

## 8 Background

At its meeting on 20 November 2013, the GNSO Council approved the Charter on the Translation and Transliteration of Contact Information PDP Working Group. (See <http://gns0.icann.org/en/council/resolutions#201311>). Accordingly a group of volunteers is being convened to provide answers to the questions posed in the Charter, consult with the Community, draft an Initial Report and finally present a Final Report to the GNSO Council.

The translation and transliteration of contact information were but two issues addressed by the IRD-WG in its Final Report. That Report recommended that the GNSO Council should request an issue report on the translation and transliteration of contact information. In the context of these issues, "contact information" is a subset of Domain Name Registration Data. It is the information that enables someone using a Domain Name Registration Data Directory Service (such as WHOIS) to contact the domain name registration holder. It includes the name, organization, and postal address of the registered name holder, technical contact, as well as administrative contact. Domain Name Registration Data is accessible to the public via a Directory Service (also known as the WHOIS service). The Registrar Accreditation Agreement (RAA 3.3.1) specifies the data elements that must be provided by registrars (via Port 43 and via web-based services) in response to a query, but it does not require that data elements, such as contact information, must be translated or transliterated.

The IRD-WG identified internationalized domain name registration data requirements in addition to the translation and transliteration of contact information. It recommended that ICANN staff should develop, in consultation with the community, a data model for domain registration data. The data model should specify the elements of the registration data, the data flow, and a formal data schema that incorporates the standards for internationalizing various registration data elements. Accordingly, in its 08 November 2012 resolution and [Action Plan](#) the ICANN Board directed staff to: 1) task a working group to determine the appropriate internationalized domain name registration data requirements, evaluating any relevant recommendations from the SSAC or GNSO; 2) produce a data model that includes (any) requirements for the translation or transliteration of the registration data, taking into account the results of any PDP initiated by the GNSO on translation/ transliteration, and the standardized replacement protocol under development in the IETF's Webbased Extensible Internet Registration Data Working Group; 3) evaluate available solutions (including solutions being implemented by ccTLDs). Thus, the results of the PDP on translation and transliteration of contact information will be considered by the working group described above for which a separate Call for Volunteers will be issued.

With respect to the two issues identified above concerning the translation and transliteration of contact information, the following additional background may be useful. On the first issue, whether it is desirable to translate contact information to a single common language or transliterate contact information to a single common script, the IRD-WG noted that, "[t]o balance the needs and capabilities of the local registrant with the need of the (potential) global user of this data, one of the key questions ... is whether DNRD-DS [Domain Name Registration Data Directory Services] should support multiple representations of the same registration data in different languages or scripts." In particular, the IRD-WG members discussed whether it is desirable to adopt a "must be present" representation of contact data, in conjunction with local script support for the convenience of local users. By "must be present" the IRD-WG meant that contact data must be made available in a common script.

In general, the IRD-WG recognized that, "the internationalized contact data can be translated or transliterated into the 'must be present' representation. As noted above, in this context, **Translation** is the process of conveying the meaning of some passage of text in one language, so that it can be expressed equivalently in another language. **Transliteration** is the process of representing the characters of an alphabetical or syllabic system of writing by the characters of a conversion alphabet." Based on this definition, and consistent with the current state of domain name registration data, the IRD-WG noted that if transliteration were desired, then the "must be present" script would be the Latin script. If translation were desired, then the "must be present" language would be English.

The IRD-WG did note that many language translation systems are inexact and cannot be applied repeatedly to translate from one language to another. Thus the IRD-WG noted that there will likely be problems with both consistency and accuracy, such as:

- Translation/transliteration may vary significantly across languages using the same script.
- Two people may translate/transliterate differently even within a language and the same person may translate/transliterate differently at different times for the same language.
- How would a registrar determine which particular spellings to use for a particular registrant? How would a registrant ever verify the correctness of a translation or transliteration, even if presented such data by the registrar or by a third organization that does the translation/transliteration?

Furthermore, the IRD-WG noted that for a given script, there may exist multiple systems for transliteration into Latin scripts. In the case of Chinese, the multiple transliteration systems are not only quite different from each other, but most of the systems use particular Latin characters to represent phonemes that are quite different from the most common phoneme-character pairings in European languages.

Finally, it is unclear whether translation or transliteration would serve the needs of the users of contact data. For example it is unclear that translating the name of the registrant and city would be useful. Would one have to translate "Los Angeles" into "City of the Angels" and translate "Beijing" into "Northern Capital"? The PDP should explore whether such translations facilitate or hinder the ability to contact the registrant.

The second issue, who should decide who should bear the burden translating contact information to a single common language or transliterating contact information to a single common script, relates to the concern expressed by the IRD-WG in its report that there are costs associated with providing translation and transliteration of contact information. For example, if a PDP determined that the registrar must translate or transliterate contact information, this policy would place a cost burden on the registrar. The IRD-WG considered several alternatives to address translation and transliteration of contact information as follows:

- The registrant submits the localized information as well the translated or transliterated information.
- The registrant only submits the localized information, and the registrar translates and transliterates all internationalized contact information on behalf of the registrant.
- The registrant only submits the localized information, and the registrars provide a point of contact at a service that could provide translation or transliteration upon request for a fee to be paid by the requester.
- The registrant only submits the localized information, and the registry provides translation or transliteration.
- The end users of the registration data translate and transliterate the contact information.

During their deliberations the members of the IRD-WG recognized that many registrants will need to access domain names in their local scripts and languages, which is the one of the primary reasons for the expansion of internationalized domain names. Therefore, the IRD-WG determined that it is unreasonable to assume all registrants – wherever they happen to be located – will be able to enter the registration data in scripts or languages other than their local script or language.

On 17 October 2012 the GNSO Council requested an Issue Report to address the three issues that were identified by the IRD-WG:

- Whether it is desirable to translate contact information to a single common language or transliterate contact information to a single common script.
- Who should decide who should bear the burden translating contact information to a single common language or transliterating contact information to a single common script. This question relates to the concern expressed by the Internationalized Registration Data Working Group (IRD-WG) in its report that there are costs associated with providing translation and transliteration of contact information. For example, if a policy development process (PDP) determined that the registrar must translate or transliterate contact information, this policy would place a cost burden on the registrar.
- Whether to start a PDP to address these questions.

The [Final Issue Report](#) on translation and transliteration of contact information was submitted to the GNSO Council on 21 March 2013 and on 13 June 2013 the GNSO Council approved the initiation of a PDP on the translation and transliteration of contact information.

The GNSO Council also requested ICANN to commission a study on the commercial feasibility of translation or transliteration systems for internationalized contact data, which is expected to be completed in time to help inform the PDP Working Group in its deliberations.