Dear Expert Working Group on Internationalized Registration Data,

ICANN Staff would like to offer the following comments on the "Draft Final Report from the Expert Working Group on Internationalized Registration Data". Overall, the report makes solid recommendations, however we would like to make a few suggestions for your consideration. These suggestions may impact implementability and consistency with other related initiatives that are currently underway.

1. **RDAP**. The report deals with internationalizing registration data for gTLD registries and registrars. One of the components of both registries and registrars is the Registration Data Directory Systems (RDDS), however one of the protocols used for RDDS, namely WHOIS, does not support internationalization, a situation noted in the very same RFC (3912) that defines the WHOIS protocol. We recommend that the Expert Working Group consider RDAP as a core requirement for internationalization, and, if appropriate, update the recommendations to be consistent with RDAP. It is also worth noting that ICANN is currently planning the move to RDAP following the publication of that protocol as an IETF standard last March.

   We suggest that the report be revised to consider RDAP as a pre-requisite to move forward with IRD and reference the use of RDAP features relevant to internationalization.

   This would also be important in light of the likely outcome of the Translations and Transliteration PDP Working Group that will recommend in its Final Report that Registrants should be able to obtain internationalized registration data in machine-readable form; something that RDAP would allow.

2. **Email Address Internationalization (EAI)**. As indicated in the report, support for Internationalized Email Addresses in email applications does not seem widely available at the moment. If the recommendation of this report to add support for EAI to registries and registrars moves forward, it would imply that contracted parties will need to update their systems to support EAI. These updates may be complicated to do at this time given the level of support of EAI in existing systems.

   It may be wise to review the EAI requirement in the context of the Universal Acceptance initiative to see if such a requirement makes sense at this moment. Perhaps the requirement could have a timeline of a few years (e.g., three to five) before requiring compliance. Another alternative may be for this requirement to be placed upon contracted parties once there is a certain uptake on existing implementations or once the major players in the email industry support EAI, etc. [Section 2.3.4, p17; section 5.2.3, p23]

3. **Repository Object Identifiers (ROIDs)**. In section 5.2.4 and Tables 5, 6, 7, and 8 of the report there is no consideration for a format for the object identifiers. It

may be worth noting that there is a format defined for Domain Name, Name Server/Host, and Contact Identifiers in the EPP standards. RFC 5730, section 2.8 describes the format of such identifiers. We recommend that the Expert Working Group specify the aforementioned format for the ROIDs. [Section 5.2.4, p 24; Tables 5, 6, 7, and 8, pp. 27-29]

4. **U-label format**. Section 5.2.6 and Table 5 require that users submit domain names and hostnames only in U-label format. In our opinion, there is no need to restrict this to only one format. We suggest that U-label support is a MUST, but also allow the use of A-label format if the users so desire. [Section 5.2.6, p24; Table 5 p27]

5. **Unicode**. Section 2.1 of the report seems to hint at the need to use Unicode. It may be a good idea to be clear on whether there is a requirement to use Unicode (more precisely, any or a specific encoding of Unicode, e.g., UTF-8). Particularly, you may want to explore requiring Unicode as the way to store data in order to potentially allow localization from the Unicode to whatever makes sense locally. By using Unicode, each character has a rich set of properties that can be used when dealing with localization and internationalization. For example, each character in Unicode has a script property. [Section 2.1, pp. 9-12]

6. **Thick Whois policy implementation**. It may be helpful for the Expert Working Group to consider whether the ongoing efforts to implement the GNSO's policy recommendations (adopted in October 2013) concerning the use of Thick Whois by all gTLD registries, existing and future, has implications for the IRD work and vice versa. For example, this policy has the implication that every single gTLD registry and registrar will use the same Whois format, which is described in the 2013 Registrar Accreditation Agreement (RAA). [Section 2.2, p13]

7. **Contact types**. Sections 5.1.1, 5.1.2, 5.2.2, 5.2.3 should be generalized to cover any contact (e.g., "Billing Contact" is quite common and there could be other types of contact used by gTLD registries). [Section 5.1.1, p20; section 5.1.2, p20; section 5.2.2, p23; section 5.2.3, p23]

8. **Delegation Signer (DS) data**. The requirement in section 5.2.8 to include DS data elements seems to go beyond current Whois requirements for both gTLD registries and registrars. It also ignores that some registries require DNSKEY instead of DS records. If the intention is to indeed extend the current WHOIS requirements, it may be worth noting this extension explicitly. In any case, the proper reference in the report in regards to format of fields should probably be RFC 7483 instead of RFC 5910. [Section 5.2.8, p 25]

9. **Contact details in domain name object responses**. Requiring the inclusion of all contact details in a domain object response as indicated in section 6.2 could have made sense in WHOIS. However, in RDAP, what probably makes more sense is that the query for a domain name returns the data for the domain name

and references for the contacts and registrant. In an RDAP context, if the server provides only the data of the domain name and references to the contacts, it gives the option to the client to either obtain that data or not. If the server provides all the data from the contacts it does not leave an option for the client, even if the client is not interested in such data. In an RDAP context it is simple to query other related objects if you need/want to. Finally, in the context of a data model (section 6.2) it probably doesn't make sense to require contact details in the response given that domain names and contacts/registrants are different objects. [Section 6.2, p26]

10. **Domain name, Registrar, and Name server object elements**. Section 6.2 of the report specifies the elements for Domain name, Registrar, and Name server objects in the context of a proposed data model. However, this proposal does not match what the Registry Agreements[1], Registrar Accreditation Agreement[2], Whois advisory[3], AWIP[4], and the Thick Whois Policy Recommendations[5] prescribe. The set of fields does not match what the aforementioned documents prescribe. It would probably make sense to either make it clear that the intention is to modify the current set of fields or revise the proposal to match the current set of required fields. [Section 6.2, p26]

11. **IDNs**. Table 5 seems to require a language tag for the domain name element of domain name objects only if they are internationalized. The cardinality indicates it is always required. You may want to either fix the cardinality to show it optional or change the text in the "Language Tag" column to indicate that the language tag is also required for ASCII names.

