Initial Report on the Translation and Transliteration of Contact Information Policy Development Process

Status of this Document

This is the initial report of the Translation and Transliteration of Contact Information PDP Working Group for submission to the community to seek public comment. A Final Report will be prepared following public comment and presented to the GNSO Council.

Table of Content

Executive Summary

Objectives and Next Steps

This is the initial report of the Translation and Transliteration of Contact Information PDP Working Group, presented to the Community to gather feedback on the various recommendations the Working Group is presenting. Following the review of any public comments received, the WG will prepare a final report to be presented to the GNSO Council for its review and possible adoption.

Mission and Scope

The Translation and Transliteration of Contact Information Policy Development Process (PDP) Working Group is concerned with the way that contact information data – commonly referred to as 'Whois' – are collected and displayed within generic top-level domains (gTLDs). According to the Charter (see also Annex A), the PDP Working Group "is tasked to provide the GNSO Council with a policy recommendation regarding the translation and transliteration of contact information. As part of its deliberations on this issue, the PDP WG should, at a minimum, consider the following issues:

- 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 [of] translating contact information to a single common language or transliterating contact information to a single common script?"

In relation to the first question, the Charter notes "text requests and content returned by Domain Name Registration Data Services (such as WHOIS) are historically encoded using US-American Standard Code for Information Interchange (ASCII). This is a character-encoding scheme originally based on the English alphabet. While the WHOIS protocol does not specify US-ASCII as the exclusive character set for text requests and text content encoding, the current situation is that no standards or conventions exist for all WHOIS protocol implementations to signal support of character sets other than US-ASCII."

The second 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."

Finally, the Charter also encouraged the Working Group to consider the following issues related to its two core charter questions:
- What exactly the benefits to the community are of translating and/or transliterating contact data, especially in light of the costs that may be connected to translation and/or transliteration?

- Should translation and/or transliteration of contact data be mandatory for all gTLDs?

- Should translation and/or transliteration of contact data be mandatory for all registrants or only those based in certain countries and/or using specific non-ASCII scripts?

- What impact will translation/transliteration of contact data have on the WHOIS validation as set out under the 2013 Registrar Accreditation Agreement?

- When should any new policy relating to translation and transliteration of contact information come into effect?

Approach taken by the Working Group

The Translation and Transliteration Working Group convened its first meeting on 19 December 2013. The Working Group prepared a work plan, which has been reviewed on a regular basis, and revised when necessary. Also, Constituency and Stakeholder Group statements with regard to the Charter questions (see Annex A) were solicited. This request was also directed to other ICANN Supporting Organizations (SOs) and Advisory Committees (ACs) and a summary of responses can be seen in the public comment review tool. The Group prioritized discussing the community input received, to understand better the arguments brought forward by various stakeholders. This is also the reason that the Group decided to create a straw man proposal to drive forward the debate on whether or not it is desirable to translate/transliterate. This proposal provided a focal point to the Group's discussion and was updated on a regular basis.

Membership

Name	Affiliation*
Amr Elsadr	NCUC
Anthony Oni	NCUC
Ching Chiao	RySG
Chris Dillon (co-Chair)	NCSG
David Cake (Observer)	NCSG
Dennis Tan Tanaka	RySG
Edmon Chung	RySG
Ephraim Percy Kenyanito	NCUC
Jennifer Chung	RySG
Jim Galvin	SSAC

Name	Affiliation*
Jonathan Robinson (Observer)	RySG
Justine Chew	Individual
Mae Suchayapim Siriwat	GAC
Patrick Lenihan	NCUC
Peter Dernbach	IPC
Petter Rindforth	IPC
Pitinan Kooarmornpatana	GAC
Rudi Vansnick (co-Chair)	NPOC
Sarmad Hussain	SSAC
Vinay Kumar Singh	Individual
Volker Greimann (Observer)	RrSG
Wanawit Ahkuputra	GAC
Wolf-Ulrich Knoben	ISPC
Yoav Keren	RrSG
Zhang Zuan	NCUC

