CLAUDIA RUIZ: Good morning, good afternoon, and good evening to everyone. Welcome to the NARALO monthly teleconference call on Monday, the 11th of January 2021 at 2000 UTC.

On the call today, we have Eduardo Diaz, Alfredo Calderon, Bill Jouris, Dana Perry, David Mackey, Glenn McKnight, Gordon Chillcott, Herb Waye, Jose Lebron, Judith Hellerstein, Ken Herman, Marita Moll, and Maureen Hilyard. From staff we have Heidi Ullrich, Naela Sarras; and myself, Claudia Ruiz, on call management.

Before we begin, I would like to remind everyone to please state their name for the transcription purposes, and to please keep your microphones muted when not speaking to prevent any background noise.

Thank you all very much. And with this, I turn the call over to you, Eduardo.

EDUARDO DIAZ: Thank you so much. Thanks, everyone, for being here today. It is the first meeting of 2021, and things are going forward.

First of all, thanks for taking the time to be here today. This is Eduardo. I'm the current chair of NARALO, and I'm talking to you directly from San Juan, Puerto Rico which is very breezy, 80 degrees, and sunny.

I would like to start with a reminder of the changes implemented to our NARALO monthly meetings agenda due to the COVID-19 pandemic. A

few months back, the region approved a strategic outreach and engagement plan for fiscal year 2021. The plan changes its focus from outreach and engagement to engagement only. And with engagement, we mean membership engagement.

The idea is to increase membership participation not only for this space, but for the space in the Consolidated Policy Working Group and the At-Large Operations, Finance and Budget Working Group—or what is called the OBF Working Group—meeting as well.

So, what we want is people to engage here and also work and participate in those groups. The CPWG, for those you that do not know, is for policy discussion and the OFB-WG is for finances, budget, and other operational issues.

This series of webinars are in an effort in that direction. In other words, engagement and capacity building. This presentation is number six in the series and is being put together by Bill Jouris. Like I mentioned before, all admin reporting like ICANN policy updates, NomCom regional ALS updates, ABR issues, and things like that will be done via the NARALO newsletter, ICANN ALAC announcements, or just specific e-mails to that effect.

Ultimately, we will use this space for anything that requires regional discussion in a group environment if necessary, so anybody that wants to bring something up, please send me an off-list e-mail and we will put it in the agenda.

Any important announcements will be done at the beginning of the calls. Also, if you are in need to announce anything of immediate

importance, let me know beforehand to include it in the agenda just like I just said.

To finish, make sure you are counted as present in these calls since we are using the aggregated total as an engagement metric. And also, please stay a couple of minutes after the question-and-answer sessions to fill out a mini-survey that is going to be done before the end of the call by staff.

And after saying all of this, I want to again encourage you to participate in the weekly CPWG meetings every Wednesday and the OFB-WGs when they meet.

So, before I end my introduction, I want to call Judith to provide an update on the Additional Budget Request process which was extended. Judith?

JUDITH HELLERSTEIN: Yes. I am the appointed secretary. [And we think] so far we had one proposal that was put in under RALO proposals for the additional budget request. Now, we still have two more days, and staff extended the period that people can apply, so this is for any special projects that you want to do in your region that start in July and run the fiscal year. And if you have some idea, please get in touch with me as soon as possible so I can walk you through because we also do need approval from NARALO.

The one that is approved from NARALO is the Virtual School of Internet Governance, but if they are more ALAC-wide, then there's a different

approval process. Jonathan Zuck is putting together an ABR for an ALAC-wide proposal, and so we are open to many different ideas. But remember you have to get it in within two days to be considered.

I'm happy to run through with any issues, and maybe staff can post a link in the chat to the page that describes the instructions. And feel free to get in touch with me [with you] on any questions.

And with that, I'll turn it back over to Eduardo.

EDUARDO DIAZ: Thank you so much, Judith, for the update on the ABR. So, you have two days to fill out this form, put it in if you need projects or something that meets the criteria once we get the link put in the chat.

So, this is all I have to say at the beginning today. And let's move next to the presentation. Bill, you have the floor. I hope all of you like. I think it's going to be a very interesting presentation that Bill is going to do. So, thank you so much.

