Interim Report from the Expert Working Group on Internationalized Registration Data

STATUS OF THIS DOCUMENT

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

I. 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 as "The IRD Working Group") was formed to recommend the requirements for internationalized registration data (IRD) to be displayed in WHOIS and produce a data model for the IRD that matches the requirements.

The IRD Working Group began its deliberation by examining its scope and developing a methodology. The key findings are:

- The IRD Working Group has focused its analysis on the internationalization of registration data, not the localization of the data. 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). 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's analysis focuses on data element categories rather than a list of specific data elements. A complete list of data elements changes with time and may vary between registries. In contrast, data element categories are more stable and their requirements will have broader applicability.

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

- **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 that he/she could not 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 standards.

Based on these principles, the IRD Working Group proposes the following internationalization requirements, which are supported by a consensus of its members, except where there is more than one proposal listed:

Data Categories	Example Data Elements	Proposed Requirement
Name, organization of registrant, technical and administrative contact	Registrant Name, Registrant Organization	Free-form text.
Registrar Name	Sponsoring Registrar	Free-form text. The name of the sponsoring registrar should be the official name in the RAA with ICANN, in whichever language(s) or script(s).
Address of registrant, technical and administrative contact	Registrant Address1, Registrant Address2, Registrant City, Registrant State/Province	[Proposal 1] Free form text. The language(s) or script(s) of an address should be appropriate for the region that 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.
Postal Code of registrant, technical and administrative contact	Registrant Postal Code	Free form text.
Country of registrant / technical and administrative contact	Registrant Country	The country and territory names should be from the United Nations Group of Experts on Geographical Names, Technical Reference Manual for the Standardization of Geographical Names, Part III Names of Countries of the World;
Country Code	Registrant Country Code	The country code should be from ISO 3166 part 2 code list.

Data Categories	Example Data Elements	Proposed Requirement
Status	Domain Status	The text value of the domain status should conform to EPP specification defined in RFC 5731 section 2.3.
Phone Fax Numbers	Technical Contact Facsimile Number, Technical Contact Phone Number	The phone and fax numbers should comply with ITU E.164 standard.
Email	Technical Contact Email, Registrant Email, Administrative Contact Email	Email address 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	For the IANA ID, it should be the numeric ID assigned by the IANA for a given registrar or registry. For other elements no internationalization standards

		should be applied.
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 or values described in RFC 5910.
URLs	Referral URL, Registrar URL (registration services)	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 during the submission. For display, require directory services to display both U-label and the corresponding A-label for all domain names.
Dates	Last Transferred Date, Domain Last Updated Date, DS creation date	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.

Next Steps

This Interim Report will be published in the Public Comment Forum until 22 May 2014. Upon the closure of the Public Comment Forum, the IRD Working Group will finalize the proposed requirement based on community input. In anticipation of the finalized requirement, the IRD Working Group will also produce a data model that will be published in a Final Report to be delivered to the ICANN Board.

II. BACKGROUND

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

The WHOIS Policy Review Team, in its final report, highlights the needs to define requirements, data models and evaluate solutions for internationalized registration data, 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 are 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 transliterations of contact information are needed. If so, specify who should bear the burden.

This report is an interim report related to the first effort.

III. 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. These are explored in detail below.

3.1 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).

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, and proposes the following distinction below:

Localization of Registration Data and Directory Service refers to the adaption of the directory service to meet the language, cultural, and other requirements of a specific data consumer group. For example, the following could be the localized directory service (WHOIS) output of an IDN "ドメイン名例.JP " for Japanese registries, registrants, and end-users. Both the description of the data element, and the data element itself is in a form suitable for local usage.

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
[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)
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
[Web Page]
[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
[FAX 番号]	03-5215-8452

From the above example, we note that for domain name registration data directory service, localization can entail customization related to:
Numeric, date and time formats that complies 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
One important point to note is that for a given set of registration data, there may be a need multiple localized versions as the consumer of the data is the global Internet audience. For example, the above data could be localized to English speaking audience, as shown below. In the example shown below, Japanese labels like "登録年月日", "住所", and "電話番号" are respectively localized to "Created on", "Postal Address", and "Phone", as well as only the English translation / transliteration of the address is shown.

