

## DRAFT: Translation and Transliteration of Contact Information Policy

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119, which is available at <http://www.ietf.org/rfc/rfc2119.txt>.

### Scope

This Policy SHALL apply to all ICANN-accredited registrars and gTLD registries.

### Announcement – Effective Date

TBD

All provisions detailed herein are dependent on the availability of the Registration Data Access Protocol (RDAP) as a production service in the gTLD space and corresponding Extensible Provisioning Protocol (EPP) extensions to accommodate the display of transformed data. As of Q3 2017, RDAP is in a pilot phase, and all EPP extensions required to implement this policy have not yet been defined.

Those provisions requiring an EPP extension have been emphasized in the policy text below.

### Definitions

- 1.1. **Contact Information** refers to data fields corresponding to a single contact (registrant, administrative, technical and billing are roles of a contact with respect to given domain name). Contact information includes the following data: Contact ID, Contact Name, Contact Organization, Contact Address, City, State/Province, Country, Contact Postal Code, Contact Phone, Fax, E-mail. It can also correspond to a single registrar and includes the following data: Registrar ID (conforming to the IANA registrar-ids registry), Contact ID of Registrar, Registrar Administrative Contact ID, Registrar Technical Contact ID, Registrar Billing Contact ID, Registrar URL, Registrar Creation Date, and Registrar Last Updated Date (see Implementation Notes)
- 1.2. Registrars MAY gather language data from registrants in a manner of their choosing.

## 2. Requirements for language tags within the Registration Data Access Protocol (RDAP) (for those parties opting to display transformations of RDDS data in RDDS outputs):

- 2.1. **[EPP extension required]** If a transformation of registration data is performed and displayed in an RDDS output, an RFC 5646-compliant language tag **MUST** be displayed alongside the original registration data provided by the registrant for each data element requiring one according to the IRD data model provided in the Implementation Notes.
- 2.2. **[EPP extension required]** An RFC 6497-compliant language tag **MUST** be displayed alongside each data element for any transformed registration data to indicate the source of and type of transformation performed. In cases of transliteration, the RFC 6497-compliant language tag **MUST** include the transliteration standard that was used.
- 2.3. **[EPP extension required]** The entity that has performed any transformation of registration data that is displayed in an RDDS output **MUST** be identified when displaying a transformation of contact information. The accepted values for these entities are limited to:
  - Registrant
  - Reseller
  - Registrar
  - Registry
  - Other
- 2.4. Language tags for each data element requiring them per the IRD data model **MUST** contain at least the primary language and script subtags per RFC 5646.
- 2.5. A script subtag **MUST** be valid for all Unicode code points detected by automated means. Registries **MUST** validate that the Unicode code points in the entered data match with the corresponding script subtag.
- 2.6. If the language is not known, the primary language subtag "und" **MUST** be used in the language tag.
- 2.7. If the script cannot be determined (e.g., in cases of mixed scripts that do not conform to mixed-script usage within a language) the script subtag "Zyyy" **MUST** be used in the language tag.
- 2.8. Private use language tags as described in [RFC 5646, Section 2.2.7](#) **MUST NOT** be used to generate language tags.

Implementation Notes for detailed data output model).

**Transliteration** refers to the practice of writing or printing a letter or word using the closest corresponding letters or symbols of a different alphabet or language.

**Translation** refers to the practice of converting the meaning and sense of words and/or text from one language to another.

**Transformation** in the context of this policy refers to translation OR transliteration.

**RDDS** is the acronym for Registration Data Directory Service.

**IRD Data Model** refers to the schema detailing which data elements require an RFC 5646-compliant language tag. This model applies only to those registries and registrars who voluntarily opt to display transformed data in RDDS outputs. It can be found in the [Implementation Notes](#).<sup>1</sup>

**The following provisions apply to both gTLD Registries and Registrars:**

- 2.9. It is OPTIONAL for registries and registrars to perform transformations of contact information in RDDS data fields.
- 2.10. It is OPTIONAL for registries and registrars to display transformations of contact information in RDDS data fields.
- 2.11. Registries and registrars MAY support any language and script for registrants to input registration data. Any data input into such a service MUST conform to standards in the Registrar Accreditation Agreement (RAA), Registry Agreement (RA), relevant Consensus Policy, Additional WHOIS Information Policy (AWIP) and any other applicable policies. Registries and registrars MUST validate entered registration data for correct format and accuracy in accordance with the aforementioned Policies and Agreements (see [A script](#) subtag MUST be valid for all Unicode code points detected by automated means. Registries MUST validate that the Unicode code points in the entered data match with the corresponding script subtag. on script validation).

---

<sup>1</sup> This data model is a product of the [Internationalized Registration Data Working Group's Final Report](#), and has been incorporated into the Translation and Transliteration of Contact Information implementation per [Board Resolution 2016.03.10.07](#): "Resolved (2016.03.10.07), the President and CEO, or his designee(s), is directed to work with the implementation review team for the new consensus policy on translation and transliteration to consider the IRD Working Group's data model and requirements and incorporate them, where appropriate, to the extent that the IRD's recommendations are consistent with, and facilitate the implementation of the new consensus policy on translation and transliteration."

