



ICANN | 54

Dublin



18-22 OCTOBER 2015





WHOIS Review Team Internationalized Registration Data Expert Working Group

Margie Milam | ICANN 54 | 18 October 2015

- Welcome – Margie Milam, ICANN
- Final Report from the Expert Working Group on Internationalized Registration (WHOIS)
 - James Galvin, Afilias and Jody Kolker, GoDaddy
- GNSO Final Report on the Translation and Transliteration of Contact Information Policy Development Process
 - Rudi Vansnick and Chris Dillon, Co-Chairs GNSO WG
- RDAP: Enabling Internationalized Registration Data
 - Francisco Arias, ICANN
- Panel Discussion: Next Steps for the IRD work:
James Galvin, Jody Kolker, Chris Dillon, Rudi Vansnick,
Margie Milam and Francisco Arias

Final Report from the Expert Working Group on Internationalized Registration (WHOIS)

James Galvin, Afilias

Jody Kolker, GoDaddy

Background Expert Working Group

The Whois Review Team [Internationalized Registration Data Expert Working Group](#) was chartered to determine the requirements for internationalized registration data, and produce a data model that matches the requirement.



Formed as part of the effort to implement WHOIS review team recommendations 12-13



Approach: Group data by categories, Separate internationalization vs. localization, Articulate a set of principles to guide discussion of requirements



Recognized on-going efforts in other areas (e.g. GNSO PDP Translation and Transliteration, Directory Services EWG, IETF WEIRDS working group)

WHOIS Policy Review Recommendation

Determine appropriate Internationalized Domain Name Registration (IRD) data requirements.

- Submission
- Directory Services
- Storage
- Transmission

Principles of Internationalization

Specific principles identified to guide the development of internationalization registration data:

- User Capability Principle
- Simplicity and Reusability Principle
- Extensibility

Two High Level Requirements

IRD Working Group proposed two high level requirements for community consideration:

- **Registrants should only be required** to input registration data in a language(s) or script(s) with which they are most familiar under ordinary circumstances
- 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
 - "tagging" is expressly intended to reflect a requirement that it be possible to know with deterministic certainty the language(s) and script(s) used by the data; it is not prescriptive of a solution

IRD Technical Considerations Identified

Lack of Internationalized Support in Technical Protocols

- EPP (Extensible Provisioning Protocol) Issues
 - Lacking language and script attribute
 - Lacking conversion-mechanism attribute

WHOIS Issue

- RFC3912. WHOIS Protocol Specification is not capable of handling “UTF-8” characters consistently, as it has “no mechanism for indicating the character set in use”

IRD Technical Considerations con't

Encoding of data requires "standard" languages and scripts

- Necessary to support "tagging"
- Registry/registrar changes to "store tags"

Workflow changes are required at registrars

- Potential interactions with registrants
- Postal address requirements

Internationalized email addresses

- Lack of adoption
- Operationally not backward compatible

Localization vs. Internationalization

Localization: the adaptation of registration data to meet the language, cultural, regional and other requirements of a target data consumer group:

- ❑ Numeric, date and time formats complies with local usage patterns
- ❑ Localized label for data elements
- ❑ Localized data (names, addresses)

Internationalization: the adaptation of registration data to *enable* easy localization for target audiences that vary in language, culture or region.

- ❑ Ensure data elements represented and transmitted in standard forms
- ❑ Ensure that data is appropriately encapsulated and tagged to allow localization



Proposed IRD Data Model

- The **domain object** corresponds to a single Registered Name.
- The **contact object** corresponds to a single contact (registrant, administrative, technical and billing are roles of a contact with respect to given domain name).
- The **registrar object** corresponds to a single registrar.
- A **nameserver object** corresponds to a single registered nameserver.

Internationalization

The display of registration data entails the following:

- Designing and developing in a way that removes barriers to localization.
- Providing support for features that may not be used until localization occurs.
- Enabling code to support local, regional, language, or culturally related preferences.

High Level Requirements Adopted

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

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

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.

