WHOIS Task Force

TRANSCRIPTION
October 30 2006, 10:30 EST, 15:30 UTC, 16:00 CET

Note: The following is the output of transcribing from an audio recording of the WHOIS task force call on 30 October 2006. 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. The audio is also available at:
http://gnso-audio.icann.org/WHOIS-200610306-tf.mp3

http://gnso.icann.org/calendar/#oct
Invited Speakers: **Dave Piscitello** - Security and Stability Advisory Committee (SSAC)Fellow
:http://www.icann.org/committees/security/information-gathering-28Sep2006.pdf
Mike Zupke - Registrar Liaison Manager
http://www.icann.org/announcements/announcement-06oct06.htm
http://www.icann.org/compliance/reports/registrar_update_oct06.html.

Attendees: Jordyn Buchanan Task Force chair**,** Ross Rader, Tom Keller Registrar C., Steve Metalitz IPC

Staff: Maria Farrell, Policy Officer, Glen de Saint Géry, Secretariat

(Jordan): Okay, I'm back. And hopefully, now, the recording's going to start and we can…

(Dave Piscitello): (You're welcome). Is the quality of the audio from my end okay? Because (there's) - (we're having) a little bit of an echo.

(Jordan): Yeah, I think it's probably not 100% ideal, but it's certainly (unintelligible).

(Dave): Let me see if a can just pick the phone here (instead of speaker).

Hold on.

Is that better?

(Jordan): Well, that is a little better, yeah.

(Dave): Right. Why don't I just do this? I just have to re - coordinate the logistics on my desk here.

(Dave): So, this will work best if you have a copy of the presentation in front of you. And Glen sent the hyperlink. This report is - well, presentation is on the ICANN Web site under the SSAC documents hyperlink. And sort of a little bit of background, you know, the report is entitled Information Gathering Using the Domain Name Registration Records.

Information gathering is a term in security that is commonly used by both security professionals and attackers in describing basically the sort of investigatory or surveillance phase that is a precursor to an attack. It's also used in the forensics world as a term that, you know, represents trying to find out, you know, or gather evidence.

Evidence is always a tough word to use because evidence has a legal connotation, but trying to gather information about an incident to determine what happened. So it's part of what's called an incident response. We're going to use it more in terms of the surveillance and target acquisition context today.

So, the objectives of this particular effort that I undertook after you asked me, Steve some license for pouring some time and attention into, you know, into this matter was to essentially approximate the extent to which personal information might be extracted from domain name records as a starting point. So by no means does it suggest that the domain name registration records are in and of themselves entirely useful, you know, in a vacuum to go often and find address. They may be, but that's not necessarily the way that people would use them.

Steven Metalitz: (Dave), this is Steve Metalitz. Did you want to have questions at the end or could I interrupt with one now?

(Dave): Well, I think (that) I'm personally happy to make this interactive. So, you know, the only constraint is that the most questions we ask, the less I'll be able to get in. So, why don't you just some questions and do that, yeah.

Steven Metalitz: Okay. I just wonder what went up to this work that you've done. Were you asked by the Committee to do this? Or what was the step prior to your initiating this study?

(Dave): Well, the step - I guess part of this was sort of a growing sense that, you know, a fair number of people in the Committee talk a great deal about the, you know, the presence or absent of personal information in the Whois without a whole lot of concrete numbers or any sort of researcher data on, you know, the extent to which personal information is actually present.

So, I've had talks with the policy (staff), I had talks with SSAC, I had spend, you know, a fair amount of time in the (audience) in a number of the Whois and GNSO sessions in Marrakesh. And I concluded, you know, maybe it's worth my while to go off and just take a look at this. But there was no - there's no motive other than let me see what the numbers look like.

Steven Metalitz: Okay, thank you.

(Dave): Uh-huh.

So moving on - I'm not going to go over the definition of personal contact information. But just to be clear on, you know, on Slide 3, I had to choose a definition and this is the definition that I chose.

There may be other definitions of personal contact information, but for me what I was looking for was, "is there enough information here that could lead me to find someone", you know? And I didn't want to overload it by, you know, by adding any sort of legal (leads) or any sort of contacts with, you know, with any US or international regulations or privacy laws. I just wanted to come out with a definition that was a working definition for the study.

I think more interestingly was kind of identifying the methodology so the people had a sense that this was not a scripted program that ran through 60,000 records and ignored fields that didn't make any sense or that didn't compare to a string argument.

This was a hand-and-eye exercise. So, the way that someone might go and try to find, you know, a useful target of a Web attack is very similar to the way that I actually went through the domain name registration data that I have. So I began with a potential set of targets and that set of targets was a list or a set of 5000 registration records that I had filtered down from approximately 2 million. And the filter argument was Philadelphia, Pennsylvania. And I'll explain why I chose that filter target in a minute.

Steven Metalitz: (Dave) this is Steve. I just had one brief question.

Why would someone who is planning - why do you think someone who was planning a Web attack would want to know whether they could be confident that they're registering as an individual?

(Dave): No, no. That was an analogy, Steve. I mean…

Steven Metalitz: Okay.

(Dave): …this is the same method that people use to pick out a Web attack or the methods that I applied to find out if I could learn a great deal about individuals to the extent that I might be able to visit his address or make a phone contact or otherwise communicate with them.

Steven Metalitz: Okay. But you're not making any statement about likelihood of Web attacks against the…

(Dave): No, no, no. Not at all.

Steven Metalitz: Okay, fine.

(Dave): This is just the way that people do it in the Web world.

Steven Metalitz: Great, thank you.

(Dave): Sure.