BILL JOURIS: Greetings, everyone. I gave a presentation on some of what was happening with Internationalized Domain Names a year ago if you were there. Some of this may be old news for you, but I think some of it will be new as well.

Let me say right up front that while I've been involved in the IDN project for the past five years, I am not the spokesman for the project. These

are my own observations and opinions. Nothing more. Next slide, please.

Let's start off with a little history. We're going to talk about changes to the Domain Name System and why you should be aware of them. Originally, the Internet was a U.S. government project. Unsurprisingly, domain names were created for the World Wide Web and they were done in English. That is, you could use 26 letters, A-Z, and ten digits, 0-9.

But around the turn of the century, the management of the Internet moved from the U.S. Department of Commerce to various independent international corporations—ICANN, IETF, etc. And once that happened the rest of the world asked, quite reasonably, "Why can't we have domain names in our own scripts?"

As always, the question is reasonable, but the devil is in the details. The first question, obviously is, "What symbols can be used?" That is, what is our repertoire? Because you have to program the computers to deal with them.

And the first restriction that has been applied is it has to be symbols used to write a language—a living language. So, no emoji and other non-language symbols. There may be people who can read cuneiform or Egyptian hieroglyphics, but we're not going to need to use them for domain names.

The second principle was you can't use a domain name that conflicts with what somebody else has already registered, which sounds straightforward, but it isn't. Next slide, please.

So, the purpose of the project for the internalization of domain names, or IDN, is to expand the repertoire. Some of the scripts are probably familiar to you because they're used in the language of a large number of people or major economic countries or are just historically significant languages. Others, you may barely have heard of. The good news is we've limited ourselves to living languages. You don't have to worry about something that isn't actually spoken today.

Many of the scripts use an alphabet of some kind that are used for just a single language. So, if you want to identify the repertoire used by Georgian or Lao, all you have to do is grab a grammar school textbook and look up what the letters are.

But the scripts that use characters like Chinese, Japanese, Korean. Chinese, for example, uses hundreds of characters. You could spend a lifetime learning them and never finish learning all of them. Fortunately, an expert in the language can come up with a limited subset that every educated adult can be expected to know.

Japanese is even more challenging because it not only uses characters (kanji) which were borrowed mostly from Chinese, but not one but two phonic alphabetics, Hiragana and Katakana. Or actually three phonic alphabets, but the third one (romaji) is just a subset of the English alphabet so no extra symbols there.

Some scripts, however, are used for more than one language. The Arabic script is used for Arabic, but also for Farsi which is the language of Iran and parts of Afghanistan. Cyrillic script is used for Russian, Ukrainian, Bulgarian, Serbian, etc.

The Devanagari script is used for a number of languages from South Asia. Still, [the set's limited] and the differences between the languages that use each of those scripts is relatively limited. Next slide, please.

So, the first problem that comes up is with communication. Suppose you start a company to make an inexpensive product for the local market, say in Laos or the Republic of Georgia. You register a domain name in your local language to better to communicate with your local customers like these two. So far so good.

But then your company's a success and you start moving up market, bigger and better products. And now you have an export market. To take an example, you start out making bicycles which are cheap transportation that your local customers can afford. You move up to motorbikes, and then you move up to motorcycles. Then you move up to automobiles and you're exporting worldwide. And if you think that can't happen, that's the corporate history of Honda Corporation.

So, now you're spread worldwide and you want to communicate with your customers worldwide. Can you tell what these domain names are? If you look at the two on the screen, can you even tell which one's Laotian and which one's Georgian?

Again, if you think that can't happen, another example would be South Korea. In the early 1950s, South Korea was a poor, rural country, and what little industry it had was trashed from years of war. Today, you've all heard of the brands Samsung, Subaru, Hyundai. If the Koreans can do it, there's no reason the Laotians and the Georgians can't do it, too. Next slide, please.

In addition to the simple cases, we also have the Latin script. Thanks to centuries of Christian missionaries, first to cross Western Europe a couple thousand years ago and then following more recent colonization only 500 years ago, the Latin alphabet is used for dozens and dozens of languages on every continent. There are over 400 living languages using the Latin script; still the same couple of dozen basic characters to combine with over 20 diacritic marks on various letters.

