

**WHOIS Proxy/Privacy Reveal & Relay Feasibility Survey Webinar
TRANSCRIPTION
Wednesday 15 August 2012 at 20:00 UTC**

Note: The following is the output of transcribing from an audio. Although the transcription is largely accurate, in some cases it is incomplete or inaccurate due to inaudible passages or transcription errors. It is posted as an aid to understanding the proceedings at the meeting, but should not be treated as an authoritative record.

Coordinator: Thank you very much. Over to you, Barbara.

Barbara Roseman: All right. Welcome today. We're using this opportunity and an earlier webinar that we did this morning - or earlier in the day - to review the results from Whois Study 4 which was on proxy privacy reveal and relay feasibility.

And this was a study that was taken up a couple of years ago at the behest of the GNSO who have determined that they need more solid data in order to suggest any potential changes to the Whois policy and practices.

Many of the considerations have to do with the ability to reach the domain users as part of that. And so, you know, we wanted to explore how the proxy and privacy providers fit into that discussion.

This study began with one set of directions and moved to a different set based on results that were found at the very beginning of attempting to perform the initial study. So it was intended to be an in depth study into communication relay and identity reveal requests sent for gTLD domain names registered using proxy and privacy services.

However it morphed into something else and I'm going to turn this over to Lyman to make the presentation - Lyman Chapin of Interisle - and to go through the results that they found. Lyman.

Lyman Chapin: Yes, thank you, Barbara and welcome to everyone who's come out to join us today. As Barbara said this study was undertaken in part because the initial expectation of the GNSO Council with respect to a study of privacy and proxy registered relay and reveal requests showed that there was - it was very difficult, if not impossible, for people to respond because there were substantial uncertainties concerning both the data that might be available from such a study and the willingness and ability of important principals to participate.

And as a result when the RFP for the original proxy and - excuse me - privacy and proxy relay and reveal study was published back in September of 2010 what we found was that the responses - there either were no response to the RFP or the responses basically said we can't estimate the study cost or duration because there are too many uncertainties.

And as a result the GNSO decided instead to first undertake a feasibility survey to try to fill in these gaps, to try to find out, for example, how much and what kinds of data would be available if a full study were undertaken and who the participants might be.

Would they, for instance, represent a sufficiently broad cross section of the community that's involved in both operating and querying these privacy and proxy services? So the GNSO Council in April of 2011 authorized a feasibility study. And Interisle Consulting Group was selected to conduct that study.

We originally set the study up in such a way that anyone who responded - and we recruited respondents in addition to leaving the survey open to anyone who wanted to participate - we set it up in such a way that people could respond without revealing either their identity or any of the data that

they might have available to ICANN. So this was a survey that was undertaken independent of ICANN with those confidentiality guarantees.

We started with a Web-based survey. We used a survey tool that we translated - we presented in five different languages. We gave people the option of taking the survey online in any of the languages that you see listed here.

And we deliberately sought respondents from as broad a range of geographical regions and as broad a range of interests as possible while, as I said, keeping the survey open to anyone who wanted to respond so it was not limited just to the people that we solicited. But we did spend quite a bit of time and effort in outreach.

The questions covered, as you might expect, the way in which the potential participants, the people responding to this survey, thought that they could contribute to a full study. By full study here we mean a study intended to satisfy the objectives that the GNSO Council originally had when they prepared the RFP back in 2010.

We were particularly interested to know what kinds of data could be made available by relay and reveal originators, relay and reveal processors and other intermediaries such as registrars. And also what level of willingness and ability those potential participants had if they were invited to participate in a full study.

We conducted the online survey - oh, excuse me - from September, 2011 through the end of October. It was hosted on a server controlled by Interisle in order to ensure that we could guarantee that - to the potential respondents that if they wished their responses could be kept confidential.

And we ended up with 168 people who completed the online survey which gave us a pretty good data set to work with. We supplemented that with

follow up interviews both with some of the people who had responded to the survey who perhaps had suggested that they had additional information that they might be willing to share with us and also to ensure that we had an opportunity to talk with some people who may or may not have taken the survey.

