

Recommendations 5-9: Data Accuracy

Implementation Briefing for RDS2 Review Team



TABLE OF CONTENTS

RDS/WHOIS1 REVIEW TEAM'S RECOMMENDATIONS 5-9: DATA ACCURACY	3
Recommendation 5	3
Recommendation 6	3
Recommendation 7	3
Recommendation 8	3
Recommendation 9	3
ICANN Board Action	4
ICANN Board's Rationale for Board Action	4
DATA ACCURACY IMPLEMENTATION ACTIVITIES	5
WHOIS Accuracy Reporting System (ARS)	5
Information regarding requirements for registrants to provide accurate information, and the consequences of providing inaccurate information	13
2013 Registrar Accreditation Agreement (RAA) and PDP	13
ESTIMATED ICANN ORG TIME SPENT FOR IMPLEMENTATION	13
USEFUL LINKS	14

RDS/WHOIS1 Review Team's Recommendations 5-9: Data Accuracy

Recommendation 5

ICANN should ensure that the requirements for accurate WHOIS data are widely and pro-actively communicated, including to current and prospective Registrants, and should use all means available to progress WHOIS accuracy, including any internationalized WHOIS data, as an organizational objective. As part of this effort, ICANN should ensure that its Registrant Rights and Responsibilities document is pro-actively and prominently circulated to all new and renewing registrants.

Recommendation 6

ICANN should take appropriate measures to reduce the number of WHOIS registrations that fall into the accuracy groups Substantial Failure and Full Failure (as defined by the NORC Data Accuracy Study, 2009/10) by 50% within 12 months and by 50% again over the following 12 months.

Recommendation 7

ICANN shall produce and publish an accuracy report focused on measured reduction in WHOIS registrations that fall into the accuracy groups "Substantial Failure and Full Failure", on an annual basis.

Recommendation 8

ICANN should ensure that there is a clear, unambiguous and enforceable chain of contractual agreements with registries, registrars, and registrants to require the provision and maintenance of accurate WHOIS data. As part of these agreements, ICANN should ensure that clear, enforceable and graduated sanctions apply to registries, registrars and registrants that do not comply with its WHOIS policies. These sanctions should include de-registration and/or de-accreditation as appropriate in cases of serious or serial non-compliance.

Recommendation 9

Board should ensure that the Compliance Team develop metrics to track the impact of the annual WHOIS Data Reminder Policy (WDRP) notices to registrants; metrics should be used to As per (1) above, the Board will initiate a policy on the purpose of the gTLD WHOIS service, and this will help drive the principles behind privacy/proxy develop and publish performance targets, to improve data accuracy over time; if this is unfeasible, Board should ensure that an

alternative, effective policy is developed and implemented that achieves the objective of improving data quality, in a measurable way.

ICANN Board Action

- ⦿ The Board directs the CEO to:
 - a) proactively identify potentially inaccurate gTLD data registration information in gTLD registry and registrar services, explore using automated tools, and forward potentially inaccurate records to gTLD registrars for action
 - b) publicly report on the resulting actions to encourage improved accuracy.
- ⦿ The Board directs the CEO to ensure that WHOIS information pages make clear the requirements for registrants to provide accurate information, and the consequences of providing inaccurate information.
- ⦿ The Board continues to support the RAA negotiation process to find ways to improve WHOIS accuracy, and as per (1) above is initiating a PDP to reform the WHOIS policy to support the objectives and balance the concerns of the multi-stakeholder community.

ICANN Board's Rationale for Board Action

- ⦿ As per actions related to Recommendation 3 (Outreach), the ICANN portal for gTLD WHOIS services will make clear the requirements for registrants to submit accurate information, and the risk that their names may be cancelled if the information is not accurate.
- ⦿ ICANN will report on current levels of accuracy from the recent data studies, and will track and report on improvements.
- ⦿ ICANN already has an enforceable chain of contracts. The gTLD registrar agreement includes sanctions that include de-accreditation if a registrar fails to respond to reports of inaccurate WHOIS information.

Data Accuracy Implementation Activities

This briefing document provides the Review Team with a summary of data accuracy work and links to additional detailed information.

WHOIS Accuracy Reporting System (ARS)

In November 2012, the ICANN Board approved a series of improvements recommended by the 2012 WHOIS Review Team to the manner in which ICANN carries out its oversight of the WHOIS Program, in response to recommendations compiled and delivered by the 2012 WHOIS Review Team, under the Affirmation of Commitments (AoC).