So, in addition to the domain name registration records, you know, I've decided I'd, you know, that I would also follow the - their methodology and I'd go and I'd use as many publicly accessible resources to collect different credit information and sort of tie together things I was drawing out of registrant and the administrative contact information.

And the goal here was to essentially profile or piece together as much information about or derived from, you know, initially from record to - see, there was a high confidence that the registration record could – or said to be a personal contact or a business contact.

Just for you, you know, for some, you know, anecdotal information, there are plenty of publicly accessible resources beyond those that I used. And if I had wanted to spend a fair amount of money and invest, you know, the same kind of time that a private investigation might employ, I could use things like, you know, the Federal Election Committee FTP Site, searches on (Dot Net).

Searches - and (Dot Net) is actually a pay portal where you then have access to 36,000 databases that are not visible in Google searches but are visible through this pay portal. And they have significantly more information -- personal property records, licenses, criminal records and the like. And they're used by law enforcement investigators, private investigators and other people who can go fairly far into a personal - a person's history -- they can get information from the newspaper clippings and things like that. And so that - I didn't use those resources because I thought I was - that was going a little bit too far. It was also fairly expensive.

So really, you know, what most investigators tend to do is use the information like Whois as a, you know, as a first step, much like as they would if they would go in Google and use that as a first step in a, you know, criminal investigation. And similarly, I'm just looking at the first step kind of situation.

And if you're interested in, you know, the domain names of some of these places that you can get this information, I'm more than happy to share my list with you. Just send me an email, okay?

And the resources that I did use are listed on Slide 5. So I, you know, obviously, I gathered registration records using a bulk CD purchase from, you know, from someone, you know, who offered them off the Web. And they were acquired in bulk using Whois protocol. At least that's what they say they do.

Then, I also used real estate database, and I'll explain, you know, how I used that, Internet, you know, telephone directory, I'll explain how I used that as well; search engines; aerial photographs of, you know, locations within Philadelphia to distinguish a business residence from an obvious (Cul-de-sac) in a, you know, a nice, little quiet suburban part of, you know, of the City of Philadelphia; emaps to, you know, to also compliment the aerial photographs; companies and (unintelligible) directory called (unintelligible); Web sites; and then, you know, I lived in Philadelphia for almost 30 years, and so I used my knowledge of the area as sort of, you know, additional confirmation. So I know some of the neighborhoods, you know, well enough where I can look at that address and, you know, I might have a friend who lived a block away. Or, I did, you know, I shop at this particularly or I know this particular area's downtown, commercial, not-a-soul-lived-there, you know, kind of environment.

So using all those recourses -- and I'll explain the matching, you know, process in a moment -- I tried to classify the results in the following, you know, broad categories.

There's a personal contact. And what that person will contact was was an individual where the registration's name was an individual's name and the other fields contain some sort of personal contact information, okay?

There was also sort of a subcategory here where it was a home-operated business and the registrant operated under a fictitious name like (mary'sembroidery.com). And so, that was not a personal name, but everything else pointed to an apartment in a building in a residential area in a residential suburb, Philadelphia.

Business contacts were obviously the opposite. This is a place where people, you know, have employees, you know, operate under a fictitious name, have some administrative hierarchy in their information technology department and it was very obvious that was not, you know, a one- or two-man operation.

The other categories were domain name business. This were companies that I, yeah, that I saw a fairly large number of domain names owned by - and I went, you know, did a little bit more research into the company and discovered that they were in the business of buying and speculating on domain names in one form or another.

Domain name proxy agents, these - like other, you know, other areas, there are number of ISPs and a number of companies that will eventually register domain names on behalf of their subscribers or the customers. People who do Web posting do this themselves you know? So those people fell into the category of domain name proxy agents.

Then there was a - there were also records that were, you know, provided inconclusive data. Either there was - there were insufficient information in many of the fields, or the information was completely inaccurate, or, you know, the aerial photograph was, you know, was not very well distinguished, you know, but, you know, essentially not - that there weren't a sufficient number of matches in my criteria to actually come up with a conclusion that this is a personal contact or a business contact.

So, if you go to Slide 7, Slide 7 is essentially, you know, even though I show a picture of a fingerprint, that's more, you know, just for the purpose of, you know, of providing some, you know, some distraction from just more and more slides more and more words.

But essentially, the philosophy is the same. What I was looking for is there's a large number of things that, you know, of matching criteria. If you look at the ten boxes on Slide 7 and Slide 8, these are the kinds of questions I asked on each record that I examined. And so, I went through the 5000 records a record at a time. I went through 5000 records and I, you know, looked up the address in the real estate database. So I looked up neighboring addresses or blocks in the database. I looked up reverse telephone directories to see if they, you know, it'd result to, you know, to an actual individual or business.

I used search engines to find out, you know, something about the site based on the domain name or based on the individuals. I looked at the aerial photographs, you know, that the Google Earth provides to decide that, oh my God, this is a two-storey brown house right in the heart of the neighborhood, you know, that is clearly residential, or here is a nice little colonial on a cul-de-sac in, you know, in - near the suburb of Philadelphia -- so it actually still has a Philadelphia street address. I went to the Web site. You got a sense of whether or not this is a small company or a large company.

The things that I - that Steve Crocker and ICANN Legal and others thought were off, you know, off the plate in terms of pursuing, you know, pursuing this information to verify we're actually calling people or sending them email or driving up and knocking on the door. So I wasn't going to actually complete the, you know, the process and make contact. I wanted to see if I could get a high confidence that contact could be made.

Similarly, with businesses -- if you go a Slide 8 -- since we’re kind of running out on - out of time, again same kinds of questions, the same kinds of pursuit on each and every record.