12 Data Categories Identified

Developed 12 data categories that cover all of the known data elements:

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

Example of WHOIS Output

Localized for Japanese audience

ドメイン情報:
[ドメイン名] ドメイン名例.JP
[ドメイン名] XN--ECKWD4C7CU47R2WF.JP
[登録者名] エグザンプル株式会社
[ネームサーバ] ns01.example.co.jp
[ネームサーバ] ns02.example.co.jp
[登録年月日] 2001/08/09
[有効期限] 2008/08/31
[状態] Active
[最終更新] 2007/09/01 01:05:05 (JST)
公開連絡窓口:
[名前] 日本 太郎
[電子メールアドレス] taro@example.jp
[郵便番号] 101-0065
[住所] 東京都千代田区西神田三丁目 8 番 1 号
 千代田ファーストビル東館 13F
[電話番号] 03-5215-8451
[FAX 番号] 03-5215-8452

Localized for English-speaking audience

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

Technical Challenges for Current System

- Registrars need to be able to detect, validate and verify the script and language in use. This functionality does not exist in the current registrar workflow.
- Changes and harmonizing of data models is needed in Registry Agreements, Registrar Accreditation Agreement, WHOIS advisory, AWIP, and the Thick WHOIS Policy Recommendation.
- GNSO PDP on Translation/Transliteration of contact data policy implications for IRD need to be addressed, including significant adoption of Registration Data Access Protocol (RDAP).

IRD Recommended Next Steps

- Implementation dependent on alignment with GNSO's PDP on Translation/Transliteration of contact data.
- Need appropriate follow-up to review the broader policy implications of the Report as it relates to other GNSO policy development work on Whois issues.
- Requirements should not apply until significant uptake in the adoption of Registration Data Access Protocol (RDAP).
- A transition plan for the registry and registrar adoption of internationalized email address should be identified.
- Data models should be harmonized with current Registry Agreements, Registrar Accreditation Agreement, Whois advisory, AWIP, and the Thick Whois Policy Recommendations.



Image credit: www.dkit.ie

- [GNSO Final Report on the Translation and Transliteration of Contact Information Policy Development Process](#)
- [Final Report](#) from the Expert Working Group on Internationalized Registration (WHOIS)
- [IETF Web-extensible Internet Registration Data](#) (WEIRDS) working group registration data access protocol (RDAP) RFC 7480 to 7485

Requirements for contact data elements

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

Requirements for other data elements

Data Category	Example Data Elements	Requirement
Status	Domain Status	The text value of the domain status should conform to the EPP specification defined in RFC 5731 section 2.3.
Phone and Fax Numbers	Technical Contact Facsimile Number, Technical Contact Phone Number	The phone and fax numbers should comply with the RFC 5733.
Email addresses	Technical Contact Email, Registrant Email, Administrative Contact Email	Email addresses should comply with RFC 5322 and its extension in RFC 6532 (see section 3.2) for internationalized email addresses.
Identifiers	Registrar ID, Registrant ID, Sponsoring Registrar IANA ID, Domain ID	Registrar IANA IDs are assigned by IANA. Other identifiers should conform to format set forth in section 2.8 of RFC 5730.
DNSSEC Information	DS Key Tag 2, DS Key Tag 1, Digest Type 1, DS Maximum Signature Life 2, Algorithm 2, Digest Type 2, Algorithm 1, Digest 2, DS Maximum Signature Life 1, Digest 1	Elements should conform to format and values described in RFC 5910.
URLs	Referral URL, Registrar URL (registration services)	Elements should conform to standards set forth in RFC 3986 and RFC 3987.
Domain Names	Domain Name, Whois Server, Name Server	Where registrant provides a domain name, registrants to provide domain name in either U-label (preferred) or A-label format [RFC5890] during the submission. For display, require both U-label and the corresponding A-label at all times.
Time and Dates	Last Transferred Date, Domain Last Updated Date, DS creation date, Domain Expiration Date	Date and time elements should conform to formats specified in RFC3339, and represented in UTC with no offset from the zero meridian.

