

Public Comments-Report Template (v3.0)

Overview:

This template is being provided to assist Staff in the preparation of a report that summarizes and, where appropriate, analyzes community comments. Please save the document in either *.doc or *.pdf format and submit to: web-admin@icann.org. For presentation consistency and to preserve formatting, all Staff Reports will be uploaded to the forum in PDF format; text reports will no longer be supported.

Instructions:

- **Title:** Please enter the exact title that was used in the original Announcement.
- **Comment Period:** Enter the original Open Date and Close Date/Time (*Format: Day Month Year, e.g., 15 June 2013; Time should be expressed in UTC*). Please note if any extensions were approved, e.g., “Extended to Day Month Year [UTC Time]”.
- **Prepared By:** This field will accommodate a situation where a report is developed by an individual or group other than the principal Staff contact, e.g., a Working Group.
- **Important Information Links:** Do not enter any information in this section; Web-Admin will provide the appropriate links.
- **Section I: General Overview and Next Steps:** Please use this area to provide any general summary or highlights of the comments and indicate the next steps following publication of the report. (*Note: this field will auto-text wrap*).
- **Section II: Contributors:** Please use the tables provided to identify those organizations/groups and individuals who provided comments. It is not necessary to identify “spammers” or other commenters who posted off-topic or irrelevant submissions. In addition, if there is a large number of submissions, it is acceptable to characterize the respondent communities rather than attempt to list them individually in tables.
- **Section III: Summary of Comments:** This section should provide an accurate, representative, and thorough review of the comments provided. As the disclaimer explains, this is a summary only of those contributions which the author determines to be appropriate to the topic’s purpose. Authors are cautioned to be conscious of bias and avoid characterizing or assessing the submissions. If an analysis of the comments is intended, please use Section IV below. (*Note: this field will auto-text wrap*).
- **Section IV: Analysis of Comments:** Please use this section for any assessments, evaluations, and judgments of the comments submitted and provide sufficient rationale for any positions that are advocated. If an analysis will not be undertaken or, if one will be published subsequently, please add a note to that effect in this section. (*Note: this field will auto-text wrap*).

Note: You may also utilize, for this section, the Public Comment Issue Tracking Checklist template, which is available at: <https://community.icann.org/x/d67hAg>.

Translations: If translations will be provided please indicate the languages below:

--	--	--	--	--	--	--	--	--	--	--	--	--

Report of Public Comments

Title:	Draft Final Report from the Expert Working Group on Internationalized Registration Data															
Publication Date:	18 May 2015															
Prepared By:	Steve Sheng															
<table border="1"> <tr> <td colspan="2">Comment Period:</td> </tr> <tr> <td>Comment Open Date:</td> <td>12 March 2015</td> </tr> <tr> <td>Comment Close Date:</td> <td>21 April 2015</td> </tr> </table>		Comment Period:		Comment Open Date:	12 March 2015	Comment Close Date:	21 April 2015	<table border="1"> <tr> <td colspan="2" style="text-align: center;">Important Information Links</td> </tr> <tr> <td colspan="2" style="text-align: center;">Announcement</td> </tr> <tr> <td colspan="2" style="text-align: center;">Public Comment Box</td> </tr> <tr> <td colspan="2" style="text-align: center;">View Comments Submitted</td> </tr> </table>	Important Information Links		Announcement		Public Comment Box		View Comments Submitted	
Comment Period:																
Comment Open Date:	12 March 2015															
Comment Close Date:	21 April 2015															
Important Information Links																
Announcement																
Public Comment Box																
View Comments Submitted																
Staff Contact:	Steve Sheng	Email: steve.sheng@icann.org														
Section I: General Overview and Next Steps																
<p>As part of the broader effort to implement the recommendations from the ICANN Whois review team, the internationalized registration expert working group produced the draft final report that recommended requirements and data models for internationalized registration data (IRD).</p> <p>Following the input from this public comment forum on the final report, the working group will produce the final report for ICANN Board consideration.</p>																
Section II: Contributors																
<p><i>At the time this report was prepared, a total of [number] (n) community submissions had been posted to the Forum. The contributors, both individuals and organizations/groups, are listed below in chronological order by posting date with initials noted. To the extent that quotations are used in the foregoing narrative (Section III), such citations will reference the contributor's initials.</i></p>																
Organizations and Groups:																
Name	Submitted by	Initials														
ICANN Staff	Francisco Arias	ICANN Staff														
Business Constituency	Steve Delbianco	BC														
Intellectual Property Constituency	Greg Shatan	IPC														
Registry Stakeholder Group	Paul Diaz	RySG														
Registro .it	Rita Forsi	Registro .IT														
Individuals:																
Name	Affiliation (if provided)	Initials														
Chris Dillon	Co-Chair GNSO PDP On Translation and Transliteration in individual	CD														

	capacity	
Greg Shatan	President, IPC in individual capacity	GS

Section III: Summary of Comments

General Disclaimer: This section is intended to broadly and comprehensively summarize the comments submitted to this Forum, but not to address every specific position stated by each contributor. Staff recommends that readers interested in specific aspects of any of the summarized comments, or the full context of others, refer directly to the specific contributions at the link referenced above (View Comments Submitted).

