ICANN84 | AGM – ccNSO: Joint Tech Day and DNS Abuse Session
Monday, October 27, 2025 – 15:00 to 16:00 IST

CLAUDIA RUIZ

Hello and welcome to the ccNSO Tech Day session. My name is Claudia Ruiz and I, along with my colleague Andrea Glandon, are the participation managers for this session. Please note that this session is being recorded and is governed by the ICANN Community Participant Code of Conduct, the ICANN Expected Standards of Behavior, and the ICANN Community Anti-Harassment Policy.

Please observe the following guidelines to participate in this session. I will also post them in the chat for your reference. During this session, questions or comments submitted in chat will be read aloud if put in the proper form as noted in the chat. Interpretation for this session will include English, Spanish, and French. If you would like to speak during this session, please raise your hand in Zoom.

When called upon, virtual participants will be given permission to unmute in Zoom. On-site participants will use a physical microphone to speak. Please state your name for the record and the language you will speak, if speaking a language other than English, and speak at a reasonable pace to allow for accurate interpretation. Thank you, and with that, I will now hand the floor over to Nick Wenban-Smith, Chair of the DASC. Thank you.

NICK WENBAN-SMITH

Gracias, Claudia. Thank you, Claudia, and welcome everybody this afternoon. Bank holiday in Dublin, so thank you for taking your day off to come and listen to this. I think it's going to be a very exciting and engaging presentation. So, without further ado, just to recap here, obviously, the instructions.

The first one, to speak slowly and clearly, is actually a specific one they put in for my benefit, so I have to remember to speak slowly and clearly. Please, obviously, feel free to join the online Zoom room. You can have online commentary, chat, comments, questions, and we will track that. I'm logged in. The headsets, if you have some issues with the audio, the headsets are actually much better for hearing clearly what everybody is saying.

And, yeah, if you do want to speak in other languages, then start off your intervention by stating which language so that the interpreters can adjust and prepare. Thank you very much.

So, a huge thank you to the Tech Day organizing group for having us here for an hour this afternoon. The specific remit of the session is around the increasing use of artificial intelligence in combating DNS Abuse. Just for repeating context for people who weren't aware, we do an annual survey every two years of the ccTLDs globally, and use of AI went from pretty much zero in 2022 to pretty much 33% in 2024, and when that survey is repeated in 2026, I expect it to be north of 50%, so it's a really topical and important trend.

So, very pleased here to be joined by three ccTLDs who are going to give us a little snapshot of the adoption of artificial intelligence in three different ccTLDs. So, I'll allow my speakers each to introduce themselves, and then we'll get into the presentations. Just so everybody understands the format, there'll be three presentations, about 12 minutes, a little time for some specific questions on those presentations, so keep your interventions succinct and relevant to that presentation, please.

And at the end, I hope we'll have sort of 10 minutes or so for a more community discussion and a two-way engagement process where we can start to get some ideas from other people who haven't had a chance to speak about what they see the future for AI is. I'll allow my speakers to present themselves now.

PHILIP STRUYF: Good afternoon, everybody. My name is Philipp Struyf, and at the beginning of the year, I joined EURid as a Program Manager, Abuse Prevention.

JAKE VINCET: Hi all, my name is Jake Vincet, I'm Nominet's Registry Policy Manager, and I've been looking at our DNS Abuse strategy recently.

CRYSTAL PETERSON: Hello, my name is Crystal Peterson, and I'm a Senior Director of Key Accounts and Operations for GoDaddy Registry.

NICK WENBAN-SMITH

Fantastic, and could I just remind my speakers to speak slowly and clearly for the benefit of the non-native English speakers in the room and the translations team. Thank you. So, I do have a couple of introductory slides I just remembered. And this is just a recap of the DASC rationale and our mandate is set out in our terms of reference.

It's not to set policy, but it's to raise awareness and understanding, provide an open and constructive dialogue, and fundamentally to showcase good examples and best practices for the benefit of the whole community. And I say that specifically in the context of things I was hearing on Saturday for the smaller ccTLDs who are less well-resourced and want to get some of that experience. Recap of our activities, I referred to the survey previously.

All of our sessions are recorded. If you're interested in topics such as tools and measurements, the gTLD regulatory regime for abuse reporting and action, the role of maintaining accurate registration data in DNS Abuse and online scams and financial crime, then those sessions are all in the archive and you're very free to go back and look at those. So, I'll hand over now to Flet.

PHILIP STRUYF

Thank you and good afternoon, everyone. My goal today is to guide you through AI-powered domain name abuse detection at your ID. And after a brief introduction, I will highlight our challenges and

explain our challenges, what we were facing, and explain why we chose to implement machine learning and AI for our pre-delegation checker.

Next, I will detail the development and the current implementation of our prediction system, and then I will share the progress we made in offering the functionalities to the broader registry community. And as Nick said at the end, I will gladly address your questions.