[Domain Name]	XN--ECKWD4C7CU47R2WF.JP
[Registrant]	Example Corporation
[Name Server]	ns01.example.co.jp
[Name Server]	ns02.example.co.jp
[Created on]	2001/08/09
[Expires on]	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

One can imagine the data above can be localized to other languages / scripts.

Internationalization of Registration Data and Directory Service refers to a registration data directory service application and its data that enables easy localization for specific data consumer groups that vary in culture, region, or language.

Internationalization in this context entails:

- 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 or 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 this interim report, the IRD Working Group's proposals focus on internationalization of registration data, not localization.

3.2 Categories of Data Elements

There are over 150 data elements currently outputted 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 Data). 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 13 data categories that cover all of the known data elements.

- Names (includes names of organization, registrants, and registrars)
- Postal Addresses
- Statuses
- Phone / Fax Numbers
- Email Addresses
- Country Codes
- Country Names
- Identifiers
- DNSSEC Information
- URLs
- Domain Names
- Time and Dates
- Languages

IV. 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 a set of data elements, the capability of the average user should be the constraining factor. Users should not be burdened with tasks that he/she could not 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.

V. 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.

5.1 Registrants should only be required to input registration data in a language(s) or script(s) that they are skilled at.

Per user capability principle, a cornerstone assumption of an internationalized system is that a registrant must be able to use the language and script 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.

The application of this principle led the IRD Working Group to propose at least two requirements:

- A registry must be able to accept and store any language or script that might reasonably be expected to be used in the 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.
- 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. A direct consequence is that a registrant may have a limited set of registrars from which to choose since they must find one that supports the language and script in which they are most skilled.

5.2 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 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 it can be canonicalized, which cannot be done unless we are certain of what it represents.

VI. PROPOSED REQUIREMENTS FOR INIDIVIDUAL DATA CATEGORIES

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

6.1.1 Name, organization of registrant, technical and administrative contact

- **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 essential to give registrants the ability to submit information in the language/script of his choice, as many users will be monolingual (user capability principle).

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

  In terms of registry support, RFC 5733 has a method to allow localized language characters into the data elements. The contact object allows two elements for postal address information (internationalized and localized). They must be defined with a "type" attribute (i.e. "int", "loc") and

each type allows a certain type of characters to be used, where "loc" is the most inclusive one because it allows "unrestricted" UTF-8.

The WG notes that if translation or transliteration is employed, the method described in RFC 5733 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 considered out of the scope for this WG.

The WG did 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 allowed to submit internationalized registration data, whether to require that users submit a corresponding single common script version of the internationalized registration data?
- If registrants are required to submit a single common script version of the internationalized 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 the internationalized 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 an internationalized version of their registration data, should there be a requirement to detect whether both literally match each other?

- If there are two versions of the registration data, which version should be considered primary or authoritative if there is a mismatch?
- 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?

6.1.2 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
- **Proposed Requirement:** The Working Group explored the following proposed requirements for the address element:

  - Proposal 1: Free form text. The language(s) or script(s) of an address should be appropriate for the region that 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.

  The IRD Working Group was unable to reach consensus on which proposal is the best. In the sections below, pros and cons of each approach is highlighted. The IRD Working Group wishes to solicit community feedback on this issue, and will make a determination in its final report, based on community input.

- **Rationale:**

Pros and cons for proposal 1: In proposing this requirement, IRD Working Group members understands that it is possible (or even common) for a registrant to reside in a region in which he/she is not skilled in the language/script for that region. (e.g., a Chinese speaker living in America or Middle East that does not know much English or Arabic). However, this is still a useful requirement to propose for the following reasons: 1) an address is closely associated with the region and the country it is in. 2) One of the use cases discussed by the IRD Working Group is to send postal mail based on registration data records. Thus, to maximize the possibility of delivery, it is best to have the address in a language a script that can be understood by the postal delivery agents in the region.

