

# 2022-02-10 Registration Data Accuracy Scoping Team-Meeting #17

The Registration Data Accuracy Scoping Team call will take place on **Thursday, 10 February 2022 at 14:00 UTC for 90 minutes.**

For other places see: <https://tinyurl.com/mv9h95yv>



## PROPOSED AGENDA

1. Welcome & Chair Updates (5 minutes)
2. Measurement of accuracy (60 minutes)
  - a. Review of existing data sources - role in existing current situation & identifying possible gaps (see <https://mm.icann.org/pipermail/gnso-accuracy-st/2022-February/000253.html>)
  - b. Review input received on [google doc \[docs.google.com\]](#), see page 25 - How and by whom can it be measured whether current goal(s) of existing accuracy requirements are met?
  - c. Confirm next steps

3. Confirm action items & next meeting (Thursday 17 February at 14.00 UTC)

## BACKGROUND DOCUMENTS



## RECORDINGS

[Audio Recording](#)

[Zoom Recording](#)

[Chat Transcript](#)

GNSO transcripts are located on the [GNSO Calendar](#)



## PARTICIPATION

[CRM Attendance](#)

**Apologies:** Lori Schulman, Marika Konings (staff), Roger Carney

**Alternate:** Owen Smigelski



## Notes/ Action Items

### Action Items

1. Scoping Team to review the Support Staff [list of data sources for accuracy](#) and confirm that nothing is missing.
2. Scoping Team to work asynchronously via the email list to further discuss the definition of accuracy and what changes might be necessary. (For current working definition/construct, please refer to p. 12 of this [Google Doc \[docs.google.com\]](#).)
3. Still outstanding for BC, GAC, ISPCP, and SSAC reps: Scoping Team members who have not yet completed the assignment to consider what is needed and from whom to obtain information identified as necessary to measure whether current goals are met. Additionally, scoping team members to begin identifying specific ways in which measurement can be undertaken. (see **page 25** at <https://docs.google.com/document/d/11msexuoqWSUsFj8ZjVvWF-XHpcMJntWH/edit?pli=1#> [docs.google.com]). For groups that referenced ARS, please consider ICANN org's recent memo in your response: <https://mm.icann.org/pipermail/gnso-accuracy-st/2022-January/000236.html>.

**Registration Data Accuracy Scoping Team – Meeting #17**

Thursday 10 February at 14.00 UTC

1. Welcome & Chair Updates (5 minutes)

- Continuing the work of last week
- Many stakeholder groups have completed their homework, still waiting BC, GAC, ISPCP, SSAC
- If we finish this part of our work early today, we will discuss the potential of gathering UDRP decision data that has been discussed on the list

2. Measurement of accuracy (60 minutes)

a. Review of existing data sources - role in existing current situation & identifying possible gaps (see <https://mm.icann.org/pipermail/gnso-accuracy-st/2022-February/000253.html>)

- Further to Support Staff's action item, Support Staff has compiled a separate list of data sources that have been referenced.
- Many (if not all) were included as part of the briefing materials; however, now all are included in a dedicated document.
- Scoping Team to please review and ensure that this is complete.

b. Review input received on [google doc \[docs.google.com\]](https://docs.google.com), see page 25 - How and by whom can it be measured whether current goal(s) of existing accuracy requirements are met?