DNRD-DS Proposed Model for the Domain Object

Data Element	Format	Min length	Max length	Cardinality	Language Tag (RFC 5646)
Domain Name (Internationalized)	RFC 5890	1	255	1	Required if it is U-label or A-label.
Domain ID	Freeform text	1	255	1	n/a
Referral URL	RFC 3986 / 3987	1		1	n/a
Updated Date	RFC 3339		32	{0,1}	n/a
Creation Date	RFC 3339		32	1	n/a
Registry Expiry Date	RFC 3339		32	1	n/a
Sponsoring Registrar IANA ID	Registrar ID registry	1	255	1	n/a
Domain Status	RFC 5731		32	{1,11}	n/a
Registrant ID	Freeform text	1	255	1	n/a
Admin ID	Freeform text	1	255	1	n/a
Tech ID	Freeform text	1	255	1	n/a
Billing ID	Freeform text	1	255	1	n/a
DS created	RFC 3339		32	{0,1}	n/a
DS Key Tag	RFC 4034, 5910			{0,2}	n/a
Algorithm	RFC 4034, 5910			{0,2}	n/a
Digest Type	RFC 4034, 5910			{0,2}	n/a
Digest Value	RFC 4034, 5910			{0,2}	n/a
DS Maximum Signature Life	RFC 4034, 5910			{0,2}	n/a

DNRD-DS Proposed Model for the Nameserver Object

Data Element	Format	Min length	Max length	Cardinality	Language Tag (RFC 5646)
Nameserver ID	Freeform text	1	255	1	n/a
Host Name	RFC 5890 (both A-label and U-label)	1	255	1	n/a
IP Address	RFC 0791/RFC 5952			{0, ..}	n/a
Registrar ID	Freeform text	1	255	1	n/a
Referral URL	RFC 3986 / 3987	1		1	n/a
Creation Date	RFC 3339		32	1	n/a
Last Updated Date	RFC 3339		32	{0,1}	n/a
WHOIS Server	RFC 5890 (both A-label and U-label)	1	255	1	n/a

DNRD-DS Proposed Model for the Contact Object

Data Element	Format	Min length	Max length	Cardinality	Language Tag (RFC 5646)
Contact ID	Freeform text	1	255	1	n/a
Registrar ID	Freeform text	1	255	1	n/a
Contact Name	Freeform text	1	255	{0,1}	required
Contact Organization	Freeform text	1	255	{0,1}	required
Contact street	Freeform text in a language or script appropriate for its region.	1	255	{1,3}	required
Contact City	Freeform text in a language or script appropriate for its region.	1	255	1	required
Contact State / Province	Freeform text in a language or script appropriate for its region.	1	255	{0,1}	required
Contact country / Territory	ISO 3166 part 2 code list	2	3	1	n/a
Contact Postal Code	Freeform text	1	255	{0,1}	n/a
Contact Phone	RFC 5733		64	1	n/a
Contact Phone Ext	RFC 5733		64	{0,1}	n/a
Contact Fax	RFC 5733		64	{0,1}	n/a
Contact Fax Ext	RFC 5733		64	{0,1}	n/a
Contact Email	RFC 5322 / 6532		255	1	n/a

DNRD-DS Proposed Model for the Registrar Object

Data Element	Format	Min length	Max length	Cardinality	Language Tag (RFC 5646)
Registrar ID	Freeform text	1	255	1	n/a
Contact Organization	Freeform text. Name of the registrar should be the official name in the RAA with ICANN, in whichever language(s) or script(s).	1	255	{0,1}	required
Contact street	Freeform text in a language or script appropriate for its region.	1	255	{1,3}	required
Contact City	Freeform text in a language or script appropriate for its region.	1	255	1	required
Contact State / Province	Freeform text in a language or script appropriate for its region.	1	255	1	required
Contact country / Territory	ISO 3166-1 alpha-2	2	2	1	n/a
Contact Postal Code	Freeform text	1	255	1	n/a
Contact Phone	RFC 5733		64	1	n/a
Contact Fax	RFC 5733		64	1	n/a
Contact Email	RFC 5322 / 6532		255	1	n/a
Registrar Admin Contact ID	Freeform text	1	255	1	n/a
Registrar Technical Contact ID	Freeform text	1	255	1	n/a
Registrar URL	RFC 3986 / 3987	1		{0,1}	n/a



