLDs are within just a handful of queries per month of each other.

But it gets weirder. So if you look at Google, that looks like a pretty straight line, when in fact, they're all identical. Yes.

UNIDENTIFIED MALE: [inaudible]

TIMOTHY APRIL: It looked … Warren and I talked about this yesterday because I hadn't looked at their data until I was sitting next to him last night. So they have a whole bunch of data. That way, I pulled up Afilias just to see what theirs looked like. At the beginning, it looks very similar to what we saw with New Star and then I pull out a whole bunch of the other stuff and so there's that same gradient in the middle there, but there's this other section that looks exactly flat. When I look at it, it's a whole bunch of dot-brands that have identical numbers of queries. No idea what's

going on and those, if I look back through the months, they look identical. Patrik?

PATRIK FÄLTSTRÖM: Well, when I have nothing to do in the evenings, I go to each one of them, their webpage, an equal number of times, of course.

UNIDENTIFIED MALE: I was actually thinking it was an academic study. I guess that's probably the most likely explanation. It's just monitoring queries, right? Especially the brands, there's nothing going on.

UNIDENTIFIED MALE: The other possibility is that since it's all the same WHOIS server, they're reporting the aggregate numbers rather than actually going back and breaking them down for what was actually queried.

Let me just say that –

UNIDENTIFIED MALE: [It's] part of the story.

UNIDENTIFIED MALE: Oh yeah.

TIMOTHY APRIL: If you want to explain the Charleston Red one, yeah.

WARREN KUMARI: Yeah, so please this is definitely not for sharing. But can you quickly flip back to Charleston pictures? So yeah, that's perfectly flat and the reason for that – actually, Tim pointed it out. I wrote up an e-mail and sent it, and then immediately after that, saw Richard Roberto who is one of our registry folk, sitting two rows in front so we poked him. We now have a bug open on this. It is the WHOIS server counts the number of WHOIS queries and is not currently breaking it down by the TLD. And if you go back to the very first slide, I pointed them at this and there is now an ongoing discussion on whether one is actually required to break it up by TLD because this is –

UNIDENTIFIED MALE: [Inaudible] TLD.

WARREN KUMARI: Nope. It says "A report shall be provided per gTLD using the API blah, blah blah." But the actual data does not actually say that it must be broken down by TLD. Believe me. I'd be having the

same view and we've got people. They're not shouting and screaming back and forth.

UNIDENTIFIED MALE: I have the mic, SSAC 105.

WARREN KUMARI: There is stuff further down and it is slightly ambiguous and I mean, obviously people interpret it in the way that was easiest for them. It's being worked on but there is still some back and forth and I'm not sure. Maybe do you have different back ends for the brands?

UNIDENTIFIED MALE: For all of the brand things, they're all in the same complex so yeah, it could be that we've got something stupid going on too. But no, we're pretty good about [by] TLD stuff is the way we do things so I wouldn't expect that but I won't swear to anything until I ask.

WARREN KUMARI: So all of our other stats are by TLD except the WHOIS where they were like, "They screwed that up."

TIMOTHY APRIL: So yeah, there's Google doing bad things, Afilias. I went back through the data and it shows all of those brands have identical data going pretty far back.

WARREN KUMARI: It's too spread to be [inaudible].

TIMOTHY APRIL: And that's all I had for slides, so I can actually … I was going to show the actual thing I wrote. So sorry to Geoff. I can send you the code.

UNIDENTIFIED MALE: Much more fun data, [inaudible] data.

TIMOTHY APRIL: So overall, I wrote … Why do I have two pointers on my screen? That's fascinating. They did for a second there. They weren't going in the same direction.

UNIDENTIFIED MALE: [Inaudible]

TIMOTHY APRIL: Damn it, Warren.