We couldn't necessarily tell because some people did, in fact, take advantage of the opportunity to remain anonymous. But whom we knew to have good perspectives or knowledge about the way in which registrant information is maintained behind proxy and privacy services.

We learned a few things in conducting the survey and the follow up interviews that is not so much interesting from the standpoint of the conclusions of our feasibility survey but will be useful to the design and the implementation of any follow on study should the GNSO decide to undertake one.

In particular we discovered that a substantial amount of effort is necessary to actually contact the registrars and the proxy and privacy service providers that participate in the handling of relay and reveal requests. This required a considerable amount of effort; much more than we had expected.

We were somewhat surprised that only 11 out of the total 168 survey responses were in a language other than English. And it would be tempting to conclude from this that any follow on study could be conducted entirely in English but of course we can't leave out the - or rule out the possibility that even though we offered the survey in five different languages it was still not convenient for people whose native language is not English to participate.

The geographical distribution of the responses to the online survey were somewhat predictable, a little bit disappointing from a coverage standpoint; 60% from North America, 30% from Europe. Most of the rest of the responses were from Asia-Pacific; extremely low turnout from regions other than North America and Europe.

And this was despite an extensive outreach program which included very aggressive and targeted efforts undertaken by the survey team during the ICANN meeting in Dakar which took place during the period in which the survey was open and respondents could go online and take the survey.

So even with that kind of outreach and even outreach centered in areas that were underrepresented we found that most of the interest in this topic seems to be concentrated in North America and Europe.

I do have a few slides that I'll show you that summarize the findings of the survey. I will not - I'm not going to go into tremendous detail. Feel free to, if you have questions about this we'll have plenty of time at the end to ask questions about details. And I'll be happy to provide those.

One of the most interesting and in retrospect, perhaps, predictable findings with respect to responses from different constituencies is that the people who are interested in obtaining registrant information, people who make relay and reveal requests, were far more likely to say that - were far more likely to participate in the survey and also to say that they would be willing to participate in a full study than the people on the other side of the process, the people who provide proxy and privacy services and who receive or respond to the relay and reveal requests that seek information - that seek registrant information.

That's probably not surprising. The people who seek information probably are more highly motivated; they have specific reasons to be perhaps frustrated at the moment by some of the difficulties they encounter in obtaining the information. The people who are interested in making it more difficult to obtain that information probably are quite comfortable with that.

So this is something that would be, you know, that should be taken into account if we think about what a full study should - how a full study should be designed.

The - because we noticed that in the online survey we made a deliberate attempt in selecting people for the follow-on interviews to strike a somewhat more even balance between those who make these requests and those who receive and respond to them.

Among both requesters and responders, both the relay and reveal requesters and the organizations and individuals who provide proxy and privacy services, there was a general sense - less than 50% but still a sizeable fraction of the population that responded - would be interested in participating. Not surprisingly a very large percentage of all three groups said that they would be interested in the results.

And in the - for the other three questions it's probably interesting to note that, again, as we saw earlier the people who are requesting information, who are sending relay and reveal requests attempting to obtain applicant - or sorry, registrant information - were more likely to say that a full study would be beneficial either to themselves, their organizations or their clients and customers or to the Internet community at large.

The things that we discovered would be negative factors, meaning things that would deter potential participants from participating in a full study, were for the most part on the provider side had to do with revealing confidential information, whether their own or client information.

On the requester side some of the same concerns were evident but to a far lesser degree. So requesters tended to say, for instance, that the time and effort required to participate in a study would be a negative, a deterrent whereas in the case of, for example, revealing client confidential information 60% of the privacy and proxy service providers felt that that would be a

deterrent, something that would make it difficult or impossible for them to participate.

On the positive side things that we could do, for instance, if we were designing a full study, things that we could do that would encourage people to participate - the two that stand out here particularly for the providers of the proxy and privacy services are privacy guarantees and the ability to control access to what they might consider to be sensitive data. And that's a theme that recurred many times during the course of this study.