Overall comments:

- CD commented that scripts where letters can be read in more than one way or which do not use spaces to define word boundaries (Japanese falls into both of these categories) will be the most resistant to automated transliteration/translation.
- Registro .IT has no particular comments to make, but fully agrees with the principles expressed in the report.
- The RySG strongly recommends using existing RFCs to address registry policy decision and technical adjustments to enable consistent and reliable passing of registration data in native characters. Specifically, it recommends a standard policy on the use of the EPP “loc” and “int”. Internationalized e-mail address should be passed end-to-end, RFC 5733 supports internationalized characters, but RFC 5322 does not support internationalized email address formats.
- ICANN staff noted that current WHOIS protocol cannot consistently support internationalized registration data and recommended the Expert Working Group consider revise the report to consider RDAP as a pre-requisite to move forward with IRD and reference the use of RDAP features relevant to internationalization. (RFC7480-7485).
- ICANN staff noted the level of support for EAI in email applications does not seem widely available at the moment and suggested the working group to consider review the EAI requirements in the context of Universal Acceptance initiative to see if such a requirement makes sense at the moment. It further recommended a requirement 3-5 years before requiring compliance.

Proposals for internationalizing contact data:

BC's view:

- The BC supports the principle that registrants should only be required to input registration data in languages (scripts) with which they are skilled. However, as stated in our previous comment, the BC believes that, to balance the needs of domain name owners, registrars, and users of Whois, it is important that non-ASCII character sets are supported but not required, meaning that registrants may use ASCII or non-ASCII scripts according to their skill set.
- The BC supports the principle that unless explicitly stated otherwise, all data elements should

be tagged with the language(s) and script(s) in use, and this information should always be available with the data element. However the BC adds that localized tagging does not replace standard USASCII data labels, but rather accompanies it. This is the appropriate balance between the needs of the local registrant, registrar and registry, and the global needs of the users of Whois data.

- The BC notes that IRD Group selected Proposal 1 instead of the proposal 3 BC supported. However the BC does not object to proposal 1, except that BC thinks that proposal 3 and proposal 1 achieves similar goals.
- The BC did not understand how free form text in the contact data elements would be problematic to registrars under RAA 2013, as currently registrars only required to validate phone numbers and emailing addresses.

RySG's view:

- RySG believes that the purpose of any domain name registration data, internationalized or not, is to find reliable contact information for any given domain name; and the role of registries is to ensure that the DNRD can be set and propagated end to end in native characters without any transformation. The RySG recommended that the requirements are set to meet the needs of the general users or registrants of domain names. Such users, in the context of domain name registration data, should only be required to input contact information in a language or script of their choice.
- RySG is supportive of an internationalized registration data model in which end-to-end registration data can use native characters. RySG believes that the needs of users like law enforcement and trademark agents represent special and not general users of registration data so it may not make sense for all users to subsidize the fulfillment of their needs. These needs should be met by a different system that could use domain name registration data as a source.

In response to RySG's input, GS strongly disagrees:

- RySG's statement reflects a significant lack of understanding of the concerns, needs and even identity of those who use registration data. This issue does not concern "trademark agents", but concerns trademark owners. Virtually every business in the world (including non-profits) is a trademark owner. Trademark owners can be large or small – or even individuals – and they are in every country in the world. In addition to trademark owners, copyright owners are equally concerned with intellectual property.
- The number of these users is numbered in the millions, if not hundreds of millions. These cannot be "special users."
- GS left it to global law enforcement to comment on their ability to use government funds to pay for registration data. For larger trademark owners and copyright owners, this could be a significant expense, which could well be passed on to consumers. For smaller trademark and copyright owners, this could be the difference between being able to defend their works and businesses, and letting piracy, counterfeiting and cybersquatting put them out of business.