Some of the things that, you know, that were sort of cues in business records - that will distinguish some cues in, you know, in residential records, for example, was, if there was an apartment number, it’s very likely that this was an individual. Whereas, if it was a suite number, it was very likely that this was, you know, a business operating in a professional office building of some sort.

The other thing that I did was use proximity. If you looked at the - through the database, you know, I might not have actually been able to find that the, you know, that the home or the particular address was actually on the market. But a home on the same block might have been on the market or on a block next to it. I could then and do some more examination using some mapping and things to say yes, this is, you know, it’s very clearly a, you know, a residential street; it’s very likely that this is, you know, a resident.

So, where are we now in my slides I don’t know where I am….

I think I’m a little out of sequence here on my slides.

So, we’re getting a little bit deeper into some of the findings, some of the information. In the records that I had, I was able to use about 4400 of the 5000 filtered records. That means that the others were just not sufficiently accurate or had inconclusive information to actually be useful.

Out of that 4400, as you see in the pie chart, you know, the vast majority of the records that I had in this particular case were com, net and org. Was it intentional? It's just the way it came out.

So, if you look at some of the findings on Slide 10, about 9% of the contacts based on registration contact fields alone were personal contacts, 56% were business contacts under the criteria and definition I had before, then you see domain name records, domain name proxy service and home-operated businesses.

Then, going on to Slide 11, what I tried to do here was simplify the findings because if you look at the pie chart it gets a little bit complicated. But one of the things that I combined here is the personal contacts from the home-operated business because I do end up getting to an individual. So, about 13.4 % are there. If you combine - so as business that are not in the domain name business but are business contacts, you get about 55%. And then, the domain name businesses where about 21.6%

Up until this point, what I had done was only use the information in the registry - registrant field. There should not be registration field. So, registrant field. Sorry, we'll fix that.

So now, I've decided that I would expand the investigations to say, well, what happens if (Joe) registers the domain name, joe.com, but he, you know, he has a friend, (Harry), who runs a Web server. And (Harry) put himself down, you know, as the administrative contract. And, you know, are there other - so to put - the question was, are there other individuals that might, you know, might be present in the domain name registration records than the registrant proper?

So, if you go back and you go through the records that contain personal contact, I found, you know, a few more that, you know, that were different individuals, more that contained, you know, the contact information or identified the business contact, and more that contains inconclusive data in the administrative field. So this would be administrative contact, administrative address administrative, administrative email and the like.

I didn’t go through the technical contact. Anecdotally, the technical contact looks like (it would) be the most well-maintained field, which is not surprising because it’s the one, you know, the one that most people, you know, rely on in operations when things go wrong.

Okay. Slide 13 just is some illustration of the - (probably illustrations of), you know, whether or not the registration name contained a first name and surname, or the administrative contained a first name and surname.

Slide 14. This is sort of a summary of the degrees of inaccuracy or incompletion. So when, you know, when I looked at the records, one of the things that was frustrating was that, you know, God, there are an awful lot of phone numbers missing here. So 25% of the records that I found useful where missing, you know, registrant phone number; eighty-seven percent were missing registration facts; 10% were missing admin contact, 11% were missing admin contact email, 12% were missing admin contact address and 50% were missing admin contact facts.

As a note and sort of an interesting subplot here, the bulk, you know, the service that I bought the CD from doesn’t give you the registrant email addresses. So, you can interpret that in a couple of ways. I seriously doubt whether they’re doing it because they are interested in keeping registration - registrant privacy. On the other hand, it might be that what they’re doing is using those in a different product (sign-on) them for people who want to put together a list, (unintelligible) and email. And in fact, that same provider, you know, at the time when they were in operation were selling email lists.

Steven Metalitz: (Dave), this is Steve. Can I ask another question here?

(Dave): Sure.

Steven Metalitz: Are you - if this is coming from registrant -- excuse me -- from registrar data, the registrars aren’t required to make available phone and fax number for the registrant. Or in - so, did you factor that in?

(Dave): I didn’t make any factors at all. All I’m doing is telling you what I have…

Steven Metalitz: Okay.

(Dave): …okay? So, I mean, honestly, Steve, I don’t have an agenda here. I wanted to (give the number), yeah.

Steven Metalitz: No, no. I understand. I’m just asking because some of this data is not what would necessarily be in the registrar, Whois data.

(Dave): Yes, that’s quite possible. And, you know, since I was taking the same tact that somebody who was trying to use what was publicly accessible was taking, I’ll have to take that into account.

And I’m not saying that, you know, that even the registration fact wasn’t available. It really wouldn’t have made any difference for me to try to go (cite) whether or not, you know, whether or not I was going to actually make contact with an individual.

You know, in the cases that I - that’s - where I was able to draw a conclusion, that was probably not as useful. Using the email wasn’t going to actually, in all likelihood, you know, get me that much closer to an individual, so.

Steven Metalitz: Okay, thanks.

(Dave): All right?

Steven Metalitz: Thank you.

(Dave): Yeah.

Okay. So some of the conclusions, you know, all we were trying to do here was offer a set of findings that, you know, that people might be able to mull over when they're discussing or debating, you know, whether or not there are some privacy implications in the way that, you know, the public is able to access records Whois services today.

What we were able to show in the study was about one and seven records provides additional information to identify personal contact.

And similarly, about one and seven contained institution information that quickly distinguished whether the contacts are businesses or individuals. The, you know, the other 84%, 86% appear to be business contacts of one form or another.

So, you know, in addition to those conclusions, some of the things that we, you know, that I was able to draw was that there's an awful lot of incomplete data. And what - you've got a very, very useful data point, Steve, by saying this is not necessarily the extent to which records are accurate in registrar and registry databases but, you know, to the extent to which records are accurate that are made available to the public.