So, my name is Philip, and at the beginning of the year, I joined .eu as a program manager, abuse prevention, and at your ID, we envision a safe and reliable top-level D to support the European digital single market. A key focus in achieving that goal is data quality, and we collaborate closely with our registrars and registrants to make sure that the data in our registry database are accurate. And this is both a priority as well as a legal obligation for your ID.

And the graph shows the evolution of the number of Flet suspicious domain names since the launch of Uridity, which is our internal data quality system, and it kind of serves as the central hub in our abuse prevention landscape, where all triggers, like our prediction system, converge, and where the KYC, the verification processes, start.

So, the goal is clear, but it's not easy to achieve. As you know, DNS Abuse is threatening us, and although the definition of DNS Abuse has been widely discussed and will never be perfect, I generally try

to stick to the categories and the definitions outlined by the ICANN Security and Stability Advisory Committee, and I guess I won't have to elaborate on them for this audience.

Now, the graph from the ICANN DNS Security Threat Mitigation Program highlights that spam, and we all know that it is considered DNS Abuse when it is used as a vector or a delivery mechanism for other types, has always been predominant. And these are the red portions of the bars on the graph.

Now, this validates our initial focus on detecting patterns in the registration data of domain names used for spam runs, and we performed this initially manually, and as you can imagine, it was both slow, labor-intensive, and ineffective due to the hit-and-run nature of those spam runs. However, it led to the foundation for the initial development of our prediction system.

And after extensive fine-tuning, and there the focus was to limit the amount of false positives, the system was deployed gradually, and we also switched to proactive measures, at that time being the delayed delegation of domain names.

Now, our current deployment is fully integrated with the .eu registration platform, and triggering the system now leads to immediate suspension of newly registered domains, and that is completely in line with the procedures as outlined by our internal data quality system. And soon, we aim to enhance and leverage the

system by opening it up to other registries, so they can benefit and contribute to the model, but more on that later.

So, our Abuse Prevention Early Warning System, abbreviated to APUSE, provides nearly real-time predictions of potential abuse. It has been developed in collaboration with the Catholic University of Leuven, and some scientific publications are available on our website. It is patented, and it has been proven effective for urethritis for many years.

Now, the prediction system analyzes domain name data at registration, so it is a pre-delegation check, and depending on the predictor model, it takes into account registration data, so the data that we get from the registrars, and or registration metadata, and that is data that we calculate from the registration data itself. So, for instance, we do an address check, or we calculate the reputation of the associated name servers with the domain name registration. We have currently three predictor models in place.

We have a reputation-based classification model, a similarity-based clustering model, which is shown here, and then we also have a third naive model that is focusing on burst registrations. Now, the models are trained continuously using both external sources and an internal feedback loop, and their output merges into a single domain prediction controller that is accessed by EURid via the reputation web service, as shown in this diagram.

So, where are we now? The current setup includes the machine learning part, as explained before, and a set of rules that check and

detect patterns and keywords where machine learning is unnecessary. So, the system runs as a loop of checks, each of which can be enabled or disabled, and also assign the weight.

When a set threshold is reached, the prediction changes to malicious, and this combination and configuration of components allows easy fine-tuning and optimization. For example, we can easily check for keywords related to the COVID-19 pandemic or the war in Ukraine, as we are requested by the European Commission, and this without wasting computational effort.

So, as mentioned and as shown on the graph, the system has been gradually deployed and effective for many years, but I want also to point out that it's only one building block in a larger abuse prevention landscape. We also perform post-delegation checks and analysis using crawling and thread feeds, and there we rely heavily on collaboration with third parties.

Now, the current on-premise deployment is, as I mentioned, fully integrated with the .eu registration platform, and that comes with its dependencies, like for instance, we have a specific message broker system. But obviously, it also limits the feedback loop that we have in place to the .eu domain name registration data, and that's why we explored ways to extend the system by opening it to other registries, allowing them both to benefit and to contribute.

So recently, we developed a proof of concept for APEWS as a service, and our focus there was threefold. First focus was to have

a containerized deployment, eliminating all EURid dependencies on European cloud provider.

Second focus point was the use of REST APIs to simplify integration, and the third focus was Data Anonymization for storing and training machine learning components to ensure data integrity. That led to the high-level design shown here, and the main idea is to have a shared APEWS system connecting multiple registries.

This system is to be hosted on European clouds to ensure compliance with EU data and security standards, and registries can score domain, provide score domain names, provide training data, and adjust their request rule sets, all via secure REST APIs. And the ultimate goal is, of course, to improve the APEWS detection accuracy through collaborative data contributions.