*ALAC – At-Large Community

RrSG – Registrar Stakeholder Group

RySG – Registry Stakeholder Group

CBUC – Commercial and Business Users Constituency

NAF – National Arbitration Forum

NCUC – Non Commercial Users Constituency

IPC – Intellectual Property Constituency

ISPCP – Internet Service and Connection Providers Constituency

NCSG – Non-Commercial Stakeholder Group

The Statements of Interest (SOI) for the Working Group members can be found at:

https://community.icann.org/x/WDd-Ag

The attendance records can be found at: https://community.icann.org/x/VlF-Ag

The email archives can be found at: http://forum.icann.org/lists/gnso-contactinfo-pdp-wg/

Deliberation and Recommendations

This section provides an overview of the deliberations of the Working Group. It is intended to serve as a record of the discussion and analysis of the Working Group, in support of the recommendations made in the following section.

During its initial discussion the Working Group identified a number of further issues and questions that are directly linked to the Charter questions, including relevant taxonomies. Details can be found on the Working Group's wiki page: https://community.icann.org/x/WwmuAg.

The Working Group decided to define clearly what is meant by 'contact information', relying on the Final Issue Report on the Translation and Transliteration of Contact Information that is based on the definition in the Registrar Accreditation Agreement 2013: "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."[1]

Deliberation on the two main Charter questions
Is it desirable to translate contact information to a single common language or transliterate contact information to a single common script?

A key issue that emerged early on in the Group's discussion was the agreement that their recommendation should bear in mind that the main purpose of transformed[2] data is to allow those not familiar with the original script of a contact information entry, to contact the registrant. This means that the accuracy of contact information data that are entered and displayed is paramount. There remains however some divergence in the Working Group about whether the need for accuracy is an argument in favour of transformation or not –

---

[1] See also: https://community.icann.org/display/tatcipdp/1+What+is+contact+information+and+What+Taxonomies+are+Available
[2] 'Transformed' is used throughout this Report, meaning 'translated and/or transliterated'; similarly 'transformation' is to mean 'translation and/or transliteration'.

and this is also reflected in the public comments received (see 'Community Input' below). At this stage, the Working Group has decided to summarise the discussion and put them to the community to gage support levels:

Working Group's arguments supporting mandatory transformation of contact information in all generic top-level domains

- Mandatory transformation of all contact information into a single script would allow for a transparent, accessible and, arguably, more easily searchable[3] database. Currently all data returned from the Whois database in generic top level domains (gTLDs) are provided in ASCII and such uniformity renders it a very useful global resource. Having a database with a potentially unlimited number of scripts/languages might create logistical problems in the long run.
- Transformation would to some extent facilitate communication among stakeholders not sharing the same language. Good communication inspires confidence in the Internet and makes bad practices more difficult. At this stage ASCII/English are the most ~~effective~~common script/language choices. However, it should be noted that already today many users of the Internet do not share English as a common language or the Latin script as a common script. The number of these users will grow substantially as internet access and use is continues to expand across countries/continents and so the dominant use of English might deter participation of those not confident in or familiar with it.
- For law enforcement purposes, when Whois results are compared and cross-referenced, it may be easier to ascertain whether the same registrant is the domain holder for different names if the contact information are transformed according to standards.

---

[3] p.11
The AGB defines "searchable" on p.113:
A Searchable Whois service: Whois service includes web-based search capabilities by domain name, registrant name, postal address, contact names, registrar IDs, and Internet Protocol addresses without arbitrary limit. Boolean search capabilities may be offered. The service shall include appropriate precautions to avoid abuse of this feature (e.g., limiting access to legitimate authorized users), and the application demonstrates compliance with any applicable privacy laws or policies.

Commented [CD1]: More examples in these sections are welcome.

Commented [CD2]: Added in response to feedback on 20 Nov. 2014.

Commented [CD3]: There is disagreement about whether it is easier to search data in one language.

Commented [MZ4]: Whois isn't exactly intended to be searchable (although some companies aggregate Whois data so that it can be searched). I wonder if they mean that it's more readily interpreted? "Searching" suggests you could query for a registrant and all the domain names with his name would be presented. "Querying" is more what people do with Whois when they enter a domain name and results appear. Perhaps it doesn't matter, but just thought I'd point that out.