As part of these improvements, ICANN committed to proactively identifying potentially inaccurate generic Top Level Domain (gTLD) WHOIS contact data and forwarding potentially inaccurate records to gTLD registrars for investigation and follow-up. To accomplish these tasks and address Governmental Advisory Committee (GAC) concerns on WHOIS accuracy ICANN initiated the development of the WHOIS Accuracy Reporting System (ARS)—a framework for conducting repeatable assessments of WHOIS accuracy, publicly report the findings, and provide data to the ICANN Contractual Compliance team to follow up on potentially inaccurate records with registrars.

ICANN designed the ARS to be organized into three Phases based on the types of validations described in the [SAC058: SSAC Report on Domain Name Registration Data Validation](#) (Syntax, Operability, and Identity) published in March 2013 by the Security and Stability Advisory Committee (SSAC).

1. **Syntactic Validation** refers to the assessment of data with the intent to ensure that they satisfy specified syntactic constraints, conform to specified data standards, and are transformed and formatted properly for their intended use. For example, if the data element is expected to be an email address, is it formatted as an email address? In general, it is expected that syntactic validation checks would be entirely automated and could be executed in-line with a registration process, follow up information reviews, and whenever registration data changes.

[ARS Phase 1](#) established syntax accuracy testing designed to assess the contact information of a WHOIS record by comparing it to the formats specified by the applicable contractual requirements of the RAA. Syntax tests are performed on all nine individual contact information fields in a record (i.e., email address, telephone number, and postal address for the Registrant, Administrative, and Technical Contacts) and compiled as an entire record. The resulting data is analyzed to produce statistics of syntax accuracy for WHOIS contact information across subgroups such as gTLD, Region, and RAA type (i.e., 2009 RAA or 2013 RAA4).

2. **Operational Validation** refers to the assessment of data for their intended use in their routine functions. Examples of operational validation include 1) checking that an email address or phone number can receive email or phone calls; 2) checking that a postal address can receive postal mail;

[ARS Phase 2](#), begun in 2015, added Operability Testing. The WHOIS ARS team has worked with vendors to define Operability criteria, and during each subsequent cycle has tested a representative sub-sample of collected records for Operability.

3. **Identity validation** refers to the assessment that the data corresponds to the real world identity of the entity. It involves checking that a data item correctly represents the real world identity for the registrant. In general, identity validation checks are expected to require some manual intervention.

ARS Phase 3. ICANN has researched identity verification or validation, but due to cost and feasibility issues, ICANN is not currently pursuing this path. ICANN org is [seeking](#) information regarding commercial services that focus on global address validation as part of the [Across-Field Address Validation](#) work.