- 2.12. Registrars who opt to gather language and script data from registrants for transfer to and validation by registries SHOULD reach an agreement with relevant registry operators to populate and validate the underlying data for language tags in accordance with the provisions outlined in [above](#) and those Policies and Agreements detailed in Registries and registrars MAY support any language and script for registrants to input registration data. Any data input into such a service MUST conform to standards in the Registrar Accreditation Agreement (RAA), Registry Agreement (RA), relevant Consensus Policy, Additional WHOIS Information Policy (AWIP) and any other applicable policies. Registries and registrars MUST validate entered registration data for correct format and accuracy in accordance with the aforementioned Policies and Agreements (see [A script](#) subtag MUST be valid for all Unicode code points detected by automated means. Registries MUST validate that the Unicode code points in the entered data match with the corresponding script subtag. on script validation).
- 2.13. **[EPP extension required]** If a registrar opts to gather language and script data from a registrant to populate the underlying data for a language tag, the registrar MUST provide the language data to registries to populate language tags in RDDS outputs for each data field requiring a language tag according to the IRD data model provided in the [Implementation Notes](#).
- 2.14. If a transformation is carried out and displayed in an RDDS output, the original data MUST be displayed along with the transformed data. If multiple transformations are carried out, registries and registrars MAY display the various versions of transformed data at their discretion, so long as the original data is displayed.
- 2.15. If a transformation is carried out and displayed in an RDDS output, the original and transformed data elements MUST be accompanied by a language tag per the requirements set out in [Requirements for language tags within the Registration Data Access Protocol \(RDAP\) \(for those parties opting to display transformations of RDDS data in RDDS outputs\)](#): and according to the IRD data model provided in the [Implementation Notes](#).
- 2.16. Registries and registrars MAY use automated means to detect scripts.
- 2.17. It is RECOMMENDED that the latest version of Unicode be utilized for entry of registration data.

### 3. The following provisions apply to Registrars only:

- 3.1. Registrars MAY allow registrants to provide language data for the registration data elements requiring a language tag in the IRD data model.

- 3.2. Registrars MAY allow—but MUST NOT REQUIRE—registrants to provide transformations of registration data elements requiring a language tag in the IRD data model.
- 3.3. Registrars MAY gather language data from registrants in a manner of their choosing.
4. **Requirements for language tags within the Registration Data Access Protocol (RDAP) (for those parties opting to display transformations of RDDS data in RDDS outputs):**
  - 4.1. **[EPP extension required]** If a transformation of registration data is performed and displayed in an RDDS output, an RFC 5646-compliant language tag MUST be displayed alongside the original registration data provided by the registrant for each data element requiring one according to the IRD data model provided in the Implementation Notes.
  - 4.2. **[EPP extension required]** An RFC 6497-compliant language tag MUST be displayed alongside each data element for any transformed registration data to indicate the source of and type of transformation performed. In cases of transliteration, the RFC 6497-compliant language tag MUST include the transliteration standard that was used.
  - 4.3. **[EPP extension required]** The entity that has performed any transformation of registration data that is displayed in an RDDS output MUST be identified when displaying a transformation of contact information. The accepted values for these entities are limited to:
    - Registrant
    - Reseller
    - Registrar
    - Registry
    - Other
  - 4.4. Language tags for each data element requiring them per the IRD data model MUST contain at least the primary language and script subtags per RFC 5646.
  - 4.5. A script subtag MUST be valid for all Unicode code points detected by automated means. Registries MUST validate that the Unicode code points in the entered data match with the corresponding script subtag.
  - 4.6. If the language is not known, the primary language subtag "und" MUST be used in the language tag.

- 4.7. If the script cannot be determined (e.g., in cases of mixed scripts that do not conform to mixed-script usage within a language) the script subtag "Zyyy" MUST be used in the language tag.
- 4.8. Private use language tags as described in [RFC 5646, Section 2.2.7](#) MUST NOT be used to generate language tags.

## Implementation Notes

1. Data Model for RFC 5646-Compliant Language Tagging for Original and Transformed RDDS Outputs (to be utilized by those parties who voluntarily opt to transform original data AND display the transformed data in RDDS outputs)

Table 1: RDDS Model for Domain Object<sup>2</sup>

Data Element	Format	Min length	Max length	Cardinality	Language Tag (RFC 5646)
Domain Name (Internationalized)	RFC 5890	1	255	1	Required if it is U-label or A-label.
Domain ID	Freeform text	1	255	1	n/a
Referral URL	RFC 3986 / 3987	1		1	n/a
Updated Date	RFC 3339 (Date and time in UTC as specified in RFC3339, with no offset from the zero meridian).		32	{0,1}	n/a
Creation Date	RFC 3339		32	1	n/a
Registry Expiry Date	RFC 3339		32	1	n/a
Sponsoring Registrar IANA ID	Registrar ID registry (The Registry is available at: <a href="http://www.iana.org/assigments/registrar-ids/registrar-ids.xml">http://www.iana.org/assigments/registrar-ids/registrar-ids.xml</a> .)	1	255	1	n/a

<sup>2</sup> “**Domain Object**” corresponds to a single Registered Name. Each domain object includes the following data: Domain ID, Domain Name, Sponsoring Registrar, Domain Statuses, all contact information (including all details) with at least one each of: Registrant, Administrative, Technical that are instances of the contact object below; All nameservers associated with this domain; Domain Registration Date; Domain Expiration Date; Domain Last Updated Date, and other relevant information regarding the domains (e.g. DNSSEC).