Now, regarding the roadmap of APEWS as a service, the core development and the establishment of AI models are complete, and the lift and shifts involved removing all EURid-specific components and replacing them with easy-to-use REST APIs, and data anonymization is also put in place. And basically, we are now ready to open our proof of concept to interested registries to request feedback on current and missing features. We have still some features on our radar to be developed, like bulk scoring and web interfaces, but we would like to welcome input on priorities from partner registries as well.

Now, estimating the costs and effort for a registry to integrate with APEWS is difficult, because it depends on the desired integration

level and the internal system technical architecture. Nevertheless, we outlined three possible scenarios, but in the interest of time, I won't go into detail, but my colleagues and I will be around to discuss these further if you're interested.

So, that concludes my presentation, and we would like to welcome collaboration to build an AI-powered European shared shield against domain name abuse. If you're interested, please reach out, or you can surf to APEWS.eu and request access to both our documentation websites and the proof-of-concept system itself, and like Nick said, if there are any questions, feel free. Thank you.

NICK WENBAN-SMITH: Thank you very much. I think APEWS might be my new favorite ICANN acronym. So, very, very interesting presentation. If there is any, there's a quick question I noticed in the chat from, there's a couple of questions. Luc Seufer, do you want to speak?

LUC SEUFER: I wrote it down because Claudia has a better voice than me. Yeah, I wanted to know if the identity of the European cloud service provider could be disclosed. Is it an actual EU-based provider, or is it a foreign, non-EU one that has clusters in the EU?

PHILIP STRUYF: No, no. The goal is really to have an EU-based hosting provider. We still have to do the final selection. We already have, of course, relationships with European hosting providers, and for the POC, we

have selected one, but for the production system, we will do a specific tender and selection.

NICK WENBAN-SMITH: And then, I've got a follow-up question from Andrei. Andrei, do you want to speak it? Yeah, please.

ANDREI KOLESNIKOV: Yeah, short question. Which neural network was used and trained to achieve maximum results? Which platform did you use? What kind of network?

PHILIP STRUYF: It's not really, it's basically enhanced statistics. It's not--

ANDREI KOLESNIKOV: So, you didn't use the neural engine? No. Okay.

NICK WENBAN-SMITH: Thank you. I don't see any other questions in the Zoom room, but I see several hands here. One, two, three, and from left to right, I didn't see which one was first.

MACIEK PIASECKI: Hello, Maciek Piasecki, first time ICANN Fellow. I wanted to know what's your false positive and false negative rate is, and how do you measure that, and if there's any appeal process?

PHILIP STRUYF The focus on false positives was when we did, when we have it as a dry run. Afterwards, it's very hard because once you, when you immediately suspend the domain names, you're not sure anymore if it was a false positive or not. Our current situation is that we ask the registrants of the suspended domain names to do an identity verification. So, in the end, it helps to improve the data quality, and it's the number of false, if even if the number of false positives would be very high, it wouldn't hurt us.

NICK WENBAN-SMITH That was the question I was going to ask. Next, please, behind you. Yes, sir.

SAKSHAM JAIN Hi, my name is Saksham, and I'm part of the NextGen program. My question is, can you speak a little about the post-delegation process at .eu? So, especially for someone who's had a running website, for example, for several years, and then it gets hacked, what does .eu do about it?

PHILIP STRUYF So, we do both. We rely on third-party thread feeds, and we also do crawling on existing websites, and as I said, it's integrated into the data quality platform. So when there is a trigger in the post-delegation phase, just as it's in, just as APEWS function as a pre-delegation trigger, the people are referred to the data quality

program, EURidity,,, and there they are requested to perform the KYC process via a number of available methods.

SAKSHAM JAIN: But is the current process manual, or does it do everything automatically, the crawling and everything?

PHILIP STRUYF: I have to give a mixed answer there. The thread feeds are dealt with manually, partly. The crawling is done automatically.

NICK WENBAN-SMITH: Thank you. And thank you. There's one other question, and then I'll close the queue for now, and we'll move on to the next presentation.

SAMANEH TAJALIZADEHKHOOB: Thank you for the presentation, Samaneh from ICANN Org. You've mentioned that you maximized the model for the purpose. Can I know what have you maximized the model for? Is it accuracy? Is it less false positives? There are a couple of metrics you can maximize the model.

PHILIP STRUYF: When we had it run as a dry run, we tried to minimize the number of false positives, and that was the main goal before we switched to proactive measures.

NICK WENBAN-SMITH Thank you, Samaneh. Thank you for joining the session. It's good to have you with us again. I've got one last question.

KAREN ROSE Thanks very much. Karen Rose from Interisle Consulting Group. Thank you very much for the presentation. Do you run the system on internationalized domain requests, and if so, is there any difference in performance?

PHILIP STRUYF We have it run for both the internationalized domain names in .eu, but also in the other scripts. But I must say the number of domain names in other scripts is very limited. The more data you have, the more powerful it gets.

