
LAUREEN KAPIN: We'll get started. I'm hoping we'll have more folks join us as the second path. We're going to start off by taking a look at the work plan, which was circulated last week. I wanted you to get the use of folks on the phone about that work plan, particularly if you have any suggested changes or any thoughts or concerns about how things are scheduled.

I will note, just for folks to highlight, that this [tease-up] on October 15th or a draft report – and then have several milestones along the way, including a development of findings for our various topics, which is going to take place throughout the summer and early September, and then, in turn, recommendations really to be hammered out by the end of September so that we can produce a draft report by mid-October.

With that in mind, I want to open it up for discussion.

DREW BAGLEY: This is Drew.

LAUREEN KAPIN: That's a deafening silence. Someone chime in, please.

DREW BAGLEY: Can you hear me?

LAUREEN KAPIN: Yes.

Note: The following is the output resulting from transcribing an audio file into a word/text document. Although the transcription is largely accurate, in some cases may be incomplete or inaccurate due to inaudible passages and grammatical corrections. It is posted as an aid to the original audio file, but should not be treated as an authoritative record.

DREW BAGLEY: I was wondering, what does this contemplate in terms of analysis? For example, for the June deadline before that face-to-face meeting, we would already have obtained whatever data we wanted and then, through our broader CCT meetings, decided on how we want to spend the rest of research funds and whatnot by that point? Or what would we be analyzing at that point, I guess?

LAUREEN KAPIN: The way I'm viewing this is that we have certain data already. Whether this includes all the data we'll need, I think that's really going to be case-specific because I certainly can conceive that in certain situations we have data already and in others we do not. We certainly may not have all the data we need by early June, which is around the corner. So when I say "analysis," to me that means we're actually using what we have and engaging in analysis during that meeting. In fact, I see that as one of the primary opportunities of the June meeting.

But maybe the ICANN staff can chime in here to make sure that we are understanding this correctly.

MARGIE MILAM: Hi, Laureen. It's Margie. I think that's correct. This is intended to at least give you a roadmap for looking through some of the issues. It doesn't necessarily mean that you're done on a particular topic in June 6th or 7th, and you may actually end up with a hypothesis, for example, that you may want to investigate through additional data. But I think it's just a way of trying to organize the thinking so that you know that you got the topics that you want covered addressed in some way.

The other thing I wanted to point out, too, is the need for additional data. The sooner you can identify that for us the better, because we do have a process internally to be able to contract with vendors for data, go to contract and budgeting issues. So the sooner you can identify that, the better.

LAUREEN KAPIN:

Okay. Other questions or thoughts about the draft work plan?

Okay. Well, hearing none, I'm going to then conclude that, for now, this is a good draft framework that we will attempt to be guided by. And to the extent that we need to adjust, we'll raise these issues as soon as we can so we can adjust the work plan.

Let's talk about, then, our second topic: data requests. Now that folks have begun delving into the reading list and thinking about their topics in particular, do folks have anything flagged for data requests?

Okay. I know Carlos is going to raise some issues when he talks about his presentation regarding the domain abuse report, so I'll leave that to him. But I wanted to also hear from Karen because she had raised some issues about the public interest commitments and thoughts about a study that could measure the impact of those.

Karen, do you want to talk about that?

KAREN MULBERRY:

Sure. I'll do that. Thanks, Laureen. As you're finalizing what sort of data you are looking for to work with, I wanted to raise one metric, one of the recommended metrics from the IAG-CCT, that we haven't yet

figured out to address. I wanted to raise it with this group, as it seems the most relevant.

The metric was titled, “A qualitative comparison of Question 18 of the new gTLD application with the current actual use of the gTLD.” Question 18 in the application, for those who don’t recall, was called Mission and Purpose, and it was a pre-text set of questions that asked the applicant to define their mission and purpose for their TLD and asked what sort of social benefits or public benefits it would be expected to provide and how it work to mitigate against potential costs.

The answers weren’t evaluated as part of the evaluation, so they didn’t go to whether an applicant past or failed in the evaluation. But the intent of including those questions was to use them for later assessments or analysis as to plans and measures that applicants were putting in place or intending to put in place at the time that they applied and how those things might have changed over time.

So we’ve held off starting to do that qualitative study, wanting to figure out the way to do it. There’s a few different ways that you could go about doing that. Also, trying to establish what would be a reasonable time frame that you’re going to look at when you look at the applications from 2012, at what point of later registry operation do you want to make some sort of comparison?

On the reason that some of the PICs came into this topic, our thinking process a little bit was that one of the ways you can look at this is over time. So if you look at what the applicant wrote when they applied – and as part of the process there were GAC early warnings. The process [inaudible] for applicants to establish additional commitments that they

wanted to make would become binding in their registry agreements, so some of the applicants voluntarily did let them [inaudible]. If you're taking a particular TLD or a sample of TLDs over time, the PICs are one area that you could look at in terms of what was changed or added or modified in terms of how they operated.