Commented [MZ5]: Just wanted to confirm that we're certain this is the case? Maybe someone from tech services (Francisco, Gustavo, Ed, et al.) could verify? If it's true that all data is in US ASCII, it would seem that this is actually an argument against mandatory transformation (since it's already being done).

Commented [CD6R5]: Current rules says that in the current system they are in US ASCII, rather than transformed.

Commented [CD7]: Changed as a result of discussion on 20 Nov. 2014.

Formatted: Font: +Body (Calibri)

- Mandatory transformation would avoid possible flight by bad actors to the least translatable languages[4].

Working Group's arguments opposing mandatory transformation of contact information in all generic top-level domains

- Consistent transformation of contact information data across millions of entries is very difficult to achieve, especially because of the continued globalisation of the Internet with an increase in users ~~that are not familiar with~~whose languages are not based on the Latin script. A Domain Name Relay Daemon (DNRD) should display what the client enters. Original data should be authoritative, verified and validated. Interpretation and transformation may add errors.

- Mandatory transformation into one script could be problematic for or unfair to all those interested parties that do not speak/read/understand that one script. For example, whereas transformation from Mandarin script to a Latin script might be useful to, for example, law enforcement in countries that use Latin scripts, it would be ineffectual to law enforcement in other countries that do not read that Latin script.

- A growing number of registered name holders do not use Latin script, meaning that they would not be able to transform their contact information themselves. Therefore, transformation would have to take place at a later stage, through the registrar or the registry. Considering the number of domain names in all gTLDs this would lead to considerable costs not justified by benefits to others and be detrimental to accuracy[5] and consistency – key factors for collecting registered name holders' contact information data in the first place.

> **Commented [MZ8]:** I am not sure if this is what was intended. As more people use the Internet, more people will become accustomed to the Latin script. I think what is meant is that there will be an increase in the number of users whose languages are not based on the Latin script.
>
> **Commented [CD9R8]:** Wording changed accordingly. Both statements are true.

> **Commented [MZ10]:** There are several references to LEA that might be relevant more applicably. When I read this part of the paper, my general impression was that this was being driven by law enforcement only, so maybe the references could be broadened (unless this really is just a law enforcement issue).
>
> **Commented [CD11R10]:** Specific examples welcome.

---

[4] However, it should be noted that transformation tools may not exist for such languages and so transformation would need to be manual until they did. It would be difficult to limit languages to e.g. only the UN ones or some other subset.

[5] "Accuracy" as used in the "Study to Evaluate Available Solutions for the Submission and Display of Internationalized Contact Data" June 2, 2014:
"There are at least three kinds of use the transformed contact data in the DNRD may have in another language or script (based on the level of accuracy of the transformation):
1. Requiring accurate transformation (e.g. valid in a court of law, matching information in a passport, matching information in legal incorporation, etc.)
2. Requiring consistent transformation (allowing use of such information to match other information provided in another context, e.g. to match address information of a registrant on a Google map, etc.)

> **Formatted:** English (United Kingdom)

- The usability of transformed data is questionable because registered name holders unfamiliar with Latin script would not be able to communicate in Latin script, even if their contact information was transformed and thus accessible to those using Latin script.

- It would be more ~~efficient~~ convenient to allow registration information data to be entered by the registered domain holders in their local script and the relevant data fields to be transformed[6] into Latin script by either the registrar or the registry. This would provide greater accuracy than transformation and it would provide those wishing to contact name holders to identify their email and/or postal address. A similar method is already in place for some of the country code top level domains (ccTLDs):

```
Domain Name: example.cn
ROID: 20050505s10001s11652376-cn
Domain Status: clientDeleteProhibited
Domain Status: clientUpdateProhibited
Domain Status: clientTransferProhibited
Registrant ID: agent2899-0
Registrant: 中信安控科技股份有限公司
Registrant Contact Email: zxakkj@163.com
Sponsoring Registrar: Canada 001 names Ltd.
Name Server: dns8.66.cn
Name Server: dns9.66.cn
Registration Date: 2005-05-05 05:38:46
Expiration Date: 2022-05-05 05:38:46
DNSSEC: unsigned
LAHO-2819:~ lars.hoffmann$ whois test.cn
Domain Name: test.cn
ROID: 20030312s10001s00063170-cn
Domain Status: clientDeleteProhibited
Domain Status: clientUpdateProhibited
Domain Status: clientTransferProhibited
Registrant ID: xq317v49978fop
Registrant: 北京慧思德科技有限公司
Registrant Contact Email: info@wisdom.com.cn
Sponsoring Registrar: 北京新网数码信息技术有限公司
Name Server: ns15.xincache.com
Name Server: ns16.xincache.com
Registration Date: 2003-03-17 12:20:05
Expiration Date: 2020-03-17 12:48:36
DNSSEC: unsigned
```