To keep things a little bit under control, for the Latin script we took those language which are official languages of countries or of states, provinces, or regions within a country plus languages which aren't official languages but are spoken by over a million people. It's arbitrary, but it kept things from getting even further out of hand than they were.

Note that we may get told to consider more languages when our report comes up for public comment, something that I think will happen this spring. Of course, last year I thought it would happen this past spring, so we'll see. But I'm hopeful that in March or April we will be up for public comment. Next slide, please.

In addition to the language that it's used for, unlike most scripts, Latin uses a lot of different fonts. Here are three of them: Times New Roman, Arial, Courier. And that can make a difference. We looked at these three fonts just to get some idea of which symbols look like which other symbols.

Just to give you a flavor of what kinds of problems you can have, here are examples of the letter "a" in two different fonts or the letter "g" in two different fonts. If you're used to using the Latin alphabet, you look

at those and go, “Well, of course,” and you would never give it a second thought.

But the letter “a”, for example, the first one looks nothing like any Greek letter, but the second looks like a Greek “alpha”. So, do you have a conflict there?

And that’s not to mention the challenges you can get with serifs. On the last line, we have the letter “c” with serif and we have the letter “c” with the hook. Now, in the Arial font that would be easily distinguished. In the Times New Roman font or the Courier font, not so much. Next slide, please.

Earlier, I was a little bit rude about Japanese for having multiple phonetic alphabets, but before we get too snarky, we ought to realize that English does, too. We just call them uppercase and lowercase. In some cases, the symbols are essentially identical. And in some cases, they’re at least very similar. And in some cases, they’re totally different. Next slide, please.

Latin isn’t actually alone in this. Armenian, Cyrillic, and Greek also have upper and lowercase. As you can see in these examples, again sometimes the letters are very similar in upper and lowercase. And in some cases, they’re totally different. These are, as far as I know, the only four languages that do upper and lowercase. It’s probably coincidence that they’re also the four living language derived from old Phoenician a couple of thousand years ago.

The folks here at ICANN and folks at the IETF that write standards are generally extremely clear that domain names are strictly lowercase.

Unfortunately, outside the microscopic slice of the population that's in these organizations, Internet users have a couple decade's experience telling them that uppercase and lowercase can be used interchangeably in domain names.

Their browsers actually convert everything to lowercase before invoking the Domain Name System, but they neither know nor care. Why does that matter? Because the IDN project, when they were giving us instruction, said, "Ignore uppercase when deciding whether two letters are variants." That is, whether they are letters that can be confused with each other.

Next slide, please.

I should have put this one earlier. I mentioned that there are some 20 diacritics. Here are all of them. Now unless you're a linguist, you probably have never encountered most of these. My sense is that any given language may use four or five of them. If you're familiar with a half dozen languages, you might even know half of them, but the others are going to be things you've never seen before.

Some of them you can tell apart, some of them are a real challenge. If you look—let's see, where's my favorite example? The tilde—oh, there we are. The breve and the caron on the left-hand column. Now one of those, the breve, is basically an arch, a curve.

The caron, on the other hand, is two straight lines meeting at a flat angle. And if you increase the size until the underlying letter is a centimeter high, you can usually tell them apart. On the other hand, if

you are looking at them in a normal font size, nobody I know can tell the difference.

So, that's the sort of situation that we had to be aware of when we were looking at variants. It's especially bad when you have letters where you have diacritics that you didn't even know existed. If you don't know that diacritics like a macron below the letter or a dot below the letter even exist, you're not going to look and you're not going to notice them.

Next slide, please. There we go.

As I say, any given language only uses a few of them. English. All of us will probably say, "Oh, no. English doesn't use any." Well, almost. As you can see here, I have spelled "naïve" the way it is technically, officially supposed to be spelled which is to say with a diaeresis, a double dot over the "i". Of course, since that isn't an option on the standard keyboard, none of us use it that way. You'll find it that way in magazines and newspapers, things that traditionally were typeset, but not anywhere else.

