

WHOIS review team internationalized registration data expert working group update

#ICANN49



Working Group Status

- Formed as part of the effort to implement WHOIS review team recommendations 12-13
- 9 members with expertise in linguistics, Unicode, registrar and registry operations, policy, internationalization and localization in applications
- Weekly meeting since mid September 2013
- In final review of interim report, expect to publish shortly after ICANN Singapore meeting

#ICANN49



Working Group Charter

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

Working Group Approach

- Recognize on-going efforts in other areas (e.g. GNSO PDP Translation and Transliteration, Directory Services EWG, IETF WEIRDS working group)
- Group data by categories
- Separate internationalization vs. localization
- Articulate a set of principles to guide discussion of requirements

#ICANN49



Localization vs. Internationalization

- **Localization:** the adaptation of registration data to meet the language, cultural, regional and other requirements of a target data consumer group:
 - Numeric, date and time formats complies with local usage patterns
 - Localized label for data elements
 - Localized data (names, addresses)

Localization Example

Domain Information: [ドメイン情報]

[ドメイン名]	ドメイン名例.JP
[Domain Name]	XN--ECKWD4C7CU47R2WF.JP
[登録者名]	エグザンプル株式会社
[Registrant]	Example Corporation
[ネームサーバ]	ns01.example.co.jp
[登録年月日]	2001/08/09
[有効期限]	2008/08/31
[状態]	Active
[最終更新]	2007/09/01 01:05:05 (JST)

Contact Information: [公開連絡窓口]

[名前]	日本 太郎
[Name]	Taro, Nihon
[電子メールアドレス]	taro@example.jp
[郵便番号]	101-0065
[住所]	東京都千代田区西神田三丁目 8 番 1 号 千代田ファーストビル東館 13F
[Postal Address]	Chiyoda First Bldg. East 13F, 3-8-1 Nishi-Kanda Chiyoda-ku, Tokyo 101-0065, JAPAN
[電話番号]	03-5215-8451

#ICANN49



Localization vs. Internationalization

- **Internationalization** the adaptation of registration data to *enable* easy localization for target audiences that vary in language, culture or region.
 - Ensure data elements represented and transmitted in standard forms
 - Ensure that data are appropriately encapsulated and tagged to allow localization
- The WG focuses on internationalization, rather than localization

Principles

- **User Capability Principle:**

In defining a requirement for a particular data element or a set of data elements, the capability of the average user should be the constraining factor. Users should not be burdened with tasks they can not reasonably be expected to complete.

- **Simplicity and Reusability Principle:**

Where possible, existing standards that are widely used in registry and registrar systems for handling internationalized data should be applied. Where simpler standards exist for internationalization, they should be preferred rather than the more complex rules.

Proposed Requirements

- **High level requirements**
- **Requirements for name and address information**
 - Registrant, technical and administrative contact name, organization, address, country code, country, postal code
 - Registrar Name
- **Requirements for other data categories**
 - Status, phone/fax numbers, email, identifiers, domain names, date, DNSSEC information

High level requirements

- **Registrants should only be required to input registration data in the language(s) or script(s) that they are skilled at.**
 - A registry and registrar must be able to accept and store any language or script that might reasonably be expected to be used in the target market.
 - A registry is motivated to incentivize registrars to provide supporting content and input methods in the languages and scripts that are essential to the registry's target market.

High level requirements – cont.

- **Unless explicitly stated otherwise, all data elements are tagged with their language(s) and script(s) and this information is always available with the data element.**
 - For possible translation / transliteration purposes
 - For canonicalization in searching purposes

Requirement for Names and Addresses

Data Category	Proposed Requirement
Name, organization of registrant, technical and administrative contact	Free form text
Registrar Name	Free form text. The name of the sponsoring registry should be the official name in the RAA with ICANN, in whichever language(s) and script(s).
Address of registrant, technical and administrative contact	<p>[proposal 1] Free form text.</p> <p>[proposal 2] Free form text. The language(s) or script(s) of an address should be appropriate for the region that it is located.</p> <p>[proposal 3] Free form text. The script(s) of an address should be the same as the script of the TLD or in US-ASCII.</p>

Requirement for Names and Addresses

Data Category	Proposed Requirement
Postal Code of registrant, technical and administrative contact	Free form text
Country of registrant, technical and administrative contact	UNGEGN Part III, Names of Countries of the World;
Country Code	ISO 3166 part 2 code list

Requirement for other data elements

Data Category	Proposed Requirement
status	The text value of the domain status should conform to EPP specification defined in RFC 5731 section 2.3
Phone / fax numbers	The text value of phone/fax numbers should comply with ITU E. 164 standard.
Email	RFC 5322 and RFC 6532 (section 3.2) for internationalized email address
Identifiers	For sponsoring registrar's IANA ID, it should be the numeric value assigned by the IANA registry. For other identifiers, no standards should be applied.

Requirement for other data elements

Data Category	Proposed Requirement
URLs	Conform to standards set forth in RFC 3986 for URIs and RFC 3987 for IRIs.
Domain Names	For submission, only require U-label. For display, require directory services to display both U-label and the corresponding A-label.
Dates	Date and time elements should conform to formats specified in RFC 3339, and represented in UTC with no offset from the zero meridian.
DNSSEC information	DNSSEC elements should conform to values described in RFC 4034.

Next steps

- Draft interim report published for public comment after ICANN Singapore meeting
- Work on data model
- Final report for public comment by ICANN London meeting



Questions?

#ICANN49

