Charter for WHOIS Review Team Internationalized Registration Data Expert WG

1. Background

Domain Name Registration Data (DNRD) [1] refers to the information that individuals or organizations ("registrants") submit when they register a domain name. Domain name registrars or registry operators collect these data, and some of the data[1] is made available for public display or for use by applications. The data elements that registrants must submit to ICANN Accredited Generic Top Level Domain (gTLD) registrars (and accurately maintain), are specified in the ICANN Registrar Accreditation Agreement (RAA). For Country Code Top Level Domains (ccTLDs), the operators of these TLD registries define their own data elements or follow their country's policy regarding the request and display of registration information.

Much of the currently accessible domain name registration data (DNRD) is encoded free form in US-ASCII. This legacy condition is convenient for WHOIS service users who are sufficiently familiar with languages that can be submitted and displayed in US-ASCII to be able to use US-ASCII script to submit DNRD and make and receive WHOIS queries using that script. However, these data are less useful to the WHOIS service users who are only familiar with languages that require character set support other than US-ASCII for correct submission or display. It is important to note that the latter (underserved) community is likely to continue growing and now appears likely to outnumber the former in a matter of years [2].

The WHOIS Policy review team, formed as a result of the Affirmation of Commitments (AoC), was charged to "assess the extent to which the WHOIS policy is effective and its implementation meets the legitimate needs of law enforcement and promotes consumer trust." In its final report, it highlights the needs to define requirements, data models and evaluate solutions for internationalized registration data, with the following recommendations:

> "ICANN should task a working group within six months of publication of this report, to determine appropriate internationalized domain name registration data requirements and evaluate available solutions; at a minimum, the data requirements should apply to all new gTLDs, and the working group should consider ways to encourage consistency of

---

[1] The quantum and nature of the data to be displayed varies according to registry or registrar's rules and policies.

approach across the gTLD and (on a voluntary basis) ccTLD space; working group should report within a year."

"The final data model, including (any) requirements for the translation or transliteration of the registration data, should be incorporated in the relevant Registrar and Registry agreements within 6 months of adoption of the working group's recommendations by the ICANN Board. If these recommendations are not finalized in time for the next revision of such agreements, explicit placeholders for this purpose should be put in place in the agreements for the new gTLD program at this time, and in the existing agreements when they come up for renewal." [3]

The ICANN Board adopted an Action Plan [4] in response to the WHOIS Review Team's Final Report, that instructs Staff to implement these recommendations.

At a high level, ICANN need to define the following requirements for internationalized registration data.

Submission –

1. Whether to allow users to submit internationalized registration data?
2. If users are allowed to submit internationalized registration data, what languages or scripts are registrars or registry operators expected to support?
3. If users are allowed to submit internationalized registration data, whether to require that users submit a corresponding single common script version of the internationalized registration data?
4. If users are required to submit a single common script version of the internationalized registration data, are users expected to submit a translated version, a transliterated version, or "either" (providing there is a convention or method to distinguish between the two)?
5. If users are required to submit a single common script version of the internationalized registration data and the user is unfamiliar with or unable to submit such a transformation, are registrars or registry operators expected to provide assistance (and if so, how would such assistance be manifested)?
6. If users are required to submit a single common script version and an internationalized version of their registration data, should there be a requirement to detect whether both literally match each other?
7. If there are two version of the registration data, which version should be considered primary if they mismatch?

Directory Services

1. Whether an end user can specify the language when querying IRD corresponding to it
2. If yes, how the user can specify the language when querying IRD

3. Which data should be displayed if there is no IRD corresponding to the specified language
4. Whether the user can query IRD with a internationalized string other than US-ASCII
5. If yes, whether IRD corresponding to the language of the queried string should be displayed
6. To define the degree of representation of internationalized registration data and the issues associated with it.
7. To provide a functional framework for display of entities agreed upon.

Storage
1. Thin registries: what data set is required for thin registries to store?
2. Thick registries: what data set is required for thick registries to store?
3. Data escrow: From whom and what shall be sent to an ICANN approved escrow provider?

Transmission

1. Should there be any requirement on the representation of the data in transmission?

2. Working Group Deliverables

The Whois Review Team Internationalized Registration Data WG will:

- Define the requirements for internationalized registration data (IRD)
- Produce a data model for the IRD that matches the requirement

In doing so the working group is expected to build on the previous community efforts on this issue, most notably:

- SSAC-GNSO Final Report on Internationalized Registration Data [2]
- SAC 054: SSAC Report on Domain Name Registration Data Model [7]

In addition, the WG will consider the work products of the following working groups, cooperating with them where possible and appropriate:

- GNSO PDP on Translation and Transliteration of Contact Names.
- ICANN Expert WG on next generation directory services. (ICANN EWG)
- IETF WG on Web-Extensible Internet Registration Data (WEIRDS)

The result of the WG product will go through the ICANN public comment process[2] to ensure broad input is received. During to the technical nature of the work, the IETF would also be invited to comment. It will form a basis for further policy development and contractual discussions in this area.