Spanish uses three. French has actually got five. So, you are probably all at least aware of these even if you don't happen to speak some of the languages. For instance, if you speak Spanish but you don't happen to know of a language that uses a macron—a straight line over the "n"—can you tell that you're looking at a macron and not a tilde if you don't even know it exists?

Or is “cafe” spelled where the “e” has a dot above it visibly different from “cafe” with a grave accent over it? There are people who can tell, but I’m not one of them and I suspect you aren’t either.

Next slide, please.

This is just to give you an idea of how bad it can get. It’s just one letter—admittedly the worst case—but it’s over 30 variations on the letter “o” plus a letter called an “s” which is vaguely related to a “d” but looks a lot like an “o” with some of those diacritics.

Next slide, please.

So, “what you see is what you get” works only if what you see and what you thank you see are the same. We need a mechanism to make sure that the domain name that you see is actually the one you thank you’re seeing.

Our mechanism is to define variants and, as I’ll go into a little later, confusables to constrain domain name registrations. The basic rule of thumb is, if you want a domain name and it only differs from an existing domain name by one or more variants, then no, you can’t have it.

Next slide, please.

So, when we compare characters and classify them, we have to decide. Are they variants? Are they confusable? Are they different? For top-level domain names, if what you propose as a new TLD only differs from an existing one by variants, it’s going to be blocked automatically in the software. If you don’t have that, then it will go to a manual review by a

Similarity Review Team, and they will try and look and see if there's anything that could be confusable.

Of course, the folks on the review team, like the folks on the various language teams, are all experts. They're steeped in the various symbols, so they will undoubtedly find some items that are confusable. But how well they'll be able to pick out what a normal user will notice is another question.

And then we get to second- and third-level domain names, and we'll talk about those a little more later.

Next slide, please.

Just for an example, consider this domain name. When you look at it, you probably, on first sight, say "www.test.joke." Except that the "k", it turns out, isn't a "k". It's actually a "k" with a horn which only occurs in Houssa. So, if you don't happen to speak that particular East African language, you've never seen such a thing.

And the difference, to my eye at least, is about the same as looking at one of the fonts that uses serifs, like Times New Roman. And it just looks like a funky serif.

The other item is that the letter "j" isn't actually a letter "j". It's a letter "i" with an ogonek, and that only occurs in Lithuanian. So, if you don't happen to know Lithuanian, oops, you've never seen that before. What you see is something that looks like a letter "i" and then extends down below the line and curves which is the definition of a letter "j". The point is, it's really hard to see things that you don't even know exist.

When we were doing the analysis in the Latin generation panel to decide what were variants, we were looking at things side by side and the problem is all of us have spent three or four years immersed in these symbols. We know all of them. We know all the diacritics. We know all the combinations.

And, as I say, we were looking at them side by side which is a luxury that the average user won't have. So, our criteria for what's a variant is extremely narrow.

Next slide, please.

Part of the problem is that there's a continuum on how different things are. Here's a whole lot of things that look like .com but aren't. As I mentioned, we have a mandate to ignore all caps. If you look at .COM in all caps, if you see that normally, you'll go, "Right. That's just .com." Except that if it was actually in Cyrillic, it looks the same but it's, as far as the computer is concerned, a totally different name.

Next slide, please.

So, when we came to look at cross-script variants, we had some challenges. Sometimes things are pretty simple. Greek, Cyrillic, and Armenian scripts all come from the same source, so they've got lots of very similar characters. The Latin letter "p" and the Cyrillic "er" look the same. The Greek letter "nu" and the Latin letter "v", etc.

Obviously, if you have domain names in one script that just use letters that look like those in a different script, you need to stop those.

Next slide, please.

But in addition to those related scripts, you've also got some symbols that are pretty generic—a vertical straight line, a circle, a crescent opening in different directions. And those turn up in totally unrelated languages.

Again, you have to say, "Well, how much of a problem is this going to be?" And you can look at those and go, "Well, yes, but probably there's not going to be a domain name that is just made up of those particular symbols." The trouble with that is, there is already a top-level domain name which is, in Latin letters, ".ooo" which means you have a potential problem with all of those other languages.

Next slide, please.