One thing that emerged from both the survey and the interviews that we hadn't necessarily expected - two key features; first of all almost everyone we talked to and many of the people who provided additional comments in their responses to the online survey made it clear that they were interested in a full study only if there were some guarantee that this time around something would actually be done.

There was quite a bit of frustration expressed about the way in which attempts had been made over a fairly large number of years now to solve the problem of managing access to registrant data.

And almost across the board uniformly with people that we talked to their sense was whatever you do if you're going to go to the trouble of conducting a full study and if we're going to go to the trouble of participating in it there had better be some concrete tangible result. We don't want this to be just another study that gets conducted, some nice data get published and then nothing happens.

Now that of course is something that's difficult to accommodate. It's very difficult to design a study of any kind that can guarantee the outcome. But I think that this is something that the GNSO and the Council and the community should take into account as they decide if and how to conduct a study and how to present it to potential respondents if they do.

It was also clear that particularly - and perhaps for obvious reasons - on the proxy and privacy service provider side there was a very strong sentiment that they would have no interest in participating if the study appeared to be simply an effort to find bad guys and call them out.

If this were - if a study were to be presented or were to come across as pretty much just a witch hunt, you know, where you framed all the questions in such a way that it was clear that people who were quote, hiding behind these proxy and privacy services were bad actors, that you would be very unlikely to get any of the people in the business of providing those kinds of services to participate in a study.

So that suggests that any full study that's contemplated in the future should be careful to be neutral with respect to the issue of whether or not there should be ways to make it difficult for people to obtain registrant information.

And then of course, as I've said before, having good privacy and confidentiality guarantees would make it possible or easier for many of the potential participants to provide data.

However even in cases where adequate privacy and confidentiality guarantees were in place it would still be difficult for many of the people who said that they had data to provide the kind of data that were anticipated in the original GNSO RFP. That is individually identifiable data that would enable you, for instance, to track an individual relay or reveal request from its originator all the way through the system and back.

Aggregated data were readily available. Individually identifiable data are not likely to be readily available even with privacy and confidentiality guarantees. So that's going to loom large as we get to the conclusions.

The first conclusion of the survey is that a full study could provide some, but not all, of the data anticipated by the GNSO. And by that I mean that, for instance, some data might be available only in aggregated or non-individually identifiable form.

And some data that the GNSO expressed an interest in obtaining, such as examples of actual individual relay and reveal requests and how they were processed through the system, might be either very difficult to obtain or might be obtainable only in a way that made them unrepresentative, in other words statistically not valid because of the very limited set of participants who would make that information available.

So a full study could be conducted; it might not provide some - it might not provide all of the data that the Council originally anticipated being able to obtain by conducting the study.

However it would be well received by people on all sides of the registrant information access debate. That was a bit of a surprise. What we found was that although you would expect people who make relay and reveal requests to be eager to have such a study because they, in many cases, have difficulty obtaining information that they need, we also found that many proxy and privacy service providers were also eager to have a study conducted because they believed that their operations were completely legitimate and that a relatively small number of bad actors were giving them a bad name.

And so they actually wanted to see a study conducted because they felt that some popular misconceptions about proxy and privacy services could be dispelled by a rigorous study that showed that in fact most of those providers were operating completely legitimate and above board and only a few bad actors were giving everybody else a bad name.

The third conclusion is that the - both the quantity and the quality of the data that could be obtained by a full study would be improved by ensuring

confidentiality and convenience for the participants. But it would not entirely overcome the reluctance of some participants to contribute.

So what that means is that we could get good data if we included confidentiality guarantees and we designed the studies - the full study - so that it was convenient for participants, didn't require a lot of time and effort.

But we would still find asymmetry in the ability and willingness of different constituencies to contribute. There's a very big difference in our - in the survey results between the willingness of people who make relay and reveal requests to participate and the willingness of people who process them; the people that run the proxy and privacy services.