- Accurate transformation is very expensive and these recommendations could effectively shift the costs from those requiring the work to registrars, registrants or other parties. Costs would make things disproportionately difficult for small players. Existing automated systems for transformation are inadequate. They do not provide results of sufficient quality for purposes requiring accuracy and cover fewer than

---

3. Requiring ad hoc transformation (allowing informal or casual version of the information in another language to provide more general accessibility)"
Both accuracy and consistency would suffer if large number of actors, for example, registrants, were transforming contact information.
[6] "Transformation" on its own is used to mean to refer to contact information, not fields, in this report.

Commented [ER12]: Not sure I understand this point in light of the bullet point above that says that transformation at a later stage by the registry or registrar would be detrimental to accuracy and consistency. How should these two points be read together?

Commented [ER13]: Why is this option not considered the same as transformation?

Commented [CD14R13]: Footnote added to clarify.

Commented [CD15]: Would this approach be practical for gTLDs?

Formatted: Font: Calibri, 10 pt, Complex Script Font: 10 pt

100 languages. Developing systems for languages not covered by transformation tools is slow and expensive, especially in the case of translation tools. For purposes for which accuracy is important, transformation work often needs to be done manually.[7] For example the translated 'Bangkok' is more useful internationally than the transliterated 'krung thep'. However, the transliterated 'beijing' is much more useful than the translated 'Northern Capital'. Automated systems would not be able to know when to translate and when to transliterate.

- Another consequence of the financial burden of transforming contact information data would be that the expansion of the Internet and provision of its benefits became more difficult, especially in less developprivileged regions that are already lagging behind in terms of internet access and often don't use Latin-based scripts.

- It would be near impossible to achieve high levels of accuracy in transforming a very large number of scripts and languages – mostly of proper nouns – into a common script and language. For some languages standards do not exist; for those where there are standards, there may be more than one, for example, for Mandarin, Pinyin and Wade Giles.

- Mandatory transformation would require validation of both the original and transformed contact information every time they change, a potentially costly duplication of effort. Responsibility for accuracy would rest on registrants who may not be qualified to check it.

Current state of discussion

The arguments for and against mandatory transformation have been discussed among Working Group members but there is no clear consensus at this stage. Therefore, the Working Group has decided to use this Initial Report to put forward alternative recommendations and encourage community feedback which will be reflected in the final report.

The various draft recommendations currently under consideration by the Working Group are as follows.

---

[7] See: *Study to evaluate available solutions for the submission and display of internationalized contact data* for further information https://www.icann.org/en/system/files/files/transform-dnrd-02jun14-en.pdf.

Commented [ER16]: less developed?

Commented [CD17R16]: Adopted.

Commented [ER18]: high levels of accuracy?

Commented [CD19R18]: Adopted. Added sentence about standards.

Commented [MZ20]: I think an additional argument could be added to say something like "There is not currently any evidence available that new translation and/or transliteration requirements will result in benefits comparable to the costs of implementation." This is a point registras have repeatedly made to the Board, so I'd be surprised if no one said it during the WG discussions...

Commented [CD21R20]: Point made on p.12 "not justified by benefits to others".

Commented [MZ22]: I wonder if it would be useful to add a sentence that says something like "if no consensus is reached, the status quo will be maintained."