On top of that, we have in-script variants. The schwa and the "turned E" are particularly nice because they are, pixel for pixel, identical. The "turned E" occurs only in one of the West African languages. And why in the world anyone decided it wasn't the same as a schwa, I don't know. But as far as the computer's concerned? Two different symbols.

As we saw with the "joke" earlier, you have lots of cases where if you don't know something exists, you have no chance of noticing that you're seeing something different.

Next slide, please.

Another little issue is underlining. Most word processing software and most browsers will automatically change the color of a domain name and underline it to make it easier for you to see it. But that becomes an

issue if you have a below-the-line diacritic because the line will go right through it or at least obscure the fact that it's there.

Here are a bunch of examples with various below-the-line diacritics. If you can actually spot that the diacritic is there, give yourself marks for exceptional eyesight. However, in most browsers and in most (but not all) word processing software—and obviously not in this particular word processing software that I used to make the slides—anything that extends below the line gets a couple of blank pixels on either side. So, if you have a macron below, the line comes up to it, stops. You have the little straight line that is the macron. It stops again. And then the line continues.

And so, the decision was made by my fellow members of the Latin panel that a "reasonably careful user" would notice. So, none of those are variants. Merely confusable.

Next slide, please.

Consider this particular example. About an hour south of me in Silicon Valley is the city of San Jose, California. Now, suppose they had taken the domain name of www.sanjose.gov. In point of fact, they used www.sanjoseca.gov—but just suppose.

Now suppose the city of San José, Costa Rica wants a domain name. They can't use www.sanjose.gov because that's already taken. But, no. They say, "Being a Spanish-speaking country, we will of course spell San José correctly. (That is to say with an acute accent over the 'e'.) And www.sanjosé.gov, for us, is not taken."

It's certainly true that a computer will recognize instantly that that's different, but whether a human being will notice or not is another question. It's not a question here of malice. It's just, things can get confusing.

Then there's the converse case. The city of Munich, as you may know, is the capital of Bavaria in Germany—except it isn't. Any German will tell you that the name of the city is München. So, here we have two visibly different names for the same place. And there probably should be a rule to consider how we deal with that, but to the best of my knowledge, there isn't.

Next slide, please.

Now we get to things where there actually is malicious intent, DNS abuse. Here's an example of something that came up a couple years ago.

There's a British budget airline call Easy Jet. Their domain name is www.easyjet.com. Then a couple of years back, somebody registered another domain name. The only difference is the "j" has been replaced by a letter "i". And they then proceeded to scam a bunch of would-be Easy Jet customers because spotting the difference in those domain names isn't all that easy.

Eventually, that was figured out and they were shut down, but it's obvious that there's a problem there. Nonetheless, I would note that that domain name is advertised today as being for sale. As we heard last month on the presentation on the secondary domain name market, you can sell domain names yourself. And that's what the owner of

www.easy-I-E-T.com is apparently doing although it's not obvious, as least to me, what non-fraudulent use there might be for that domain name.

Next slide, please.

Now, you can ask, "How much of that goes on?" Well, here are a bunch of examples which I've put in uppercase just to make it a little easier to see. Would anyone like to bet how many of those other domain names are actually registered to the Facebook company? My guess is none. This isn't quite the same kind of DNS abuse as the Easy Jet example.

Here, what usually has happened is somebody has registered the domain name and then they're going to go to the domain name holder and say, "You know, you really ought to buy this from me so it doesn't cause your customers confusion. Give me 100 times what I paid for it, and you can have it."

And given that even that amount of money is barely even a rounding error in the big corporation's budget, even though it's big money to you, they may well buy it.

Next slide, please.

One other risk we have is false equivalence. Domain names have parts, and as I mentioned all of the work that has been done so far has to do with the top-level domain names even though most of the examples I was showing were second-level domain names because I found it a little easier to generate them.

The two different parts are different in a couple of ways and they have different rules and different procedures. New TLDs are only created occasionally. They're developed by ICANN. But new SLDs happen daily, and the rules for them are mostly set up by the individual registries and registrars.

Last fall, ICANN put up a document for public comment, the "Reference Label Generation Rulesets for the Second Level." It talks particularly about blocking second-level domain names which differ by only one or more variants. The thing is, it doesn't actually specify what a variant is.