Pros and cons for proposal 2: The IRD Working Group member who proposed this requirement believes that proposal 1 poses too much burden on registrars, as the registrars would have to accept every language or script in the world. Such a requirement, coupled with RAA requirement for validation is very difficult to implement. Proposal 2 uses IDN as a signal of user demand for internationalized registration data. It strikes a balance between supporting internationalized data for IDN registrations, where most of the demand for internationalized registration data lies, and a Romanized version of the registration data, in case the registrant is not familiar with the registrar's local language of support, and where most postal mail agents can deliver.

Pros and cons for proposal 3: Proponents of proposal 3 argue that the requirement in proposal 1 would mean that a list needs to exist that would link a region to a language and script, which currently does not exist; it can be difficult for registrars to accurately determine the language of the text 100% of the time, and finally the requirement for proposal 2 seems too restrictive to prospective registrants. If the registrant would like to use a

script/language that is not appropriate for the region, why should they not be allowed to?

- **Discussion:** The IRD Working Group also discussed whether it is feasible to require the address in a *single* language/script. The preliminary conclusion is that this is not feasible as script mixing is fairly common for addresses, e.g., the use of roman numerals in street addresses. Finally, the issues related to translation and transliteration outlined in 5.1.1 also applies here.

6.1.3 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 RAA with ICANN, in whichever language(s) or script(s).
- **Rationale:** simplicity and reusability

6.1.4 Country

Proposed Requirements: The country and territory names should be from the following internationally recognized list:

- The United Nations Group of Experts on Geographical Names, Technical Reference Manual for the Standardization of Geographical Names, Part III Names of Countries of the World;

6.1.5 Country Code

- **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

6.1.6 Postal Code

- **Data Elements Covered:** Registrant Postal Code, Administrative Contact Postal Code, Technical Contact Postal Code
- **Proposed Requirement:** Free form text.

6.2 Requirements for other data elements

6.2.1 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.

6.2.2 Phone/ Fax Numbers

- **Data Elements Covered:** 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., Registrant Phone Number Ext.

- **Proposed Requirement:** The phone and fax numbers should comply with ITU E.164.2005 as defined in RFC 5733.

6.2.3 Email

- **Data Elements Covered:** 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.

- **Discussion:** Internationalized email addresses as specified in RFC 6532 have not been widely adopted as of this writing. The working group is expressly interested in feedback from the community on whether or not this is a significant issue.

  The IRD Working Group observes the international standards are backward compatible, so there is no syntactic issue. Operationally however, there are issues to be considered. Specifically, if a registrar accepts an internationalized email address, its internal email systems that use this data may need to be updated.

  In addition, third parties that consume this data will have similar issues.

  The IRD Working Group is expressly interested in feedback from the community on whether or not a transition period would be useful. During this transition period a registrant could be permitted to enter either or both an internationalized email address or not. If there are other suggestions for how to use the transition period the IRD working group welcomes comments from the community.

6.2.4 Identifiers

- **Data Elements Covered:** Registrar ID, Technical Contact ID, Registrant ID, Administrative Contact ID, Sponsoring Registrar IANA ID, Domain ID

- **Proposed Requirement:** No internationalization of this data element is required. For the IANA ID, it should be the numeric ID assigned by the IANA for a given registrar or registry.

- **Discussion:** Identifiers are usually registry or registry internal objects, therefore no internationalization standards should be applied.

6.2.5 URLs

- **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.

6.2.6 Domain Names

- **Data Elements Covered:** Domain Name, Whois Server, Name Server

- **Proposed Requirement:** Where a domain name is provided by registrant, only require registrants to provide domain name in U-label during the submission**.** For display, require directory services to display both U-label and the corresponding A-label for all domain names.

- **Discussion:** In current implementations, it is common to have both the U-label and A-label form 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.

