

Draft Final Report from the Expert Working Group on Internationalized Registration Data

STATUS OF THIS DOCUMENT

This is the Draft Final Report from the WHOIS Review Team Internationalized Registration Data Expert Working Group (IRD Working Group) recommending requirements for internationalized registration data.

EXECUTIVE SUMMARY	4
1. BACKGROUND	7
2. SCOPE AND METHODOLOGY	9
2.1 Localization vs. Internationalization	9
2.2 Categories of Data Elements	13
2.3 Technical Considerations	13
2.3.4. Internationalized email addresses	17
3. PRINCIPLES FOR INTERNATIONALIZATION OF REGISTRATION DATA	18
4. PROPOSED HIGH LEVEL REQUIREMENTS.....	18
5. PROPOSED REQUIREMENTS FOR INDIVIDUAL DATA CATEGORIES	19
5.1 Requirements for organization names, personal names, and postal addresses	19
5.1.1 Name, organization of registrant, technical and administrative contact.....	19
5.1.2 Postal address of registrant, technical and administrative contact	20
5.1.3 Registrar Name.....	22
5.1.4 Country / Territory	23
5.2 Requirements for other data elements.....	23
5.2.1 Status	23
5.2.2 Phone/ Fax Number	23
5.2.3 Email Address	23
5.2.4 Identifier	24
5.2.5 URL	24
5.2.6 Domain Name.....	24
5.2.7 Time and Date.....	25
5.2.8 DNSSEC	25
6. Proposed Data Model	25
6.1 What is included in a data model.....	25
6.2 Proposed Data Model.....	26
7. Future Work	30
8. Related Work	31

9. References.....	33
Appendix A: Compilation of current and proposed data elements	34
Appendix B: Examples of the Data Model.....	42
B2. Example XML Schemas of the Data Model	45
B2.1 IRD Domain Schema	45
Appendix C: Working Group Composition and Biographies of members	50
Appendix D: Working Group Response to Public Comments on the Interim Report.....	54

EXECUTIVE SUMMARY

As part of the broader effort to implement the recommendations from the ICANN WHOIS Review Team, the WHOIS Review Team Internationalized Registration Data Expert Working Group (hereinafter referred to as “The IRD Working Group”) was formed to recommend submission and display requirements for internationalized registration data (IRD) and produce a data model for the IRD that matches the requirements.

To aid its deliberation, the IRD Working Group developed three principles of internationalization. These are:

- **User Capability Principle:** In defining a requirement for a particular data element or category of data elements, the capability of the data-submitting user should be the constraining factor. Such users should not be burdened with tasks that cannot be completed under ordinary circumstances.
- **Simplicity and Reusability Principle:** Where possible, existing standards that are widely used for handling internationalized data should be applied. Where simpler standards exist for internationalization, they should be preferred rather than more complex standards.
- **Extensibility** - Where possible, the data model should be able to be easily extended to tailor to the evolution of data elements displayed by directory services for various TLD registries and registrars.

Based on these principles, the IRD Working Group proposes two high level requirements: First, **registrants should only be required** to input registration data in a language(s) or script(s) that they are skilled at; second 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.

The working group then categorized the common registration data elements into twelve groups and proposed the following internationalization requirements for each category, which are supported by a consensus of its members. These are summarized in two tables below: Table 1 - Requirements for contact data elements, Table 2 - Requirements for other data elements.

Table 1: Requirements for Contact Data

Data Category	Example Data Elements	Proposed Requirement
Personal names and organizational names	Registrant Name, Registrant Organization, technical and administrative contact name	Free-form text
Registrar Name	Sponsoring Registrar	Free-form text. The name of the sponsoring registrar should be the official name in the Registrar Accreditation Agreement (RAA) with ICANN, in a language and script appropriate for the region in which the registrar is located.
Postal Address	Registrant Address1, Registrant Address2, Registrant City, Registrant State/Province, Registrant Postal Code	Free form text, in languages and scripts appropriate for the region in which the address is located.
Country / Territory	Registrant country or territories code	ISO 3166 part 2 code list

Table 2: Requirements for other data elements

Data Category	Example Data Elements	Proposed Requirement
Status	Domain Status	The text value of the domain status should conform to the EPP specification defined in RFC 5731 section 2.3.
Phone and Fax Numbers	Technical Contact Facsimile Number, Technical Contact Phone Number	The phone and fax numbers should comply with the ITU E.164 standard.
Email addresses	Technical Contact Email, Registrant Email, Administrative Contact Email	Email addresses should comply with RFC 5322 and its extension in RFC 6532 (see section 3.2) for internationalized email addresses.
Identifiers	Registrar ID, Registrant ID, Sponsoring Registrar IANA ID, Domain ID	No internationalization standards apply; unique values are assigned by an appropriate entity.
DNSSEC Information	DS Key Tag 2, DS Key Tag 1, Digest Type 1, DS Maximum Signature Life 2, Algorithm 2, Digest Type 2, Algorithm 1, Digest 2, DS Maximum Signature Life 1, Digest 1	Elements should conform to format and values described in RFC 5910.
URLs	Referral URL, Registrar URL (registration services)	Elements should conform to standards set forth in RFC 3986 and RFC 3987.
Domain Names	Domain Name, Whois Server, Name Server	Where a domain name is provided by registrant, only require registrants to provide domain name in U-label [RFC 5890] during the submission. For display, require both U-label and the corresponding A-label [RFC 5890] at all times.

Time and Dates	Last Transferred Date, Domain Last Updated Date, DS creation date, Domain Expiration Date	Date and time elements should conform to formats specified in RFC3339, and represented in UTC with no offset from the zero meridian.
----------------	---	--

Finally, in Section [6. Proposed Data Model](#), the working group proposes a data model for community consideration, based on the requirements articulated above.

The proposed requirements and recommendations in this report are supported by a consensus of the working group.

1. BACKGROUND

Much of the currently accessible domain name registration data (DNRD) (previously referred to as WHOIS data) is encoded free form in US-ASCII script. 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 script support other than US-ASCII for correct submission or display.

The WHOIS Policy Review Team, in its final report, highlights the need to define requirements and develop data models 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

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. Subsequently a set of related efforts were formed to implement the WHOIS review team recommendations. These are:

- An expert working group to determine the requirements for internationalized registration data.
- A commissioned study to evaluate available solutions for internationalized registration data.
- A Policy Development Process (PDP) to determine whether translation or transliteration of contact information is needed. If so, specify who should bear the burden of the transformation.

This final report is the result of the first effort. An earlier version of the report was published for public comment in April 2014.¹

¹ See <https://www.icann.org/public-comments/ird-interim-2014-04-14-en>.

2. SCOPE AND METHODOLOGY

The IRD Working Group is chartered to:

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

In defining the requirements for internationalized data, the IRD Working Group first discussed the approach and methodology and drew two distinctions: between localization vs. internationalization, and between data elements vs. categories of data elements. Each of these is discussed below.

With respect to the question of localization vs. internationalization, the Working Group decided that if there was a choice to be made in its recommendation the choice would be decided in favor of internationalization. One effect of this decision is that some of the recommendations present technical challenges, in part because of legacy deployments and in part because general solutions do not currently exist. The last subsection below discusses the technical challenges identified by the working group.

2.1 Localization vs. Internationalization

Localization refers to the adaptation of a product, application, or document content to meet the language, cultural, and other requirements of a specific target market (a locale) or purpose. *Internationalization* is the design and development of a product, application, or document content that enables easy localization for target audiences that vary in culture, region, or language.

The IRD Working Group applied the above definitions to registration data directory service (formerly known as WHOIS), and observes the following distinction:

Table 3: An Example of localized WHOIS output for Japanese audience

ドメイン情報:	
[ドメイン名]	ドメイン名例.JP
[ドメイン名]	XN--ECKWD4C7CU47R2WF.JP
[登録者名]	エグザンプル株式会社
[ネームサーバ]	ns01.example.co.jp
[ネームサーバ]	ns02.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)
公開連絡窓口:	
[名前]	日本 太郎
[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
[電話番号]	03-5215-8451
[FAX 番号]	03-5215-8452

Localization of Registration Data in a Directory Service refers to the adaption of the directory service to meet the language requirements of a specific target market (a locale) or purpose. For example, Table 3: An Example of localized WHOIS output could be the localized directory service (WHOIS) output of an IDN “ドメイン名例.JP” for a Japanese audience.