The assumption, as far as I can tell, is that everybody will just use the variants generated by the various groups of so-called language experts. Being on one of those panels, I'm not sure how expert we really are, but ignore that for the moment.

All of those panels have been under a mandate to only—only—think about TLDs. So, we've used a very narrow definition of what constitutes a variant, the justification being that actual filtering to avoid confusion will be done manually by the Similarity Review Panel. But for second-level domains, there are no such panels. And given the volume of registrations, manual review probably isn't feasible.

So, the situations aren't actually equivalent, but we seem to be assuming that they are and so we're opening up lots of possibilities for abuse there.

Next slide, please. There we go.

Remember this one? Now, a few of you may have picked up on the fact that the .com is slightly different. But I'm betting that most of you, when you did, looked at it and said, "Oh. They put .com in capitals. Well, that's a little weird, but so what?" Except the .com in the second domain name isn't .com. It's Cyrillic.

And that opens up serious DNS abuse possibilities. If you can register that Cyrillic top-level domain name and set yourself up as a registry, you have got an enormous instant market. Everybody who has a .com domain name is going to need to come to you and get one in your top-level domain as well and pay you for it in order to keep their customers from getting confused. And if they don't, somebody will come along and use it for fraudulent purposes. And it doesn't matter to you. You're still getting paid for it.

So, maybe the Similarity Review Panel will come on and catch it in the manual review. Then again, since everybody knows that domain names are lowercase, maybe they won't. Fun times.

Next slide, please.

That's just a high-level look at what's going on at the moment. There's lots of detail that hasn't even been touched on here. There's lots of detail that I'm not even aware of. But if you have any questions, I'll be happy to try and answer them.

EDUARDO DIAZ: Anybody have a question? Please go ahead, David.

DAVID MACKEY: Hi. Thanks, Bill. I find it fascinating. One of the questions that I had in my mind when you were going through the slides is the aspect of combining characters from different languages in the same domain name. Is that part of the problem—where, when you add in these characters from the different languages and then people start mixing the characters from the different languages to get a human interpretation that might be confusing for one language?

Is that part of the problem? Or is that relevant to what you've been looking into?

BILL JOURIS: I believe there is a restriction that says, "Within any particular part of the domain name, within the top-level domain name or within second-level domain name, you're restricted to a single script. But there's no problem if you want to have your top-level domain name in the Latin script and your second-level domain name in Chinese. And, in fact, there are a whole bunch of those registered already. There is a little bit of help there, but only a little.

DAVID MACKEY: Okay.

BILL JOURIS: Does that help?

DAVID MACKEY: It does. It's still … Well, it is what it is. I need to wrap my head a little bit more around it. Can I ask another question?

BILL JOURIS: Yeah. Just a second.

Can you pop back to the previous slide, please? Let me field one other one and then I'll get to you.

Judith Hellerstein asked which was Laotian and which was Georgian. The way I can tell, the only way I can tell, is that since Georgia was part of the Soviet Union, I used the Cyrillic .com for theirs. So, the first one is Laotian and the second one is Georgian.

Okay, David?

DAVID MACKEY: Okay. Thanks, Bill. A question I have is related to the technology that's underneath the characters that you've been talking about. We've mostly been looking at an end user looking at the domain name in front of a browser.

Underneath the hood, can some of the problems be separated to something that you would find in the DNS query algorithm itself as opposed to, say, the DNS system setup that's in behind the actual browser? Is there a difference or is it only relevant to the end user visibility of the domain name?

BILL JOURIS: I'm not sure I understand the question, but if I'm interpreting it correctly, the computer doesn't have a problem figuring out what

symbol it's looking at. All of these symbols get translated into something called Unicode which is a four-digit hexadecimal number for each of the various symbols. And that's what actually gets sent off to the DNS system.

DAVID MACKEY: That makes sense and answers my question.

BILL JOURIS: Okay.

DAVID MACKEY: It really is strictly confusion from the human to interpreting the domain name. The system itself doesn't matter because everything is Unicode. So, it doesn't matter from a computer's perspective. Thanks.