    Additionally, you may want to consider separating the Domain Name field from an (optional) IDN field as it is done in the Whois advisory, i.e., the domain name element is either the ASCII name or the A-label format of the domain name; the IDN element would be the name in U-label format and only present in case of IDNs. [Table 5, p27]

12. **Country codes**. Per the gTLD registry agreement, the 2013 RAA and the Whois advisory, the element country/territory for contacts is an ASCII two-letter code and, therefore, does not need a language tag. [Table 7, p28]

13. **Language tags**. The report requires language tags for various elements of both Contact and Registrar objects in Table 7 and Table 8 and one element in the domain name object in Table 5. Our reading of the RDAP RFCs indicates that you can only have language tags per domain name object instead of tags per each

---

[1] https://www.icann.org/resources/pages/registries/registries-agreements-en

[2] https://www.icann.org/resources/pages/approved-with-specs-2013-09-17-en

[3] https://www.icann.org/resources/pages/registry-agreement-raa-rdds-2015-04-27-en

[4] https://www.icann.org/resources/pages/policy-awip-2014-07-02-en

[5] http://whois.icann.org/sites/default/files/files/Thick%20Whois%20-%20Final%20Report%20-%2021%20October%202013.pdf

element. However, since the report only requires a language tag for one of the domain name object elements, it would seem that the object language tag could be used for that purpose.

Regarding the Contact and Registrar objects, having language tags for each of the Name and Organization fields should not be a problem. However, the Street, City, State/Province, and Country/Territory fields do not seem to have separate language tags per the jCard/vCard standards, which is what is used to represent individuals and organization contact details. Therefore, there could only be one language tag for the 4-tuple of the mentioned fields. On the other hand, given the hierarchical relationship between the four elements it would seem that having this technical restriction should be workable. We would recommend modifying the report to require only one language tag for the set of mentioned fields. [Table 7, p28; Table 8, p29; RFC 7483; RFC 6350; RFC 7095] (Thanks to Jean-Francois Tremblay from Viagénie for helping identify and confirm these issues.)

14. **Language tag use cases**. Table 7 and Table 8 require the use of language tags for certain elements of the Contact and Registrar objects, respectively. As with most changes in any system, the expected gain is what justifies the cost incurred by the change. It would be interesting to know the use cases and the benefits that the Expert Working Group has in mind for requiring the language tags, since it was not immediately obvious from the report.

    For example, requiring language tags for Contact object elements would mean: requiring the registrant to provide this data (three language tags per contact object), updates to the registrar interface offered to the registrant to be able to capture these, updates to the registrar EPP client, updates to the EPP server of the registry, updates to the SRS data base and related systems (e.g., web interface if offered by a registry), and developing EPP extensions.

    An example use case could be the ability to match registrants/contacts, which could have the benefit of allowing law enforcement officials to identify names owned by bad actor. In this case an officer could use the contact id or do a comparison of the name of the contact, either exact match or approximate, without needing to transform the registration data.

    We would recommend explicitly stating the costs, use cases, and benefits related to the use of language tags. [Table 7, p28; Table 8, p29]

15. **WHOIS server element**. The element WHOIS Server has been defined as a hostname (RFC952 and RFC1123) in the Whois advisory. Also, if you were to agree on the consideration for RDAP, this field will not apply anymore. [Table 6, p28; Table 8, p29]

16. **Sponsoring Registrar**. Table 6 requires a sponsoring registrar, as opposed to a sponsoring registrar id that Table 5 requires. Table 7 does not include a

sponsoring registrar id. It may be better to standardize Domain Name, Name Server, and Contact objects to include (only) a sponsoring registrar id. [Tables 5, 6, 7, p27-29]

17. **Cardinality of various elements**. Tables 5, 6, 7, and 8 define the allowed cardinality of the elements for each object. It would seem that various object elements do not match what is allowed in other contexts, e.g., Table 5 allows a maximum of 4 Domain status values while RFC 5731 allows up to 11, Contact State/Province is required by Table 7 when there are countries that do not have such subdivisions, and Contact Postal Code it is also not used everywhere. [Tables 5, 6, 7, and 8, pp. 27-29]

18. **Consideration of the Next Steps for the IRD Report.** Since the Board convened the IRD WG, the Final Report should be delivered to the Board to determine the next steps for this work. The question of whether those recommendations are ripe for implementation or require further analysis or policy work should be addressed.

    In Staff's view, implementation of several of the recommendations should be deferred pending the outcome of the GNSO's active PDP on Translation/Transliteration of contact data. This would ensure that the collective package of recommendations is consistent. In addition, since the Report includes many recommendations that may have policy implications, Staff notes that the IRD-WG could suggest that the Board send the IRD Final Report to the GNSO for appropriate follow-up. For example, depending in part on the timing of the delivery of the Report, the GNSO Council might consider requesting that its existing PDP Working Group on Translation/Transliteration examine these recommendations as part of its overall recommendations from the PDP. Alternatively, the GNSO Council could convene a follow-up effort to review the broader policy implications of the Report as they relate to other GNSO policy development work on Whois issues.

Thank you for the opportunity to comment on your draft report. Please let us know any questions you may have for us.

Regards,

ICANN Staff