NICK WENBAN-SMITH It's a great question, because it's a question we had. It came up in our prep session, and it's a very interesting answer as well. Okay, over to our next presenter.

JAKE VINCET Cool, thank you. I'm going to talk about Domain Watch. Domain Watch is a tool built in-house by Nominet, and it's one of the tools we've got for combating DNS Abuse. Before I jump into it, I should

say I can't take any credit for its development. That's done by Nominet's data science and insights team.

Domain Watch has been running for around seven years now, and we've iterated on it several times. What is Domain Watch, then? It's our anti-phishing initiative to further increase the security of .uk. It's designed to identify domains shortly after registration that are clearly deceptive, and it's a mixture of rules and machine learning models that look for combinations of typos on key brands and keywords. Like we heard in the previous presentation, the rules help us effectively deal with issues that we need to solve quickly, and COVID is a really good example of where we've used those rules. I thought we'd take a look at a bit of an example here.

Every registered .uk domain is checked against the Domain Watch models. I've got some examples listed on the screen there. You can see those that are scored pink are above a certain threshold, and we set that score to 9.1, and we've iterated on that score over time, looking at the false positive rate and optimized it several times. At the moment, it flags about 20 domains a day, and out of that, there's about a 50% suspension rate.

We do lift some suspensions, and a good example of that might be if a security researcher has registered a domain for a phishing exercise, and we're able to validate who they are and what their use case is. On those examples, you can see there that PayPan.co.uk is clearly very similar to PayPal, so that would tend to score quite highly, so that would be flagged and be sent to one of our analysts.

Similarly, towards the bottom, you've got tax-gov.uk there, so .uk has a second-level domain of gov. uk, so, again, that looks deceptive, as if it's looking like it's the gov.uk SLD. So, we've been refining that score over time, and we do have a concept of a near-miss queue, and that's where it hasn't breached the 9.1 score, but our analysts will go in and have a look at it there.

So, the model has been trained on multiple threat feeds, both open source and commercial, and those include phishing feeds and DJs, so they're looking for random strings, so you can see on the example there, we have a bunch of numbers, and .co.uk, it scored quite highly, but in this case, didn't quite trigger the threshold. So, one of the challenges our insights team have is kind of getting hold of good, reliable abuse data, so one of the things we've been thinking about for our use of AI is getting it to generate large data sets for us, particularly built around an organization or a campaign, and using that to train the models.

So, I thought I'd talk a little bit about the process. So, we have kind of a human-in-the-loop interaction, so the domain gets registered and immediately scored by Domain Watch. If the threshold is breached, it's sent to one of Nominet's analysts for review. They'll take a look at the findings, they'll have a look at the registrant data, and if they mark it as safe, it's a BAU process, it gets delegated and so on, but if they're not happy with it, the domain is suspended, and notification gets sent to the registrant, and we go through an ID verification process and understand their use for that domain, and

again, the analyst will make a decision there on whether or not we should allow the domain to be registered.

So, I thought I'd just show some numbers at a high level. So, we consider Domain Watch to be a really effective tool in combating domain abuse within Nominet, and we plan to kind of iterate and continue working on it further. So, in a year, we suspended over 4,000 domains via Domain Watch, and I've pulled out some kind of interesting categories there of data, so we saw quite a few related to login pages, and there's some examples of the domains there.

We saw some Barclays Bank was quite a high-profile target there, and similarly with HMRC, which is the UK's tax office. So, that's Domain Watch at a high level, but I did want to spend a little bit of time talking about a new tool that we've been using recently called Genie.

So, Genie is an AI-assistive feature of a product called Databricks, which is an analytics platform that Nominet has recently migrated to. Our use of Genie really, really is in its infancy, and we're kind of exploring what the use cases for this might look like, but what Genie does allow us to do is translate business questions into analytical queries, so that allows business users can query data that wouldn't typically have SQL knowledge or access to the database.

So, the way I think about Genie is kind of ChatGPT style. We could find the information elsewhere, we do have it elsewhere, but we can get it quickly without the need to do development work.

So, thinking about what a really basic use case could be for Genie, so here you can see I asked it a question to show me the domains which have been suspended more than once, which is kind of an interesting question to ask it. It replays the question back to us and returns the results. You can see there that I can show the code, so I can see the query that gets generated.

I think looking at this, you can see we've clearly got a number of domains there that have been suspended roughly 11 times, and I think that kind of begs the question of why has that domain been suspended so many times? We could also ask it to return key characteristics of the domains which have been suspended and to help kind of further use that data to train our domain watch model.

So, having a look at some slightly more advanced use cases for it, we've asked it is there any unusually high registrations for a registrar compared to their average? So, we've got a list there. You can see that daily registrations for a particular registrar peaked at over 1,000, but they normally have less than 150. So again, that on its own doesn't necessarily mean that there is domain abuse there, but it is kind of a leaping off point for us to go off and investigate, but also thinking about how could we use this going forward, kind of blue sky, could we use it to pivot on registrant data for associated domain checks?