So I have this nice [inaudible] database that has all of the data from all of those reports in it – I don't now what I just did – that has this web app that I wrote that is terrible. So if I pull up donuts from … Hey, let's look at today's date. That looks a little fishy. For some reason dot-works and dot-ltd seem to be the weird children here. There's also dot-travel has some absurdly low number of queries. But when you look at this data, it's not as clean as when John and I were looking at it in San Francisco anymore. But if you take the Earth [inaudible] of all of the records that I didn't exclude, so not dot-works, dot-travel, or dot-ltd, all of the numbers fall within 1% of the mean.

When we did it that time, they were exactly plus one and minus 1% of the mean and if you divided the number of TLDs in half, if you pick that item, it was exactly the mean. So I had no idea what the heck was going on. If you plotted the distribution from the middle, it created this very nice bell curve. The only theory I had was that donuts was picking a random number in the middle and doing enough jitter to make it look like the numbers weren't identical. It seemed like … I don't want to call it malicious but that's the best word I can come up with.

Yeah, Warren?

WARREN KUMARI: You'll get that exact same distribution if you take random numbers on either size of a zero and add it to an offset of what somebody might think a reasonable number would be.

TIMOTHY APRIL: Yeah, and when I was doing it, the means … When I was doing it with Google Docs, so I don't know if Google Docs was just … Oh yeah, the mean was an even number.

UNIDENTIFIED MALE: A whole number.

UNIDENTIFIED MALE: Yeah, exactly.

ROD RASMUSSEN: Have you calculated yet, queries per TLD based upon the number of domains in the zone? Because some of these TLDs are really large and some of them are really small but somehow they're giving the exact same numbers.

TIMOTHY APRIL: I have not calculated that yet.

ROD RASMUSSEN: So the size of the TLD and what's in it isn't making a difference in a lot of these cases which is what you would think would be happening, obviously.

TIMOTHY APRIL: I've also looked at … Well, I can do it again where I go to donuts and I pick November, remove the three less interesting ones. And if you just happen to glance – I'll make it bigger so it's not as big of an [eye] chart. So just pick your favorite position on this chart and we'll go back a couple months and you'll see that they're all different every month. The ordering is random, as best I can tell.

UNIDENTIFIED MALE: That's weird.

TIMOTHY APRIL: Yeah, and the peaks sometimes move like the LTD and there's one in the middle here it was an outlier in the previous month and now it's not and it kind of moves around.

WARREN KUMARI: That could be you have people who [inaudible].

TIMOTHY APRIL: Yeah, it could be.

WARREN KUMARI: [inaudible]

TIMOTHY APRIL: Yeah.

WARREN KUMARI: I know [inaudible].

TIMOTHY APRIL: Does anyone have their favorite registry they want to go look at and see if they're doing weird things?

UNIDENTIFIED FEMALE: Funny you should ask.

UNIDENTIFIED MALE: It sounds fun.

UNIDENTIFIED MALE: [inaudible]

TIMOTHY APRIL: That's donuts.

CentralNic. Yeah, so I remember looking at CentralNic and thinking it looked fairly normal. There's minor noise down at the bottom here.

ROD RASMUSSEN: While Tim's playing with that, did you have more you wanted to … ?

TIMOTHY APRIL: That's all I had.

ROD RASMUSSEN: Danny's seen his data.

DANNY MCPHERSON: Actually, if I could say …

TIMOTHY APRIL: Here's Danny's. Anyone want to guess where dot-com is in that [inaudible]?

DANNY MCPHERSON: Yeah, right. So one of the things I'll say, we looked at this a couple of times based on when John pointed out and then I was

talking to Tim and our monitoring, it was per TLD, like 80 queries a second for monitoring and I'm like, "Wow, that's a lot," and then forgot that we, all the time, move. We have WHOIS running 80 or 100 sites and we query each one of those second and we look for [positive]. And so that's why those numbers for us are so massive with that and there is some skew at the end. But anyway, so that's what the really, really large numbers for Verisign are.

TIMOTHY APRIL: So if anyone wants the code or the SQI database, let me know. I'm happy to share it.

UNIDENTIFIED MALE: How big is [inaudible]?