So, you know, so one or the other, the only comment that I'm making is that there's an awful a lot of, you know, an awful a lot of incompleteness and inaccuracy in the information that I have. And that's, you know, you can treat that in any way you'd like. And if I had better registration - or if - there is certainly an open question as to whether I, you know, if I had received the same 5000 records from the registries, whether the information would have been more accurate. You know, I can't argue that.

The - but, you know, if its true that there was a preponderance of registration that - where the registrant withholds or does not provide facts and that has become a less and less useful bit of information, you know, one question we, you know, we could ask is, is there any purpose in continuing to ask for a fact?

You know, sort of a side answer to that is how often, well, you know, our facts are used to contact the registrant. And so, you know, these are just questions that come up when you look at the data. They're not necessarily criticisms. You know, I don't know how often registrars contact, you know, contact individual registrants with a fax. I certainly have never received one. But that's not necessarily, you know, something that we'd expect.

Tom Keller This is Tom Keller I might (be having a) question about your methodology.

(Dave): Sure.

(Tom Keller One question I have is whether you only applied all your methods to domain names that are in the area of Philadelphia and if you have tried to use the same kind of methodology with domain names such as in other parts or the world?

(Dave): I…

(Tom Keller And (if you're saying that that) would have the same result (on this)?

(Dave): Okay. So, answering the first part, just to give you a sort of a measure of the labor to do it the way that I did it, you know, I spent about 100 hours of my time and other (eyeballs), you know, in order to do this. So, it is - this is a very labor-intensive piece, you know, piece of work because for every record, I basically had to go through each and everyone of those steps of Google search and Google Earth and, you know, real estate databases. And in many cases, it wasn't just one query - it was several queries -- to try to get a high confidence level.

(Tom Keller Uh-huh.

(Dave): So this is a fairly labor-intensive way to go about doing it.

And I chose Philadelphia because I had, you know, I had the advantage of knowing something about the area.

(Tom Keller Uh-huh.

(Dave): The, you know, if I were to go Zurich, I would be disadvantaged because I don’t know very much about it even though I’ve visited there a couple of times. So I wouldn’t really be able to say to you some of their tools that I had.

However, I do think that someone who is a native of Zurich might be able to do the, you know, or conduct the same study, and my suspicion is that they would come up with approximately the same kinds of result.

(Tom Keller (Okay).

(Dave): But I cant make that claim.

The other thing is that, you know, the information I had was largely net, com and organization. And so, I don’t know if registrant behavior is any different in, you know, in one of the country codes, (PLDs).

Does that answer you question, sir?

Tom Keller Well, yes it does. And then, what I think is it’s kind (of interesting in that you’re seeing this the same way. I mean, having looked at the methodology, I don’t want to (criticize) at all, but just the way you figured out whether this is a private person (put on) certain factors which wouldn’t working at least in (Germany). I know that for sure. But looking at residential areas or business areas, because there's a total mix in Germany at least of these things, it might be harder and totally different to find out whether a person is a person or business.

(Dave): Well, if they were someone who was a 30-year native of Düsseldorf...

Tom Keller Yes.

(Dave): …then would you agree that that person would have a relatively good, you know, good insight into whether neighborhoods where residential or business?

Tom Keller Well, we don’t - that’s exactly what I'm trying to say, is that - and for example, (let’s say, they have) - there’s 100 - 280,000 living there and we don’t really have a residential area or business area -- that’s all intermingled into one and its really hard to tell from just having the streets to figured out whether that could be a borough or not.

For example, the borough I’m sitting in right now, all the works look like a residential area which is combined with business as well. So that’s no real way of telling, you know, whether this is a person or company.

I, you know, I don’t, you know, criticize all that.

((Crosstalk))

(Dave): And that’s a good…

Tom Keller … - yeah.

(Dave): …data point. But again, obviously, you know, living an American and living with the sort of, you know, sort of American style of residences, you know, neighborhoods or residential neighborhoods are, you know, urban or suburban, and its fairly easy to distinguish.

And I could show you pictures, you know, that of streets because I now have, you know, some cache on my desk of around 16 megabytes of photos that I looked at that you would unquestionably recognize as, you know, your stereotypical American, you know, two-storey colonial home. And maybe, that's not as easily applied in, you know, in Germany as it is here.

Tom Keller Yeah.

Well, I guess the only thing I really want to express is that it would be really hard to apply the conclusion you have or you made on you report about what happens in the area of Philadelphia to the rest of the world. Well, especially if it comes to data (correctness) and stuff like that, there never have been any kind of researches that have shown how, for example, the Germans or the Swiss or the Italians behave in terms of data correctness or not.

And I know there have been a couple of studies, one of Ben Edelman if I'm not correct - if I'm not mistaken, and I would be surprised that you would find the exact the same behavior. For example, Germany is very - people are very bureaucratic (in some ways), look at the domain registration as an official act and, therefore, might be inclined to look a little more after their domain name and the data that's attached to it.

I'm not saying that because I want to, kind of give preference to the Germans. I'm just saying that it might be different depending on the place that we're just looking at. And if you have a sample of 5000, somewhere in the world does not have to correspond with how it is somewhere else.

So I - the only thing I really want to put forth is that you have to be very, very careful in making conclusions that it might be probably be the same situation somewhere else.

(Dave): Okay. And that's all legitimate input. As I said, you know, I don't mean to be a tunnel visioned American…

((Crosstalk))

Tom Keller I'm not saying (that if) you have to start somewhere, probably you would have started here. And so, I'm just doing that for the record I guess

(Dave): Well, I think that's - yeah, like I said, that's a very valuable input. And, you know, if I were to go and try to do a study in Germany, I probably would think of, you know, be more familiar and hope that I would try to become more familiar with, you know, the way the people live and behave there. And I just - and that's one of the reason why I chose Philadelphia, because I was…

Tom Keller Uh-huh.