So, as I said earlier, our use case really is in its infancy, but, you know, some questions, we found some interesting results and it's allowed us to go off and run some other investigations. And then finally, we've also got another good use case we found for Genie to try and help us work smarter when it comes to domain abuse and free our analysts up. And that's kind of collating reports and intelligence from a variety of sources and using Genie to analyze the reports and extract key data. So, that produces findings in a kind human consumable format for us for action. That is something we're kind of currently working on at the moment.

So, as I said at the start, Genie for us really is in its infancy and we're still looking at the underlying data and kind of training it. But hopefully I can come back and present on this in the future and kind of update you all on how we got on. I'd be happy to take some questions. Does any…?

NICK WENBAN-SMITH: I just have one, two, three, and there's nothing in the chat at the moment.

MACIEK PIASECKI: Maciek Piasecki, ICANN Fellowship. I wanted to know when you get to the ID check, you check that against what database, especially with foreign IDs. And if there's a further appeal procedure or clear way of court action.

JAKE VINCET Yes, we've got my tech ID verification check. So, we outsource that to a third party. So they'll let us know if they've passed. But yes, we do have appeals processes around.

NICK WENBAN-SMITH Second hand was, there we go.

FURKAN ÇOLHAK Thank you for presentation. I'm Furkan Çolhak, ICANN from fellowship. And I have a question about the data protection. You said that we're using LLMs for doing this kind of operations, but you're sharing your zone file kind of zone file. And I think there's a missing part about some data protection because you can't use third party large language models inside. So, you need to use Open Source. So, it needs some computational power. Thank you for presentation again.

NICK WENBAN-SMITH Cool. Yeah, let's chat about that after. Thank you. Yeah, Jake can talk to the torch that, I'm sure. So, I'm Nominet's data protection officer. So, another hat on. And finally, Luc, I think it was your hand.

LUC SEUFER Luc Seufer, yes. Could you expand on the role of the analyst review in the domain watch process? Because are those like trained

lawyers, like kind of bypassing the DRS by having a unilateral review of the cases or are they just looking at obvious DNS Abuse?

JAKE VINCET

Yeah, so this is completely separate to our kind of DRS process. So, we've got a group of threat analysts or DNS Abuse analysts in house. So, when they're reviewing the cases, they're looking at kind of who the registrant is, the registrant data. Do we know them? Are they kind of a security company previously, et cetera? Did that answer your question?

LUC SEUFER

Kind of. I'm not reassured, but yes. Thank you.

NICK WENBAN-SMITH

Thank you. And I don't see anything, comments or anything in the chat or any questions. So, let's move over to our third presentation. We're exactly on time, by the way. So, thank you very much, everybody, for keeping your points short and your presentations spot on. Over to you, Crystal.

CRYSTAL PETERSON

Awesome. Thank you. So, again, my name is Crystal Peterson, working with GoDaddy Registry for .us. And I'm excited to share with you today on some of our proactive approaches for DNS prevention, DNS Abuse prevention. Just a quick state of the domain of the size of what we're looking at is around 2.5 million

domain names, having a 57% blended renewal rate over the past year.

So, with combating DNS Abuse before it begins, as we all know, DNS Abuse is a persistent threat. Our traditional reactive approaches we feel are no longer fully sufficient. And so, we wanted to look to grow our threat management service to include a proactive model using AI powered threat intelligence and cross TLD pattern recognition.

So with reactive detection, it means that abuse is only mitigated after harm has occurred. So, we see that attackers can exploit gaps across TLDs. And our solution is to look to detect patterns for threats and temporarily block potentially malicious labels before they're registered, effectively helping to stop a patterned attack before it begins within our TLD.

So what we're looking to do is have a multi-layered system that includes historic alert integrations, real-time threat feeds from multiple sources, AI analysis and agent building, and then a blocking gateway platform. And this will allow us to act before any abuse reaches internet users.

So some of the goals that we're looking to achieve with this AI powered proactive prevention is to look at a 20% reduction in malicious registrations, have a less than 5% false positive, also to be able to have 75% faster abuse response times because we have

such a reduction in actual harmful abuse. So, what we're looking to do is a three-phase approach to implementation.

We are currently right now within our proof of concept. So, to be able to deploy this across .us and first focusing on phishing detection amongst all of the definitions of DNS Abuse. And then we want to expand into the other types. Also, we want to establish baseline metrics against the goals that I just shared. And then we are in developing the appeals process for anything that is blocked as well.

From there, we would move into a phase two, which is optimizing upon what has already been built. And then ultimately our goal is to be able to expand this across other portfolios of TLDs that are managed, both TLDs and ccTLDs, and look to be able to have some intelligence sharing across the platform and partnership integration.

