2022-08-26 IDNs EPDP - Meeting #49

**Meeting date changed to avoid conflict with GNSO Council meeting

The call for the IDNs EPDP team will take place on **Friday, 26 August 2022 at 13:30 UTC for 90 minutes.**

For other places see: https://tinyurl.com/4tm58pm6

PROPOSED AGENDA

1. Roll Call & SOI (2 minutes)
2. Welcome & Chair Updates (5 mins)
3. Continued discussion of small group outputs on String Similarity Review and objection processes (80 mins)
4. AOB (3 mins)

BACKGROUND DOCUMENTS

EPDP Team Meeting #49 Slides - String Similarity & Objection Processes.pdf

PARTICIPATION

Attendance

Apologies: Lianna Galstyan, Farell Folly, Joseph Yee

RECORDINGS

Audio Recording

Zoom Recording (including audio, visual, rough transcript and chat)

GNSO transcripts are located on the GNSO Calendar

Notes/ Action Items

Welcome & Chair Updates

- Reminder to review Group 3 text (31 August)
- Acknowledgment of input from RySG, expressing concerns and suggested edits to Group 2 text.
- Team will not be meeting the week before and after the ICANN meeting.
- Leadership team giving thought to how to utilize sessions at ICANN75. Considering utilizing a process flow/process map to help understand context of recommendations and potentially stress test them.

Continued discussion of small group outputs on String Similarity Review and objection processes

- Slides were developed to help provide the rationale for why the hybrid model was determined to be the most appropriate to mitigate misconnection risks.
- Slide 3 – reminder of hybrid model.
- Slide 4 – background papers that help provide rationale for mitigating risk as much as possible. Includes RFC 5891, RFC 6921, SAC089, and the staff paper.
- Slide 5 – recap of the small team's discussions. Notes reservations about including non-requested allocatable and blocked variant labels. Also notes that the hybrid model is identified to meet the singular goal of risk mitigation of failure modes which are 1) denial of service and 2) misconnection. The small team did not consider implementation complexity and deferred to EPDP Team.
- Slide 6 – real world example of "denial of service" risk
- Suggestion the denial of service in this context is misleading and should be something like "no expected service provided". However, notes that the phrase is a known term and used in SAC060.
- Slide 7 – real world example of "misconnection" risk.
- https://www.icann.org/en/system/files/files/sac-060-en.pdf - "Recommendation 7: Should ICANN decide to implement safeguards, it should distinguish two types of failure modes when a user expects a variant to work, but it is not implemented: denial of service versus misconnection."
- Notes that the misconnection risk in SAC060 is in the context of variant set but misconnection risk described here seems to be in the context of crossing variant sets/name spaces.
- String similarity evaluation and objections are determined by people/evaluation panels, not strictly a technical perspective. Essentially, the hybrid model allows for identifying the misconnection risks but does no guarantee a certain outcome. Relying on a less restrictive model ensures that the risk will never be identified.
- Misconnection risk can be motivated by string similarity. The end-user ending up at the wrong web site can potentially be exploited.
- Reiterates that use of denial of service in SAC060 could be confusing versus distributed denial of service in SAC008.
- Confirms that terms used by this EPDP will be included in a glossary. And the term does seem to express what happens to the end-user – the user cannot access the service.
- Slide 8 – potential consequences of misconnection. Arriving at the wrong site can result in credential compromise and accidental exposure of information. And if the confusability is known by bad actors, it can be maliciously leveraged.
- Slide 9 – Why should blocked variants be part of the string similarity review? Notes that while the blocked variants cannot be delegated, they do exist as words or terms in real life. End-users may not understand the distinction between a blocked variant versus a real world.
- Slide 10 – Misconnection involving blocked variants. A blocked variant served as a bridge to misconnection in this scenario.
- Slide 11 – Next steps are risk analysis (e.g., likelihood and impact of failure modes, especially the misconnection risk); operational impact of the hybrid model; cost/benefit of the hybrid model.
- Slide 12 – 2012 New gTLD Program Implementation Review, which provides ICANN Org review of the program (including the String Similarity review). Helps to illustrate the scope of review in 2012, which did not have to contend with the hybrid model. This included two non-exact match IDN contention sets that were identified as a result of simplified and traditional Chinese characters. The volume of unique strings resulted in over a million combinations requiring evaluation.
- Assumption that SubPro should have mitigated timing issue (e.g., results published only two weeks prior to objections deadline). Helps to consider impacts from the increased complexity that would come from the hybrid model.
- IDN variants should be introduced successfully. Notes that cost and complexity are passed on to applicants based on the revenue neutral approach to the program.
- Confusability, or lack thereof, is certainly a factor in potential success. The String Sim process will exist again, but perhaps IDN variants can be evaluated in a different manner – keep an open mind.
- Question - Noting there were concerns from members on the hybrid model previously, do those concerns still exist?
- The question may need to be considered in the context of the "next steps".
- At a minimum, could compare complexity of the hybrid model versus level 2 (primary TLD and requested allocatable variants).
- Confirmation that the presentation helped illustrate why the "blocked variants matter". Need to understand if impact to cost would be prohibitive.
- Notes that comparing blocked variants versus blocked variants would be the most conservative and would result in exponential complexity. The hybrid model is intended to catch most cases, but lessen the complexity. The small team considered the hybrid model to be a compromise.
- Notes that the hybrid model requires human analysis, which is where the operational impact comes from. Highlights that not all blocked variants are well formed.
- Added that there are also mixed script blocked variants. Perhaps a hybrid model 2 would only include blocked variants that are valid labels.
- Confirms that the tool could produce sets that remove ill formed and mixed script variants.
- If two applicants apply for TLDs that are variants of each other, they would automatically be put in a contention set.
- Notes that fully mitigating risks introduces a risk in itself, by preventing opportunity.
- Important to consider the operational impact even of just level 2. Could the recommendation be that the evaluation be done in a layered way? Start preliminarily with just the primary labels, then work your way downward through allocatable and blocked variants.
- Points out that once something is blocked, it can end up being blocked for a long time. Therefore, should be careful what this group blocks.
- On the other hand, need to be sure not to block too little as there is little to no way to remedy that situation.
- Suggestion that recommendation could concentrate on the goal to be achieved rather than the "how" (e.g., the hybrid model).
- Seems that there appears to be comments leaning towards the hybrid model. Reminder that the policy level recommendations and/or hybrid model (e.g., the implementation, or the how) can be included in the initial report, which allows for course correction based on comments and potentially new information.
- For now, we'll set aside the risk analysis or cost/benefit analysis.

AOB

- None