TIMOTHY APRIL: Nine mg, I think. It's a terribly annoying process to go and grab them all because to get each CSV, you have to parse their webpage in a special way and download it and then I import it. It takes about three hours of hitting the ICANN website with 10 different web crawlers.

ROD RASMUSSEN: Okay, do you have a question, Jay?

[JAY]: Not a question. I was just going to say that hopefully that will be sorted out by the ICANN open data initiative because that's one of the early ones that people asked for to go onto that at some point, so it should be the case from the public thing.

And secondly, I think you've now given Warren an idea to go back to his people and say, "No, we don't have to do different reports. All we've got to do is a jitter across each one of them into a nice normal curve and present that.

TIMOTHY APRIL: You want a [Paredo] distribution. Those are harder to attack.

UNIDENTIFIED MALE: Something I would emphasize is this is 100% public data. We scraped this off of the public reports on the ICANN website. There's no secrets here other than the fact that if you decide to turn the data sideways, it's remarkably flat.

UNIDENTIFIED FEMALE: You just keep wishing that it made more of an effort.

ROD RASMUSSEN: I want to do a couple of things since we're about seven minutes away. I want to follow-up on this. Greg, you got something?

GREG AARON: I was going to say, so the reason we started doing this is we wanted to know what the effect was before the temp spec and after the temp spec and were people using WHOIS differently. And it turns out we have no idea because the data is polluted. And one of the things we're seeing is the effect of having a bunch of TLDs in one server. That's probably what you're seeing.

Registries also make test transactions, monitoring transactions using EPP but they don't count them. But it looks like they're counting them in here, so we've got to figure out a solution because we're collecting this data but it's worthless.

Yeah, so I was the one who asked the original question. How many WHOIS queries are there per day? That's actually really important if you're going to run something like the TSG has come up with is knowing your order of magnitude of queries, you might be getting at some point. And the numbers right now is over 200 billion a month, so roughly 7 billion a day, something like that. So anyways, that's interesting but we don't know if that's actually real. And then of course, there's a bunch of monitoring and all that.

And then we want to look at effects and then some other things. So net on this is that we did some outreach. We obviously talked to Danny. I've talked to Donuts and I'm trying to get Tim and one of their engineers together because I talked to Alan Woods and he's like, "I took this as far as I could and I need a technical person to work on this," and obviously, we engaged Warren on this one. It looks like there's some more as well. I gave a heads up to Jamie Hedlund that this was going on and that we were back-channeling this instead of saying, "Hey, the registries are evil. You're lying about their data." The donuts one was very troubling given it looked like they were artificially playing with the numbers. The other ones look like the flat one. It was like, "Okay, that's just a …" Well, now we know we need lawyers involved apparently.

Yeah?

UNIDENTIFIED MALE: To that point, we do internally in our reporting have good queries versus synthetic transactions that we have and we obviously decouple those by source distribution and so forth, and that's where some of our per source rate limiting comes in, although that's also per site so if you're good, you can distribute and get a lot more queries and [inaudible] rate limit, but we don't think people were [inaudible] hitting our rate limits.

But anyway, so we do have that other data if anybody is interested, I could see if there's a way to share some of that. I don't think that would be a problem, not a concern.

ROD RASMUSSEN: I think we actually need to have some clarity around what to report.

UNIDENTIFIED MALE: Absolutely.

ROD RASMUSSEN: Yeah. Okay, Tim.

TIMOTHY APRIL: Just to give you numbers of what you were asking a second ago, there were 211 billion queries in November of last year. That's 82,000 a second.

UNIDENTIFIED MALE: Unless that's the dot-com one, obviously that's huge. There is an opportunity here because within six months, everybody has to deploy RDAP and they're going to have to start counting those queries but we have an opportunity to start counting them in a

regular, predictable fashion and have some good data going forward.

So maybe one of our goals here is to make sure that everything's being done in a similar fashion and we've got some good data for the future.