In trying to study or assess this, you can look at the stages as a whole in terms of looking at a set of applications and what were the common themes that applicants were writing about at the time versus what sort of themes or commonalities came out in the registries that did make voluntary PICs. If you look at a point in time of registry operations, what are the buckets or common practice that we're seeing? We can do some comparisons there.

So this doesn't address everything there is to ask or examine about the PICs. You recall, I'm sure, that there was a letter from the Board specifically asking the review team to do some examination on the PICs, which include several items that are part of the registry agreement for everybody, as well as the applicant-specific ones that I'm stressing here.

But I wanted to note that in terms of – you were thinking about your data requests so that if there was any interest in having us pursue this metric or type of analysis around looking at that part of the application, we could incorporate that inner planning for our work.

Thanks.

LAUREEN KAPIN: Thanks, Karen. I want to make sure I understand what you were describing because it seems to me there are a couple different issues at play.

One is what was initially stated by applicants in their application in response to the criteria of the Guidebook for making applications in terms of what their mission was; looking at that and then I assume comparing that to the actual practice, and if this is a successful applicant and it's been delegated, doing a comparison of what they indicated the mission was and then what they're actually doing in the real domain world, so to speak. Am I understanding that first part correctly?

KAREN MULBERRY: Yes.

LAUREEN KAPIN: Okay. Can you flesh that out for me? I understand it would be very easy to look at what they said in their application, but what would you look at to assess what's going on in the real world? Or the virtual world, so to speak.

KAREN MULBERRY: Yes. That's a good question and something that we've explored as a different way you could try to assess actual practice in light of what they originally wrote. Obviously, it's going to be a qualitative approach. Obviously, one place that would be relevant would be the websites of the registry operator, what sort of registration policies and privacy policies or [Whois] policies or things that they have in place.

You can also look at how they're marketing or promoting either the TLD or the audience for it or the way that their TLD is run to the extent that it's there and discussion about the operator's company or organization and their values. You can look at any complaints, either that ICANN has received or that have been raised about the registry operator's performance.

So those are some of the things you could look at. I'm sure there are more, and we've tried to articulate some of those. We've tried to sketch out a proposal for how you could go about doing this type of analysis, and we can circulate that to you as well if you'd like.

LAUREEN KAPIN:

I think that would be useful. The other issues that I heard you raise – and this was probably the preliminary, instigating issue with the public interest commitments themselves because certain applicants voluntarily added public interest commitments, and then of course there were the public interest commitments added to the registration agreement. I think the issue you raised was, for the actual practice, what point in time do you start looking?

KAREN MULBERRY:

Yes.

LAUREEN KAPIN:

Have you reached any preliminary thoughts about that, or basically it's something that you're still grappling with?

KAREN MULBERRY: On the timing? Well, I think you can make an argument for any point in time. Once the registry is operating, they're operating, and that still could continue to change even after you look at it. So what might make sense is to figure out how long this type of analysis is going to be relevant. You could look at what they wrote in their application versus what's happening 20 years from now. So is this a repeated thing where you want to check the results on a periodic basis, or is this a select – I don't know – date five years later, or a particular date where you think it will – the types of operations will have reached a steady state from once they've launched.

LAUREEN KAPIN: Got it. Do folks have thoughts on this proposal from Karen? Karen will circulate her plan to flesh this out a little bit to us. But just based on Karen's helpful discussion here, do folks have insights into possible ways to study the impact of the public interest commitments and also this more general notion of whether applicants' mission descriptions actually matches with what's being done and practiced?

Okay. I'm not seeing hands or hearing folks chime in, so I think it would be helpful here, Karen, to actually see this proposal. I think that might generate more discussion on how to flesh it out.

But I will say, at least speaking for myself, that I think this is a really important topic to study and gather data on. I appreciate you raising the issue so that we can pursue how to best gather the data on this point.

KAREN MULBERRY: Thank you. Certainly – oh, go ahead.

DREW BAGLEY:

I was just going to quickly chime in and say, yes, I think it's great and I think that we'll be able to use this data, along with a lot of other data we're gathering, to see if there's correlations between the types of public commitments used, how effective they were, etc., and the size of the TLD, the amount of money that the TLD was auctioned for, and a bunch of other trends. Then we can tie that into trends we're seeing with DNS abuse and whatnot.

So I think what you're doing will help in different ways with all kinds of things we're looking at.

LAUREEN KAPIN:

Good. Thank you, Drew. With that, I think we're going to move right into our presentation. I want to thank Brian for joining us and also being part of the presentation team today.