From the above example, we note that for domain name registration data directory services, localization can entail customization related to:

- Both the description of the data element, and the data element itself
- Numeric, date and time formats that comply with local usage patterns (e.g., 2007/09/01 01:05:05 (JST))
- Localized label of the data elements (e.g., ドメイン名, 電子メールアドレス)
- Localized data (e.g., 東京都千代田区西神田三丁目 8 番 1 号), if available

The same data could be localized for an English speaking audience, as shown below. In this example, Japanese labels like "登録年月日", "住所", and "電話番号" are respectively localized to "Creation Date", "Postal Address", and "Phone", and the English translation / transliteration of the postal address is shown.

Table 4: An example of WHOIS output localized for English-speaking audience

Domain Information:	
[Domain Name]	XN--ECKWD4C7CU47R2WF.JP
[Registrant]	Example Corporation
[Name Server]	ns01.example.co.jp
[Name Server]	ns02.example.co.jp
[Creation Date]	2001/08/09
[Update Date]	2008/08/31
[Status]	Active
[Last Updated]	2007/09/01 01:05:05 (JST)
Contact Information:	
[Name]	Taro Nihon
[Email]	taro@example.jp
[Web Page]	
[Postal code]	101-0065
[Postal Address]	Chiyoda First Bldg. East 13F, 3-8-1 Nishi-Kanda Chiyoda-ku, Tokyo 101-0065, JAPAN
[Phone]	03-5215-8451
[Fax]	03-5215-8452

It is essential to understand that the original data could have been in either form: Japanese or English. Since each form can be derived from the other through translation and transliteration, each form represents an internationalized form. The critical point is that data is sufficiently internationalized if it can be adapted to any other local form needed by any of the defined purposes of registration data.

The effect of this observation is that no single internationalized form must exist. Every language and script combination could be an internationalized form. To be an internationalized form it must be specified in a sufficiently self-consistent way to ensure there is enough information available for translation and transliteration tools to be successful. (See Section 4.3 for more details).

Internationalization in the context of the display of registration data entails the following:

- Designing and developing in a way that removes barriers to localization. This includes but is not limited to such things as enabling the use of Unicode while ensuring the proper handling of legacy character encodings where appropriate, taking care over the concatenation of strings, and avoiding dependence in code of user-interface string values.
- Providing support for features that may not be used until localization occurs. For example, adding markup in your Document Type Definition (DTD) to support bidirectional text, or for identifying language and script.
- Enabling code to support local, regional, language, or culturally related preferences. Typically this involves incorporating predefined localization data and features derived from existing libraries or user preferences. Examples include date and time formats, local calendars, number formats and numeral systems, sorting and presentation of lists, and handling of personal names and forms of addresses.
- Separating localizable elements from source code or content, such that localized alternatives can be loaded or selected based on the user's international preferences as needed. In the case of directory service output, instead of hardcoding the localized labels into the source code, a directory service would be better served if these labels can be stored in separate files that can be loaded based on the user's international language or script preference.

2.2 Categories of Data Elements

There are over 150 data elements currently displayed by various gTLD registries' directory services, (see Appendix A) as well as other proposed data elements (e.g., from the ICANN Expert Working Group on Registration Data²). It is also likely that these data elements could change over time. Thus the IRD working group has chosen to group data elements into categories and focus the requirements within these categories. The IRD Working Group developed 12 data categories that cover all of the known data elements.

- Personal name and organization name
- Registrar name
- Postal Addresses
- Country / Territory
- Status
- Phone and Fax Numbers
- Email Addresses
- Identifiers
- DNSSEC Information
- URLs
- Domain Names
- Time and Dates

2.3 Technical Considerations

The requirements defined in this document pose several technical challenges for the current system of collecting, transmitting, storing, and displaying registration data. In this section, we outline these challenges.

2.3.1 Lack of Internationalized Support in Technical Protocols

² See <https://www.icann.org/en/system/files/files/final-report-06jun14-en.pdf>

EPP (Extensible Provisioning Protocol) Issues

Lacking of a language and script attribute

One of the central requirements in this document is that registrants can submit the registration data in a language and script with which they are familiar. To enable consumers of the data to translate or transliterate the data, language and script tags are required along with the registration data.

Such requirements are only partially supported by the current EPP. RFC 5733, which describes an EPP contact mapping, allows for the submission of localized contact information in UTF-8 format, with the type attribute set to "loc". The "type" attribute is used to identify the two forms. If an internationalized form (type="int") is provided, element content **MUST** be represented in a subset of UTF-8³ that can be represented in the 7-bit US-ASCII character set. If a localized form (type="loc") is provided, the element content **MAY** be represented in unrestricted UTF-8.

However, there are three issues with this feature.

First, if both a "loc" and an "int" form are present, there is no relationship indicated. The presumption is that both forms represent the same information but this is neither required nor is there an indication as to which form is the preferred form.

Second, the "type" attribute does not support the specification of a language and script tag for the contact information.

Third, the "type" attribute applies to a complete set of postal contract information elements. One of the early observations from discussions in this working group

³ See RFC 3629: UTF-8, a transformation format of ISO 10646. F. Yergeau. November 2003.

is that contact information elements can be presented with multiple languages and scripts mixed together. For example, an entity's name could be in one language and script with the postal address elements in a different language and script. Further, it is common for many languages and scripts to mix the use of the US-ASCII numerals 0-9, e.g., as part of a physical address.

Lacking of a conversion-mechanism attribute

The working group (WG) also discussed that if the data inputted by user in the local language is translated or transliterated into a “must be present” representation, then there is a possibility of two forms (one in local language and another in a "must be present" language/script) of the same record.

As the current XML schema for DNRD does not support the representation of two forms, some additional information must be added in the current XML schema to satisfy this requirement.

- To specify that the two forms of data represent the same record there should be some indication of the relationship mapping between them, e.g., which is the preferred form and which form is derived from the other.
- To specify the transformation information, e.g., whether the specific form is translated or transliterated, and the date and time of the transformation. As an example, to represent this transformation information an attribute "conversionMechanism" is added in the example of contact object in XML (Appendix B). This attribute can take values based on the transformation applied, e.g., translation or transliteration.

WHOIS Issue

On the display side, the current WHOIS protocol is not capable of handling “UTF-8” characters consistently, as it has “no mechanism for indicating the character set in use.”⁴ Thus, in order to fully support internationalized data, new protocols

⁴ RFC3912. Whois Protocol Specification. September, 2004

are needed. Currently, the IETF Web-extensible Internet Registration Data (weirds) working group is developing a lightweight registration data access protocol that supports internationalized registration data.

2.3.2 Encoding of data requires "standard" languages and scripts

The use of language and script tags for all registration data requires that a standard list and definition of languages and scripts exist.

Additionally, for postal addresses to be in the language or script appropriate for the region requires that a standard list of languages and scripts that are acceptable to a region exist. The Universal Postal Union (UPU) developed the S42 addressing standard⁵, which among other things offers the standard languages and scripts used in a given country or territory. However, less than 40 countries or territories are compliant with S42. Without a comprehensive standardized list, it will be difficult for a registrar and registry to validate the correct script and language as used in the region.

2.3.3 Workflow changes are required at registrars

Taking the language and script information from the registrant requires registrars to be able to detect, validate and verify the script in use. This functionality does not exist in today's registrar customer interface. This would necessitate changes in the registrar workflow to accommodate this requirement.

Postal regulations specify that what constitutes a valid postal address is defined by the regional authority for the region in which the address exists. Functionally, the requirements are that a country name be present (preferably in the language of the dispatching country or in an internationally recognized language) and that any other information that is present is only required to be sufficient to ensure

⁵ <http://www.upu.int/en/activities/addressing/s42-standard/compliant-countries.html>

that the final destination can be identified by the postal representative physically delivering the message.

This suggests that the western convention of requiring an address to be in the form of a name, address line 1, address lines 2, city, state or province, postal code, and country may be inappropriate in an internationalized context.

The combination of all of these observations suggests that a broad requirement to syntactically validate all postal address contact information may be impractical. Operational validation may be possible if conceptually the mechanism is simply to copy-and-paste whatever is provided by the inputting user to a physical envelope and test if the envelope is delivered.