Commented [CD23R22]: Further discussion required.

Commented [ER24]: If there is "consensus" among the public comments submitted how will the WG consider this "new" information?

Draft Recommendations Alternative #1

- The WG could recommends that it is desirable to make transformation of gTLD registration contact information mandatory.
- The WG could recommends that every current and future[8] gTLD registration data entry that is provided in non-Latin script be transformed into Latin-based ASCII.
- The WG could recommends that registrants are encouraged to submit their data in Latin script; however, if registrants are not able to provide their data in Latin script, it falls to the registrar to assure accurate transformation.
- The WG could recommend that further work be done to guide how transformation should be done. This may include, for example, the establishment of a knowledge base, the designation or creation of transliteration standards and a network of linguists covering all world languages to maintain and implement them.

Draft Recommendations Alternative #2

- The WG could recommends that it is not desirable to make transformation of contact information mandatory. Any parties requiring transformation are free to do it ad hoc outside the DNRD.
- The WG could recommends that any new Registration Directory Service (RDS) databases contemplated by ICANN should be capable of receiving input in the form of non-Latin script contact information. However, all data fields of such a new database should be tagged in ASCII to allow easy identification of what the different data entries represent and what language/script has been used by the registered name holder.
- The WG could recommends that registered name holders enter their contact information data in the language or script appropriate for the language that the registrar operates in region in which they are used.
- The WG could recommends that the registrar and registry assure that the data fields are consistent, that the entered contact information data are verified (in accordance with the Registrar Accreditation Agreement (RAA)) and that the data fields are correctly tagged to facilitate search functionstransformation if it is ever needed.

---

[8] A less burdensome alternative may be only future registrations with a retrospective conversion project for old data.

Commented [MZ25]: I reread this a couple times... I wonder if there's a way to say something more like "The working group could recommend" instead of "the working group recommends."

Commented [CD26R25]: Changed.

Commented [ER27]: Agree with this comment.

Commented [CD28]: "ICANN" changed to "DNRD" to clarify.

Commented [ER29]: Would there be a tag with information saying which script/language is used?

Commented [CD30R29]: Yes.

Commented [ER31]: What does the "it" refer to here? The domain name? Or where the registrant is located? Example, if I am an Indian-based company and I register a domain name to sell products that are primarily targeted at French-Canadians, which language should I enter my contact information?

Commented [CD32R31]: "Contact information data". "it is" changed to "they are" accordingly.

Commented [LH33]: I think the language that the registrar operates in. simply because this will mean that the market determines what languages are in use and it could avoid using 'minority' languages that are 'appropriate' for a region/country but are not in common use. The principle being: if the registrant can use/read the language the registrar operates in, then she should also be able to submit her information data in that same language/script.

Commented [CD34R33]: Adopted.

Commented [CD35]: Changed as a result of my suggestion on mailing list after 20 Nov.

Formatted: Font: +Body (Calibri), 10 pt, Complex Script Font: 10 pt, English (United Kingdom)

- The WG could recommends that if registrars wish to perform transformation of contact information, these data should be presented as additional fields (in addition to the local script provided by the registrant), to allow for maximum accuracy.
- The WG could recommend that the field names of the DNRD be translated into as many languages as possible.

Community Input

In accordance with the PDP Manual, the Working Group reached out to ICANN's Supporting Organizations and Advisory Committees, as well as to the GNSO Stakeholder Groups and Constituencies to gauge their input on the Charter questions. Community feedback is of particular importance to the work of this WG because of the binary nature of the over-arching charter question of whether or not to recommend mandatory transformation of contact information data. The call for input was sent out to the leadership of the SO/ACs and SG/Cs on 4 February 2014.[9] A reminder was sent out to all community groups on 3 March 2014 and the Working Group also encouraged community feedback at its presentation to the GNSO during the weekend session preceding ICANN 49 in Singapore and during its face-to-face meeting at the same event.

Overall, the Working Group received feedback from the GAC representatives of Thailand, China, and the European Commission (all representing communities that rely on non-Latin scripts)[10], the Intellectual Property Constituency (IPC), the At-Large Advisory Committee (ALAC), and the Non-Commercial Stakeholder Group (NCSG).[11] A summary of the contributions can be found in the SO/AC and SG/C outreach review tool and the full-length submissions are published on the WG's wiki page.