(Dave): …comfortable that I could apply the methodology with some convincing results.

Jordyn Buchanan: Dave, its Jordyn here, can I ask quickly, what - you mentioned that 600 of the records you looked at - or no, (456) were entirely unusable. What sorts of factors would make a record entirely unusable?

(Dave): Oh, okay. Well, I think, you know, fields not present, fields that had telephone numbers of 111-111-1111, fields that had names like ‘NA’ or ‘private’. So, I mean, you know, just information that literally was not, you know, not correct. And, you know, that, well, you can tell that some of these records where submitted through input forms, like a Web form, where it may have done some character check, but it didn’t say, you know, this is not a, you know, this does not look like a legitimate telephone number, so I'm rejecting what you typed in until - and you cant register the, you know, register a domain until you give me valid input.

You know, our process is not, you know, in many of the registration processes, there isn’t a significant rigor performed on input form validation, you know? People and, by and large in the Web, you type in almost any tentative number, you end up, you know, end up getting an Accept on the Web form. And that’s the kind of thing that people tend to abuse, you know? Same thing with emails, we don’t go and verify the email in a lot of situations by, you know, by building in a script that require that you bounce back or reply before your registration (is accepted), which, by the way…

:

(Dave): …I think would be wonderful but, you know…

Jordyn Buchanan: Sure. So it (seems like) these are the records you might - we may characterize as sort of being facially incorrect? Is that a…

(Dave): Yeah. Incorrect or incomplete, yeah. I mean, that - if there just, you know, if there wasn't enough there. And in some cases, there was enough there but, you know, you just couldn't tell. And the same, you know, the neighborhood, you know, was ambiguous. You know, I went and I looked and there were other apartments buildings in the neighborhood, but there were also, you know, was a pizza building and a, you know, and automobile mechanic on the same block. And so, you know, what do you do?

You know, I did not want to conclude one way or the other because they didn't want bias the result on my own speculation. So I can - I (tucked) that record into the inconclusive data record.

Jordyn Buchanan : Right. But that's in the distinct bucket from the unusable ones

(Dave): Right.

Jordyn Buchanan: You've got an unusable bucket and incomplete ones.

(Dave): Right.

Steve Metalitz: And (Dave), you also have this next bullet, Missing Information Used to Classify a Contact. Now, is that the same thing as, I mean, where does the domain name proxy service bucket fit here? You had 562 of those. Did you treat those as missing information or…

(Dave): No, no. That was basically a company that actually identified itself as a domain, you know, as an ISP. And, you know, when I went and investigated their Web site and their service, one of the things that they did was they claimed that, you know, they would register a name for you so that you would…

Steve Metalitz: Yeah, yeah.

(Dave): …your personal information would not be exposed.

Steve Metalitz: Okay. So, you don't know anything about the registrants, the true registrant, if you will, in that situation? But that's a separate category?

(Dave): Right. Right. So…

((Crosstalk))

(Dave): …there's a, you know, if you can't get to, you know, that's the whole purpose of the proxy…

Steve Metalitz: Right.

(Dave): …right? You can't get to, you know, to the registrant, you know? Through that service.

Steve Metalitz: Okay.

(Dave): Yeah.

Ross Rader: Ross here.

(Dave): Uh-huh.

Ross Rader: How are you doing?

(Dave): I'm okay, but I am now ten minutes past 8.

Ross Rader: I'll be quick then.

(Dave): Yeah.

Ross Rader: Yeah, I'll be real quick.

I'm not quite sure I understand the source of the data. Did you - is the - was - did you actually go to the Whois or did you…

(Dave): No, no. I got - actually, (Ram Mohan) had purchase six CDs with 20-some odd million registration records from a Web company, or a Web advertising company…

Ross Rader: Okay.

(Dave): …that sold the data in bulk.

Ross Rader: So we're not actually talking about Whois data here?

(Dave): So we're talking - no, we're talking about registration records.

Ross Rader: Well, but…

(Dave): And presumably - and what's the purveyor of the records claim was that, you know, that this is information collected using the (bulk of it).

Ross Rader: But we're not talking about authoritative data here?

(Dave): We are not talking - well, it may be authoritative data, but it wasn't received - it is not authoritative data. It was received through a third party.

Ross Rader: Okay. Okay, great. Thanks.

(Jordan): Okay, it sounds like we're - we've run a bit over time already. So, if there aren't any last question, we'd like to wrap things up, (Dave), and thank you for your time and (let him breathe).

(Dave): Well, you know, thank you for giving this opportunity. And if there is something else that you'd like, you know, like to do, or you want to revisit this during the ICANN meeting in Sao Paulo, you know, just give me some advance notice so I can make some time…

(Jordan): Great.

(Dave): …okay?

Maria : Thanks, (Dave).

(Dave): All right. Bye-bye.

(Jordan): Okay. We'll now, we have our second report today. This is the ICANN - the compliance report that the ICANN staff has put together. And I think (Mike) is going to be doing that this morning.

Mike Zupke : Yeah, I'm happy to be here today. (Jordan), thanks for having me.

What I'd like to do is kind of give you a real high-level overview, and then kind of give you some of the details that weren't explicitly published. I think that probably what you're more interested in and I think you're capable of reading what was already posted on the Web site.

First, I should just kind of clarify that I wasn't involved in the registry part of the audit. And so, I, you know, I apologize that I'm not really knowledgeable about the specifics of that. I'd be happy to try and get questions answered for you if you, you know, if there are things that you find that were lacking in the reports that was published. I'd be happy to, you know, try and track those sorts of things down. I can do that by email. Or if you just want to let me know in this call, I'm be happy to do that.