2.3.4. Internationalized email addresses

Internationalized email addresses as specified in RFC 6532 have not been widely adopted as of this writing.

The IRD Working Group observes the international standards are backward compatible, so there is no syntactic issue. Operationally however, there are significant issues to be considered.⁶ Specifically, if a registrar accepts an internationalized email address, its internal email systems, and/or that of registrants, may need to be updated to send or receive messages from an internationalized email address.

Third parties that consume this data may have similar issues. Specifically, email is by definition a store-and-forward service.⁷ Email messages are routinely transferred through one or more mediators in a message handling service on their way to their destination. Each of these actors along the delivery path may have components that need updating.

⁶ See RFC 6530 for the description of operational challenges.

⁷ RFC 5598. Internet Mail Architecture D. Crocker [July 2009]

3. PRINCIPLES FOR INTERNATIONALIZATION OF REGISTRATION DATA

The IRD Working Group agreed upon the following principles to guide its deliberations:

- **User Capability Principle:** In defining a requirement for a particular data element or category of data elements, the capability of the data-submitting user should be the constraining factor. Such users should not be burdened with tasks that cannot be completed under ordinary circumstances.
- **Simplicity and Reusability Principle:** Where possible, existing standards that are widely used for handling internationalized data should be applied. Where simpler standards exist for internationalization, they should be preferred rather than more complex standards.
- **Extensibility** - Where possible, the data model should be able to be easily extended to tailor to the evolution of data elements displayed by directory services for various TLD registries and registrars.

4. PROPOSED HIGH LEVEL REQUIREMENTS

The IRD Working Group proposes that these requirements apply to all categories of data elements at all times, unless explicitly documented as not being applicable.

4.1 Registrants should only be required to input registration data in a language(s) or script(s) with which they are skilled.

Per the user capability principle, a cornerstone assumption of an internationalized system is that a registrant must be able to use the language(s) and script(s) in which they are most skilled. Thus a registrant must not be required to use any specific language(s) or script(s) beyond what would be used in their ordinary daily routine.

4.2 A registry must be able to accept and store any language or script that might reasonably be expected to be used in their target market.

Note that this is distinct from the languages and scripts they support for domain names. As a practical matter this probably means they must accept every language or script for contact information, i.e., accept whatever the registrar provides.

4.3 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.

There are at least two reasons why it is essential that all data elements be tagged. First, it is not possible to properly translate or transliterate the data unless we are certain of what it currently represents. Second, it is not possible to properly do searching on the data unless the data and the query are consistent, which cannot be done unless we are certain of what they represent.⁸

5. PROPOSED REQUIREMENTS FOR INDIVIDUAL DATA CATEGORIES

5.1 Requirements for organization names, personal names, and postal addresses

5.1.1 Name, organization of registrant, technical and administrative contact

⁸ See “Study to Evaluate Available Solutions for the Submission and Display of Internationalized Contact Data”, Available at: <https://www.icann.org/en/system/files/files/transform-dnrd-02jun14-en.pdf>

- **Data Elements Covered:** Registrant Name, Registrant Organization, Administrative Contact Name, Administrative Contact Organization, Technical Contact Organization, Technical Contact Name
- **Proposed Requirement:** This should be free-form text.
- **Rationale:** User capability principle
- **Discussion:** It is unrealistic to think that the average Internet user is able to write in a language or script different from his/her native tongue, therefore it is essential to give registrants the ability to submit information in the language or script of his or her choice. This is consistent with the user capability principle explained above.

As registrars are in direct contact with the registrant, and inline with the reasons outlined above, they should allow localized data to be provided. However, the extent of registrars' support of internationalized registration data is also a business decision for the registrar. There may be languages or scripts a registrar may not support at their own discretion.

The WG notes that if translation or transliteration is employed, the "type" attribute described in RFC 5733 (see section 2.3.1) could be used to maintain both the original input data and the transformed data. However, as there is a GNSO PDP on this specific issue, further discussion is considered out of scope for this IRD Expert WG.

5.1.2 Postal address of registrant, technical and administrative contact

- **Data Elements Covered:** Registrant Address, Registrant City, Registrant State or Province, Administrative Contact Address, Administrative Contact City, State/Province, Technical Contact Address, Technical Contact city, State/Province, Registrant Postal Code, Administrative Contact Postal Code, Technical Contact Postal Code
- **Proposed Requirement:** The postal address should be free-form text in a language and script appropriate for the region in which it is located.

- **Rationale:**

The Working Group explored the following proposed requirements for the postal address elements, including asking the community for comments about each in its Initial Report:

- Proposal 1: Free form text. The language(s) and script(s) of an address should be appropriate for the region in which it is located.
- Proposal 2: Free form text. The script(s) of an address should be the same as the script of the TLD or in US-ASCII.
- Proposal 3: Free form text.

Based on the community input from the Initial Report and continuing deliberation within the working group, Proposal 1 is chosen.

Proposal 3 suggests that the only requirement is to accept whatever a registrant submits for postal information. However, the WG considered the purpose of the information in determining the internationalization requirements. The purpose of postal information is to be able to contact, via postal mail, the represented entity. A requirement to accept any data would be expected to make the validation of the data, a recent requirement added to registrars, problematic in the best case. Validation requirements are still evolving at this time but it is certain that a complete lack of structure would be a significant and substantial burden to complying with validation requirements.

Proposal 2 suggests that the script table used would be extremely restrictive. To require US-ASCII would be in conflict to the objective of internationalizing registration data. The script table used by the TLD is necessarily restrictive for security and stability reasons, which would also

be in conflict with the requirement to internationalize registration data if code points needed for names and locations were not available. In addition, some regional areas recognize and accept multiple languages and scripts, and limiting a TLD to a single script would be in conflict with the requirement to internationalize registration data.

Given that the purpose of postal information is to be able to contact the entity represented, postal regulations require that the contact information be sufficient to be useful to the local delivery agent. The user at that location is the best suited to know exactly what that information should be. The local delivery agent could reasonably be expected to understand their local language(s) and script(s) (user capability principle). At a minimum, a user that has relocated to a region without knowledge of the local language or script could reasonably be expected to copy-and-paste their postal address information to ensure it was operationally valid.

Section 2.3 describes a technical issue with this requirement. This WG recognizes that the technical issue would need to be resolved before this requirement could be fully implemented. The purpose of this requirement is to state where we should be. Developing a transition plan and solving the technical issue are future work necessary to achieving the goal of internationalizing registration data.

5.1.3 Registrar Name

- **Data Elements Covered:** Sponsoring Registrar
- **Proposed Requirement:** Free-form text. The name of the sponsoring registrar should be the official name in the Registrar Accreditation Agreement (RAA) with ICANN.
- **Rationale:** simplicity and reusability

5.1.4 Country / Territory

- **Data Elements Covered:** Registrant Country Code, Administrative Contact Country Code, Technical Contact Country Code
- **Proposed Requirement:** The country code should comply with ISO 3166 part 2 code which can be found at <https://www.iso.org/obp/ui/#search/code>.

5.2 Requirements for other data elements

5.2.1 Status

- **Data Elements Covered:** domain status
- **Proposed Requirement:** The text value of the domain status should conform to EPP specification defined in RFC 5731 section 2.3.

5.2.2 Phone/ Fax Number

- **Data Elements Covered:** Registrar Abuse Contact Phone, Technical Contact Facsimile Number, Technical Contact Phone Number, Technical Contact Phone Number Ext, Administrative Contact Phone Number Ext, Registrant Facsimile Number, Registrant Phone Number, Administrative Contact Facsimile Number Ext, Technical Contact Facsimile Number Ext, Administrative Contact Phone Number, Administrative Contact Facsimile Number, Registrant Facsimile Number Ext. and Registrant Phone Number Ext.
- **Proposed Requirement:** The phone and fax numbers should comply with ITU E.164.2005 as defined in RFC 5733.

5.2.3 Email Address

- **Data Elements Covered:** Registrar Abuse Contact Email, Technical Contact Email, Registrant Email, Administrative Contact Email

- **Proposed Requirement:** Email address format should comply with RFC 5322 and its extension in RFC 6532 (see section 3.2) for internationalized email addresses.