But we are going to start with Drew for his continuation of discussion on the Secure Domain Foundation's business case for proactive anti-abuse. Then we're going to move to Carlos to give some brief remarks about ICANN's very useful draft report for New gTLD Program safeguards. That will be followed by Brian walking us through that report a little more. Then we're going to end with Fabro's remarks on the KnujOn 2016 Internet limbo report concerning issues of consumer trust. So that's going to be the game plan today.

Lastly, before we conclude, we're going to take up two procedural issues. One, we have the results of our poll, our Doodle poll. Thanks, everyone, for responding. We'll talk about that. Then I noticed that Alice

raised a question in the remarks about whether we want to spend the time for our sub-team call. So I'll raise that as an issue, too.

Before we get to all that, we're going to head to our presentations now for materials that are on our reading list. Drew, I'll let you take it away, but I also do want to highlight that our great ICANN staff has continued to populate our reading list. So please go back to the reading list and see the extra materials there. We'll figure out a way to get presentations included on those materials as well.

With that, take it away, Drew.

DREW BAGLEY:

Yes. I'm only going to speak for just a couple more minutes because I went over the gist of what the report said two weeks ago. There's some takeaways here that I think are relevant to our tasks at hand. One of the findings in the report about the difference between being proactive and reactive or being really complaint-driven – I think, as we analyze data and we're looking at how many domain names in a particular TLD's zone file are phishing domains or whatever we're going to look at with DNS abuse, it's also important to keep in mind how quickly they disappeared and whatnot.

If you look at the slide that says "Volume Matters," one of the findings in this report was that one of these registrars that did not have as much volume as a proactive registrar did only was complaint-driven, meaning that they only acted on DNS abuse once a complaint actually came to them. They had so many complaints, and they had nearly half as many

as a much larger registrar, despite the fact that they manage only 20% of domain names.

When we're looking at the total number of domain names in a zone file and we're looking at the total number of domain names associated with DNS abuse, we'll also want to keep in mind perhaps what that registrar and even the registry is doing, if we drill down and look at registrars as well in the new TLD space; if they were doing any pre-screening of their registrants or if they were only taking care of things after the fact, because that can tell a story that might be slightly different than whether or not the new TLD safeguards are making a difference in the anti-abuse versus maybe the individual actions or a registrar or a registry and how they vet their actual customers.

The other thing along those lines that I wanted to point out about this was that, when we're looking at all of the data and trying to understand these safeguards, we should come up with a way to look at how the RAA 3.18 provision is dealt with in terms of what registrars in the new TLD arena are doing and how they interpret that, because that's interpreted pretty broadly in terms of what their requirements are to deal with DNS abuse.

We should also look at Specification 11, which Laureen I imagine is an absolute expert on. But we should definitely make sure we're looking at those because the perceived cost of complying with those and the size of the registrars in this study seemed to affect how proactive or reactive these different entities were being.

So we might find that a smaller and less financed registry perhaps is not doing as much and perhaps is not requiring as much from their

registrars vis-à-vis their relationship and perhaps even registrars in the new TLD space that are smaller or interpreting the 2013 RAA agreement in a different way than those that are bigger and more robust. So I was just thinking that those are other things we want to look at.

With that, I yield the floor to Lauren and then Carlos.

UNIDENTIFIED FEMALE: Lauren, you might be on mute.

LAUREEN KAPIN: I apologize. Thank you. I was asking to the air but not to the group if folks have any questions for Drew based on his presentation or the underlying report in general.

Okay. Hearing none, I'm going to shift then to the DNS abuse report and start with Carlos and then move on to Brian.

Carlos, do you want to take it away? It looks like Brenda is uploading your slide.

Carlos, just as a test, I'm not hearing you, so I want to make sure that we can hear you.

CARLOS RAUL GUTTIEREZ: Hello?

LAUREEN KAPIN: Okay. Now we hear you.

CARLOS RAUL GUTTIEREZ: Can you hear me now? Hello? Can you hear me now?

LAUREEN KAPIN: We're hearing you twice, Carlos.

Now we can't hear you at all.

CARLOS RAUL GUTTIEREZ: Hello?

LAUREEN KAPIN: That sounds good.

CARLOS RAUL GUTTIEREZ: I hope that you can hear me. I'm sorry. I had two mics at the same time, and then I turned off the wrong one.

To make it quick, I didn't do an analysis of this study. I just wrote some notes. Okay? The report is really interesting. I'm going to concentrate on the first part, and then Brian will talk about the project to produce some deeper analysis.

The assumption, of course, is that the expansion of the gTLDs would erode trust. This study focused on the safeguards that were written into the agreements – so to avoid that. The first question when I look at the safeguards is where the safeguards came from and if those will be effective, of course, and if there are other remedies, other safeguards, and, if so, to say – and the report doesn't talk about [exposed] remedies

like just taking the website down, which is the usual view of the regulators.