The Translation and Transliteration of Contact Information PDP Working Group reviewed and discussed the contributions received in great detail. As pointed out above, the binary nature of the charter questions meant that community feedback was particularly valued during the WG's efforts so far. Where relevant and appropriate, information and suggestions derived from the various contributions were considered and have been included in 'Deliberation and Recommendations' above.

---

[9] See Mailing list archive: http://forum.icann.org/lists/gnso-contactinfo-pdp-wg/
[10] Within the EU Greece and Bulgaria use Greek and Cyrillic scripts respectively.
[11] The Working Group also received a contribution from the International Federation of Intellectual Property Lawyers (FICPI). However, as this first call for community feedback is not a public comment but rather an outreach to SO/ACs and SG/C, the contribution was acknowledged but not given the same weight as other submissions. The Group noted, however, that FICPI is encouraged to contribute to the forthcoming public comment period and if they do not do so, the Group will consider its existing contribution more thoroughly at that point.

Background

Extract from the [Final Issue Report](#)

In April 2009 ICANN's Security and Stability Advisory Committee (SSAC) issued SAC 037, *Display and usage of Internationalized Registration Data: Support for characters from local languages or script.* In this document, the SSAC examined how the use of characters from local scripts affects the Internet user experience with respect to domain name registration data submission, usage, and display. The SSAC made three recommendations:

1. That ICANN's Board of Directors task the GNSO, Country Code Names Supporting Organization (ccNSO), and the SSAC to form a working group to study the feasibility and suitability of introducing display specifications or standards to deal with the internationalization of registration data.
2. That ICANN host a workshop on the internationalization of registration data during the next ICANN meeting (June 2009, Sydney).
3. That ICANN should consider the feasibility of having applications that query registration data services incorporate "standard" internationalization functionality.

ICANN's Board of Directors acted on Recommendation 1 by approving a resolution (2009.06.26.18) requesting that the GNSO and the SSAC, in consultation with staff, convene a working group to study the feasibility and suitability of introducing display specifications to deal with the internationalization of registration data.[12] Subsequently, the SSAC and the GNSO formed the IRD-WG to study the issues raised by the ICANN Board.

In November 2010 the IRD-WG produced an Interim Report requesting community input on several questions relating to possible models for internationalizing Domain Name

---

[12] See ICANN Board Resolutions, 26 June 2009, "Display and Usage of Internationalized Registration Data," http://www.icann.org/en/minutes/resolutions-26jun09.htm#6

Registration Data.[13] On 03 October 2011 the IRD-WG posted a draft Final Report for a 45-day public comment period.[14]

After considering the public comments received, on 07 May 2012, the IRD-WG submitted a Final Report to the GNSO Council and the SSAC for consideration.[15]

The SSAC approved the Final Report in May 2012. At its meeting on 27 June 2012 (in Prague) the GNSO Council passed a motion by which it approved the delivery of the Final Report to the Board.[16] In its motion, the Council also agreed to review the recommendations in the Final Report and to provide to the Board its advice with regard to those recommendations that may have policy implications.

At its meeting on 17 October 2012, the GNSO Council approved a motion accepting the IRD-WG Final Report recommendations.[17] The motion included the following clauses that resulted in the development of this Final Issue Report:

> "WHEREAS the GNSO Council has reviewed the Final Report and considers that while expecting the ICANN Board to respond to the SSAC-GNSO joint letter, the Recommendation 2, translation and transliteration of contact information of IRD, of the Final Report requires timely action at the policy level which involves collaboration among domain name registrant, registrar, and registry.
>
> "RESOLVED, the GNSO approves the Final Report and requests the ICANN Staff to prepare the IRD Issues Report on translation and transliteration of contact information (IRDIR-Rec2). The Issue Report should consider 1) whether it is desirable to translate contact information to a single common language or transliterate contact information to a single common script; 2) who should bear