But I do have some pretty detailed specifics on the registrar. And so, I would like to just kind of give you the overview first, and then I'll give you kind of details that weren't really in the report.

What we reported on - in the registrar are part of this compliance update. And I think, you know, I don't want to overstate what the report was. This is a fairly routine thing that we were doing. You know, registrar renewals happen. Well, for each registrar, it happens every five years, and so we conduct an audit at that time.

And so, one of the things we want to do is to provide some feedback to the Community about that process and I think we try doing that periodically. In the past I think that a lot of this data was included in our reports that were (amended) theMoU. A part of this is just being, well, trying to be, you know, presenting information to the Community, and that what this should be read as.

What we had were 20 registrars up until the time of this report that had a (accreditation) agreements that were expiring or had already expired and, you know, were either had renewed or in some process of renewal. And so, the details that I've got are sort of an outline of what happened with those 20 registrars.

Of the 20, we saw 17 registrars who we’re able to renew. And some of them are still kind of in that process of returning the agreement that have been approved for renewal. There were two who voluntarily chose not to renew and there was one who was declined a renewal accreditation agreement as a result of the compliance audit.

Through this audit, we saw registrars modify their policies or practices in 18 instances. We also kept track of areas or issues in which we had received clarification on a policy or practice by a registrar. It wasn’t always clear whether what they were doing was clarifying or whether they were modifying a practice. So, this - I mean, this isn’t exactly a scientific reporting, but I'm happy to, you know, try and clarify this as we can so that you have an understanding of where we are in that.

We found that seven registrars, as a result of the renewal process, modified their contact data. And that was either their ICANN contact data or their public contact data, which is published on an InterNic site. And just as a matter of interest, we collected roughly $57,000 in overdue invoices as a result of the renewal process.

What I'd like to do then is give you a little bit of specifics on the areas that we saw, this policy or program changes (on the) registrars. I think that probably, obviously, the task force (in most instances sort the) Whois data areas. But there were roughly nine or ten different areas that I can kind of give you the overview on.

First was generally the UDRP. We asked a few questions -- how UDRP complaints are handled and what sorts of things registrars were doing to ensure the names are not getting transferred during the course of the dispute. Two registrars changed their (policies or) practices as a result of the audit to more appropriate adhere to the requirements of UDRP.

Let's see here.

We had - one of the areas that the renewal application involves, the part of the registrar accreditation agreement which requires registrars to obtain reasonable assurance of payment before activating a registration. We had no registrars who were required to change their policy or practices. In relation to that, however, we had five registrars who clarified their practices so that we were able to determine if they were in compliance.

Three registrars modified their practices related to the EDDP. This is the Expired Domain Deletion Policy requiring that registrars delete a name within 45 days of its expiration if it's not renewed by or on behalf of the registrant.

Five registrars clarified their practices, but none were required to modify their practices in relation to Whois data accuracy. And specifically, the questions that we're asking are how do you handle complaints about accuracy of Whois data, how many do you receive.

Generally, you know, we thought of the registrars who were applying for renewal had been acting responsibly in their obligation to take reasonable steps to ensure that Whois data is accurate.

Three registrars modified their registration agreements related to the jurisdiction or venue provisions of - that are required by the registrar accreditation agreement. Specifically, what that sets out is that legal disputes can be brought by third parties related to use of a domain name either in the jurisdiction of the registrar or in the jurisdiction of the registrant.

There were - just one clarification relating to transfer policy. There were policy changes required there. One registrar modified its form of authorization that's used to obtain authorization for transfers.

The RNAP -- and I'm forgetting what that stands for. Just give me a second here.

RNAP - Restored Names Accuracy Policy, that it. We had - no registrars were required to change their practice; two clarified their practices. Specifically, the RNAP requires that before a name is restored after having being deleted for inaccurate data, the registrar has to take steps to ensure that the accurate - that the data that’s provided upon the redemption is actually accurate.

The WDRP, this is Whois Data Reminder Policy, we had one registrar change its practice to come into compliance, and three clarified their practices. Specifically the WDRP requires registrars to notify to all registrants of gTLD names once a year of the details of the registrations. Specifically, they are required to present in - their Whois data - remind them that providing inaccurate data would be ground for termination of their registration.

Looks like the largest area that we saw in activity here was in Whois access terms. And specifically, we reviewed the access agreements for all registrars’ Port 43 Whois service and Web-based Whois service. And we require that their terms are in alignment with the terms of the RAA.

One of the common themes that we saw was that registrars where more restrictive than they were permitted to under the RAA. And while we don’t specifically have a position on the amount of data or the privacy or any of these issues, what we are enforcing is the requirements of the RAA which specifically would permit a person to use Whois data theoretically for postal mail. And a lot of registrars had prohibited that practice in their terms of access. And then, like I said, the RAA would require or would preclude a registrar from forbidding use of their Whois data for postal mail.

In the RAA of record retention, all registrars we found were in compliance although we did have five registrars clarify to us what their programs were. Specifically, some registrars had notified us or had indicated they retained a record indefinitely, and we just clarified that meant at least eight years or at least the term of their RAA plus three years.

And finally, the last item we looked at was registration agreements between resellers and registrants. Now, we didn't see (that) any policy changes or practices were required to be changed as a result of the renewal process. There were just two clarifications there.

So that's kind of the detailed analysis of the renewals part of our compliance update. Before I move on to the non-renewal part of this, were there any questions about how that all came about?

Steven Metalitz: Yeah. (Mike), this is Steve Metalitz.

(Mike): Yeah?

Steven Metalitz: One of the first things you just mentioned -- you obviously went through very quickly -- was the UDRP.