But I think the first thing that caught my attention is that we're based on a very early definition of abuse developed by the GNSO, as it should be. So we have a clear definition defined in 2010 on what abuse is. It's pretty general in my view. The report does a great job pinning it down to real abuse cases and quotes from different studies [inaudible] the level of abuse that has been measured so far and quotes different studies.

Also it caught my attention that there are different groups that exist working on abuse, long-standing groups. There is one that I know very well which is not specifically listed in the lower part of this slide, which is the one by the mobile operators. It has a terrible name, [mwalk] or something like that, a mixture of letter and numbers and so on. But I know they write very interesting reports from the point of view of mobile operators.

So I think this is a very important point. There are in the community different working groups, different structures, following this issue, so it's very good to map them and go a little bit deeper on what each of those groups do, because I think, over the long term, it will be more useful to collaborate with these groups from the point of view of ICANN.

Drew already mentioned that the core of the trust issue is written down in Specification 11. The reason I quote it again in full is because in the second section, it clearly says that they [did] expect the registry operators to be the ones producing the statistics, and that was comment I was going to make as Drew was speaking. Who produced the

reports, and is it reasonable to expect that they have different strategies? When you mentioned there are some proactive ones and ones who are not proactive, the immediate question I have is: are they both fulfilling Specification 11 in the same way, or should we expect that they all perform in the same way?

The report makes very often the reference to the CCT review, and I take them as recommendations to go a little bit deeper. Every time that the report mentions that, we should take it very seriously.

It also makes a specific recommendation on Specification 11. Is there enough, or should we have a deeper discussion on Specification 11, or a wider interpretation on Specification 11?

In my way to read these, we come to the specific list of the nine safeguards that are now signed by all new gTLD registry operators. As expected, they are very, very dear to the registry operators. For me, that raises the question of what happens if their views are not at that level. What if their views are of end users or what if their views are being done at a different level, at the retail level or the user level?

So, again, a mental note that so far Specification 11 [is] only at the registry operator level.

In general, the report focuses on ways to follow up on the effectiveness of the safeguards. It makes the same note I just mentioned, that there are two different points in time for abuse; one at the registration level, and of course the safeguards are very focused on the registration level, but there are many cases of abuse that are being done after the website was created and so on.

The plot thickens when we come down to the proposal to develop indicators. I think it is right to ask these questions because, when we go back to our famous 70 or 66 matrix and we look to Chapter 3 on the main abuse, those are really just raw statistics; quantity of incidents of this and quantities of incidents of that, number of complaints, and quantity of takedowns.

I think that the question of our composed index or an index, in any case, is very appropriate. The challenge is the index itself. Can somebody understand our rate of abuse? Or could we get used to a simple number? Let's say, "Okay. This domain name has a 20% rate of abuse," compared to another one, which has a lower or a higher rate of abuse. This is, of course, from the point of view of statistics. Very, very important issue, and I think that's where we should focus next with Brian and how feasible it is to come down to a very simple but powerful index in terms of abuse.

In any case, the report says that it is up to a review team to decide the scope and method of inquiry into DNS abuse mitigation efforts, and then it jumps into the real innovative part of the report that I think Brian will be best able to address. We will hear from him.

In any case, the question is always as we have been discussing over the last three months: where do we get that data from? Who is going to produce or keep these time series of all these abuses that are listed in our long report on the matrix? Is it going to be an external source? Is it going to be the registry operators day by day? And at what point will we have a series solid enough to do research?

This is my proposed model for the working group so we can [sail across] correctly. If somebody has seen the film, I'll invite you for a beer in Washington.

Thank you.

LAUREEN KAPIN:

Thank you, Carlos. I appreciate that, including your film reference. I want to hand it over to Brian to go into the report in more detail. But before I do that, I want to commend Brian because this report reflects considerable amounts of research and effort which is catalogued very clearly and is also a great reference to us because the underlying source materials here are good things for us all to be looking at.

What I also want to emphasize to everyone is that, because it's going to be up to us to give guidance on how to proceed here, although Carlos is the person specifically assigned to this paper to give us a preview, I really want everyone to read Brian's report because we are going to have to decide the best way going forward. It will be impossible to do that unless you have the report as a foundation.

With that, let me hand it over to Brian.

[inaudible] in front. I see Brian is telling me I'm over-crediting him. So to everyone who helped create this report, my kudos to you.

BRIAN AITCHISON:

Thanks very much, Laureen. Can everyone hear me all right?

CARLOS RAUL GUTTIEREZ: Yes.

BRIAN AITCHISON: Great. It looks like we have the presentation up. Let me make sure it scrolls and does what I want. Okay.

Carlos did a great job of summing up the report. I'm going to focus mostly on method, and I'll breeze through some of the stuff that Carlos already talked about.