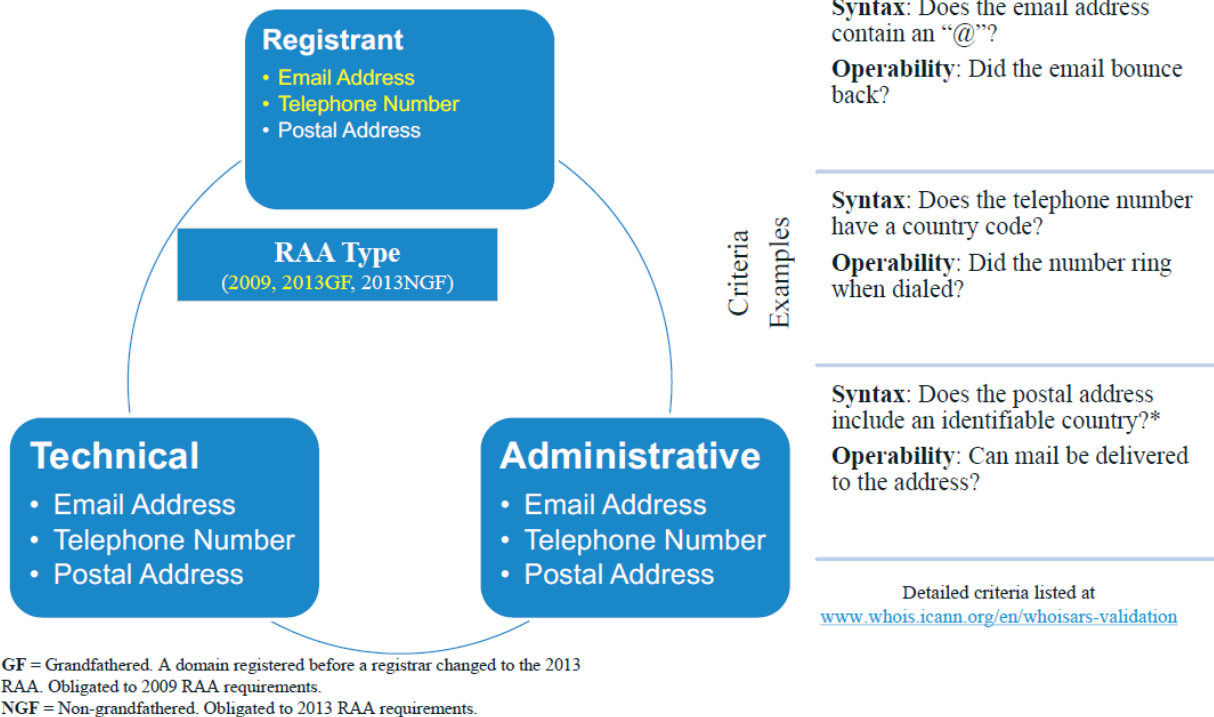
Phase 1 and Phase 2 Reporting

Since the 2013 RAA was produced concurrently with the SSAC's work, ARS Phase1 was able to use the requirements of the Registrar Accreditation Agreements (RAA) to derive the syntax accuracy criteria, ICANN sought to determine whether WHOIS records are meeting format and content requirements.

The WHOIS Accuracy Reporting System was launched in December 2015, and [ARS reports](#) are produced every June and December. When a report is published an announcement is posted. During the first 4 Cycles of Phase 2, a webinar to explain and discuss the report was held. Beginning in cycle 5 the webinar has been indefinitely postponed due to cost and low turnout. The reports are also translated.

- ⦿ December 2015 [report](#), webinar [presentation](#), webinar [recording](#)
- ⦿ June 2016 [report](#), webinar [presentation](#), webinar [recording](#)
- ⦿ December 2016 [report](#), webinar [presentation](#), webinar [recording](#)
- ⦿ June 2017 [report](#), webinar [presentation](#), webinar [recording](#)
- ⦿ December 2017 [report](#)

Contact types, modes, and testing criteria



For each report ICANN collects a sample of approximately 200,000 WHOIS records, of which 12,000 records are checked for syntax and operability accuracy. These 12,000 sample records are specifically selected to be reflective of the entire WHOIS population.

Potentially inaccurate records are identified through the syntax and operability checks sent to ICANN Compliance for follow-up. Following the internal review, ICANN Contractual Compliance will assess the types of errors found, as well as the type of follow-up required with registrars. Depending upon the nature of the error, tickets may be processed as WHOIS inaccuracy complaints or WHOIS format complaints. All WHOIS ARS complaints will follow the Contractual Compliance Approach and Process (see Recommendation 4). However, ICANN Contractual Compliance will continue to give priority to complaints submitted by community members. For more information on Compliance’s methods, please see the written briefing for Recommendation #4.

One of the challenges with the ARS process is that it takes approximately four to five months between when the sample population is polled to when the potentially inaccurate records are available for Compliance’s follow-up. The result is that some records sent to Compliance are outdated. ICANN is trying to narrow this four to five month timeline as much as possible, but it the process of looking at the population, polling the sample size and the actual syntax and operability checks, it still will remain a laborious process.

ARS Findings for Phase 2, Cycle 5 published December 2017:
Phase 1 Data

- ⦿ Shows the rates of syntax accuracy of WHOIS contact information over several dimensions, focusing on rates of conformance by contact mode (Email, Telephone or Post) to the requirements of RAAs (2009 RAA or 2013 RAA).
- ⦿ Approximately 99 percent of email addresses, 90 percent of telephone numbers, and 89 percent of postal addresses met all of the baseline syntax requirements of the 2009 RAA, as shown in the table below:

Table Ex2: Overall³ gTLD Accuracy to 2009 RAA Syntax Requirements by Mode

	Email	Telephone	Postal Address	All Three Accurate
All Three Contacts Accurate	99.6% ± 0.1%	90.2% ± 0.5%	88.9% ± 0.6%	81.5% ± 0.7%

The leading causes of syntax and operability nonconformance in the various subgroups are examined and explained in [Main Findings](#) and in [Appendix B](#) and [Appendix C](#).