---

[13] See Interim Report of the Internationalized Registration Data Working Group at: http://gnso.icann.org/issues/ird/ird-wg-final-report-15nov10-en.pdf.
[14] See Draft Final Report of the Internationalized Registration Data Working Group at: http://gnso.icann.org/issues/ird/ird-draft-final-report-03oct11-en.pdf.
[15] See Final Report of the Internationalized Registration Data Working Group at: http://gnso.icann.org/en/issues/ird/final-report-ird-wg-07may12-en.pdf.
[16] See https://community.icann.org/display/gnsocouncilmeetings/Motions+27+June+2012.
[17] See https://community.icann.org/display/gnsocouncilmeetings/Motions+17+October+2012.

the burden and who is in the best position to address these issues; and 3) whether to start a policy development process (PDP) to address those questions."

As noted above, the 'contact information' references in this Final Issue Report 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 the 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 know as WHOIS service). This protocol is a client-server, query-response protocol. The RAA (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 defined Domain Name Registration Data as information that registrants provide when registering a domain name and that registrars or registries collect. The RAA (RAA 3.3.1) specifies the data elements that must be provided by registrars (via Port 43 and via web-based services, such as WHOIS) in response to a query. (For ccTLDs, the operators of these TLDs set policies for the request and display of registration information.)

As the SSAC noted in SAC051 *SSAC Report on WHOIS Terminology and Structure*, "The term "WHOIS" is overloaded, referring to protocols, services, and data types associated with Internet naming and numbering resources, i.e., domain names, Internet Protocol (IP) addresses, and Autonomous System Numbers (ASNs)."[18] The Report further notes that WHOIS can refer to any of the following:

1. The information that is collected at the time of registration of a domain name or IP numbering resource and subsequently made available via the WHOIS Service, and potentially updated throughout the life of the resource;

---

[18] See SAC051: SSAC Report on WHOIS Terminology and Structure at http://www.icann.org/en/groups/ssac/documents/sac--051-en.pdf.

2. The WHOIS Protocol itself, which is defined in RFC 3912 (which obsoletes RFCs 812 and 954); or

3. The WHOIS Services that provide public access to domain name registration information typically via applications that implement the WHOIS protocol or a web-based interface.

The SSAC recommended in its report that the terms Domain Name Registration Data Directory Service (rather than WHOIS) should be used when referring to the service(s) offered by registries and registrars to provide access to (potentially a subset of) the Domain Name Registration Data.

To 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 the IRD-WG members discussed is whether a Domain Name Registration Data Directory Service, such as the WHOIS, should support multiple representations of the same registration data in different languages or scripts.

The IRD-WG noted that much of the currently accessible domain registration data are encoded in US-American Standard Code for Information Interchange (US-ASCII). US-ASCII is a character-encoding scheme originally based on the Latin script. This legacy condition is convenient for WHOIS service users who are sufficiently familiar with languages that can be displayed in US-ASCII.

However, US-ASCII data are less useful to the community of Domain Name Registration Data Directory Service users who are only familiar with languages that require character set support other than US-ASCII. It is important to note that this community is likely to continue growing. Thus accommodating the submission and display of internationalized registration data is seen as an important evolutionary step for Domain Name Registration Data Directory Services such as the WHOIS.

Commented [MZ36]: I am not sure if this is what was intended. As more people use the Internet, more people will become accustomed to the Latin script. I think what is meant is that there will be an increase in the number of users whose languages are not based on the Latin script.

In general, the IRD-WG recognized that internationalized contact data can be translated or transliterated into the "must be present" representation. By "must be present" the IRD-WG meant that contact data must be made available in a common script or language. 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. 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 considered five models to address the translation and transliteration of domain name registration data contact information, but it was unable to reach consensus on a single model.[19] However, it recognized that the translation and transliteration of contact information had policy implications, and thus its Final Report contained the following recommendation:

Recommendation 2: The GNSO council and the SSAC should request a common Issue Report on translation and transliteration of contact information. The Issue Report should consider whether it is desirable to translate contact information to a single common language or transliterate contact information to a single common script. It should also consider who should bear the burden and who is in the best position to address these issues. The Issue Report should consider policy questions raised in this document and should also recommend whether to start a policy development process (PDP).