- **RySG:** This assignment was a challenge because Ry Reps are still struggling against which standard to measure accuracy against – there is no current agreed-upon definition.
- There are different ways to ask Rrs to report on their accuracy rates, such as ask Rrs to track email bounces from WDRP notices. One challenge here is that we, as an accuracy scoping team, do not have a way of compelling registrars to provide this data.
- Could have a third-party audit of existing data, but DPAs would need to be in place. While this is possible, there are challenges with cross-jurisdictional transfers of data.
- With respect to third-party audits and the issue of DPAs, instead of doing a random sampling, would it be better to focus on domains that show up in OCTO reports – that may, on its face, have a higher chance of passing the legitimate interest test
- Disagree – rather than looking at names that are so-called bad is not a comprehensive look at the ecosystem. Important to have a holistic look into the DNS and not just zero in on alleged bad actors.
- Doing scoping on what has already been identified as bad isn't going to get us very far. Think that there could be a carefully-crafted and narrow DPA exclusively for this purpose.
- There is a DPA between registries and registrars. This should be doable.
- The purposes have to be defined narrowly and the DPA needs to be crafted carefully – hard to say if this would be OK without seeing the DPA. Any transfer of data across borders (as seen in recent Google analytics decision) leaves some question as to how this could be organized. A third party should probably be used, rather than ICANN doing it. Ultimately, this will not be the easiest task.
- The group has been looking at data on the number of complaints ICANN gets. There is a difference of opinion on why that number has dropped between pre-GDPR and post-GDPR. As we're looking at easy data sets to get to, if we are looking at illegal activity, that may give us a heightened legitimacy for access to the data.
- What is the standard for measurement – unclear if this is about how to measure something or what thresholds exist for how good something is or how bad something is. There are basically two kinds of numbers that stick out: (1) number of domains viewed as abusive vs. (2) percentage of registrations per registry that are being used for malicious purposes – these two are different numbers. For new gTLDs, the set of registrations is relatively small, but the percentage for some is relatively high, rather than .com – which has a large number of abusive domains but a small percentage overall. Need to be careful about where we focus our attention. The drop in complaints from ICANN Compliance does not translate into a positive signal that accuracy rates have changed for the better.
- This discussion started way back when discussing voluntary contributions from registrars since ICANN cannot compel registrars to provide data. A DPA that would need to be voluntarily signed by certain registrars has very little merit. It's unlikely to be signed by bad actors. If there is no benefit to a registrar, why would they spend money doing this? ICANN could certainly decide to do audits of various sorts. Many companies have set up data protection authorities in certain jurisdictions to avoid cross-border data flow. In terms of audits of just domains that have shown abuse -this is a fine idea. Why is it that we want data? No need to contact all 200,000,000 people but may want to contact the ones that have some relevance. Targeting may not give us a holistic picture, but it could give us a view into accuracy vs. registrant data. It may be accurate as data, but not contact information for that registrant.
- The scope of this group is not abusive domain names, but looking at domain names involved in illegal activity could be the safest and legally viable data sampling to do.
- Do not see how abuse stats b/w TLDs relate to accuracy. Abuse is not what we are looking at regardless of what the domain name is used for. New gTLDs are, in general, more regulated than legacy TLDs. If new gTLDs have high abuse whereas .com has a moderate number of abusive domain names, that leads to the conclusion that more regulation does not lead to less abuse, or, the more regulation = greater abuse.
- The two most abused gTLDs, 2 account for over 40% of abuse
- If it seems like we need to get more data, this is perhaps a way to do this. Suggest that Rys and Rrs put their heads together and see what data we have, who could be the aggregator of data and come up with something that could be a light scope.
- It looks like only 3% of abuse deals with trademark infringement – this deals with specific automated detection system, and the detection system excludes malicious abuse. What does compromised mean? Does that go beyond malicious?
- CPs, during ICANN 73, will have a session about abuse and distinguish if it is about maliciously registered domain names or compromised domain names.
- Appreciate that this group of Rrs and Rys would go back and think about what could be done, but it's the age-old problem – the problem is usually not with the Rrs and Rys that participate in ICANN – especially not the registrars that are represented here. If all Rrs do not participate, the data would be skewed.
- Just because something is abusive is very prejudicial – this makes the assumption that inaccurate data means abusive data – this should be looked at more holistically.
- **IPC:** Shares opinions submitted by others such as ALAC – concerned about time to determine whether goals are met. In terms of AFAV, this would be helpful if this were enforced. GAC used the definition for domain name registration data for the ARS – this has been suspended, but its definition did also include identity (not just operability and syntax). Identity verification should be considered. ICANN org should be the entity that audits compliance. There were previous references to the NORC group that was the prototype for ARS and it included validation and they private and public sector validation experts. Perhaps we could use these experts again.

- When it comes to how to measure accuracy, there is a suggestion to have ICANN compliance do an audit that accuracy measures are being followed. Is that correct?
- Thought validation or verification experts from the public and private sector might be useful to assist ICANN in this exercise.
- **ALAC** – This is a different question – confused by how the scoping team is supposed to be running studies – if it is, the scoping team is short on time.
- If there are recommendations on what we are supposed to do, we can adjust the timeline accordingly.
- Group seems to be conflating whether people are adhering to the RAA and whether information is accurate. The RAA is very targeted and only looks at accuracy for some fields for new and updated names. We are not in a position to compel registrars to do a full audit of their own data. If there is a belief that this is out of scope and that ICANN does not have the right to do it, we need to get clarity on that.
- Most domain names fell under the 2013 RAA due because they were transferred or deleted. Only a small subset of domain names remained under the grandfathered rules – also names registered a long time ago usually do not present an accuracy issue
- If the group comes up with a mechanism for measuring accuracy, it is important to get the results of this to get to the other parts of the assignment
- What would this look like from a process standpoint?
- Procedurally, once this group has identified the guardrails or scope of any type of study, this would need to be documented and presented to the GNSO Council first, and this would need to be approved by the Council – it would then be submitted to the Board or Org (this presumes that it would be a sizeable resource). Ultimately, that request would be evaluated by the Board and Org to determine feasibility and funding. As it stands right now, this would be an unplanned expense – the planning and financing team would have to determine where the funding would come from. This may not be the exact process; this is how the process worked for EPDP Phase 1 and 2.
- The third charge to the scoping team – undertake an analysis based on data from assignment 2 (the measurement of accuracy). It's critical to have good data that we can take concrete and actionable steps against.
- It could take over six months if we commission a study like the NORC study, assuming it's all possible and viable to do that, this would be a very significant amount of time. In terms of what RDS review team found, understanding is that there are still significant domain names that are not treated as new and not subject to 2013 RAA requirements.
- If we must pause and wait for results of a study in order to get important data, that would be worth it.
- Have heard the argument that we need data to provide concrete and helpful recommendations. One option would be to stall work. One problem is to check the magnitude of the problem. The fact that there is an inability to check whether there is an accuracy issue – for that, we do not need data – we need to think of ways to tackle this issue.
- Group could work asynchronously on the definitional issue on the mailing list.

a. Confirm next steps

3. Confirm action items & next meeting (Thursday 17 February at 14.00 UTC)