5.2.4 Identifier

- **Data Elements Covered:** Registry Domain ID,, Registry Technical Contact ID, Registry Registrant ID, Registry Administrative Contact ID, Registrar IANA ID
- **Proposed Requirement:** No internationalization of this data element is required. Unique values are assigned by an appropriate entity.

5.2.5 URL

- **Data Elements Covered:** Referral URL, Registrar URL (registration services)
- **Proposed Requirement:** The URL should conform to standards set forth in RFC 3986 and RFC 3987.

5.2.6 Domain Name⁹

- **Data Elements Covered:** Domain Name, Whois Server, Name Server
- **Proposed Requirement:** Where an internationalized domain name is provided by registrant, only require registrants to provide domain name in U-label during the submission. For display, always display both U-label and the corresponding A-label for all domain names at all times.
- **Discussion:** In current implementations, it is common to have both the U-label and A-label forms stored. The U-label should be present for human recognition and readability, and its punycode encoded A-label should be present for technical usage and backward compatibility reasons.

⁹ Including IDN Variant Domain Names where appropriate

5.2.7 Time and Date

- **Data Elements Covered:** Last Transferred Date, Domain Last Updated Date, Domain Expiration Date, Domain Registration Date, Last Updated by Registrar, DS creation date
- **Proposed Requirement:** Date and time elements should conform to formats specified in [RFC3339], and represented in UTC with no offset from the zero meridian. For example, 1990-12-31T23:59:60Z.¹⁰

5.2.8 DNSSEC

- **Data Elements Covered:** DS Key Tag, Algorithm, Digest Type, Digest, DS Maximum Signature Life
- **Proposed Requirement:** Elements should conform to formats / values described in RFC 5910.

6. Proposed Data Model

6.1 What is included in a data model

By default, a data model should include:

- The list of data elements (or categories) represented or transmitted for a given registration data directory service. For our purpose, we surveyed a typical gTLD registry DNRD-DS (DNRD Directory Service) profile. We chose gTLD as an example because all of them have a contract with ICANN and their DNRD-DS is well specified. This model does not apply to ccTLD.
- Encoding of the data element represented (by default this should be UTF-8, but it is possible to specify other encodings).
- Markup language or text data format where the model can be represented (e.g. XML, JSON or other markup languages)

¹⁰ Z is a suffix which, when applied to a time, denotes a UTC offset of 00:00; often spoken "Zulu" from the ICAO phonetic alphabet representation of the letter "Z".

6.2 Proposed Data Model

The model is separated into the following objects: domain, contact, registrar and nameserver.

- The **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).
- The **contact object** corresponds to a single contact (registrant, administrative, technical 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.
- The **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 URL, Registrar Creation Date, and Registrar Last Updated Date.
- A **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.

Proposed requirements and relevant technical standards for data categories specified in Section 1 is applied for each of the data elements enumerated above as explained below.

Table 5: DNRD-DS Model For the Domain Object

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 ¹¹		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 ¹²	1	255	1	n/a
Domain Status	RFC 5731		32	{1,4}	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
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
Digest Type	RFC 4034, 5910			{0,2}	n/a
Digest Value	RFC 4034, 5910			{0,2}	n/a
DS Maximum Signature Life	RFC 4034, 5910			{0,2}	n/a

¹¹ Date and time in UTC as specified in [RFC3339], with no offset from the zero meridian.

¹² The Registry is available at: <http://www.iana.org/assignments/registrar-ids/registrar-ids.xml>.

Table 6: DNRD-DS Model For the Nameserver Object

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
Sponsoring Registrar	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

Table 7: DNRD-DS Model For the Contact Object

Data Element	Format	Min length	Max length	Cardinality	Language Tag (RFC 5646)
Contact 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	1	required
Contact country / Territory	ISO 3166 part 2 code list	1	255	1	required

Contact Postal Code	Freeform text	1	255	1	n/a
Contact Phone	ITU E 164		64	1	n/a
Contact Phone Ext	ITU E 164		64	{0,1}	n/a
Contact Fax	ITU E 164		64	{0,1}	n/a
Contact Fax Ext	ITU E 164		64	{0,1}	n/a
Contact Email	RFC 5322 / 6532		255	1	n/a

Table 8: DNRD-DS Model For the Registrar Object

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
Contact Postal Code	Freeform text	1	255	1	n/a
Contact Phone	ITU E 164		64	1	n/a
Contact Fax	ITU E 164		64	1	n/a
Contact Email	RFC 5322 / 6532		255	1	n/a
Registrar Admin Contact ID	Freeform text	1	255	1	n/a
Registrar Technical Contact ID	Freeform text	1	255	1	n/a
Registrar WHOIS Server	RFC 5890 (both A-label and U-label)	1	255	1	n/a
Registrar URL	RFC 3986 / 3987	1		{0,1}	n/a

7. Future Work

As a registration data directory service evolves, it is natural that new categories of data elements will come into existence, and new data types will be needed. As these new data categories / elements are proposed, they need to inherit requirements of related data categories from which they are derived, and additional internationalization requirements of new data types needs to be carefully considered by ICANN as part of the policy or implementation process.

The issue of translation and transliteration is currently under consideration in a GNSO Policy Development Process. The WG note the following issues related to translation and transliteration and recommend the GNSO PDP to consider:

- If registrants are allowed to submit localized registration data, what languages or scripts are registrars or registry operators expected to support?
- If registrants are required to submit a single common script version of their localized registration data, are users expected to submit a translated version, a transliterated version, or “either” (provided there is a convention or method to distinguish between the two)?
- If registrants are required to submit a single common script version of their localized 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)?
- If registrants are required to submit a single common script version and the localized version of their registration data, should there be a requirement to detect whether both versions are equivalents of each other?

- If there are translated / transliterated versions of the registration data along with the original data, which version should be considered primary or authoritative if there is a mismatch?¹³
- If translated / transliterated versions of the data are required, how will data be maintained simultaneously in multiple languages/scripts? Should there be additional meta-level information?
- For company and individual names, should translation or transliteration be required?

8. Related Work

In this section, The IRD Working Group wishes to acknowledge previous community work that the working group builds upon, and current on-going work that the working group is in close coordination with.

The issue of supporting internationalized registration data was first called out in RFC 4690¹⁴ by the Internet Architecture Board in 2006 and by SAC 037,¹⁵ “Display and usage of internationalized registration data” in 2009.

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.

¹³ Please note that it is possible that there are two versions of the registration data. For example an organization could provide 2 separate addresses, each in its respective appropriate language, each handling inquiries appropriately in the language of the given address.

¹⁴ J. Klensin and P. Fältström, “RFC 4690: Review and Recommendations for Internationalized Domain Names (IDNs),” Network Working Group, Internet Engineering Task Force, Internet Society, September 2006, <<http://www.ietf.org/rfc/rfc4690.txt>>.

¹⁵ Security and Stability Advisory Committee, “SAC037, Display and usage of Internationalized Registration Data, Support for Characters from Local Languages or Scripts,” 21 April 2009, <<http://www.icann.org/en/committees/security/sac037.pdf>>.

The SSAC Report on the Domain Name Registration Data Model (SAC 054) focuses on the information that is associated with a domain name from the time a registration is created – and the domain name is thus “instantiated” in the domain name system (DNS) – until the registration expires. It proposes a generic data model that is structured, and extensible. In this document, the WG builds on the data model presented in SAC 054 to accommodate requirements for internationalized registration data.

The IETF Web-extensible Internet Registration Data (weirds) working group is developing a lightweight registration data access protocol that supports internationalized registration data. It is currently in its final stages of issuing the protocol standards. During the IETF protocol work, the 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.

The GNSO PDP on the Translation and Transliteration of Contact Names is focusing its attention 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. The working group has released an initial report for the community’s consideration [6].

This IRD Working Group differs from the GNSO PDP in that it will focus on which data elements need to be internationalized and the requirements of that internationalization. However, close coordination with the Translation and Transliteration of Contact Names work is essential to avoid conflicts and duplication of work.

