# Translation and Transliteration of Contact Information

## Implementation Kick-Off Meeting

Brian Aitchison | 19 July 2016

# Agenda

- Personal Introductions

- Introduction to Translation and Transliteration project

- Review of Working Group Recommendations

- T/T Scenarios and Examples

- T/T Requirements and Deliverables

- T/T Timeline

# T/T Introduction

**Goal of T/T PDP**

To determine how to best facilitate the entry of contact information into domain name registration data and directory services by non-English speakers and users of non-ASCII scripts.

**PDP Guiding Questions**

Is it desirable to translate [WHOIS] contact information to a single common language or transliterate contact information to a single common script?

Who should bear the burden of translating contact information to a single common language or transliterating contact information to a single common script?

**Terminology**

<u>Translation</u> – conversion of word/phrase from one language to another; meaning/semantic level function e.g. "lawyer" in English to "abogado" in Spanish

<u>Transliteration</u> – conversion of text from one script to another e.g. "адвокат" in Cyrillic to "advokat" in Latin script

<u>Transformation</u> – all of the above

# T/T Recommendation 1

It is not desirable to make transformation of contact information mandatory.

Any parties requiring transformation are free to do so on an ad hoc basis outside WHOIS or any replacement system, such as the Registration Data Access Protocol (RDAP).

If not undertaken voluntarily by registrar/registry (see Recommendation #5), the burden of transformation lies with the requesting party.

# T/T Recommendation 2

Whilst noting that a WHOIS replacement system should be capable of receiving input in the form of non-ASCII script contact information…

… **its data fields should be stored and displayed in a way that allows for easy identification of what the different data entries represent and what language(s)/script(s) have been used by the registered name holder.**

# T/T Recommendation 3

The language(s) and script(s) supported for registrants to submit their contact information data may be chosen in accordance with gTLD-provider business models.

Regardless of the language(s)/script(s) used, it is assured that the data fields are consistent to standards in the Registrar Accreditation Agreement (RAA), relevant Consensus Policy, Additional WHOIS Information Policy (AWIP) and any other applicable polices. Entered contact information data are validated, in accordance with the aforementioned Policies and Agreements and the language/script used must be easily identifiable.

# T/T Recommendation 5

If the transformation of contact information is performed, and if the WHOIS replacement system is capable of displaying more than one data set per registered name holder entry, **these data should be presented as:**

**[1] additional fields (in addition to the authoritative local script fields provided by the registrant) and that these fields be**

**[2] marked as transformed and**

**[3] their source(s) indicated**

# T/T Recommendation 6

Any WHOIS replacement system, for example RDAP, should remain flexible so that contact information in new scripts/languages can be added and expand its linguistic/script capacity for receiving, storing and displaying contact information data.

# T/T Recommendation 7

These recommendations should be coordinated with other WHOIS modifications where necessary and are implemented and/or applied as soon as a WHOIS replacement system that can receive, store and display non-ASCII characters, becomes operational.

# T/T Example Scenario per recommendations

Example of localized WHOIS output (thanks Theo!):
https://whois.icann.org/en/lookup?name=translationtransliterationirt.com

**Base Scenario**

Russian registrant does business in China and buys domain from Chinese registrar.

Russian registrant enters WHOIS contact information into RDAP in Cyrillic script.

# T/T Scenario per recommendations

**Sub-Scenario 1**

Chinese registrar targets and accommodates Russian customers as part of their business model.  (See Rec. 3)

The decision to translate or transliterate is left to the discretion of the registrar. Chinese registrar voluntarily undertakes transformation of Cyrillic entry into local Chinese script. (See Rec. 1)

The authoritative local script—in this case Cyrillic—is presented in WHOIS output, along with (see Recs 2 and 5):

- Additional data fields to show WHOIS outputs transformed into Chinese script
- A tag indicating that a transformation was performed.
- A tag indicating which entity performed the transformation, which in this case is our Chinese registrar.
- A "language/script" tag indicating which language/script is present in initial and transformed data fields.

WHOIS data is validated for correctness of format in accordance with relevant policies and standards (see Rec. 4)

# T/T Scenario per recommendations

**Sub-Scenario 2**

Chinese registrar chooses not to perform transformation of registrant's Cyrillic WHOIS contact information (see Rec. 1).

Chinese law enforcement identifies criminal activity associated with the Russian registrant's domain and requires transformation.

Chinese law enforcement undertakes transformation of Cyrillic WHOIS contact information into Chinese language/script (see Rec 1.).

No action is required on the part of the registrar.

# T/T Requirements and Deliverables

**Bottom line**
No party is required to T/T, but RDAP must accommodate those who wish to do so.

**Summary Implementation Plan**

- Add one RDAP extension to provide additional data fields for transformed contact information. (**Recommendations 2 and 5**)
- Add one Extensible Provisioning Protocol (EPP extension) that includes tags for registration data that shows (**Recommendations 2 and 5**):
    - That a transformation was performed
    - The source of the transformation
    - The languages and scripts present in the initial and transformed fields

# Timeline

| | |
|---|---|
| **IRT Call #1: Introduction, discuss implementation plan and potential new policy language, schedule next calls** | **19 July 2016** |
| Share revised implementation plan with IRT by email (by staff) | 22 July 2016 |
| Share first draft of new consensus policy language as discussed with IRT and ICANN Legal and Compliance staff (by staff) | 29 July 2016 |
| **IRT Call #2: Discuss draft new consensus policy language and technical updates to RDAP profile** | **2 August 2016** |
| Share draft new consensus policy language/red-line document with IRT (by staff) | 9 August 2016 |
| **IRT Call #3: Discuss updated new consensus policy language with aim to publish for public comment** | **16 August 2016** |
| Share preliminary new consensus policy language for public comment for final revisions by IRT and staff (by staff) | 19 August 2016 |
| Publish new consensus policy language for public comment | 29 August 2016 |
| **IRT Call #4: Discuss public comments as necessary** | **18 October 2016** |
| Share revised policy language incorporating public comments as applicable with IRT (by staff) | 28 October 2016 |
| **IRT Call #5 (if required): Finalize all policy language** | **1 November 2016** |
| Make final revisions to policy language as necessary, sharing updates with IRT via email (by staff) and scheduling any additional calls as required | 1 November 2016 – 1 December 2016 |
| Announce policy effective date of 1 August 2017 | 1 December 2016 |

Next steps:
1. Draft and share new policy language for T/T implementation
2. IRT: Review implementation plan document and annotate as necessary

# Thank You!

## Questions?

Reach me at:
Email: brian.aitchison@icann.org
Website: icann.org

ICANN

twitter.com/icann

gplus.to/icann

facebook.com/icannorg

weibo.com/ICANNorg

linkedin.com/company/icann

flickr.com/photos/icann

youtube.com/user/icannnews

slideshare.net/icannpresentations