It's about a 2 to 1 disparity. So that asymmetry would almost certainly exist no matter what kind of confidentiality and convenience and so forth you built into the study design.

And finally the last conclusion is that if a full study were conducted observing all of the different constraints and requirements that this feasibility survey suggests would be necessary to get people to participate the results might not satisfy the original expectations of the GNSO Council - and by extension the community that contributed to its work - with respect to either statistical validity of the results or independent verifiability.

And the important point here is that if a full study - if the GNSO Council decides that it is worth going ahead with a full study it should do so understanding that it is not likely to produce the data that were originally expected when the RFP was first put together in 2010.

So I'd like to stop here for a moment and allow people an opportunity to ask questions. We'll finish - at the end of the webinar we'll finish with a summary of the next steps. But, Barbara, perhaps I'll turn it back over to you and you can see if there are any questions from our audience.

Barbara Roseman: Okay thank you. If you have a question you can indicate that either by raising your hand in the Adobe Connect room or just speaking into the, you know, to the conference bridge and I'll try to maintain a good queue. So are there any questions to this point?

Okay, Lyman, if there are no questions from the participants then I think I have a few questions I'd like to raise. Can you talk about whether there were any constituencies that may have had unique concerns about participation such as law enforcement or any of the other groups?

Lyman Chapin: Yes, as you might expect the different constituencies did have different concerns about, you know, how they would - how they would view a study - a full study and to what extent they would be able to participate.

The people who operate proxy and privacy services they were not just reluctant to provide data because they're in the business of, you know, trying to limit access to data but they, in many cases, said that their contractual agreements with their customers would prevent them from doing so even if the guarantees of confidentiality and privacy and so forth were in place.

So a number of them said that although they might be willing to share aggregated data they would find it very difficult and in many cases impossible to share any individually-identifiable data such as data concerning a specific relay or reveal request as it went through the system.

Both because they, in many cases, don't maintain that information for - again for customer confidentiality reasons, but also because in some cases their business contracts would prevent them from doing so; so that was one very specific concern.

Law enforcement agencies actually turned out to be the most likely to offer active assistance to suggest that they might be in a position to actively contribute to a study by participating in the search for sources of information.

There was some concern about data privacy regulations mostly in Europe and almost certainly related to the EU privacy directive that is a fairly stringent control on how organizations can share data about their customers and clients. But for the most part the greatest willingness to share information was found among the people who make relay and reveal requests.

Barbara Roseman: Okay. Margie has a question.

Margie Milam: Hi, yeah, this is Margie Milam on staff. I found interesting one of your findings related to the call for the study to lead to something I guess, you know, that if it were to be conducted that the participants will want to see some action resulting from it.

And I don't know, Lyman, if you're familiar with what's been going on with respect to the RAA negotiations. But as part of the RAA negotiations that are the concepts that ICANN would create an accreditation program for privacy and proxy providers the parameters of which would need to be developed.

So it seems like, you know, with that information, you know, if the next version of the study, you know, would actually proceed that we might actually be able to address that concern because it seems like some of this information might be useful in crafting the parameters of such an accreditation program. So I just wanted to raise that to your attention because it seems like the results seem to be timely given what's going on in the RAA negotiations.

Lyman Chapin: That I think would help a lot. And it's probably useful in that context to remind everybody that the survey was conducted last September and October so it's almost a year ago. And there has been a lot of progress in the RAA context

on a number of things that I think would make it much easier to conduct a full study. But you're absolutely right.

Margie Milam: Thank you.

Barbara Roseman: Okay and Lisa has a comment in the chat room. "Could you comment on the ability of a full study to examine current practices?"

Lyman Chapin: Yes, it certainly seems both from the survey data, from the online survey data and from the interviews that although various people in different constituencies had concerns about sharing data we found essentially no reluctance to share information about the way in which relay and reveal requests are processed.

So we feel quite confident that a full study would be able to create and document a fairly complete picture of how the system works today as sort of a as operating on the ground description of the way in which these services work.