- Regarding the individual contact modes, we can make the following general observations:
 - If an email is provided, it always passed all syntax accuracy tests
 - Two-thirds of the telephone numbers that failed at least one syntax accuracy check (13 percent of all telephone numbers) failed the length criteria for the applicable country
 - Postal addresses that failed at least one syntax accuracy check (23 percent of all postal addresses) were typically missing at least one required field such as postal code, state, city, or street

- ⦿ Syntax Accuracy data:

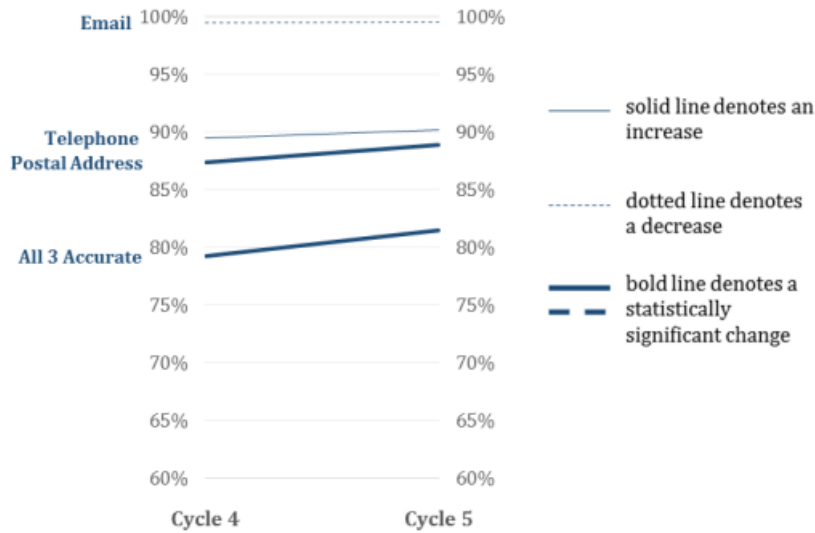
Overall Syntax Accuracy by Cycle Table, All Cycles

Table 28: Overall Accuracy by Cycle – 2009 RAA Syntax Requirements

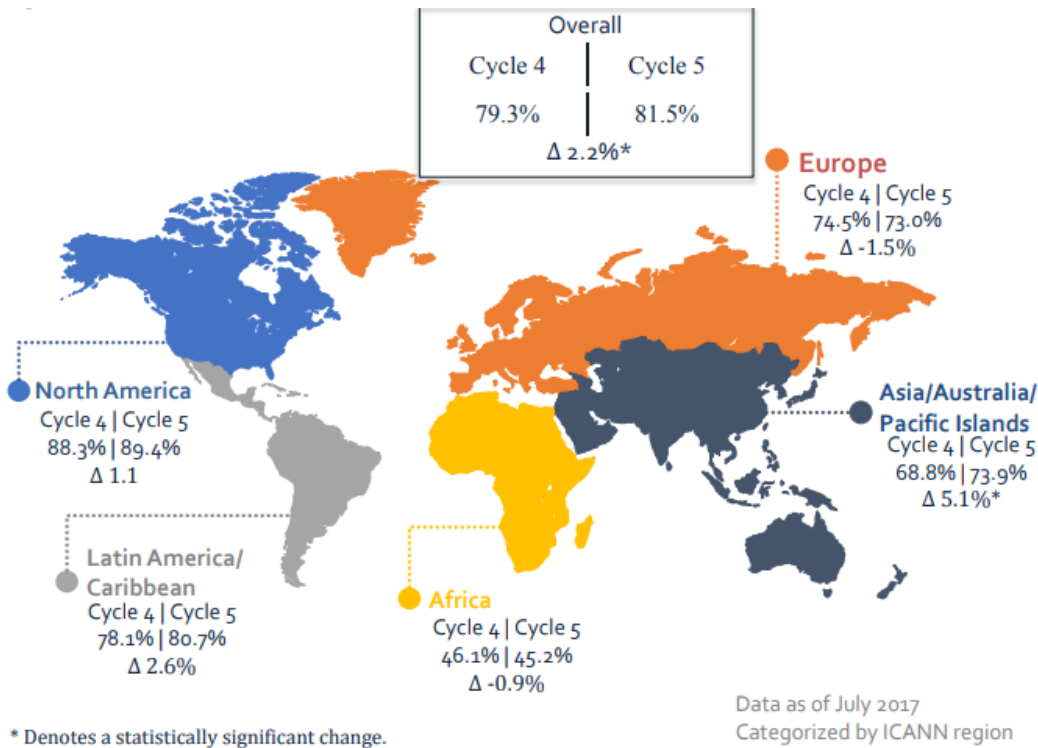
	Email	Telephone	Postal Address	All Modes Accurate
Cycle 1	99.1% ± 0.2%	83.3% ± 0.7%	86.5% ± 0.7%	73.1% ± 0.9%
Cycle 2	99.2% ± 0.2%	85.3% ± 0.6%	86.3% ± 0.6%	75.3% ± 0.8%
Cycle 3	99.6% ± 0.1%	88.5% ± 0.6%	87.0% ± 0.6%	78.0% ± 0.7%
Cycle 4	99.5% ± 0.1%	89.5% ± 0.6%	87.4% ± 0.6%	79.3% ± 0.7%
Cycle 5	99.6% ± 0.1%	90.2% ± 0.5%	88.9% ± 0.6%	81.5% ± 0.7%
Change (C5–C4)	-0.1% ± 0.2%	0.7% ± 0.8%	1.5% ± 0.8%	2.2% ± 1.0%