(Mike): Uh-huh.

Steven Metalitz: And I think you said two registrars had to modify or clarify, but what was sent out today says 11 registrars. Is this - are we talking about the same thing here or are these two different things?

(Mike): There were - I'm sorry, there were nine registrars who clarified for us what their policy or practices were to ensure that they were in compliance. There were two who modified either their policy or practices - I think the thing - well, I may have said there were a total of

11 because there were - this is probably one of those instances where there were - it - there were cases where it was ambiguous whether a registrar was actually changing its practice or whether it was explaining that its practice was correct.

We tend to give them, you know, the benefit of the doubt, and if it appears to be a clarification, you know, we're not so much (unintelligible) pound of flesh on this. You know, what we want to ensure is that, you know, going forward, registrars are adhering to the requirements. So in this case, you know, there were two specific changes; there were 11 instances where we investigated further than just simply accepting the answer that's provided by the registrar.

Steven Metalitz: Okay. And this sense also talks about two issues -- implementing UDRP decisions…

(Mike): Uh-huh.

Steven Metalitz: …and ensuring that registrants agree to the terms of the UDRP.

(Mike): Right.

Steven Metalitz: Do you have any information on which was involved there?

(Mike): Let me just look here.

I don't have specific details. But from memory, I believe that there was one case where a registrar had - where we thought that their inclusion of UDRP wasn't probably as clear as it could have been. And so, we asked them to clarify that. In the other instance, it was a case of a registrar - I think during the course of their renewal, we found an instance where they had, you know, not implemented the decision in a timely manner. And so, we included that as part of our renewal review to ensure that they had practices in place or policies in place that would ensure that there were no future problems with that.

(Ross): (Mike), (Ross) here for a follow-up if Steve's done.

Steven Metalitz: Yeah, go ahead.

(Ross): You were talking about the use or the sort of the restrictions that registrars were putting on the use of data for marketing purposes.

(Mike): Uh-huh.

(Ross): And you had mentioned that postal mail was something that would be permitted under the registrar accreditation agreement?

(Mike): You know, if you look at the RAA, there are two different places (where ideally) Whois data can be used.

(Ross): Yup.

(Mike): And in the - let me just pull it up here.

At Paragraph 3.6, the agreement talks about bulk access to data, and that was modified by the Whois - (let's see this policy) here -- by the WMRP, the Whois Marketing Restriction Policy. And that one would actually - that one addresses postal data.

(Ross): Uh-huh.

(Mike): But as you look at Paragraph 3.5, that deals with sort of the one-off access to the Whois server rather than the bulk access to data. And that one specifically does not permit a registrar to include an exclusion for postal mail.

(Ross): Got you.

You caught me not half listening, so I was just wondering what the scope of that comment was. So thank you.

(Mike): Well, and, you know, I tend to, you know, I mean, I always have the RAA in front of me. So I tend to sometimes abbreviate more than I should.

(Ross): (Mike), you should have it memorized by now.

Thank you.

(Mike): Any other questions on the renewal process before we move on to the other piece of this?

Okay. Well, let me then go into the other piece. And this is - beginning fairly early in August, our ICANN CEO had indicated he wish for us to, you know, have statistics available in terms of what sorts of compliance efforts we're undertaking on kind of - maybe, informal is not the right word, but more of an ad hoc basis. And he wanted to track, you know, what sorts of efforts we are putting into issues, what sorts of issues we're seeing and trends and - to be able to provide feedback to the Community on that sort of thing because I think that in the past, it was something that we didn't have very good accessibility to that data. I think that as we've got more staff and resources, it's something that we're a little bit better able to do.

We're still kind of developing the methodology for this. What we've begun with is tracking the complaints - what you're looking at is specifically the complaints that I handled through receiving from a staff (member) of ICANN either who answers - the icann@icann.org (email address), (transfer questions), (email address), the registrar info which is the email address that's used to submit registrar complaints
And so, these sort of - as people get complaints through these different channels, those that are believed to address - to involve a compliance issue are forwarded for review in this in this update. So that's what - that's the data here that's included in our report.

So, in what we reported, there were roughly two months of data. And we found that there were 33 potentials compliance issues and things that were clearly non-jurisdiction -- for example, issues involving Web hosting or, you know, spam or, you know, various content-related areas -- we didn't include in this at all and it was just clearly ruled out as a non-jurisdictional sort of thing.

So we looked at 33 different issues. On average, we were able to close them within four days -- of the 31 that were closed by the time we're reporting. Eighteen were closed without any formal compliance action. This may have been something simple like a person was unable to get a name unlocked or to obtain an (op code).

In seven instances, we saw registrars change their practices or policies as a result of the compliance action. Three of these compliance tickets were referred to the transfer dispute resolution policy because they were brought by registrars or by another registrar or registrars' adherence to the transfer policy. And finally, three issues were resolved with further oversight by ICANN.

I gave a report to the registrars' constituency last week, Tuesday. And what I did for - in preparation for that, I just took my spreadsheet that I've been keeping on compliance statistics so that I had something a little bit more current. And I've got some data here that I'd like to share with you. This is as of 24 October.

So, the numbers aren't going to match - they're not going to add up to 33 in here because this isn't keyed necessarily to the report as published. This is pretty much the most up to date and accurate information I can give you. So, I'll just kind of walk you through it here.

We saw over the roughly 2-1/2 to 3 months that we are tracking these statistics, we saw 13 tickets were raised related to Whois compliance. Specifically, four of those involved Port 43 Whois availability; four of those involved Web-based Whois availability; three involved population of Whois fields as specified in the registrar accreditation agreement; and two involved Whois data accuracy.