Translation and Transliteration of Contact Information PDP: Final Report

Chris Dillon/Rudi Vansnick, Working Group Co-Chairs

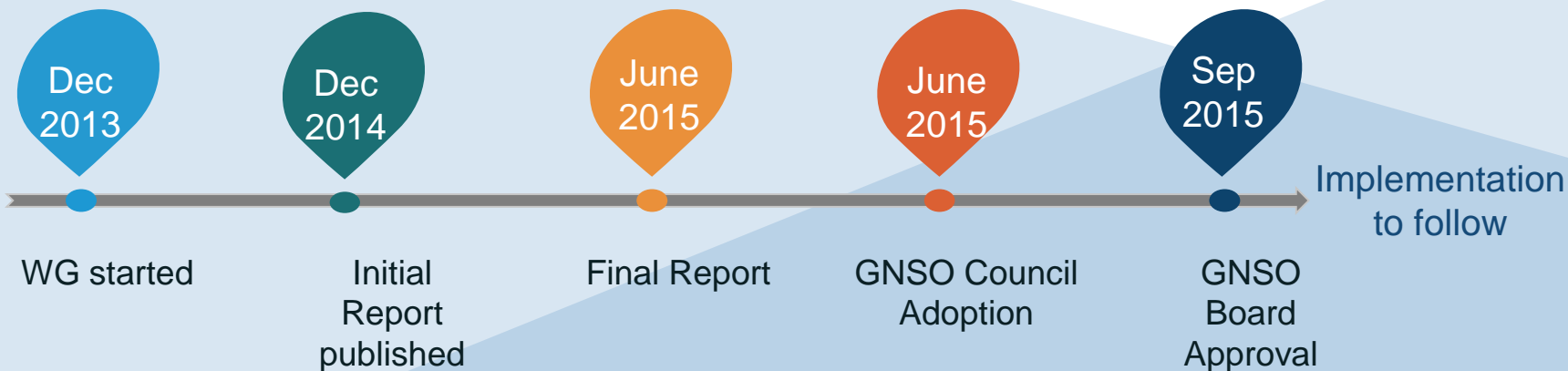
Charter Questions and Timetable

Two Charter Questions

1. Whether it is desirable to translate or transliterate contact information into a single common language or script?
2. Who should decide who should bear the burden of transforming* contact information to a single language or script?

* The WG has uses the short form 'transformation' throughout this presentation to replace the term 'translation or transliteration'.

Timeline



Issues with Mandatory Transformation

(as identified by the Working Group)

- ⦿ It would be near impossible to achieve consistent accuracy in transforming addresses (proper nouns) into a common script.
- ⦿ Manual translation is very expensive - ICANN language services pay a minimum of \$25 per translation (each new verification would have to be transformed) and accuracy and consistency remain highly challenging.
- ⦿ The WG was not convinced that transformation is 'a regular cost of doing business', due to the small number of times transformed data may be called upon, compared to the quantity of WHOIS registration datasets submitted.
- ⦿ Usability of transformed data is questionable because registered name holders unfamiliar with Latin script would not be able to communicate in Latin script.
- ⦿ Biased towards one language/script

What non-mandatory Transformation means

(as identified by the Working Group)

- ⦿ Submitted data are likely to be as consistent and reliable as possible.
- ⦿ The more consistent the data, the better searchable is a database.
- ⦿ Equal costs/opportunities for registrars and registrants regardless of their linguistic/script background.
- ⦿ Language and Script should be easily identifiable to facilitate such third-party transformation if/when necessary.
- ⦿ Consumers of contact information – those requesting data – carry the burden of transformation.

Substantive Recommendations

1

Recommendation 1

It is not desirable to make transformation of contact information mandatory. Burden of voluntary transformation lies with requestor of information.

2

Recommendation 2

Data fields are stored and displayed in a way that allows for easy identification of what individual data entries represent and what languages/scripts have been used.