Overall Syntax Accuracy by Cycle Graph, Cycle 4 and Cycle 5 only

Graph 19: Overall Accuracy by Cycle – 2009 RAA Syntax Requirements



Overall Syntax Accuracy by ICANN Regions, Cycle 4 and Cycle 5



Phase 2 Data

- ⦿ Phase 2 examines both syntax and operability accuracy of WHOIS contact information over several dimensions, focusing on rates of conformance by contact mode (Email, Telephone or Post) to the requirements of RAAs (2009 RAA or 2013 RAA).
- ⦿ Reporting on this phase started in December of 2015, with semi-annual updates taking place on June 2016, December 2016, June 2017 and December 2017. Results from the December 2017 reporting are as follows:

- 98 percent of records had at least one email or telephone number meet all operability requirements of the 2009 RAA, which implies that nearly all records contain information that can be used to establish immediate contact.
- 98 percent of postal addresses, 67 percent of telephone numbers and 94 percent of email addresses met all operability requirements of the 2009 RAA.
- 63 percent of domains passed all operability tests, for all contact types (registrant, administrative and technical) and contact modes (email, telephone and postal address), which is a 4 percent drop from Cycle 4 findings

- ⦿ Operability Data:

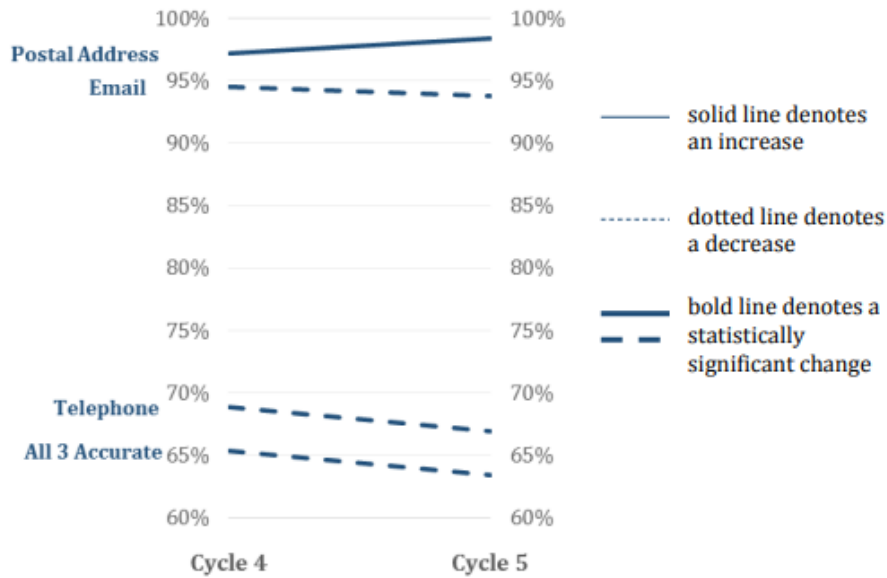
Overall Operability Accuracy by Cycle Table, All Cycles