[BILL JOURIS:] Right.

EDUARDO DIAZ: Bill, there's another question for you from Alfredo in the chat.

BILL JOURIS: Let me scroll up and find it.

EDUARDO DIAZ: I can tell you. "Is it a resolver issue when it sends it to the DNS?"

BILL JOURIS: I'm not sure I understand the question.

EDUARDO DIAZ: Alfredo, why don't you get on the microphone and ask the question?

ALFREDO CALDERON: Sure. Can you hear me?

EDUARDO DIAZ: Yes.

BILL JOURIS: Yes.

ALFREDO CALDERON: Okay. Thank you, Bill, for the presentation. My question is sort of a follow up of what David was asking. When I as an end user see a URL in the script and it is converted, it's done by the resolver. So, my guess is that the interpretation, although I may interpret it in a certain way when it gets resolved by the resolver in the DNS, I'll get something different like it's not available, it's not found, the page doesn't exist, and so forth.

Am I correct? Thank you.

BILL JOURIS: Yes, if I understand what you're saying. If you put in, for instance, something where the .com is done in Cyrillic, you think it looks like .com but the resolver will look at it and go, "No. That TLD doesn't exist." And it will send you back and error message.

It's only if the domain name that you're looking at does exist—even though it's not the domain name you think you're looking at—that you have a problem.

EDUARDO DIAZ: Bill, I do have a question.

BILL JOURIS: Yeah.

EDUARDO DIAZ: If we look at the slide that you have in the Cyrillic, which we have .com that looks like a capital "m" but it's not. I don't know. I think you answered this to David, but I probably didn't catch it. But in the second level in Cyrillic, can I put any characters there, or do they have to be Cyrillic?

BILL JOURIS: No, you can put any characters there. I mean, these two domain names, assuming the .com existed, would be perfectly valid even though they're different languages for the second level and the top level.

I'm sure that they expressed an opinion in the chat that they ought to be restricted, but given that, for the moment, only Latin letters are used for TLDs, that ship has sailed and we're not going to be able to get back to it.

I'm sorry. Yes, you had something more?

EDUARDO DIAZ: Yes. So, I can have my bank in Latin characters in front of the .com which is Cyrillic, then. Is that what you're saying? I can have that if it's possible.

BILL JOURIS: Yep.

EDUARDO DIAZ: Wow.

BILL JOURIS: [Absolutely.]

EDUARDO DIAZ: That is DNS abuse. That will be very hard to find, very hard to catch. Very.

BILL JOURIS: Yeah. Well, it will be necessary to not let that Cyrillic .com be created. For that, we're dependent on the Similarity Review Team which presumably will get set up when we get ready for the next round of TLDs.

EDUARDO DIAZ: Okay. So, are there any more questions?

I have another question. Do you foresee, going forward, using these new languages in the top-level domain to do DNS abuse in the future?

BILL JOURIS: Do you mean something like registering the Cyrillic .com?

EDUARDO DIAZ: Right.

BILL JOURIS: I don't know that there will be much of it, but human nature being what it is, I will be amazed if someone doesn't at least try it.

EDUARDO DIAZ: Okay. So, anyone have any questions for Bill? I don't see any hands. I don't hear anybody. So, I just want to say, Bill, thank you so much for putting together this presentation.

Oh, David. You have your hand up. Go ahead.

DAVID MACKEY: Hi. Actually, I'm wondering, from a DNS perspective, if Jonathan had anything to say with what he's been working on and how that might relate to his work. If not, that's fine. I'm just curious.

JONATHAN ZUCK: David, this is part and parcel of some of the things we've been talking about. So, way back in ICANN67 we talked about these types of issues and they're used primarily for phishing and pharming attacks to create end user confusion.

So, the example I used at the time that was very similar to Bill's Facebook example was a Bank of America example of several domains that had, in fact, been registered fraudulently that were just variations with very, very visibly similar—to an English speaker—to bankofamerica.com.

And so, it is a complicated question and I think we're going to have a constant struggle with the applicants and contracted parties, etc., who want to more narrowly define variants so that fewer and fewer things are handled automatically and more require human review.