3

Recommendation 3

Language(s) and script(s) supported for registrants to submit their contact information data may be chosen in accordance with gTLD-provider business models (as they need to verify).

4

Recommendation 4

Regardless of language used, data fields must be consistent, entered data must be verified, script/language used must be easily identifiable

5

Recommendation 5

If Whois replacement system is capable of multiple data set per entry, and if voluntary transformation is performed,, transformed data should be marked as such and presented as additional fields.

6

Recommendation 6

Any Whois replacement system, e.g. RDAP, must allow for new scripts/languages to be added and expand its linguistic/script capacity.

Minority View and Consensus

- ⦿ Recommendations 2-6 received **full consensus**
- ⦿ Recommendation 1 received **consensus**
- ⦿ One WG member was not able to support recommendation 1 and supplied a minority statement with regard to recommendation 1:

“Working Group member Petter Rindforth, in line with the position taken by his Constituency, the Intellectual Property Constituency (ICP), recommends mandatory translation and/or transliteration (transformation) of contact information in all generic top-level domains (gTLDs). [...] There are a number of situations where a global WHOIS search, providing access to data in as uniform a fashion as possible, is necessary for the data registration service to achieve its goals of providing transparency and accountability in the DNS.”

More Information

- ⦿ Final Report: <http://gnso.icann.org/en/group-activities/active/transliteration-contact>
- ⦿ Initial Report: <http://gnso.icann.org/en/issues/gtlds/transliteration-contact-initial-15dec14-en.pdf>
- ⦿ Redline Final Report from Initial Report
<https://community.icann.org/download/attachments/41890837/Final%20Report%20RedLine.pdf>
- ⦿ Public Comment in Initial Report: <https://www.icann.org/public-comments/transliteration-contact-initial-2014-12-16-en>
- ⦿ Webinar on Initial Report: <https://icann.adobeconnect.com/p2lzjk3zy0f/>
- ⦿ Wiki Page: <https://community.icann.org/x/FTR-Ag>