9. 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, 2013. Available at < <http://tools.ietf.org/html/draft-ietf-weirds-object-inventory-06>>
6. Internet Corporation for Assigned Names and Numbers (ICANN). (2014) Initial Report on the Translation and Transliteration of Contact Information Policy Development Process. Marina Del Rey, CA: ICANN. Available at: < <http://gnso.icann.org/en/issues/gtlds/transliteration-contact-initial-15dec14-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>

Appendix A: Compilation of current and proposed data elements

Data Elements	Source
Domain ID	Common data elements
Domain Name	Common data elements
Domain Registration Date	Common data elements
Domain Expiration Date	Common data elements
Domain Last Updated Date	Common data elements
Last Transferred Date	Common data elements
Created by Registrar	Common data elements
Last Updated by Registrar	Common data elements
Sponsoring Registrar	Common data elements
Domain Status (each status one line)	Common data elements
Registrant ID	Common data elements
Registrant Name	Common data elements
Registrant Organization	Common data elements
Registrant Address1	Common data elements
Registrant Address2	Common data elements
Registrant Address3	Common data elements
Registrant City	Common data elements
Registrant State/Province	Common data elements
Registrant Postal Code	Common data elements
Registrant Country	Common data elements

Registrant Country Code	Common data elements
Registrant Phone Number	Common data elements
Registrant Phone Number Ext.	Common data elements
Registrant Facsimile Number	Common data elements
Registrant Facsimile Number Ext.	Common data elements
Registrant Email	Common data elements
Administrative Contact ID	Common data elements
Administrative Contact Name	Common data elements
Administrative Contact Organization	Common data elements
Administrative Contact Address1	Common data elements
Administrative Contact Address2	Common data elements
Administrative Contact Address3	Common data elements
Administrative Contact City	Common data elements
Administrative Contact State/Province	Common data elements
Administrative Contact Postal Code	Common data elements
Administrative Contact Country	Common data elements
Administrative Contact Country Code	Common data elements
Administrative Contact Phone Number	Common data elements
Administrative Contact Phone Number Ext	Common data elements
Administrative Contact Facsimile Number	Common data elements
Administrative Contact Facsimile Number Ext	Common data elements
Administrative Contact Email	Common data elements
Billing Contact ID	Common data elements
Billing Contact Name	Common data elements

Billing Contact Organization	Common data elements
Billing Contact Address1	Common data elements
Billing Contact Address2	Common data elements
Billing Contact Address3	Common data elements
Billing Contact City	Common data elements
Billing Contact State/Province	Common data elements
Billing Contact Postal Code	Common data elements
Billing Contact Country	Common data elements
Billing Contact Country Code	Common data elements
Billing Contact Phone Number	Common data elements
Billing Contact Phone Number Ext	Common data elements
Billing Contact Facsimile Number	Common data elements
Billing Contact Facsimile Number Ext	Common data elements
Billing Contact Email	Common data elements
Technical Contact ID	Common data elements
Technical Contact Name	Common data elements
Technical Contact Organization	Common data elements
Technical Contact Address1	Common data elements
Technical Contact Address2	Common data elements
Technical Contact Address3	Common data elements
Technical Contact City	Common data elements
Technical Contact State/Province	Common data elements
Technical Contact Postal Code	Common data elements
Technical Contact Country	Common data elements
Technical Contact Country Code	Common data elements

Technical Contact Phone Number	Common data elements
Technical Contact Phone Number Ext	Common data elements
Technical Contact Facsimile Number	Common data elements
Technical Contact Facsimile Number Ext	Common data elements
Technical Contact Email	Common data elements
ENS_AuthId	AERO
Maintainer	AERO/ASIA/CAT/POST
IPR Name	ASIA
IPR Number	ASIA
IPR CC Locality	ASIA
IPR Applied	ASIA
IPR Registered	ASIA
IPR Class	ASIA
IPR form	ASIA
IPR Entitlement	ASIA
IPR Type	ASIA
CED ID	ASIA
CED CC Locality	ASIA
CED State/Province	ASIA
CED City	ASIA
CED Type of Legal Entity	ASIA
CED Type (Other)	ASIA
CED Form of Identification	ASIA
CED Form of ID (Other)	ASIA
CED Identification Number	ASIA

Operations and Notifications ID	ASIA
Operations and Notifications Name	ASIA
Operations and Notifications Organization	ASIA
Operations and Notifications Address	ASIA
Operations and Notifications Address2	ASIA
Operations and Notifications Address3	ASIA
Operations and Notifications City	ASIA
Operations and Notifications State/Province	ASIA
Operations and Notifications Country/Economy	ASIA
Operations and Notifications Postal Code	ASIA
Operations and Notifications Phone	ASIA
Operations and Notifications Phone Ext.	ASIA
Operations and Notifications FAX	ASIA
Operations and Notifications FAX Ext.	ASIA
Operations and Notifications E-mail	ASIA
Registration Agent ID	ASIA
Registration Agent Name	ASIA
Registration Agent Organization	ASIA
Registration Agent Address	ASIA
Registration Agent Address2	ASIA
Registration Agent Address3	ASIA
Registration Agent City	ASIA
Registration Agent State/Province	ASIA
Registration Agent Country/Economy	ASIA
Registration Agent Postal Code	ASIA
Registration Agent Phone	ASIA

Registration Agent Phone Ext.	ASIA
Registration Agent FAX	ASIA
Registration Agent FAX Ext.	ASIA
Registration Agent E-mail	ASIA
Domain Name ACE	CAT
Domain Language	CAT
Name Server ACE	CAT
Registrar ID	CAT
Whois Server	COM/NET/JOBS/ICANN
Referral URL	COM/NET/JOBS/ICANN
Created by ID	COOP
Last updated by ID	COOP
Sponsoring registrar ID	COOP
Contact Type	COOP
Host ID (each one one line)	COOP
Trademark Name	MOBI, INFO
Trademark Date	MOBI
Trademark Country	MOBI, INFO
Trademark Number	MOBI, INFO
Date Trademark Applied For	INFO
Date Trademark Registered	INFO

DNSSEC	ORG, ICANN
DS Created 1	ORG
DS Key Tag 1	ORG
Algorithm 1	ORG
Digest Type 1	ORG
Digist 1	ORG
DS Maximum Signature Life 1	ORG
DS Created 2	ORG
DS Key Tag 2	ORG
Algorithm 2	ORG
Digest Type 2	ORG
Digist 2	ORG
DS Maximum Signature Life 2	ORG
Registration Type	PRO
Registrar URL (registration services)	TEL/TRAVEL/BIZ
Sponsoring Registrar IANA ID	TEL/ICANN
Other names registered by registrant	NAME
Registrar Jurisdiction	EWG
Registry Jurisdiction	EWG
Reg Agreement Language	EWG
Original Registration Date	EWG
Registrant Company Identifier Registrant SMS/IM/Other	EWG

Contact SMS/IM/Etc	EWG
--------------------	-----

Appendix B: Examples of the Data Model

The following sections are a partial representation using XML of some of the requirements proposed in this document. They are not complete because of the technical considerations described in Section 2.3.

B.1 Example of Domain Name Object in XML

```
<?xml version="1.0" encoding="UTF-8"?>
<domain xmlns="urn:ietf:params:xml:ns:irdDomain-1.0"
  xmlns:domain="urn:ietf:params:xml:ns:domain-1.0"
  xmlns:idn="urn:ietf:params:xml:ns:idn-1.0"
  xmlns:secDNS="urn:ietf:params:xml:ns:secDNS-1.1"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="urn:ietf:params:xml:ns:irdDomain-1.0
    irdDomain-1.0.xsd">
  <name>xn--eckwd4c7cu47r2wf.test</name>
  <roid>DOMAINxn__eckwd4c7cu47r2wf_test-TEST</roid>
  <status s="ok"/>
  <registrant>jdl1234</registrant>
  <contact type="admin">sh8013</contact>
  <contact type="tech">sh8013</contact>
  <ns>
    <domain:hostObj>ns1.xn--eckwd4c7cu47r2wf.test</domain:hostObj>
    <domain:hostObj>ns2.xn--eckwd4c7cu47r2wf.test</domain:hostObj>
  </ns>
  <clID>clientX</clID>
  <crID>clientY</crID>
  <crDate>1999-04-03T22:00:00.0Z</crDate>
  <upID>clientX</upID>
  <upDate>2009-12-03T09:05:00.0Z</upDate>
  <exDate>2015-04-03T22:00:00.0Z</exDate>
  <idn >
    <idn:table>ja-JP</idn:table>
    <idn:uname>ドメイン名例.test</idn:uname>
  </idn>
  <secDNS>
    <secDNS:maxSigLife>604800</secDNS:maxSigLife>
    <secDNS:dsData>
      <secDNS:keyTag>12345</secDNS:keyTag>
      <secDNS:alg>7</secDNS:alg>
      <secDNS:digestType>1</secDNS:digestType>
      <secDNS:digest>93358db22e956a451eb5ae8d2ec39526ca6a87b9</secDNS:digest>
    </secDNS:dsData>
  </secDNS>
</domain>
```