ROD RASMUSSEN: I've got a couple things I want to run through here real quick. There's a request if … I want to come back to this, but Danny asked me to run through anything else. We have four minutes. So there's a request, and if we want to get some sort of answer to a question from any ICANN exec, we need to file that by what?

UNIDENTIFIED FEMALE: Noon tomorrow.

ROD RASMUSSEN: Noon tomorrow, so if we have a question for an ICANN exec, I'm just throwing that out there. If you have something, let the admin, let me or Julie or Kathy know and we'll fire that over.

We do need to talk more about the hijacking stuff. Were out of time today but we will have some time maybe in the wrap-up session. We have the wrap-up session. We also have our open meeting tomorrow and the meeting with the Board.

I know some of you may be traveling Thursday, but if you're here tomorrow, we want as many … If you're in Kobe, you're here please unless you have a commitment for another SSAC or ICANN-related thing that you have to do. And then also, it's really important for the wrap-up. We had some folks miss the wrap-up in Barcelona because they were doing some sightseeing and that's kind of a no-no.

The Board, as many members as you can have at the Board as well. It's in the morning. That's going to be an important meeting for us so please do that. That's all [inaudible] if you want to get going.

UNIDENTIFIED MALE: Sorry, Rod. The open data session is at the same time as the wrap-up.

ROD RASMUSSEN: Then you've got … Yeah, that's an excused absence.

UNIDENTIFIED MALE: Yeah, but I was telling everyone else as well, as frustrating as that is because I can go and report back or something.

ROD RASMUSSEN: Obviously, I'm not going to put you in detention for not being here, but it's a matter of let's be professional about things.

UNIDENTIFIED MALE: I think you should start doing that, yes.

ROD RASMUSSEN: Yeah. Yes.

UNIDENTIFIED FEMALE: Isn't this detention?

ROD RASMUSSEN: Isn't this detention? Okay. So just to finish up on the WHOIS stuff, there's a lot of questions we'd actually like to answer, effective GDPR. Can you look at patterns of WHOIS requests versus when we saw abuse at TLDs, things like that? Are there some leading indicators, things like that? There's a whole bunch of things, but you can't do it with bad data.

The other thing that this pointed out was that nobody was minding the store at ICANN either because we found this just by pretty simple stuff and the Charleston [Road] stuff should have been found a long time ago if anybody was even looking at it, right? So I brought that up to Jamie's attention. So we want to collaborate on this. We don't really want to make a big public

showing of throwing registries under the bus and this is why I talk about building capital because this is all about collaborating and fixing things rather than scoring points. So that's really what we want to do. Steve?

[STEVE SHENG]: I think probably talk to GDD folks. They have the most direct influence and contact on this.

ROD RASMUSSEN: Yeah. One of the things we want to do is actually get the scope of the problem and figure out the registries involved and let them know, point out their issues. So I didn't realize that affiliates might have a small issue and New Star might have an issue too. If everybody's got a little bit of a problem, well then, we'll go as a community and fix the problem.

So one of the things we want to do, Tim, John, myself and Greg or whoever is interested, is take a look at the data and see if we can pull out people to talk to and then there might be some outreach to do on that.

All right, so think about the hijacking stuff. We'll talk about that in the wrap-up. And we are out of time it looks like. Yes, we are. So we're going to close the meeting. Thank you, everybody, for a jam-packed and hopefully productive session, and thanks for

the lightning talks. And Tim, we forgot to give you applause at the end. That was great stuff.

Thank you. Okay, so one very important thing. Thank you to our staff for making us look good, keeping us on track, and our staff and the ICANN facility staff. Other than the thing that exploded, and literally, we had smoke in here on Saturday morning because of the power [inaudible]. That was the only glitch and it fixed itself immediately, so it was great. Okay, it didn't fix itself.

UNIDENTIFIED FEMALE: Other than that, this is like [inaudible].

UNIDENTIFIED FEMALE: Good-bye.

[END OF TRANSCRIPTION]