Table 25: Overall Accuracy by Cycle – 2009 RAA Operability Requirements

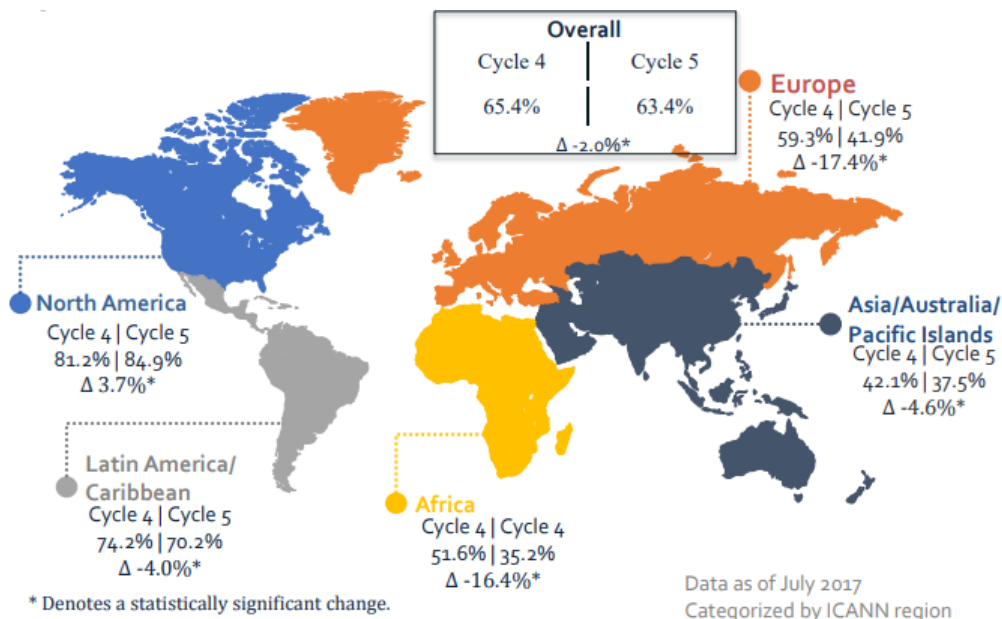
	Email	Telephone	Postal Address	All Modes Accurate
Cycle 1	87.1% ± 0.7%	74.0% ± 0.9%	98.0% ± 0.3%	64.7% ± 0.9%
Cycle 2	91.4% ± 0.5%	76.0% ± 0.8%	97.6% ± 0.3%	70.2% ± 0.8%
Cycle 3	90.1% ± 0.5%	72.4% ± 0.8%	96.8% ± 0.3%	65.1% ± 0.9%
Cycle 4	94.5% ± 0.4%	68.9% ± 0.8%	97.2% ± 0.3%	65.4% ± 0.9%
Cycle 5	93.8% ± 0.4%	66.9% ± 0.8%	98.4% ± 0.2%	63.4% ± 0.9%
Change (C5–C4)	-0.7% ± 0.6%	-1.9% ± 1.2%	1.2% ± 0.4%	-2.0% ± 1.2%

Overall Operability Accuracy by Cycle Graph, Cycle 4 and Cycle 5 only

Graph 16: Overall Accuracy by Cycle – 2009 RAA Operability Requirements



Overall Operability Accuracy by ICANN Regions, Cycle 4 and Cycle 5



[Compliance Metrics for ARS](#)

ICANN Contractual Compliance Metrics for WHOIS ARS Phase 2 Cycle 4
(Data as of 1 October 2017)

This Phase tested Syntax (the format of a record) and Operability (functionality of the information in a record).

NOTE: Ticket count may not add up to total tickets created as a ticket may reach multiple notice stages. Due to rounding and tickets being closed with multiple closure reasons, percentages may not always appear to add up to 100%.

For more about the ICANN Contractual Compliance Approach & Process, please see <https://www.icann.org/resources/pages/approach-processes-2012-02-25-en>.