Like I said essentially no one said that they would be reluctant to describe to someone conducting a full study how they operated. They might not want to say anything about, you know, the specific data that went through the system but they seem to be perfectly happy to describe the system itself.

Barbara Roseman: Okay. And Steve Sheng asks, "Do you know what kind of aggregate data they would be willing to share?" Like what would that data consist of basically?

Lyman Chapin: Well a number of people referred to data sets that are available in the public domain as part of the trial record of legal proceedings. Mostly what they're talking about when they say aggregate data is, you know, the number of reveal requests that are processed, you know, per week or per month and the

number of - and of those the fraction that are successfully processed and the fraction that they reject as being invalid.

If you're a privacy service provider you typically receive any number of requests to reveal information about the registrant who is your customer that are bogus; they're not - they're either invalid on their face because they're from sources that can't be verified or they are not presented in a way that gives the privacy service provider any reason to comply, they're not associated with a warrant or affidavit or anything else.

So a service provider might provide aggregate statistics about the number of requests that it received, the number that it was able to successfully process and the number that it was - that it rejected. So it's a - it's the kind of information that a business would maintain as a matter of course.

Very few of the respondents to either the online survey or the interviews that we conducted showed much interest in implementing new procedures or building new tools to gather data that they did not already collect as a matter of course in the operation of their business as it exists today.

So we did not find much willingness on the part of potential participants in a full study to instrument their operations with new tools that would impose on them some overhead of collecting new information. They'd be perfectly happy to give us the kinds of dashboard aggregates that they maintain already.

Barbara Roseman: And were there any regional considerations about sharing that data? I know that there's some regional and local laws that often make sharing individual data challenging.

Lyman Chapin: Yeah, the only one that appeared from the data that we collected in the survey was a few people who mentioned the EU privacy directive as a potential impediment to sharing information.

We didn't - except for the European region we did not hear from any other constituencies that they felt that there would be regulatory barriers to participation.

Barbara Roseman: Okay are there any other questions from the audience? If not why don't we go ahead and - oh wait, I see somebody typing in the chat room. Having heard these questions is there - are there any other issues that you would like to raise, Lyman?

Lyman Chapin: No just - I think it probably bears, you know, restating, not to belabor the point but it will be extremely important for the GNSO Council and the community to carefully consider the results of this survey in setting expectations for any full study that they might decide to commission.

I'm confident that with the information from this feasibility survey that we've just described that a new RFP could be constructed which would, in fact, be much more successful than the one that was issued in 2010 because we have more data available on, you know, what would be necessary to conduct the study.

But if that is - but that should be done with the understanding that the original objectives of the study that was contemplated by the GNSO Council back in 2010 are almost certainly not achievable and a different set of objectives might be but that will require some careful consideration.

Barbara Roseman: Okay thank you. That's very helpful. Why don't we move onto the next steps slide and take a look at that.

Lyman Chapin: Okay. So the next steps; the report has been published; it was published at the end of the May. A public comment period was opened through the 4th of June. And a reply period was opened and has recently - and has been extended; the reply period is still open. It's been extended to the 22nd of

August so comments on the survey and its conclusions can be submitted to the public comment area until the 22nd of August.

After that point a final report will be published. And the expectation is that the GNSO Council then in consultation with the community will use the information in that final report in order to make a decision about whether and in what way to commission a full study into relay and reveal handling for gTLD domain names.

Barbara Roseman: Okay those sound like good next steps. And I know that right now we're planning on putting this on the agenda for - the GNSO is planning on putting this on the agenda for the Council meeting in September. We look forward to having you - everyone send any comments that you have to the reply forum that's still open and will be open for another week.

Glen, can you provide information on where people can download the - view the webinar and find the slides for this presentation?

And you might be on mute, Glen.

Well these slides and the webinar link will be available on the GNSO Website via the Calendar I believe. And the original report is also available on the GNSO Website.

So thank you again, everybody, for calling in and participating. We appreciate it quite a bit. And that concludes our webinar today.

END