So, one of the things that we are looking at is the fact that this initiative disrupts abuse mitigation. It helps to improve industry reputation for our TLDs, reduce costs, as well as it helps to strengthen brand protection for brands that may be using our TLDs. It also, we believe, supports regulatory compliance and fosters collaboration between TLDs and the community itself.

So with that, thank you very much. And I can take any questions.

NICK WENBAN-SMITH: Wonderful. Thanks, Crystal. Are there any questions for Crystal? One. Furkan, please go ahead.

FURKAN ÇOLHAK: Furkan again from Fellowship. And you mentioned about computational efficiency. So, we need to lower our, you know, increase the computational efficiency. But the problem is, in my mind, we have very strong computers right now. And also, proactive solutions can be, of course, perfect. But the problem is, pitchers or any other something else, when they register, the domain can be very unrelated things, like benefit of the water, for instance. When we are just decreasing the computational efficiency, I'm just looking for why we need it. By the way, thank you for presentation. It was all perfect.

CRYSTAL PETERSON: Thank you. What we are looking to do with this particular AI-powered machine learning is from what we've seen, where we have seen patterns within TLDs, and we have seen TLD hopping. So, this is looking to be able to help solve from that perspective. Yes, to grow to other things. But this particular test and experiment is looking to be able to, how can we further and quickly detect patterns that we do see jumps from TLD to TLD and keep that from jumping into our TLD?

NICK WENBAN-SMITH I have just one question, actually, for you, Crystal, if that's okay. I understand that you identify high-risk strings and then block them preemptively from registration. So I wondered, how big is that list and can you ever get off it?

CRYSTAL PETERSON Great question. So, from how big that list is, it could be small, it could be large. If we take there's been an attack out there that has been off of tolls, right? And some of those toll patterns we found have been a certain set of string characters that stay within, and then it is a randomized set of letters or numbers.

So, that can go into a large set, but typically that is not necessarily where brands play that you want that. And so that's what we would be temporarily blocking, but we are working on an appeals process that if you do find that your ideal domain name is within that, that you would be able to appeal and to be able to get that off.

NICK WENBAN-SMITH That's a follow-up to that then. So if I want to register a domain name innocently and I go onto a registrar platform, let's say I go to GoDaddy registrar platform, and I try to register it and it doesn't work, how do you interface with your registrars to let them know which of the domains which are available or not available for registration because of the threat blocking?

CRYSTAL PETERSON

Great question. So, we are still developing some of the messages and detail around when blocked, what that would show within the EPP. So, for example, any reserved names that are reserved by the registry, there's a blanket reserved response that currently right now would be that response. We are reviewing if we need to update that response for this particular service to be this is blocked due to a potential threat.

That is one of the things we're reviewing, or do we want to showcase to bad actors the fact that we are detecting it, which can then prompt them to try to be smarter kind of thing. So, that's what we're reviewing. But for right now, currently, it would be just this has been reserved by the registry.

NICK WENBAN-SMITH

Brilliant. Thank you. That's a very clear answer. So, it is registry reserved. So, we've now got, that's exactly right, we've got 15 minutes for more general conversation or other issues or people from the floor or other questions on a more general topic of AI and use of AI across registries and specifically around DNS Abuse detection and prevention.

So, start to think about things. I'm happy to, I've got a number of questions, which I'm happy to put across all three of our presenters. I'm happy to take further questions from the audience on a more generic sort of area of the use of AI and the increasing

use of AI. And it's opportunities and pitfalls. But first, another question for you, sir.

MACIEK PIASECKI: Maciek Piasecki, an ICANN Fellow. So, a question to all of you. I know ICANN strays away from content whenever possible, but are you able to use content either in the learning process or the actual suspension? Because if a website uses, let's say a Facebook logo or a bank logo, that would, I would say clearly market might be abusive.

PHILIP STRUYF: Yeah, we use it when we use screenshots and content when it's in the security, in the thread feeds. But at this moment, it's not used to train the predictive models.

JAKE VINCET: Yeah, I think that's similar for our tools. So, we'll look for screenshots when we're providing evidence of abuse, but not within domain watch or any of the machine learning.

CRYSTAL PETERSON: I'm going to say some of the same as well. So, part of our mitigation efforts, do use screenshots and from our analysts and investigator reviews. But within the place where we are with this tool, this is using past confirmed alerts as well as the feed. So, it's not currently using content.

NICK WENBAN-SMITH Very interesting. Are there any other further questions from the floor or online? I've got Hilde, and then Michaela. Excellent, good. Hilde first and then Michaela.

HILDE THUNEM Thank you. I'm Hilde Thunem from the Norwegian registry. I was just wondering, and thank you for very interesting presentations. I know this is tech day and I'm going to ask, you know, things about legal stuff. So I appreciate that it might not be easy to answer. But considering that you're using AI, I would suspect that you have looked at the AI Act and how this relates to the European regulation on the use of AI to make decisions. Could you say something? Well, for two of you, at least. Could you say something about what conclusions you've drawn? If you know about them?