I think, you know, what I can tell you more specifically from memory is that generally, with the Port 43 availability, I think that probably half of the instances where cases where registrars' Whois appeared just not to be functioning and the other two instances were instances where the Whois service was not returning record (on specific)…

In terms of the Web-based Whois availability, we discovered that there were - four registrars were not (planning) Web-base Whois availability, and three of those - four were actually affiliated with each other through some, you know, shared ownership structuring. However, these were registrars who were not actively putting out a retail front but were more involved in the expired domain catching. And I think there may have been some misunderstanding as to what they were required to provide to the public in terms of a Whois (tool). So, they were able to, within a day or two, come into compliance and made sure they were all providing a Web-based Whois interface.

So that's kind of the area sort of specific to Whois. I'd be happy to walk you through the other areas that we saw compliance issues, but I'm sure that this is probably where you'd like to probe further if there's anything that you'd like to know more. Are there questions in particular about these Whois issues that came up?

Steven Metalitz: (Mike), this is Steve.

I just had a question about your third category there -- population of Whois data fields. Does that mean there were fields where there was no data, or as - I mean, you distinguished that from Whois data accuracy, so I wondered if…

(Mike): Okay.

Steven Metalitz: …you can explain.

(Mike): Sure. The population of Whois fields comes from Paragraph 3.3.1 of the RAA. And specifically, that enumerates certain fields that a registrar must publish. So, for example, registrant name and address, admin, technical contact, name, address, email, phone number. You know, specifically, I think in these, we were seeing cases where registrars were not including all of the required fields. It wasn't a matter of that the field is not being populated by the registrant. It was actually the registrar wasn't providing the field in its Whois output.

You know, as an example, there was one registrant who provided a registrant name and address and nothing more. And, you know, obviously, that doesn't comply with the requirements of the RAA. So we went back and insisted that they, you know, fully populate that. And that was corrected, you know, pretty easily.

I think that there is a concern by, you know, I suspect -- and I - maybe, I'm speaking without having, you know, good (facts), but I think, among sort of the smaller registrars, that they want, you know, to ensure that they are protecting the interests of the registrant. They may not be thinking about the consequences of, you know, not adhering to the requirements of the RAA, or may not even be thinking if there are specific (cases) that are set out. So I think that that's probably where most of that comes about in terms of Whois data accuracy.

These sorts of issues were things where a user (would file and say) WDPR (registrar) report about a domain name, and bring through a formal action. Or - however, you know, if there was inaccurate Whois data, and somehow bring this to a staff member's attention (if hadn't been addressed). And we saw that, you know, we would ask the registrar to explain what action had been taken and or to take the appropriate action. And so, that's coming from the provision of the RAA that requires a registrar to take reasonable steps towards ensuring data accuracy.

Steven Metalitz: Okay, thank you.

(Mike): All right. Any other questions (relative to the Whois part of this)?

All right. Well, I'll give you just kind of the brief - really brief listing of the other areas that we address to kind of give you some perspective and the scope of Whois-related complaints.

As I said, we had 13 tickets related to compliance, tickets related to Whois between roughly August 7 and October 24. We had three compliance issues involving registrars and their being overdue on invoices. We had one issue of a reseller actually using the ICANN-accredited logo without authorization or without having, you know, (entered into a licensing) agreement.

We had 20 tickets raised related to transfer policies, specifically four, related to registrars providing (off codes), five related to registrars either using or providing upon request the FOA for obtaining authorization, five issues related to locking and unlocking, there were six that were sort of uncategorized -- generally, these related to a registrant's ability to update through email address or another way to get information about the ability to transfer domain names.

We had seven issues related to the UDRP. I said seven. Yeah, seven. Four are related to registrar's implementing decisions in a timely manner, and three of them involving the locking of a domain name or otherwise preventing transfer upon commencement of (a EDRP) action.

And so, that concludes the sort of the summary of what we're seeing in terms of compliance tracking.

One of the things that we're looking at going forward is (the) emerging issues of registrars populating Whois data within 24 hours of activation of new registrations. It appears that there has been some misunderstanding about this in the Registrar Community and we're trying to do some outreach here to clarify the differences. And I think we talked about this in the last call and that I was not in that, but that registrars are required to update registries with Whois data within five days, but they're required to populate their own Whois Data within 24 hours of activation of a new registration. So that's something that we're continuing to work on going forward.

Any other questions for me?

(Jordan): Okay. Thanks, Mike.

I think we'll - we don't have any other topic (on this end) today and we were - had limited ourselves to just reviewing these two reports. So, I think that that will probably wrap up the call.

We'll meet again next week with a bunch of substantive topics. Hopefully, (I'm going to send) out somewhat more detailed in the couple of days. But I'm hoping that we can (as best get to the) - many of the substantive issues that we have left.

And I think - Maria, do you think that based on some of the comments (today that) we’ve been having back and forth, that we’re likely to see a revised version of the report between now and then as well?

(Maria): Yeah, absolutely. And I’m working on - I’m basically trying to pull together a lot more of the discussion that we had on the various proposals to make (it) a little bit more (substance) and a little bit more of a guide for, you know, for the public comments and for potential readers. So I expect to have that out for the taskforce within the next couple of days.

(Jordan): Great. So people can be - on the look out?

(Maria): Yes. And also, because it will be quite a bit more substantive. There may well be some more discussions to be had on it. So that’s a heads-up.

(Jordan): So, I think I’d like to thank (Mike) again for his time in joining us today.

(Mike): Thanks for having me.

(Jordan): And we’ll wrap up this call.

There will be a transcript of this call. And we’ll wrap up this call and meet again next week.

Steve: Thanks, (Jordan).

(Jordan): Thanks everyone.

END