> **Commented [ER37]:** Is this a direct quote? Consider indenting or putting this in quotes, if so.

The Affirmation of Commitments signed on 30 September 2009 between ICANN and the US Department of Commerce contains specific provisions for periodic review of four key ICANN objectives, including WHOIS Policy.[20] The WHOIS Policy Review Team completed its review and published its Final Report on 11 May 2012.[21] In its Final Report the Review Team echoed the IRD-WG by calling for a Working Group to be formed (Recommendations 12 and 13) to develop internationalized domain name registration requirements that would include a data model that would address, "(any) requirements for the translation or transliteration of the registration data." In addition, the SSAC further emphasized

---

[19] See Annex A: Different Models Proposed in the Internationalized Registration Data Working Group Final
[20] See Affirmation of Commitments at
http://www.icann.org/en/about/agreements/aoc/affirmation-of-commitments-30sep09-en.htm.
[21] See WHOIS Policy Review Team Final Report at: http://www.icann.org/en/about/aoc-review/whois/final-report-11may12-en.pdf

the IRD-WG's recommendation in SAC055: *WHOIS: Blind Men and an Elephant (SSAC Comment on the WHOIS Policy Review Team Final Report).*[22] In the Report the SSAC agreed with the recommendations of the Review Team on translation/transliteration of registration data and called on the ICANN Board of Directors to adopt Recommendation 2 in the IRD-WG's Final Report. The SSAC also stated that the ICANN Board should pass a resolution clearly stating the criticality of the development of a registration data policy defining the purpose of domain name registration data.

On 08 November 2012 the ICANN Board of Directors adopted several resolutions (2012.11.08.01 - 2012.11.08.02) relating to WHOIS, in response to the recommendations it received from the WHOIS Policy Review Team and the SSAC described above.[23] In particular, the Board directed the CEO to:

> launch a new effort to redefine the purpose of collecting, maintaining and providing access to gTLD registration data, and consider safeguards for protecting data, as a foundation for new gTLD policy and contractual negotiations, as appropriate (as detailed in the 1 November 2012 Board paper entitled, "Action Plan to Address WHOIS Policy Review Team Report Recommendations" – ICANN Board Submission Number 2012-11-01), and hereby directs preparation of an Issue Report on the purpose of collecting and maintaining gTLD registration data, and on solutions to improve accuracy and access to gTLD registration data, as part of a Board-initiated GNSO policy development process;[24]

The Board's Action Plan envisions the possibility of a PDP on the issue of translation and transliteration of contact information as follows: The Board directs the CEO to have Staff: 1) task a working group to determine the

---

[22] See SAC055: Blind Men and an Elephant (SSAC Comment on the WHOIS Policy Review Team Final Report) at http://www.icann.org/en/groups/ssac/documents/sac-055-en.pdf.
[23] See http://www.icann.org/en/groups/board/documents/resolutions-08nov12-en.htm#1.a
[24] See the Action Plan to Address WHOIS Policy Review Team Report Recommendations at: http://www.icann.org/en/groups/board/documents/briefing-materials-1-08nov12-en.pdf.

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 Web-based Extensible Internet Registration Data Working Group.

The Action Plan further tasks the CEO to create an Expert Working Group on gTLD Directory Services to: create material to launch GNSO policy work and inform contractual negotiations, as appropriate. Working group output is expected within 90 days and will ideally include a straw-man model for managing gTLD registration data. The working group's output form the basis for an Issues Report to accompany Board-initiated, expedited GNSO policy work that is expected to result in consensus policy that, at a minimum, addresses the purpose of collecting, maintaining and making available gTLD registration data, and related accuracy, data protection, and access issues. On 13 December 2013 the ICANN CEO announced the formation of the Expert Working Group. On 14 February 2013 ICANN announced the selection of the members of the Expert Working Group on gTLD Directory Services.[25]

---

[25] See the EWG homepage for all information, including membership, Initial Report, Status Report, and Final Report: https://community.icann.org/x/VQZlAg.