And the notion that it has to be a "careful user" is a pretty difficult bar because most of us are uncareful most of the time when it comes to this. Right? We check things very quickly, etc.

I think it took forever for people to hover the mouse over a link before they clicked it and look at what the actual link was because we had the situation where the text of the link didn't have to match the actual address of a link. Right?

And so that's used innocuously when you say, "For further instructions click here," and the "here" was a link and then it was linked to a website. But very quickly it became the case where the text of the link could just look like another link and make it look like you were clicking on the address that you wanted.

And I feel like even getting the general public to where they could engage in the discipline of checking what the link was, the hyperlink. We're not even there yet. So, the notion of a careful user, I think, is probably a bad standard because even those of us that are professionals at this, so to speak, are uncareful most of the time. So, I think it's a very big issue.

Thanks for the question, David, unless you had something more specific.

EDUARDO DIAZ: Okay. So, Glenn wants a couple minutes to talk about the NomCom. But before that, I just put a link in the chat. It's called the "phishing quiz" made by Google. Take a look at that site. It's a quiz about phishing, and you have to get your score at the end to see how many of you get 100 out of 100. There are about eight questions. They're very, very interesting, so I encourage you to link into that website and pass this quick made by google. It gives you a chance to test your knowledge about phishing.

CLAUDIA RUIZ: Before Glenn goes, I have a question for him.

EDUARDO DIAZ: Okay. Go ahead. Judith, go ahead.

JUDITH HELLERSTEIN: So, Bill, on all these characters, do you also test that screen readers can read them and interpret them? You know, devices used by …

BILL JOURIS: In a word, no. There's a whole ICANN effort called Universal Acceptance which is attempting to deal with the fact that lots of browsers and, especially, lots of e-mail systems won't handle … In some cases, they won't handle all of the various scripts, and in some cases they won't handle anything except Latin.

JUDITH HELLERSTEIN: No. I'm talking about devices used by people who have low sight. Screen readers.

BILL JOURIS: Oh, I see. I have no idea.

JUDITH HELLERSTEIN: And I would be interested in knowing if there's an effort by ICANN to also have people with those being on the test [site] so they could actually interpret them because that's a different effort and it needs to be done because then we're leaving out a population in all the countries that these won't work in. So, that was my question.

BILL JOURIS: It would be interesting to have an effort in that regard, but I'm not aware of one at the moment which may just be my ignorance. But I'm not actually aware of anybody doing any testing on any of these to say, "Does this work? Does this confuse people?" I just don't see it happening.

EDUARDO DIAZ: Okay. I have many hands now. Your question is very good, but we ran out of time.

Gregory, can you be very quick? And then we'll stop with Glenn.

But before we go, Bill, thank you so much for putting this stuff together. It's an excellent presentation. And for taking the time of doing it. So, thank you so much, Bill.

Greg. And then Glenn, please.

GREG SHATAN: Thanks. I just briefly wanted to mention that the screen reader software picks up on the tag of the language of the webpage and uses it to interpret what the actual characters will mean. This can be a problem if the language tag is not entered into a webpage, because then the screen reader doesn't know if it's reading English or Georgian. Thanks.

EDUARDO DIAZ: Thank you, Greg. And Glenn, we end with you. Thank you.

GLENN MCKNIGHT: Okay. Thank you, everybody. And, again, thank you, Bill, for a great presentation.

I just wanted to emphasize—because the NomCom application process opens up on the 18th of January, this is before our next meeting—there are nine positions that current NomCom is entertaining. Three for the ICANN Board. There are three for ALAC, but not our region. So, again, emphasize there is no position for NARALO for ALAC. That's just for the other regions. Two spots open for GNSO, and one for CNSO.

So, if you're interested and if you have any questions about eh NomCom this year, please reach out to me. Thanks.

EDUARDO DIAZ: Thank you, Glenn. There is a link in the chat to fill out the survey. We want to see if these things are going fine and just to get some feedback about what happened in today's call. So, please take the time. It takes only two minutes to fill out.

Thank you so much. And with this, we adjourn this meeting. Thank you so much for attending today's meeting. Bye.

[END OF TRANSCRIPTION]