B.2 Example of Nameserver Object in XML

```
<?xml version="1.0" encoding="UTF-8"?>
<host xmlns="urn:ietf:params:xml:ns:irdHost-1.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="urn:ietf:params:xml:ns:irdHost-1.0
    irdHost-1.0.xsd">
  <name>ns1.xn--eckwd4c7cu47r2wf.test</name>
  <roid>HOSTns1_xn__eckwd4c7cu47r2wf_test-TEST</roid>
  <status s="linked"/>
  <status s="clientUpdateProhibited"/>
```

```

<addr ip="v4">192.0.2.2</addr>
<addr ip="v4">192.0.2.29</addr>
<addr ip="v6">2001:db8::a</addr>
<clID>clientY</clID>
<crID>clientX</crID>
<crDate>1999-05-08T12:10:00.0Z</crDate>
<upID>clientX</upID>
<upDate>2009-10-03T09:34:00.0Z</upDate>
<trDate>2007-01-08T09:19:00.0Z</trDate>
</host>

```

B.3 Example of Contact Object in XML

```

<?xml version="1.0" encoding="UTF-8"?>
<contact xmlns="urn:ietf:params:xml:ns:irdContact-1.0"
  xmlns:contact="urn:ietf:params:xml:ns:contact-1.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="urn:ietf:params:xml:ns:irdContact-1.0
    irdContact-1.0.xsd">
  <id>sh8013</id>
  <roid>CONTACTsh8013-TEST</roid>
  <status s="linked"/>
  <status s="clientDeleteProhibited"/>
  <postalInfo type="loc">
    <contact:name>日本 太郎</contact:name>
    <contact:org>エグザンプル株式会社</contact:org>
    <contact:addr>
      <contact:street>千代田ファーストビル東館 13F</contact:street>
      <contact:street>西神田三丁目 8 番地 1 号</contact:street>
      <contact:city>千代田区</contact:city>
      <contact:sp>東京都</contact:sp>
      <contact:pc>101-0065</contact:pc>
      <contact:cc>JP</contact:cc>
    </contact:addr>
  </postalInfo>
  <postalInfo type="int">
    <contact:name>Taro Nihon</contact:name>
    <contact:org>Example Corporation</contact:org>
    <contact:addr>
      <contact:street>Chiyoda First Bldg. East 13F</contact:street>
      <contact:street>3-8-1 Nishi-Kanda</contact:street>
      <contact:city>Chiyoda-ku</contact:city>
      <contact:sp>Tokyo</contact:sp>
      <contact:pc>101-0065</contact:pc>
      <contact:cc>JP</contact:cc>
    </contact:addr>
  </postalInfo>
  <voice>+81.352158451</voice>
  <fax>+81.352158452</fax>
  <email>taro@example.co.jp</email>
  <clID>clientY</clID>
  <crID>clientX</crID>
  <crDate>2009-09-13T08:01:00.0Z</crDate>
  <upID>clientX</upID>
  <upDate>2009-11-26T09:10:00.0Z</upDate>
  <irdPostalInfo conversionMechanism="userinput">
    <name lang="ja"/>
    <org lang="ja"/>
    <addr lang="ja">

```

```

        <country>日本</country>
      </addr>
    </irdPostalInfo>
    <irdPostalInfo conversionMechanism="translation">
      <name lang="en"/>
      <org lang="en"/>
      <addr lang="en">
        <country>Japan</country>
      </addr>
    </irdPostalInfo>
  </contact>

```

B.4 Example of Registrar Object in XML

```

<?xml version="1.0" encoding="UTF-8"?>
<registrar xmlns="urn:ietf:params:xml:ns:irdRegistrar-1.0"
  xmlns:irdContact="urn:ietf:params:xml:ns:irdContact-1.0"
  xmlns:contact="urn:ietf:params:xml:ns:contact-1.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="urn:ietf:params:xml:ns:irdRegistrar-1.0
    irdRegistrar-1.0.xsd ">
  <id>clientX</id>
  <name>Example Inc.</name>
  <gudid>123</gudid>
  <status>ok</status>
  <postalInfo type="int">
    <addr>
      <contact:street>123 Example Dr.</contact:street>
      <contact:street>Suite 100</contact:street>
      <contact:city>Dulles</contact:city>
      <contact:sp>VA</contact:sp>
      <contact:pc>20166-6503</contact:pc>
      <contact:cc>US</contact:cc>
    </addr>
  </postalInfo>
  <voice x="1234">+1.7035555555</voice>
  <fax>+1.7035555556</fax>
  <email>jdoe@example.test</email>
  <url>http://www.example.test</url>
  <whoisInfo>
    <name>whois.example.test</name>
    <url>https://whois.example.test</url>
  </whoisInfo>
  <crDate>2005-04-23T11:49:00.0Z</crDate>
  <upDate>2009-02-17T17:51:00.0Z</upDate>
</registrar>

```

B2. Example XML Schemas of the Data Model

The following sections are a partial representation using XML of some of the requirements proposed in this document. They are not complete because of the technical considerations described in Section 2.3.

B2.1 IRD Domain Schema

```
<?xml version="1.0" encoding="UTF-8"?>
<schema targetNamespace="urn:ietf:params:xml:ns:irdDomain-1.0"
  xmlns="http://www.w3.org/2001/XMLSchema"
  xmlns:irdDomain="urn:ietf:params:xml:ns:irdDomain-1.0"
  xmlns:domain="urn:ietf:params:xml:ns:domain-1.0"
  xmlns:idn="urn:ietf:params:xml:ns:idn-1.0"
  xmlns:secDNS="urn:ietf:params:xml:ns:secDNS-1.1"
  xmlns:eppcom="urn:ietf:params:xml:ns:eppcom-1.0"
  elementFormDefault="qualified">

  <import namespace="urn:ietf:params:xml:ns:domain-1.0"
    schemaLocation="domain-1.0.xsd"/>
  <import namespace="urn:ietf:params:xml:ns:idn-1.0"
    schemaLocation="idn-1.0.xsd"/>
  <import namespace="urn:ietf:params:xml:ns:secDNS-1.1"
    schemaLocation="secDNS-1.1.xsd"/>
  <import namespace="urn:ietf:params:xml:ns:eppcom-1.0"
    schemaLocation="eppcom-1.0.xsd"/>

  <annotation>
    <documentation>IRD domain schema</documentation>
  </annotation>

  <element name="domain" type="irdDomain:domainType"/>

  <complexType name="domainType">
    <sequence>
      <element name="name" type="eppcom:labelType"/>
      <element name="roid" type="eppcom:roidType"/>
      <element name="status" type="domain:statusType" maxOccurs="11"
        minOccurs="0"/>
      <element name="registrant" type="eppcom:clIDType" maxOccurs="1"
        minOccurs="0"/>
      <element name="contact" type="domain:contactType"
        maxOccurs="unbounded" minOccurs="0"/>
      <element name="ns" type="domain:nsType" maxOccurs="1"
        minOccurs="0"/>
      <element name="clID" type="eppcom:clIDType"/>
      <element name="crID" type="eppcom:clIDType" maxOccurs="1"
        minOccurs="0"/>
      <element name="crDate" type="dateTime" maxOccurs="1" minOccurs="0"/>
      <element name="upID" type="eppcom:clIDType" maxOccurs="1" minOccurs="0"/>
      <element name="upDate" type="dateTime" maxOccurs="1" minOccurs="0"/>
      <element name="exDate" type="dateTime" maxOccurs="1" minOccurs="0"/>
      <element name="trDate" type="dateTime" maxOccurs="1" minOccurs="0"/>
      <element name="authInfo" type="domain:authInfoType" minOccurs="0"/>
      <element name="idn" type="idn:idnDataType" maxOccurs="1" minOccurs="0"/>
    </sequence>
  </complexType>
</schema>
```

```

        <element name="secDNS" type="secDNS:dsOrKeyType" maxOccurs="1"
            minOccurs="0"/>
    </sequence>
</complexType>

</schema>