NICK WENBAN-SMITH Just to clarify, I believe post Brexit, there's only one ccTLD here in the EU. Sorry.

HILDE THUNEM I am really sorry, Nick.

NICK WENBAN-SMITH I've still got PTSD, as you know, about all of this stuff. Philip?

PHILIP STRUYF: We've looked into this. The final review is still ongoing. But we, as I said, we have invested a lot of efforts into the data anonymization that is stored for the training of the models. So, that if you want to do the prediction, the data that needs to be stored in order to facilitate the training can be stored anonymously so that it's in line with the GDPR.

NICK WENBAN-SMITH: Brilliant. Thank you. Michaela, last but not least. We might well take up the rest of the time with this topic, I suspect.

MICHAELA SHAPIRO: Yeah, no, it's such a fun topic. Michaela Shapiro, for the record. Thanks for letting me crash this session. I'm here really just interested in the kind, and this will probably vary across each of your respective companies, but just curious whether you're envisioning with these kinds of tools, any kind of collection of additional data, or is the idea that this is all based on data that you already have access to?

And perhaps this could be taken up separately, but I'm very interested in the data and anonymization tools that you're using. But that could also be after the session. Thank you.

CRYSTAL PETERSON: I can start. So, data anonymization, as well as then, are there any further data points that are being collected as well? So, from respect, I'm going to go from the first one, which is any extra data

points. With what we're analyzing right now, we are analyzing past historic confirmed alerts and feeding that into our agents, but also current threat feeds.

There wouldn't be any additional data points we would be collecting from that. The registry itself has its own data points, but the points that we're looking to be able to derive are pattern-based and also where we are seeing confirmed abuse happening within other TLDs and making sure that that is protected within our TLD.

So, I would say that was a long answer to say no, but the registry itself does have other data points. And then anonymized data within the data feeds, which are public that can be purchased and or some of them are free based on which data feeds they are. Those data feeds are coming to us and we are ingesting, so they are mainly anonymized already that we get those sources.

And then we match that for the TLDs that we operate and support. We match that against internal data, but for TLDs that we wouldn't operate, that would be all anonymized based on what is given to us, as well as what we can get through either a zone file or a lookup.

JAKE VINCET Yeah, I think my answer would probably be very similar. We've obviously got the registrant data, but that's only looked at manually by the analysts once the main watch has flagged it on the string itself.

PHILIP STRUYF: Yeah, so in addition to what we get from the registrant data, we also do calculations. For instance, the reputation of the name servers attached over the past 15, 30 or 60 days so we do a number of calculations, but everything starts with the data that we get from the registrants.

NICK WENBAN-SMITH: Thank you. Actually, I thought you were going to ask whether any of the panelists had thought about the implications of human rights and whether there was a human rights impact assessment test. But you can ask that in a bit if you like. I've got Kristian actually first, and then Lucas, if you want to ask another one. Kristian first.

KRISTIAN ØRMEN: Thank you. My name is Kristian and I'm from .se. Question for .eu, I think the AP system is quite interesting and I saw on your website that we could try the proof of concept. So, a technical question, since we will soon be removing postal addresses from our registry and I saw that postal address was one of the data points the systems are using, would we be able to try out the system without using those data points or would you require the whole data set in order to test it?

PHILIP STRUYF I think the system will take into account whatever data we get, but we can check it with the developers of course if you're interested. Thank you.

NICK WENBAN-SMITH So, there's one question in the chat, it's from John McCormack. Pricing data would be a good input. TLD hoppers often follow discounts. Domain names registered as part of a discount might have a higher risk separate from registrations with stolen and or bad credentials. So I guess the question for the panelists is with the analytics tools that we have at our disposal can you also overlay that with price promotions and deep discounts at the retail registrar level?

JAKE VINCET Yeah, I guess that's a really good point and I guess something we've been thinking about with domain watches, how do we move it to the next level? I don't think that's anything we'd considered but definitely something to take away.

CRYSTAL PETERSON I think from our perspective for the actual detection of patterns and other abusive trends, pricing data is not something that's being input. However, that being said, pricing data against alerts and confirmed activity is something that we do tend to look at but it's not necessarily in the AI-powered proactive system that we're

looking to do too. But as we have confirmed alerts that is part of what we do and we analyze.

NICK WENBAN-SMITH: So, are there any other questions in the chat or from the floor? You, sir, first.

SAKSHAM JAIN: Hi, this is Saksham again from the NextGen Program. I don't know who wants to answer this, but a lot of the systems that have been talked about they kick in during the registration or the pre-registration time. Is any work being done for post-registration stuff? So, for example if a domain has been around for a few years and then it gets affected by DNS Abuse, is anything being done to make that process faster of getting the domain restored?