Ticket Processing Summary & Breakdown by Process Step	
---	--

Total tickets created	4,681
Total tickets closed	4,084
Tickets in process as of 1 October 2017	597

Tickets closed before 1st notice	2,498
Tickets that went to a 1st notice	1,668
Tickets that went to a 2nd notice	175
Tickets that went to a 3rd notice	15
Tickets that went to enforcement	0

Closure Reason for Tickets Closed Before 1st Notice	
---	--

WHOIS data when ticket processed different from sampled WHOIS data	47.4%
Domain not registered when ticket processed	19.4%
WHOIS format issue identified for 2013 Grandfathered Domain	13.6%
Domain suspended or canceled	12.3%
Known Privacy/Proxy service	7.3%
Other (remaining closure reasons representing less than 0.5% of cases)	0.1%

Closure Reasons for Tickets Closed After at Least One Notice	
--	--

Domain suspended or canceled	72.6%
WHOIS data changed or updated	14.9%
Registrar corrected WHOIS format	5.9%
Registrar verified that sampled WHOIS data is correct	5.5%
Registrar demonstrated compliance with RAA	0.4%
Other (remaining closure reasons representing less than 0.2% of cases)	0.4%

Information regarding requirements for registrants to provide accurate information, and the consequences of providing inaccurate information

Please see written briefing for Recommendation 3 Outreach regarding the whois.icann.org portal and registrant education effort.

2013 Registrar Accreditation Agreement (RAA) and PDP

Please see written briefing for Recommendation 10: Privacy and Proxy, for more information regarding the 2013 RAA negotiations.

Estimated ICANN Org Time Spent for Implementation

Implementation of recommendations 5-9 resulted in the work of the 2013 RAA implementation (please see written briefing for Recommendation 10: Privacy and Proxy, for ICANN org's time estimate) and the WHOIS Accuracy Program Specification. The main work to implement the WHOIS Accuracy Reporting System took approximately four (4) months, and 1.25 FTEs. The ongoing annual operation of the WHOIS ARS, consisting of two report cycles, requires approximately 0.25 FTE; this does not include the efforts of the Compliance Team in remediating any issues found as a result of the report analysis.

Useful Links

RDS/WHOIS REVIEW TEAM 1

- ⦿ WHOIS Policy Review Team Final Report
<https://www.icann.org/news/announcement-2015-08-24-en>
- ⦿ ICANN Action Plan For the WHOIS Policy Review Team Final Report
<https://www.icann.org/en/system/files/files/implementation-action-08nov12-en.pdf>
- ⦿ Quarterly WHOIS recommendations implementations statuses updated quarterly and published on WHOIS Review Wiki page
<https://community.icann.org/display/WHO/WHOIS+Review+Implementation+Home>

2013 RAA

- ⦿ 2013 Registrar Accreditation Agreement (2013 RAA)
<https://www.icann.org/resources/pages/approved-with-specs-2013-09-17-en>
- ⦿ Registrars on the 2013 RAA
<https://www.icann.org/resources/pages/registries/registries-agreements-en>
- ⦿ Registries on the new registry agreements
<https://www.icann.org/resources/pages/registries/registries-agreements-en>
- ⦿ 2013 RAA: WHOIS Accuracy Program Specification's Public Comment Report (July 2015)
<https://www.icann.org/en/system/files/files/report-comments-2013-whois-accuracy-spec-review-20jul15-en.pdf>
- ⦿ 2013 RAA: WHOIS Accuracy Program Specification
<https://www.icann.org/resources/pages/approved-with-specs-2013-09-17-en#whois-accuracy>
- ⦿ Clarifications to the Registry Agreement and the 2013 Registrar Accreditation Agreement (RAA) regarding applicable Registration Data Directory Service (WHOIS) Specifications
<https://www.icann.org/resources/pages/registry-agreement-raa-rdds-2015-04-27-en>

WHOIS Accuracy Reporting System (ARS)

- ⦿ SAC058: SSAC Report on Domain Name Registration Data Validation, March 2013
<https://www.icann.org/en/system/files/files/sac-058-en.pdf>
- ⦿ WHOIS ARS Testing Criteria
<https://whois.icann.org/en/whoisars-validation>
and <https://whois.icann.org/en/file/whois-ars-testing-criteria-updated>
- ⦿ WHOIS ARS Reporting Materials
<https://whois.icann.org/en/whoisars-reporting>
- ⦿ WHOIS ARS Phase 1 Reporting
<https://whois.icann.org/en/whois-ars-phase-1-reporting>
- ⦿ WHOIS ARS Reporting
<https://whois.icann.org/en/whois-ars-phase-2-reporting>
- ⦿ WHOIS ARS Compliance Metrics
<https://whois.icann.org/en/whoisars-contractual-compliance-metri>