IPC views the working group's selection of proposal 1, which allows any free form text to be used for WHOIS data, with the restriction that the scripts of the address should be appropriate for the region in which it is located, is problematic to the Internet community as a whole because it:

- Allows a script to be used for WHOIS data that is at odds with other requirements for WHOIS data and with the domain names themselves;
- May be expensive and impractical for many registrars and registries to verify and validate as may (and should) become required of them;
- Focuses solely on the presumed interest of the registrant, without appropriately balancing the registrant’s needs with the impact on the main interests of many other users of WHOIS data for its intended purpose – to easily identify, locate and contact domain name registrants – and the widely-acknowledged importance of improving the quality of WHOIS data.
- Adopting proposal 1 creates the risk that bad faith actors will deliberately obscure ownership information about the domain by selecting a seldom-used language/script with no relation to the registrant or the TLD, a risk that must be weighed against the relatively rare occurrence that the registrant has a legitimate need to display ownership information in a language/script other than those representable in ASCII or in the TLD’s script.
- Adopting an entirely free form text model as proposed by the IRD Working Group would place WHOIS data at odds with other information concerning domain names that is already required to use US-ASCII. For instance, characters in domain names to the right of the “.” in the generic top-level domains (other than IDN gTLDs) are already in US-ASCII format. It would be inconsistent and impractical to tell registrants that they may use a script other than US-ASCII for WHOIS data for a domain name that already includes US-ASCII. The use of US-ASCII must already be a part of the “ordinary daily routine” of such registrants to the extent that this routine includes writing, typing, and using the gTLD domain name they have registered.

The IPC requests that proposal 2 be considered. If the IRD and the ICANN community insisted on the adoption of proposal 1, the IPC suggests the following revision: “Proposal 1: Free form text. The language and script of an address MUST MATCH THE LANGUAGE AND SCRIPT USED BY THE OFFICIAL NATIONAL POSTAL AUTHORITY OF THE COUNTRY IN WHICH THE REGISTRANT IS LOCATED.”

Comments on language tags

ICANN staff noted the following about 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 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]

- **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]

Comments on next steps from ICANN staff

- 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 EWG could suggest that the Board send the this 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.

Comments on others sections of the report:

- Page 5: The proposed requirement for Registrar Name should be consistent with the proposed requirement in Section 5.1.3 which reads: "Free-form text. The name of the sponsoring registrar should be the official name in the Registrar Accreditation Agreement (RAA) with ICANN." The sentence "in a language and script appropriate for the region in which the registrar is located" should be removed as it may conflict with the former requirement (i.e., official name in the RAA). For example, some registrars have legal names in English even

though English is not the official language of the territory they operate in. (RySG)

- Sections 2.1, ICANN staff noted that 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, the EWG 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.
- Sections 2.2, ICANN staff noted 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).
- Sections 5.1.1, 5.1.2, 5.2.2, 5.2.3, ICANN staff noted these sections should be generalized to cover any contact (e.g. Billing Contact).
- Sections 5.2.4 and Tables 5,6,7,8: ICANN staff noted that 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 5.2.6 and Table 5. On U-labels, ICANN staff commented 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 5.2.8. ICANN staff noted 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. Also the proper reference in the report in regards to format of fields should probably be RFC 7483 instead of RFC 5910.
- Sections 6.2 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.
- 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, Registrar Accreditation Agreement, Whois advisory, AWIP, and the Thick Whois Policy Recommendations 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.

- Table 5 seems to require language tag for domain name only if they are internationalized. Suggest either fix the cardinality to show it is optional or change the text to indicate language tag is also needed for ASCII names. Additionally consider separating the Domain Name field from an (optional) IDN field.
- References to RFC 5322 should be updated to state RFC 6854. (RySG)
- Table 4 is an ideal of clean data, that currently may only be produced manually and that it contains aspects of both transliteration ((e.g. 千代田 -> Chiyoda) and translation (e.g. ビル -> Bldg.). Moreover, the relationships between the original and the transformed records are complex and it is not possible to move automatically in either direction. (CD)
- Tables 5,6,7, 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.
- Table 6: 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.
- Table 7: There is blank row between contact Country / Territory and Contact Postal Code. Recommend either remove it or label it as intentional to avoid confusion; The Contact Country/Territory minimum and maximum length should be 2 to be consistent with ISO-3166 Alpha 2 codes. There is little value in requiring language tag for Contact Country / Territory Code as it is a 2 character ISO code that is always in ASCII, recommend changing it to not applicable (RySG, ICANN Staff)
- Table 8 WHOIS server element. The element WHOIS server has been defined as a hostname (RFC952 and RFC1123) in Whois Advisory. This will be obsoleted by the RDAP protocol, suggest remove it.

Section IV: Analysis of Comments

General Disclaimer: This section is intended to provide an analysis and evaluation of the comments received along with explanations regarding the basis for any recommendations provided within the analysis.

Given the public comments, it is clear that the community has not agreed on what is the best model moving forward to internationalized contact data. Thus, further community based policy discussions are needed.

Various commentators has provided detailed feedback on sections of the report, these comments will

be addressed an incorporated into the final report.