We have on your screen our agenda. It's going to be pretty quick today. We'll do an introduction. We'll go over the definition of DNS abuse. We'll talk a bit about the safeguards but then spend the majority of our time talking about points 4 and 5 here, ways to structure our research efforts and how to maybe use the model to provide a framework for all this data we're collecting, which I know is a big, difficult, and often messy exercise. That's perfectly normal by my experience.

I'll introduce myself because we haven't really met. I come from a research background. I hold a Master of Research, which I know is a weird-sounding masters, but they are out there. It's designed to give people expertise in qualitative and quantitative research design.

I also have a PhD in political science, where I was able to apply those research methods. That was the idea behind the program I was in. I was also research lead on a project to study the impact of the Internet of Things on economic growth. The reason I mention that is because dealing with these big, complex variables is something that I have some experience in. I know how hard it can be to structure all these kinds of things, so I definitely sympathize with all of you and your efforts.

You may be asking, “What if any of this has to do with DNS abuse?” As you see, the answer is method, having a design to organize all these big, complex variables. Competition, choice, and trust are very hard to pin down. There’s a lot of different definitions for them.

Also, in methods we want to emphasize primary research over secondary, which I think we’re in the process of doing. The reason I mentioned that is we have received a number of public comments already. The comment period closes tomorrow. There’s something of a misconception that research means doing a literature review and then you’re done. But I think we all know that things are going to be much more complex than that.

Also, demonstrating causality through hypothesis testing: we want to explain why the New gTLD Program may have an effect on DNS abuse or safeguards have an effect on DNS abuse. We don’t want to just describe it. So we want to move into inferential over descriptive statistics, and I’ll talk a little bit more about that.

Moving on to the second slide, I think we’re defining DNS abuse as an aspect of consumer trust. I think there’s overlap in the other areas, but I think that’s how we’re looking at it. The working definition that we crafted was “intentionally deceptive, conniving, or unsolicited activities that actively make use of the DNS and/or the procedures used to register domain names.” Very much a working definition. We agonized a lot over it, so I’m happy to alter that as you guys carry on with your work.

I’ll just move to the next slide. I don’t think much of that is too important. Now, Carlos hit on the RAPWG’s working definition of

registration and use abuse. I think we don't need to spend too much time on that, given time constraints now.

Also we have our safeguards here put together by a number of expert working groups and subject matter experts. They presented these questions to themselves, and you can see that these are the answers that they arrived at. So that's a bit of background and definitional framework.

Let's jump in the method, the real meat of this presentation. This method is fairly standard. It's more or less the scientific method. Let's just go through this list of where we want to be and see where there's white space in the sub-team's project.

Step 1 is to identify the research problem clearly and ask yourself, "What's the empirical puzzle we're trying to solve?" I just put some fairly simple propositions here. The research problem is essentially that it's unclear how effective safeguards to mitigate DNS abuse in new gTLDs have been. The empirical puzzle that we're trying to solve is that there's some indicators pointing to reduce amounts of DNS abuse in TLDs in general, while others point to increasing rates of abuse in particular TLDs. So the extent to which safeguards to mitigate DNS abuse played a role in this variation remains unclear.

That can be restated in other ways, but that's just to give an idea of how to frame something like that.

Step 2 is to review and synthesize previously published literature. I think that's in process. That's the step we're on right now. Moving into Step

3, we want to clearly specify our research questions and the hypotheses we're trying to test.

A big question I always ask myself is, "What explains variation?" You see I bolded it here. That's a common question in social science inquiry. What explains the variation in rates of abuse in different TLDs? What explains variation in abuse rates in new versus legacy TLDs? What about variation within new TLDs? These are the kind of research questions I think we're all asking ourselves.

We also want to make sure these questions and hypotheses indicate how each term is segmented and defined and how we're going to measure it, so effectiveness is a concept we're talking about a lot; competition, choice, and trust in the other areas. So we want to be very clear how we define those terms so we can get down to measuring them.

Step 4 and 5 we'll talk more about in the next slides, but let's hit more on Step 3, which is about hypothesis formulation. You'll see independent or explanatory variables and dependent and response variables. I imagine you might be having flashbacks to your high school chemistry class, but essentially this is a way to explain a cause-effect relationship, which is a standard way to approach any kind of scientific inquiry like this.

We'll see that it moves down to different levels. As we approach this, we frame this as an assumption, but I think it's more useful to frame as a hypothesis – this second point here. We're essentially hypothesizing at a very high conceptual level that the New gTLD Program has had some effect on competition, choice, and trust.

At a more nuanced level, the next level down, we're saying that the New gTLD Program has had an effect on trust in some way. Then, of course, at the level where I'm now, we're hypothesizing, essentially, that the safeguards included in the New gTLD Program has had some effect on DNS abuse, which we're assuming is a proxy for trust. And that's okay. We'll want to explain that as we move on, but that's a fair proxy.