JAKE VINCET: I'm not too sure what you mean by restored I guess in that scenario, but I guess yes, but not using AI, I think it'd be the answer to that question.

NICK WENBAN-SMITH: Luc.

LUC SEUFER: So I did not ask it so I will for Crystal for .us. When you mentioned the block list, do you have any safeguard to ensure that it's not

encroaching upon freedom of speech that you are not blocking domain names that for example I know your current administration would not like to be registered?

CRYSTAL PETERSON: Great question. So, in this, this is looking to detect patterns and also detect and look at what is being used and or registered as abusive already. However, it may happen to be in another TLD and what we don't want to have occur is TLD hopping into our TLD. So, therefore if there are confirmed abusive activities that that's the pattern that we're looking for versus one-off domain names from that perspective of trying to cut off someone registering a domain name that wouldn't necessarily be in a pattern. Does that answer your question?

LUC SEUFER: So, DNS Abuse as it's defined by ICANN, no?

CRYSTAL PETERSON: Yes, that's correct.

NICK WENBAN-SMITH: So we've got a couple of minutes left before we wrap up and I've got a couple more questions just for everybody which is, they spend a lot of time in the ICANN community on a sort of doctrinal argument over the definition of abuse. Everybody here seems to be targeting phishing which is the predominant abuse that we see.

For simplicity should we just, instead of talking about DNS Abuse, talk about phishing? Because that helps the target. And my second question is around how do we measure success of these initiatives? And I liked particularly the US had an objective in terms of numerical reduction. I wondered whether there should be some sort of benchmarking or how do you feel that these things are working in terms of measurement? Whether you have… you have a problem, you have an intervention, then do you see the things change? Or is it that we're losing the arms race with the bad actors and threat actors who are conducting these phishing campaigns? Wondered if there's any thoughts on that, those sorts of subjects.

CRYSTAL PETERSON So, from the first question about should we just talk about phishing versus others, as, as you had mentioned, phishing is one of the most prevalent and so that's where we wanted to start first because we feel it's a low-hanging fruit. But we do want to be able to expand into some of the other definitions. So whereas there's phishing and then being able to look at, okay, are there the patterns and can we look at malware or some of the other CNC and bots and stuff like that. But phishing is where we wanted to start. Second question, apologies.

NICK WENBAN-SMITH About measuring success and are we winning the arms race?

CRYSTAL PETERSON So with the measurements of success, that is currently our goals. And so, we are in the infancy stages of being able to gather and, and start the blocking and looking at that. So, we don't have any metrics. We haven't been able to gather enough data to see how we're working against our goals.

NICK WENBAN-SMITH So, I think that's for me is the acid test, which is you're blocking domains, but are you actually preventing phishing?

CRYSTAL PETERSON Exactly. Part of what we're going to be looking at to see, are we preventing this, is info checks. If we see info checks against those strings, and we see that jump, other areas that we can see as well as within the feeds in the data, do we see that jump? Do we see it missed .us, but potentially jump to somewhere else that has to get mitigated versus it was proactively blocked?

Those are some of the areas where we're looking at. Now, are we meeting our goals because we are seeing a reduction because it didn't come in in the first place.

NICK WENBAN-SMITH Thank you. Any other comments from the panelists? Great. Well, if there-- oh, go on, Samaneh.

SAMANEH TAJALIZADEHKHOOB: One last question. Samaneh from ICANN Org. I wonder because there is a large body of research on the gap between what the data and what the models speculate based on the results of the data. So, basically the use of the LLM models from getting what the tables show into what this plot means.

Since now I hear that a couple of registries adopted LLMs inside their services. Have you guys tested if the interpretations that the LLMs return are in line with what the data shows, knowing that the data, one plot can have multiple interpretations based on what we ask?

NICK WENBAN-SMITH: That's a good question. And actually, I have a question for you, Samone, which is, is ICANN Org starting to use AI and LLMs in its work? If you could give us that answer in one moment. I'll ask the panelists first. That's a negative. So, could you share any light on ICANN Org's use of AI? Are they adopting it, or are they?

SAMANEH TAJALIZADEHKHOOB: No, no. AI as in the large language models, we are not adopting yet and there is, no plans yet. We are using machine learning models to do predictive works. This is kind of the same as part of the work that EURid is doing, but not more than that.

NICK WENBAN-SMITH: I think it's really interesting. I think I'll wrap it up there since we're coming up to the hour. The final comments are a huge thank you

to my panelists for sharing their insight and expertise. And thank you very much for the engagement and questions. I think my conclusion is to quote the poet T.S. Eliot, that AI is more of a whimper than a bang when it comes to ccTLD adoption for DNS Abuse. Thank you very much for your introduction, and join me in a round of applause for the panelists.

[END OF TRANSCRIPTION]