6.2.7 Dates

- **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.[1]

6.2.8 DNSSEC

- **Data Elements Covered:** DS Key Tag, Algorithm, Digest Type, Digist, DS Maximum Signature Life

- **Proposed Requirement:** Elements should conform to formats / values described in RFC 5910.

VII. Summary of Proposed Requirements and Next Steps

The table below summarizes the key aspects of the IRD Working Group's proposed requirements.

Data Categories	Example Data Elements	Proposed Requirement

[1] 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".
1

Name, organization of registrant, technical and administrative contact	Registrant Name, Registrant Organization	Free-form text.
Registrar Name	Sponsoring Registrar	Free-form text. The name of the sponsoring registrar should be the official name in the RAA with ICANN, in whichever language(s) or script(s).
Address of registrant, technical and administrative contact	Registrant Address1, Registrant Address2, Registrant City, Registrant State/Province	[Proposal 1] Free form text. The language(s) or script(s) of an address should be appropriate for the region that 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.
Postal Code of registrant, technical and administrative contact	Registrant Postal Code	Free form text.
Country of registrant / technical and administrative contact	Registrant Country	The country and territory names should be from the United Nations Group of Experts on Geographical Names, Technical Reference Manual for the Standardization of Geographical Names, Part III Names of Countries of the World;

Country Code	Registrant Country Code	The country code should be from ISO 3166 part 2 code list.

Data Categories	Example Data Elements	Proposed Requirement
Status	Domain Status	The text value of the domain status should conform to EPP specification defined in RFC 5731 section 2.3.
Phone Fax Numbers	Technical Contact Facsimile Number, Technical Contact Phone Number	The phone and fax numbers should comply with ITU E.164. standard.
Email	Technical Contact Email, Registrant Email, Administrative Contact Email	Email address 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, Admin Contact ID, Tech Contact ID	For the IANA ID, it should be the numeric ID assigned by the IANA for a given registrar or registry. For other elements no internationalization standards should be applied.
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, Digist 2, DS Maximum Signature Life 1, Digist 1	Elements should conform to format or values described in RFC 5910.
URLs	Referral URL, Registrar URL (registration services)	Conform to standards set forth in RFC 3986 and RFC 3987
Domain	Domain Name, Whois	Where a domain name is

Names	Server, Name Server	provided by registrant, only require registrants to provide domain name in U-label during the submission. For display, require directory services to display both U-label and the corresponding A-label for all domain names.
Dates	Last Transferred Date, Domain Last Updated Date, DS Creation date	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.

This Interim Report will be published in the Public Comment Forum until 22 May 2014. Upon the closure of the Public Comment Forum, the IRD Working Group will finalize the proposed requirement based on community input. The IRD Working Group currently is working on a data model that will be incorporated into the Final Report to be delivered to the ICANN Board.

VIII. Related Work

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

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.

The IETF 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 behaviour of a subset of ccTLDs to create a profile of a registration data model.

It is also possible that the ICANN Expert Working Group on Next Generation Directory Services may suggest additional requirements for domain name registration data that will affect the internationalization of some data elements.

The IRD Working Group will consider all of these inputs as part of its work to define the data model, which will need to consider the data elements that need to be included in the data model.

The GNSO PDP on the Translation and Transliteration of Contact Names is just beginning its work. Based on the issues report [6] upon which it is based its work would appear to be focused 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.

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.

IX. 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 (draft-zhou-weirds-dnrd-ap-object-inventory-01), 2013. Available at < http://tools.ietf.org/html/draft-ietf-weirds-object-inventory-01>

6. Internet Corporation for Assigned Names and Numbers (ICANN). (2013) Final Issues Report on translation and transliteration of Contact Information. Marina Del Rey, CA: ICANN. Available at: <http://gnso.icann.org/en/issues/gtlds/transliteration-contact-final-21mar13-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: 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 expertise areas and their biographies.

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 developed and deployed helper tools and systems for operation of JP Registry system. He was also 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

Dr. Sarmad Hussain is currently a professor of Computer Science and holds the Research Chair on Multilingual Computing at Al-Khawarizmi 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.