Moving down into the levels of each safeguard, we're saying vetting registry operators has had some effect on DNS abuse. That's our hypothesis.

Now, notice I say "an effect." I think we need to very careful to not use any kind of normative language [that] doesn't have a positive or negative effect. We want to just describe an effect now, at least when we're framing hypotheses.

Here's a pictographic representation of these hypotheses. It's a conceptual map, and I encourage everyone as we gather data to think about which side of the variable or causal relationship our data points will go on. Is this an explanatory variable? Does it go on this left side of the chart? Are we trying to say that that explains or are the causal factors leading to some results or outcome on the response side of things? Can we look at the number of new and legacy domains and see that that is somehow tied or correlated to a DNS abuse response? So that's just a way of pictographically looking at it.

I'm going to move real quick here because we just got a couple more minutes. We have these potential data sets. This is all included in the DNS abuse report. There's obviously a lot more data we can talk about,

and we already have. But the methods of analysis to demonstrate causality are here on the orange part of the screen here: qualitative, quantitative. You'll notice I have surveys listed under both. Depending on the size and scope of the survey, you can apply statistical methods to survey answers. But if you want that deep drill-down analysis, you can do more of a smaller-scale qualitative study with deeper questions.

Here's how this model works or how I envision it to work. Same kind of framework; cause/effect, independent to dependent variable on the right. Let's look at the high level, where this is our hypothesis here and what kind of method we should apply.

Restating the hypothesis at this high level, the New gTLD Program has had some effect on the proportion of abusive domains. However, we choose to segment that. On the left side, we're going to have a set of descriptive statistics – Carlos alluded to this – and on the right side we're going to have a set of descriptive statistics. But there's nothing to correlate those two in a causal way. We can't say one has caused the other without inferential statistics, right?

So there's going to be a set of numbers on one side, a set of numbers on the other, and we want to, number one, calculate this correlation between the two; determine what the correlation is. But as we know, correlation isn't causation. But to get at that causal factor, we want to calculate the probability that that correlation is due to random chance. That's when we get into the ideas of statistical significance, and when we can start making some claims of causality, although we have to be quite cautious of that.

Moving quickly to the next point, it's a similar kind of level. We have lots of descriptive stats on DNSSEC deployment, just to take an example. We have lots of stats on abuse rates. We'll want to carry out the same kind of exercise to determine if there is a potential causal relationship. Remember, stats can only tell us so much.

Incidentally, John Oliver did a great thing on his show last night about scientific studies and the weaknesses of stats and correlation, so I encourage everyone to watch that. It's fun.

But then if we don't have numbers – these big, nice numbers that we can run statistical analyses on – we can fall back on qualitative inquiry. Let's just take the expedited registry security request process as an example. It's an existence versus non-. It's about a process. It doesn't generate a lot of quantitative data.

That's when we can rely on survey or interview methods to get into that causality, but you're not going to have that nice numerical causality. You have to frame your questions in a causal way, always keeping in mind what your hypothesis is and what you're trying to test as you build these questions out.

I'm sorry I can go into more detail right now just because of time. My way of conclusion is to just go through these points very quickly. Having a good research method and structure is essentially a work plan. If the methods are arranged systematically by data type and within a model like this, it takes you a long way, and you just fill in the gaps. I say "just fill in the gaps." The data gathering is the messiest, most difficult part of this, as we all know. So it's not that easy.

Always try to keep these hypotheses in mind. Whenever you're doing quantitative or qualitative work, just having some kind of arrow-diagram-type structure is great. It's very useful. Mixed methods are great. Quantitative work can give a breadth of an idea, and qualitative will give depth.

You always want to keep things as measurable as possible. These methods are applicable in all areas of the study if we wanted to use it. Also, I know we have some budgets and economics consulting or research firm is really money well spent on this because it does take a ton of legwork and some expertise, very specialized expertise that I don't have. I worked with teams like this before, and they can get a lot done in relatively short time, especially if we have a model and framework to hand to them.

So that's my presentation. Sorry I had to rush through it, but I hope it was useful. I will end it there and take any questions.

LAUREEN KAPIN:

Thanks so much, Brian. I realize we're running short of time. Maybe I'll ask people to indulge for a few extra minutes. Do folks have questions for Brian? I also might put in a request now, Brian. Are you planning to be participating in the June meeting, or is that something you're not traveling for?

BRIAN AITCHISON:

I was not committed either way. If it makes sense, I'm happy to be at the June meeting. I can talk to Karen about it.

LAUREEN KAPIN: Okay. Well, I'm just thinking that at some point, whether it's the June meeting or another telephone conversation, we probably really want to do a deeper dive with you into best ways to move forward and get a little more information about what data is available regarding certain safeguards and what approach, quantitative or qualitative, makes sense within specific safeguards. That's the sort of thing that I think we'd benefit from a more in-depth opportunity to discuss this with you.