3. Work Plan

Here is the suggested work plan for the WG. This is subject to change as required by circumstances, especially the need to coordinate and cooperate with other working groups as mentioned elsewhere.

10/30/2013 – Approve Charter

12/18/2013 – Produce Draft Requirements for Internationalized Registration Data

12/18/2013 – 2/5/2014 – Public Comment on the First Draft of Requirements

2/26/2014 - Publish final draft of Requirements for Internationalized Registration Data

3/30/2014 – Produce first draft of Registration Data Model

3/30/2014 – 5/11/2014 – Public Comment on the Registration Data Model

5/30/2014 – Produce Final Registration Data Model

4. Methodology

Some of the questions to be considered by this working group include the following.

- Which data elements need to be considered for internationalization?
- What does it mean to internationalize the data elements?
- Should data elements be multi-valued to support multiple representations of the data?

---

[2] Public comment gives opportunity for the public to comment on each substantial piece of work before it is considered for approval. Each public comment topic (opened from 1 January 2012) is subject to a Comment and a Reply period as follows: 1) The official minimum Comment period is 21 days; 2) The official minimum Reply period is 21 days; 3) If no substantive comments are received during the Comment period, then there will be no Reply period; 4) During the Reply period, participants should address previous comments submitted; new posts concerning the topic should not be introduced. More information is available at: http://www.icann.org/en/news/public-comment.

The SSAC-GNSO Final Report on Internationalization considered the question of which data elements need to be internationalized. This work product was technically-based, focused on a known set of common data elements and the standards that could apply when internationalizing the representation of each of those elements.

The IETF WEIRDS Working Group has also considered the question of which data elements are part of the set of domain name registration data elements. A survey [5] of existing registry and registrar agreements and their directory service requirements was combined with the existing behavior of a subset of ccTLDs to create a profile of a registration data model.

It is also possible that the ICANN Expert Working Group on Next Generation Directory Services may suggest additional requirements for domain name registration data that will affect the internationalization of some data elements.

This working group will need to consider all of these inputs as part of its work to define the data model, which will need to consider the data elements that need to be included in the data model.

The GNSO PDP on the Translation and Transliteration of Contact Names is just beginning its work, currently focused on developing its charter. Based on the issues report [6] upon which it is based its work would appear to be focused on the quality of the registration data, specifically whether the data should be translated or transliterated to a single common script, and who should decide who bears the burden of performing this function.

This working group will focus on which data elements need to be internationalized and the requirements of that internationalization. Close coordination with the Translation and Transliteration of Contact Names work is essential to avoid conflicts and duplication of work.

5. References

1. ICANN Security and Stability Advisory Committee (SSAC). (20011) *SSAC Report on Domain Name WHOIS Terminology and Structure* (SSAC publication No. 051). Retrieved from http://www.icann.org/en/committees/security/sac051.pdf

2.   Internet Corporation for Assigned Names and Numbers . "*Final Report of Internationalized Registration Data Working Group*", 2012. Available at http://gnso.icann.org/en/issues/ird/final-report-ird-wg-07may12-en.pdf

3.   Internet Corporation for Assigned Names and Numbers. "*WHOIS Policy Review Team Final Report*", 2012. Available at http://www.icann.org/en/about/aoc-review/whois/final-report-11may12-en.pdf

4.   Internet Corporation for Assigned Names and Numbers. "Action Plan to Address WHOIS Policy Review Team Report Recommendations." 2012. Available at:
http://www.icann.org/en/groups/board/documents/briefing-materials-1-08nov12-en.pdf

5. L. Zhou, et. al. "Domain Name Registration Data Access Protocol Object Inventory Analysis." IETF Work in Progress (draft-zhou-weirds-dnrd-ap-object-inventory-01), 2013. Available at < http://tools.ietf.org/html/draft-ietf-weirds-object-inventory-01>

6. Internet Corporation for Assigned Names and Numbers (ICANN). (2013) Final Issues Report on translation and transliteration of Contact Information. Marina Del Rey, CA: ICANN. Available at: <http://gnso.icann.org/en/issues/gtlds/transliteration-contact-final-21mar13-en.pdf>

7. ICANN Security and Stability Advisory Committee (SSAC). (2012) SSAC Report on Domain Name Registration Data Model (SSAC publication No. 054). Retrieved from http://www.icann.org/en/groups/ssac/documents/sac-054-en.pdf