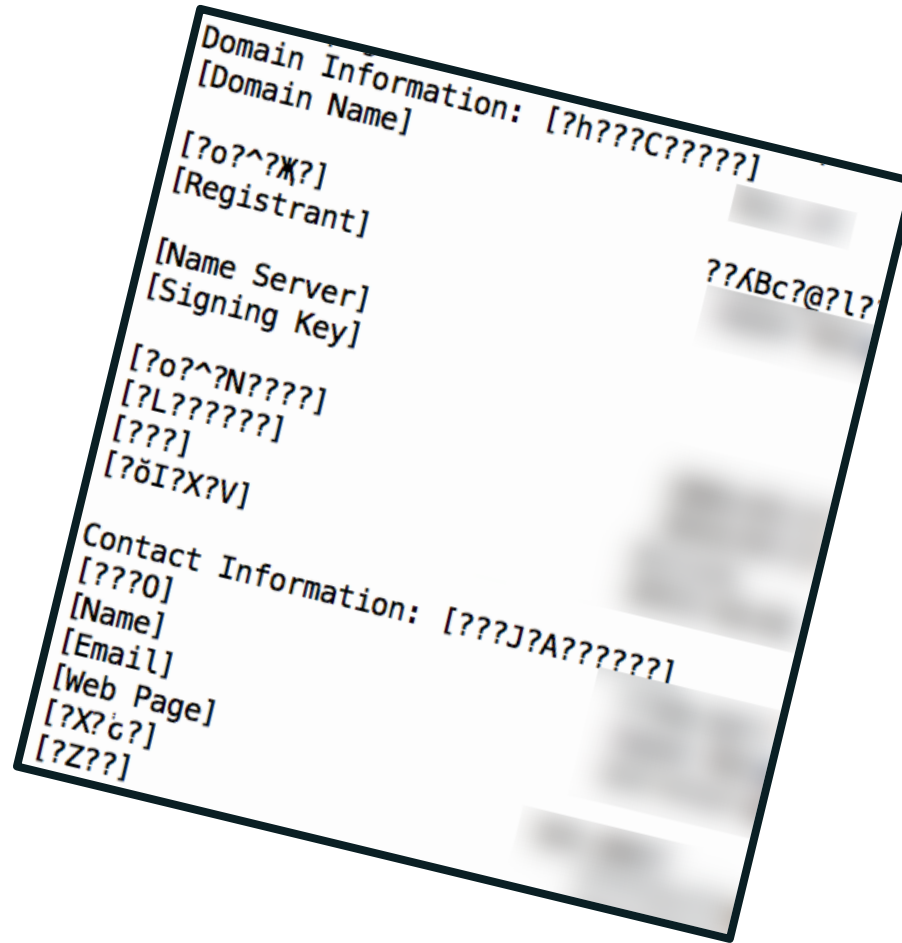


RDAP: Enabling Internationalized Registration Data

Francisco Arias, ICANN

Why WHOIS (port-43) should be replaced?

- ⦿ Not internationalized



Why WHOIS (port-43) should be replaced?

- ⦿ Non standardized format

```
ACE: [redacted] (RFC-3490, RFC-3491, RFC-3492)
Contacto Administrativo (Administrative Contact):
Nombre : [redacted]
Organizaci?n: [redacted]
Contacto T?cnico (Technical Contact):
Nombre : [redacted]
Organizaci?n: [redacted]
Servidores de nombre : [redacted]
```

```
domain: [redacted]
owner: [redacted]
ownerid: [redacted]
responsible: [redacted]
country: [redacted]
owner-c: [redacted]
admin-c: [redacted]
tech-c: [redacted]
billing-c: [redacted]
nserver: [redacted]
nsstat: [redacted]
nslastaa: [redacted]
nserver: [redacted]
nsstat: [redacted]
nslastaa: [redacted]
nserver: [redacted]
```

```
Domain Name: [redacted]
Created On: [redacted]
Expiration Date: [redacted]
Last Updated On: [redacted]
Registrar: [redacted]
URL: [redacted]
Registrant:
Name: [redacted]
City: [redacted]
State: [redacted]
Country: [redacted]
```

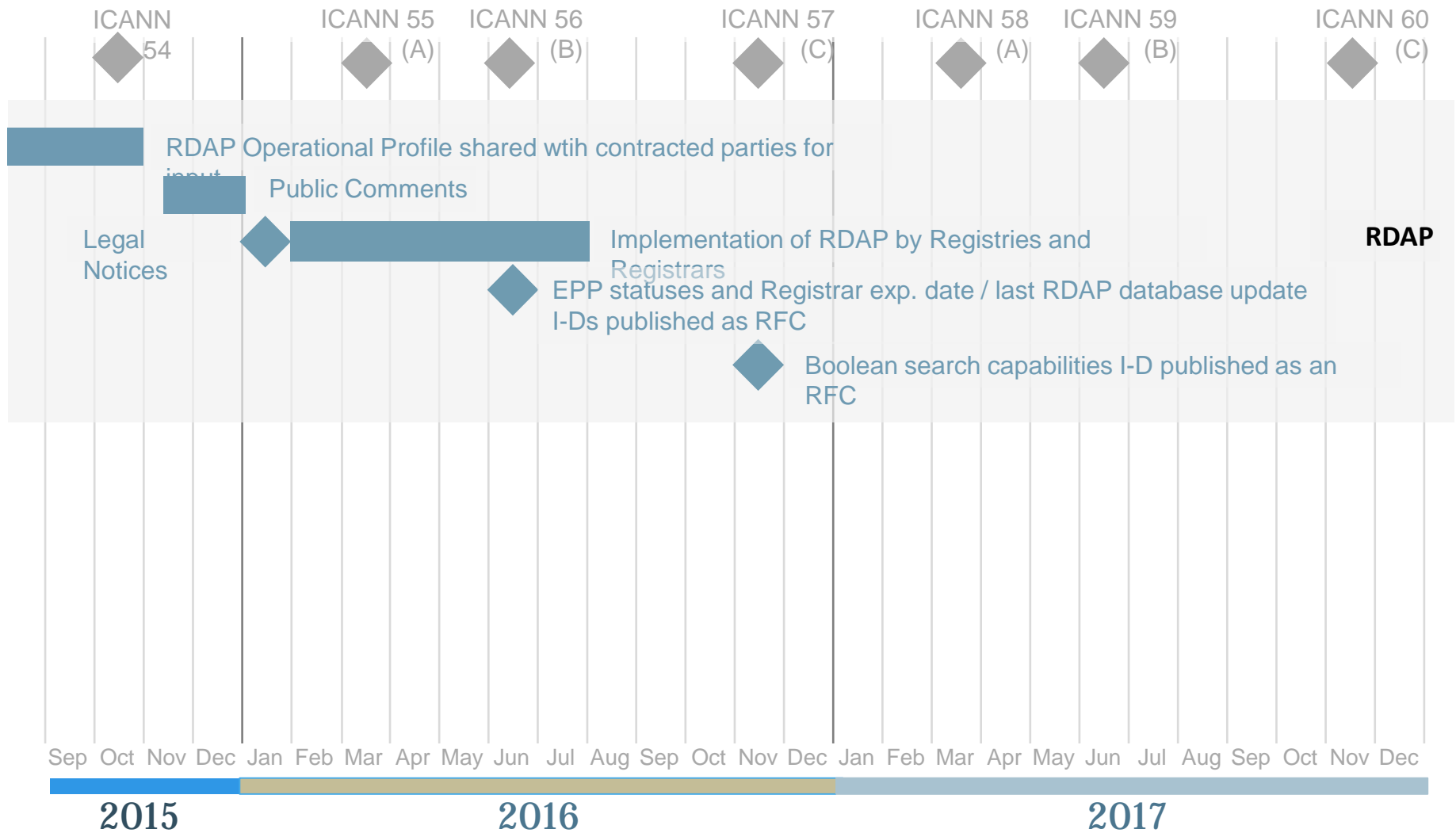
Why WHOIS (port-43) should be replaced?