```

B2.2 IRD Host Schema

```

<?xml version="1.0" encoding="UTF-8"?>
<schema targetNamespace="urn:ietf:params:xml:ns:irdHost-1.0"
    xmlns="http://www.w3.org/2001/XMLSchema"
    xmlns:irdHost="urn:ietf:params:xml:ns:irdHost-1.0"
    xmlns:host="urn:ietf:params:xml:ns:host-1.0"
    xmlns:eppcom="urn:ietf:params:xml:ns:eppcom-1.0"
    elementFormDefault="qualified">

    <import namespace="urn:ietf:params:xml:ns:host-1.0"
        schemaLocation="host-1.0.xsd"/>
    <import namespace="urn:ietf:params:xml:ns:eppcom-1.0"
        schemaLocation="eppcom-1.0.xsd"/>

    <annotation>
        <documentation>IRD host schema</documentation>
    </annotation>

    <element name="host" type="irdHost:hostType"/>

    <complexType name="hostType">
        <sequence>
            <element name="name" type="eppcom:labelType"/>
            <element name="roid" type="eppcom:roidType"/>
            <element name="status" type="host:statusType" maxOccurs="7"/>
            <element name="addr" type="host:addrType" minOccurs="0"
                maxOccurs="unbounded"/>
            <element name="clID" type="eppcom:clIDType"/>
            <element name="crID" type="eppcom:clIDType"/>
            <element name="crDate" type="dateTime"/>
            <element name="upID" type="eppcom:clIDType" minOccurs="0"/>
            <element name="upDate" type="dateTime" minOccurs="0"/>
            <element name="trDate" type="dateTime" minOccurs="0"/>
        </sequence>
    </complexType>

</schema>

```

B2.3 IRD Contact Schema

```

<?xml version="1.0" encoding="UTF-8"?>
<schema targetNamespace="urn:ietf:params:xml:ns:irdContact-1.0"
    xmlns="http://www.w3.org/2001/XMLSchema"
    xmlns:irdContact="urn:ietf:params:xml:ns:irdContact-1.0"
    xmlns:contact="urn:ietf:params:xml:ns:contact-1.0"
    xmlns:eppcom="urn:ietf:params:xml:ns:eppcom-1.0"
    elementFormDefault="qualified">

    <import namespace="urn:ietf:params:xml:ns:contact-1.0"
        schemaLocation="contact-1.0.xsd" />

```

```

<import namespace="urn:ietf:params:xml:ns:eppcom-1.0"
  schemaLocation="eppcom-1.0.xsd"/>

<annotation>
  <documentation>IRD contact schema</documentation>
</annotation>

<element name="contact" type="irdContact:contactType"/>

<complexType name="contactType">
  <sequence>
    <element name="id" type="eppcom:clIDType" />
    <element name="roid" type="eppcom:roidType" />
    <element name="status" type="contact:statusType" maxOccurs="7" />
    <element name="postalInfo" type="contact:postalInfoType"
      maxOccurs="2" />
    <element name="voice" type="contact:e164Type" minOccurs="0" />
    <element name="fax" type="contact:e164Type" minOccurs="0" />
    <element name="email" type="eppcom:minTokenType" />
    <element name="clID" type="eppcom:clIDType" />
    <element name="crID" type="eppcom:clIDType" />
    <element name="crDate" type="dateTime" />
    <element name="upID" type="eppcom:clIDType" minOccurs="0" />
    <element name="upDate" type="dateTime" minOccurs="0" />
    <element name="trDate" type="dateTime" minOccurs="0" />
    <element name="authInfo" type="contact:authInfoType" minOccurs="0" />
    <element name="disclose" type="contact:discloseType" minOccurs="0" />
    <element name="irdPostalInfo" type="irdContact:irdPostalInfoType"
      minOccurs="0" maxOccurs="2" />
  </sequence>
</complexType>

<complexType name="irdPostalInfoType">
  <sequence>
    <element name="name" type="irdContact:nameType" minOccurs="0" />
    <element name="org" type="irdContact:orgType" minOccurs="0" />
    <element name="addr" type="irdContact:addrType" />
  </sequence>
  <attribute name="conversionMechanism"
    type="irdContact:conversionMechanismEnumType" use="required"/>
</complexType>

<complexType name="nameType">
  <attribute name="lang" type="language" use="required"/>
</complexType>

<complexType name="orgType">
  <attribute name="lang" type="language" use="required"/>
</complexType>

<complexType name="addrType">
  <sequence>
    <element name="country" type="irdContact:countryType" />
  </sequence>
  <attribute name="lang" type="language" use="required"/>
</complexType>

<simpleType name="countryType">
  <restriction base="token">
    <minLength value="1"/></minLength>
    <maxLength value="255"/></maxLength>
  </restriction>
</simpleType>

```

```

<simpleType name="conversionMechanismEnumType">
  <restriction base="token">
    <enumeration value="transliteration" />
    <enumeration value="translation" />
    <enumeration value="userinput" />
  </restriction>
</simpleType>
</schema>

```

B2.4 IRD Registrar Schema

```

<?xml version="1.0" encoding="UTF-8"?>
<schema targetNamespace="urn:ietf:params:xml:ns:irdRegistrar-1.0"
  xmlns="http://www.w3.org/2001/XMLSchema"
  xmlns:irdRegistrar="urn:ietf:params:xml:ns:irdRegistrar-1.0"
  xmlns:irdContact="urn:ietf:params:xml:ns:irdContact-1.0"
  xmlns:eppcom="urn:ietf:params:xml:ns:eppcom-1.0"
  xmlns:contact="urn:ietf:params:xml:ns:contact-1.0"
  elementFormDefault="qualified">

  <import namespace="urn:ietf:params:xml:ns:irdContact-1.0"
    schemaLocation="irdContact-1.0.xsd"/>
  <import namespace="urn:ietf:params:xml:ns:eppcom-1.0"
    schemaLocation="eppcom-1.0.xsd"/>
  <import namespace="urn:ietf:params:xml:ns:contact-1.0"
    schemaLocation="contact-1.0.xsd"/>

  <annotation>
    <documentation>IRD registrar schema</documentation>
  </annotation>

  <element name="registrar" type="irdRegistrar:registrarType"/>

  <complexType name="registrarType">
    <sequence>
      <element name="id" type="eppcom:clIDType" />
      <element name="name" type="contact:postalLineType"/>
      <element name="gurid" type="positiveInteger" minOccurs="0" />
      <element name="status" type="irdRegistrar:statusType" />
      <element name="postalInfo" type="irdRegistrar:postalInfoType"
        maxOccurs="2" />
      <element name="voice" type="contact:e164Type" minOccurs="0" />
      <element name="fax" type="contact:e164Type" minOccurs="0" />
      <element name="email" type="eppcom:minTokenType" />
      <element name="url" type="anyURI" minOccurs="0" />
      <element name="whoisInfo" type="irdRegistrar:whoisInfoType"
        minOccurs="0" />
      <element name="crDate" type="dateTime" />
      <element name="upDate" type="dateTime" minOccurs="0" />
      <element name="irdPostalInfo" type="irdContact:irdPostalInfoType"
        minOccurs="0" maxOccurs="2" />
    </sequence>
  </complexType>

  <simpleType name="statusType">
    <restriction base="token">
      <enumeration value="ok"/>
      <enumeration value="readonly"/>
    </restriction>
  </simpleType>

```



```
        <enumeration value="terminated"/>
      </restriction>
    </simpleType>

    <complexType name="postalInfoType">
      <sequence>
        <element name="addr" type="contact:addrType"/>
      </sequence>
      <attribute name="type" type="contact:postalInfoEnumType"
        use="required"/>
    </complexType>

    <complexType name="whoisInfoType">
      <sequence>
        <element name="name" type="eppcom:labelType" minOccurs="0"/>
        <element name="url" type="anyURI" minOccurs="0"/>
      </sequence>
    </complexType>
  </schema>
```

Appendix C: Working Group Composition and Biographies of members

The WG gathered a broad set of participants with a diverse set of expertise in areas that includes linguistic, experience with Unicode, registry and registrar operations, ICANN policy, and internationalization and localization in applications. Included below please find a brief description of these expertise areas and the biographies of the participants.