So whether we do that in June or whether we do that in a phone call, I want to put that as an action item: that we need to set aside a time to have a more in-depth conversation with you about it.

BRIAN AITCHISON: Yeah. I totally agree, Lauren. I think just getting us in a room with a whiteboard and maybe a matrix or framework to discuss the data we have, the kind of methods we could apply, would be a very useful exercise.

LAUREEN KAPIN: Then I'm thinking that this would be something we'd want to do face-to-face, and since the next time we're meeting is in June, if we can work that into our schedule, I think that would be a great use of our time.

BIAN AITCHISON: Okay. I'll look into that, Lauren.

LAUREEN KAPIN: Okay. And if we can't do that, then let's figure out an alternative. So let's you and I stay in touch on that.

BRIAN AITCHISON: Absolutely.

LAUREEN KAPIN: At this point, questions for Brian?

BRIAN AITCHISON: Does Jonathan have his hand up?

LAUREEN KAPIN: Does Jonathan have his hand up? Yes, Jonathan does have his hand up. Thank you. Jonathan?

JONATHAN ZUCK: Yes. Hey. Can you hear me okay?

BIAN AITCHISON: Loud and clear.

JONATHAN ZUCK: Okay, great. As you might be aware somewhat the idea of data-driven analysis, so I really enjoyed your presentation and methodology that you're laying out for this study.

When you got to the slide about potentially having to switch to a qualitative analysis, I understand how that comes to pass, so to speak, but I what I wanted to do is try to contextualize that content in our intention towards continuous improvement.

So in addition to defining – I’m trying to come up with neutral terms, so just forgive me in advance for saying something more loaded than I intend. So the success or failure of the New gTLD Program in improving consumer trust – I don’t know if there’s softer words than “success” or “failure” because it’s never going to be absolute. But [one you] stated to identify the degree to which it improved or the degree to which decreased trust and then make a set of recommendations as a review team that, once implemented, there’s some mechanism that’s using the same data to determine whether or not the recommendations were successful.

So I’m curious about your ideas about how that intention would translate with a more qualitative analysis.

BRIAN AITCHISON:

Sure. In the initial stages of defining the research method and scope, I have deliberately avoided what I call normative terms of improving and success. That does not mean that we can’t have an ultimate aim of creating a report that contributes to those goals of improving and being successful.

What is important to do is to define success and improvement in very measurable terms so we can, as you said, Jonathan, continue to assess the effectiveness of safeguards in a similar, replicable fashion, which is

also a key component of the scientific method. Studies need to be replicated and able to be replicated. So I think that's totally a great objective, and I think we'll just need to work more on how we define improvement and success and concepts like that.

I hope that answers some of your question.

JONATHAN ZUCK:

It does. It sounds like we're on the same page. You don't need to make a value judgment on that, but I think the review team as a whole is starting with the assumption that areas where trust is enhanced are positive and from which it was detracted is a negative and should be the source of recommendations for improvement, right? So being able to, as you stated, define success and therefore improvement is critical, and making that measurable is critical. So that's my only hesitation about a shift to a qualitative analysis.

BRIAN AITCHISON:

Right. No, that's fair. I think we are on the same page. I would just emphasize that, as we approach the initial stages of the study, I think it's important to frame things in terms of positive, null, or negative results and accept that there can be no results from some of these quantitative and qualitative tests we do. So it's up to us to really be deliberate about how we define these efforts.

So, yeah, I think I'm just repeating what you're saying. But, yeah.

Is there a question from Fabro?

FABRO STEIBEL: Yes. Can you hear me?

LAUREEN KAPIN: Fabro? Yes.

FABRO STEIBEL: Yeah. My question is – well, it's not actually a question. It's more of a comment. It's really interesting, as they say, to try to not only describe but to [inaudible] evaluation we have a legacy opposite to new TLDs. So think this [inaudible].

I have two comments. The first one is, when we try to define what success is, there are two key issues, two key concepts, we should somewhat define. The first one is the [inaudible]. How we define abuse [inaudible] right. So we define abuse is different ways [inaudible] define we will [inaudible] as well.

The second one how we define consumers. When we define the definitions of – using this review team, we get two types of consumers. We have registrants and we have end users. This will be key, again, to define success.

Of course, we can work around that and find some good things we can model and try to not only describe but also explain.

[inaudible] stage [inaudible] describe more empty space. I'm not sure if the space on different groups [inaudible] and different data sets you have and so on. You're going to have some ways to explain. You're going to have more to describe. I'm saying that as quite [upset]. I hope you can explain more because then we will be better.