Domain Status	RFC 5731		32	{1,11}	n/a
Registrant ID	Freeform text	1	255	1	n/a
Admin ID	Freeform text	1	255	1	n/a
Tech ID	Freeform text	1	255	1	n/a
Billing ID	Freeform text	1	255	1	n/a
DS created	RFC 3339		32	{0,1}	n/a
DS Key Tag	RFC 4034, 5910			{0,2}	n/a
Algorithm	RFC 4034, 5910			{0,2}	n/a

Table 2: RDDS Model for Nameserver Object<sup>3</sup>

Data Element	Format	Min length	Max length	Cardinality	Language Tag (RFC 5646)
Nameserver ID	Freeform text	1	255	1	n/a
Host Name	RFC 5890 (both A-label and U-label)	1	255	1	n/a
IP Address	RFC 0791/RFC 5952			{0, ..}	n/a
Registrar ID	Freeform text	1	255	1	n/a
Referral URL	RFC 3986 / 3987	1		1	n/a
Creation Date	RFC 3339		32	1	n/a
Last Updated Date	RFC 3339		32	{0,1}	n/a
WHOIS Server	RFC 5890 (both A-label and U-label)	1	255	1	n/a

<sup>3</sup> “**Nameserver Object**” corresponds to a single registered nameserver. The nameserver object includes the following data: Name Server ID, Name Server Host Name, Name Server IP Addresses if applicable, Current Registrar, Name Server Creation Date, Name Server Last Updated Date.



Table 3: RDDS Model for Contact Object<sup>4</sup>

Data Element	Format	Min length	Max length	Cardinality	Language Tag (RFC 5646)
Contact ID	Freeform text	1	255	1	n/a
Registrar ID	Freeform text	1	255	1	n/a
Contact Name	Freeform text	1	255	{0,1}	required
Contact Organization	Freeform text	1	255	{0,1}	required
Contact street	Freeform text in a language or script appropriate for its region.	1	255	{1,3}	required
Contact City	Freeform text in a language or script appropriate for its region	1	255	1	required
Contact State / Province	Freeform text in a language or script appropriate for its region.	1	255	{0,1}	required
Contact country / Territory	ISO 3166 part 2 code list	2	3	1	n/a
Contact Postal Code	Freeform text	1	255	{0,1}	n/a
Contact Phone	RFC 5733		64	1	n/a
Contact Phone Ext	RFC 5733		64	{0,1}	n/a
Contact Fax	RFC 5733		64	{0,1}	n/a

<sup>4</sup> “**Contact object**” corresponds to a single contact (registrant, administrative, technical and billing are roles of a contact with respect to given domain name). The contact object includes the following data: Contact ID, Contact Name, Contact Organization, Contact Address, City, State/Province, Country, Contact Postal Code, Contact Phone, Fax, E- mail.

Contact Fax Ext	RFC 5733		64	{0,1}	n/a
Contact Email	RFC 5322 / 6532		255	1	n/a

Table 4: RDDS Model for Registrar Object<sup>5</sup>

Data Element	Format	Min length	Max length	Cardinality	Language Tag (RFC 5646)
Registrar ID	Freeform text	1	255	1	n/a
Contact Organization	Freeform text.  Name of the registrar should be the official name in the RAA with ICANN, in whichever language(s) or script(s).	1	255	{0,1}	required
Contact Street	Freeform text in a language or script appropriate for its region.	1	255	{1, 3}	required
Contact City	Freeform text in a language or script appropriate for its region.	1	255	1	required
Contact State/Province	Freeform text in a language or script appropriate for its region.	1	255	1	required
Contact country/territory	ISO 3166-1 alpha-2	2	2	1	n/a

<sup>5</sup> “**Registrar Object**” corresponds to a single registrar. It includes the following data: Registrar ID (conforming to the IANA registrar-ids registry), Contact ID of Registrar, Registrar Administrative Contact ID, Registrar Technical Contact ID, Registrar Billing Contact ID, Registrar URL, Registrar Creation Date, and Registrar Last Updated Date.

Contact Postal Code	Freeform text	1	255	1	n/a
Contact Phone	RFC 5733		64	1	n/a
Contact Fax	RFC 5733		64	1	n/a
Contact Email	RFC 5322 / 6532		255	1	n/a
Registrar Admin Contact ID	Freeform text		255	1	n/a
Registrar Technical Contact ID	Freeform text		255	1	n/a
Registrar URL	RFC 3986 / 3987	1		{0, 1}	n/a