Area of Expertise	Summary Description
Linguistics / Unicode	Linguistic experts in the specific languages/script, ideally with some knowledge in Unicode.
Registry/Registrar Operations	gTLD and ccTLD experts familiar with registry/registrar operations and standards with WHOIS and EPP.
Policy	Experts knowledgeable of ICANN's current WHOIS policy and contractual obligations
Internationalization and Localization in Application	Experts knowledgeable in internationalization and localization in applications

Registry / Registrar Operations

Dennis Tan

Dennis is the principal point of contact for Internationalized Domain Names (IDNs) within the Naming Business Division at VeriSign, Inc. As product manager for VeriSign, he oversees all product management activities for IDNs, including defining requirements for VeriSign's provisioning and resolution system for IDNs.

Prior to joining VeriSign, Dennis worked in the Telecommunications industry as well as the textile industry. He holds a Bachelor Degree in Industrial Engineering and a MBA.

Jody Kolker

Jody joined GoDaddy in 2001 as a senior developer. He designed and developed the systems for registrar – registry interactions. He is currently focusing on the architecture of GoDaddy’s Registrar Systems and managing GoDaddy’s registrar operations team.

Naoki Kambe

Naoki is an R&D staff from Japan Registry Services Co, Ltd (JPRS). He had both operation and research experience in Whois and DNS fields. In 2004 he deployed helper tools and systems for operation of JP Registry system. He was also a part of the team to operate back-end databases for Whois service. He now engages in development of new gTLD Whois server for .JPRS, including modelling registration data for the new gTLD.

Zheng Wang

Zheng is the director of Joint Labs at China Organizational Name Administration Center (CONAC). He plays a crucial role in designing and building the IDN technical solution including Whois, EPP/SRS, Data Escrow, DNS, and DNSSEC systems for CONAC. He is also a leading researcher or principal investigator of several state-funded research or engineering projects on the DNS measurement, modelling and optimization. Before joining CONAC, he served as a senior researcher on DNS operations at CNNIC. He has a doctorate in computer science from Chinese Academy of Sciences.

Policy

Edmon Chung

Edmon is serving as the CEO for DotAsia Organisation and heads the secretariat for the Asia Pacific Regional Internet Governance Forum (APrIGF). He is an inventor of patents underlying technologies for internationalized domain names (IDN) and email addresses on the Internet. Edmon has served on many global IDN related committees, including technical and policy groups, that made it possible for the introduction of multilingual domain names and email addresses on the Internet. He served as an elected GNSO Councilor from 2006 – 2010 and an elected ALAC Member from 2010-2012.

James Galvin

Dr. James Galvin is Director of Strategic Relationships and Technical Standards at Afilias, focusing on the development of strategic initiatives in all aspects of networking and the life cycle of domain names. Jim is experienced in registry and registrar operations (including WHOIS and EPP standards) as well as policy issues covering WHOIS and contractual obligations. He serves as the Vice-Chair of ICANN's Security and Stability Advisory Committee (SSAC). He was the co-chair of the GNSO-SSAC Internationalized Registration Data Working Group. Jim holds a Ph.D. in Computer and Information Sciences from the University of Delaware.

Unicode / Linguistics**Nishit Jain**

Nishit Jain is a research staff with the Centre for Development of Advanced Computing (C-DAC), the premier R&D organization for the Ministry of Communications & Information Technology in India. He is involved with various projects related to Indian Languages on Digital medium, including internationalized domain names, efficient searching algorithms for Indian languages. He is also working on an Indian government project on defining

Internationalized Domain Names (IDNs) requirements for the 22 official languages of India.

Sarmad Hussain (September 2013 – May 2014)¹⁶

Dr. Sarmad Hussain is a professor of Computer Science and holds the Research Chair on Multilingual Computing at Al-Khwarizmi Institute of Computer Science in Pakistan. He holds a doctoral degree in linguistics and his research is focused on linguistics, localization, language computing standards, speech processing and computational linguistics. He has been developing computing solutions for languages spoken across developing Asia, including standards for Unicode encoding, locale and collation.

Internationalization and Localization in Applications

Takao Suzuki

Takao recently joined GoDaddy as Senior Product Manager International with current focus on domains: international planning, strategy, internationalization, and localization. Prior to GoDaddy, he worked for Microsoft over 18 years as International Program Manager on various products including Internet Explorer, Windows, and Windows Live. Takao is a native Japanese.

¹⁶ Sarmad Hussain participated in the preparation of this Report as a WG member from September 2013 to May 2014, prior to assuming his current position as IDN program Senior Manager at ICANN.

Appendix D: Working Group Response to Public Comments on the Interim Report

On 14 April 2014, the IRD Working Group produced its Interim Report for public comment. In total 4 comments were received. The IRD Working Group has studied and addressed these comments in its subsequent meetings. Below please find the working group response to the public comment.

Comments from ICANN Business Constituency (BC)

BC supported the proposal in the IRD report (proposal 3), with additional comments:

- The BC recommends the use of free-form text for all name and address fields.
- 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. Per the Whois Accuracy Specification, phone numbers and e-mail address formats should still be validated.
- The BC supports localizing data labels based upon registrant location, but it is important that standard US-ASCII data labels still also appear.

Working Group Response:

Based on the input from public comment, and working group deliberations, the working group chose Proposal 1. Please see Section 5.1.2 Postal address of registrant, technical and administrative contact for the discussions of pros and cons of each of the proposal and the rationale for choosing proposal 1.

The working group disagrees with BC that non-ASCII character sets are optional. Per user capability principle, a cornerstone assumption of an internationalized system is that a registrant must be able to use the language(s) and script(s) in which they are most skilled. Thus a registrant must not be required to use any

specific language(s) or script(s) beyond what would be used in their ordinary daily routine.

Regarding the internationalization of the label, this is a *presentation* issue that is left for clients to decide. The *representation* of the label is US-ASCII. So by default, end user who chose to see US-ASCII label will be able to.

Comments from Verisign Inc (Verisign)

Verisign supports proposal 3, but recommended language tag be removed as a requirement, and that EAI support be made optional until EAI is widely adopted, and finally data labels should remain in ASCII and not be localized.

Working Group Response:

Based on the input from public comment, and working group deliberations, the working group chose Proposal 1. Please see Section 5.1.2 Postal address of registrant, technical and administrative contact for the discussions of pros and cons of each of the proposal and the rationale for choosing proposal 1.

Verisign cites that language tags currently are not implemented in WHOIS and EPP, the working group agrees that such technical challenge exists (see section 2.3: Technical Considerations). However, the working group is making recommendation for the end state, not the transition.

Regarding the internationalization of the label, this is a *presentation* issue that is left for clients to decide. The *representation* of the label is US-ASCII. So by default, end user who chose to see US-ASCII label will be able to.

Comments from ICANN Intellectual Property Constituency (IPC)

IPC notes that user capability principle not only should apply to registrants, but it should also apply to users of the data. The IPC preferred option 2, and raised the question why option 2 does not apply to other categories such as the name and organizational name.

Working Group Response:

The user capability principle states that in defining a requirement for a particular data element or category of data elements, the capability of the data-submitting user should be the constraining factor. Such users should not be burdened with tasks that cannot be completed under ordinary circumstances. The data submitting user is preferred over the user of the data because the former is an integral part of the life cycle of a domain name. Further, the needs of the user of the data can be met with a localization transformation of the internationalized data.

Based on the input from public comment, and working group deliberations, the working group chose Proposal 1. Please see Section 5.1.2 Postal address of registrant, technical and administrative contact for the discussions of pros and cons of each of the proposal and the rationale for choosing proposal 1.

Comments from ICANN At-Large Advisory Committee (ALAC)

ALAC reaffirms that IDN are critical for enhancing diversity and multilingualism on the Internet, and that it is important the IDN policy evolution should allow more accessibility to those who do not use the ASCII character set.

Working Group Response:

The working group agrees with ALAC's comment.