But I think we're going to [inaudible] describe more and explain less. This is just a thought. But just [inaudible] from the point of view of [do you think] that we're going to be able to explain way more or how we think we're going to define success and measures. [inaudible].

BRIAN AITCHISON:

Sure, Fabro. I agree as well. I made a distinction between description and explanation. It's a conceptual distinction, I suppose – an important one, but at some level, especially at these early stages, you're going to have to rely on descriptive stats, narratives, user descriptions of experiences that are going to – with these safeguards that might seem anecdotal. As you build that research corpus around those descriptions, whether it's quantitative or more qualitative, the way it tends to work is you start seeing patterns and you're able to formulate hypotheses around these specific relationships more clearly.

So think you're right. It was kind of hard to hear you, but at this early stage it's perfectly all right to be at that descriptive literature-review-type stage, seeing what the state of affairs, how people are talking about DNS abuse and that kind of thing, and their experiences with it.

But down the road, we'll want to get more into what I call the explanation of how, say, a safeguard contributes to a rate of DNS abuse, however defined. There might not be a nice, neat numerical rate of DNS abuse, and we'll have to just rely on a user feedback kind of metric for how effective something was.

So, yeah. I agree.

Does Carlos have a question?

CARLOS RAUL GUTTIEREZ: Yes, please.

LAUREEN KAPIN: [inaudible] Go ahead, Carlos.

CARLOS RAUL GUTTIEREZ: Thank you. Yes. I already hinted in my comments that, for me, the approach is very literal – sorry. There is some noise here. I don't know if you can hear me well.

We see just from the compliance side if the safeguards are effective or not. What I worry about is that the safeguards may be fine and very effective, but in the end there is erosion of trust.

Sorry, but this is why I mentioned it. It's very compliance-oriented. It's very [inaudible] because this is the way it's specified in the contract. But what if there is erosion? Although the safeguards look pretty fine, where is the sailor? Do we need more safeguards, or do we need other remedies?

When I go and look at the metrics, in the real world what is important is the security breaches and takedowns. This is what erodes trust in the Internet or erodes the trust in e-commerce, or erodes the trust in the judiciary, as in the case of Brazil, which took down [WhatsApp] for a few days last week and so on.

So I agree as for the documents existing. We're looking at this side of safeguards – the [inaudible] side of safeguards. In ICANN we're working

with called compliance. In the rest of the world, it's called regulation, but you cannot use that word here.

But then there is the impact and the effective remedies, and those are, of course, outside of the relationship between ICANN and their registry operators. But what if the safeguards are fine but there is erosion in trust? That might be the issue for our next study, but I don't fail to see that your proposal is very solid. But it's very literal. Thank you.

LAUREEN KAPIN:

Okay. I'm going to jump in now just because we're now over time. I'm mindful of people's time and also mindful of my own upcoming appointment.

Two quick items. Brian, consider how we may want to respond to Carlos' comment, and you can do that to the group or we can have you appear at our next conversation.

But two quick administrative issues. One, the result of our Doodle poll indicate a strong preference for the Tuesday at 13:00 UTC. If I'm mixing this up, I'll ask to staff to jump in, but believe that was the preference of most of the members of the group. So we will be shifting to that time, and I will send out a revised schedule of presentations because will need to shift a little bit.

Fabro, I'm sorry that we didn't get to you this time. We will get to you next time. That's second. I'll be sending that out. Three, Alice suggested in the notes that perhaps – and I think this conversation illustrates that – we might want to have these sub-team meetings last for a longer time. I'm proposing that we shift to 90 minutes. Unless I hear via e-mail

or the chat that folks have a really strong objection against that, I am going to start scheduling calls for 90 minutes so that we have sufficient time to get through what we need to go through.

So those are the three quick procedural items I wanted to raise. If anyone wants to have a quick response now to any of those topics, now is the time. But you can also respond via e-mail to the group, and we can continue deliberation on these topics; one, the new time, and, two, the new duration.

Okay. I see Carlos is suggesting we go up to 90 minutes for the presentations only. Maybe that's a good approach; that we do it 90 minutes for the presentations and then only go beyond 60 minutes if there's a special topic that needs attention. That's something to consider.

Any other comments? Carlos, your hand is still up but I'm assuming that's an old hand. Okay.

CARLOS RAUL GUTTIEREZ: Yes. Sorry.

LAUREEN KAPIN: With that [we're at] 15 minutes over. Does someone else want to chime in?

Okay. Then we'll conclude, and I will send out information regarding the time and duration of our next call and revised presentations. Again, if

folks have thoughts on any of these issues, we can circulate it to the group via our circulation list.

Thanks, everyone, especially the presenters and especially our guest, Brian. Thanks to the staff for their great support, and thanks, everyone, for your patience with the longer discussion today.

BRIAN AITCHISON: Thank you very much.

[END OF TRANSCRIPTION]