- ⊙ Unauthenticated
 - ⊙ Unable to differentiate between users
- ⊙ Unable to provide differentiated service
 - ⊙ The same fields are provided to all users
- ⊙ Insecure
 - ⊙ No support for an encrypted response
- ⊙ No bootstrapping mechanism
 - ⊙ No standardized way of knowing where to query
- ⊙ Lack of standardized redirection/reference
 - ⊙ Different workarounds implemented by TLDs

History on Replacing the WHOIS Protocol

- ⦿ SSAC's SAC 051 Advisory (19 Sep 2011):
 - *The ICANN community should evaluate and adopt a replacement domain name registration data access protocol*
- ⦿ Board resolution adopting SAC 051 (28 October 2011)
- ⦿ Roadmap to implement SAC 051 (4 June 2012)
- ⦿ Registration Data Access Protocol (RDAP) community development within IETF working group started in 2012
- ⦿ Contractual provisions in: .biz, .com, .info, .name, .org, 2012 Registry Agreement (new gTLDs), and 2013 Registrar Accreditation Agreement

RDAP Implementation Timeline



RDAP

If you would like to know more about RDAP join us:

Registration Data Access Protocol (RDAP) Implementation

- ⦿ **When:** Wednesday, 21 October 2015 - 12:30 to 13:45
- ⦿ **Room:** Liffey Hall 1

Panel Discussion

IRD - Next Steps

Panel Members

James Galvin, Afilias

Jody Kolker, GoDaddy

Chris Dillon, NCSG

Rudy Vansnick, NPOC

Francisco Arias, ICANN

Margie Milam, ICANN

Questions for the Panel Discussion

1. What do you think the next steps should be for the IRD Report in light of the adoption of the T&T policy by the Board?
 - a. Are there any IRD recommendations that need further policy work?
 - b. If so, which ones?
2. Are there any inconsistencies between the recommendations in the IRD report and the T&T Report?
3. If the IRD report recommendations were to become policy what technical challenges do you foresee?
4. Where are the areas where more work is needed?

A world map where the continents are defined by a complex network of white dots and thin white lines, resembling a network or data visualization. The background is a solid dark blue color.

Thank You!