

|  |
| --- |
| Proposed Subgroup Report Structure |
|  |
| Insert subtitle here (75 characters maximum) |
| Subgroup 1 WHOIS1 Rec #5-9: Data Accuracy |
| 31 March 2018 |
|  |

Table of Contents

[1 Topic 3](#_Toc510337939)

[2 Summary of Relevant Research 3](#_Toc510337940)

[3 Analysis & Findings 3](#_Toc510337941)

[4 Problem/Issue 9](#_Toc510337942)

[5 Recommendations 12](#_Toc510337943)

# Topic

[Introduce objective and explain why it was considered.]

Consistent with ICANN’s mission and Bylaws, Section 4.6(e)(iv), the Review Team will (a) evaluate the extent to which ICANN Org has implemented each prior Directory Service Review recommendation (noting differences if any between recommended and implemented steps), (b) assess to the degree practical the extent to which implementation of each recommendation was effective in addressing the issue identified by the prior RT or generated additional information useful to management and evolution of WHOIS (RDS), and (c) determine if any specific measurable steps should be recommended to enhance results achieved through the prior RT’s recommendations. This includes developing a framework to measure and assess the effectiveness of recommendations, and applying that approach to all areas of WHOIS originally assessed by the prior RT (as applicable).

The implementation of Recommendations #5, 6, 7, 8, 9 on Data Accuracy were reviewed and assessed here after by this subgroup.

# Summary of Relevant Research

[Include methodology that was employed, list of relevant materials, briefings received, reading, input from meetings (as appropriate)]

The feeds for this subgroup's review are from 4 sources:

(1) Background documents on WHOIS1 Implementation Briefings (listed on wiki workspace page [https://community.icann.org/display/WHO/WHOIS1+Rec+%235-9%3A+Data+Accuracy](https://community.icann.org/display/WHO/WHOIS1%2BRec%2B%235-9%3A%2BData%2BAccuracy))

(2) Writing briefing ICANN and answers from SMEs on specifically raised questions

(3) Views exchanged during the Review Team's plenary calls and subgroup calls

(4) Open source research on Internet about the impact of data accuracy enforcement on communities.

# Analysis & Findings

[Provide overview of Review Team Findings(including materials of reference).]

A handful of measures are either in effect or have been taken by ICANN to progress Whois accuracy since prior WHOIS review. (1) A WHOIS Informational Website has been established as a Whois policy documentation, to educate registrants on WHOIS, their rights and responsibilities, and to allow Internet users to submit complaints on Whois inaccuracy. (2) The 2013 RAA introduced contractual obligations for registrars to validate and verify Whois data upon registration. (3) ICANN is in the midst of developing a WHOIS Accuracy Reporting System (referred to as the ARS), proactively identify potentially inaccurate gTLD registration data; explore using automated tools, and forward potentially inaccurate records to gTLD registrars for action. (4) The Whois Data Reminder Policy (WDRP), adopted by ICANN as a consensus policy on 27 March 2003 and is in effect till today, requires a registrar to present to the registrant the current Whois information at an annual basis, and remind the registrants to review their Whois data, and make any corrections.

3.1 Implementation review of Recommendation 5 - WHOIS requirements for accurate data widely and pro-actively communicated

A WHOIS Informational Website has been established as a Whois policy documentation, to educate registrants on WHOIS, their rights and responsibilities, and to allow Internet users to submit complaints on Whois inaccuracy. It's explicitly required by ICANN for the registrants to be solely responsible for the registration and use of the domain name registered, and must provide accurate information for Whois data publication, and promptly update this to reflect any changes. These requirements are both elaborated on WHOIS Informational Website and in 2013 RAA as below.





The 2013 RAA obligates each Registrar to publish on its website(s) and/or provide a link to the Registrants' Benefits and Responsibilities Specification. ICANN’s Contractual Compliance Team checks to determine whether registrars are publishing this information and follows up to bring the Registrar into compliance if it is not doing so.

The 2013 RAA clearly outlined that registrant's willful breach of Whois accuracy policy above will lead to suspension and/or cancellation of the registered domain name (see below).



The Whois accuracy policy was depicted both in 2009 RAA and 2013 RAA, it is assumed every registrant has been exposed to the above responsibilities, although it's not clear whether or how the above responsibilities have been included in the registration agreement between registrar and registrant. In other words, it is uncertain whether the above responsibilities are actually enforceable through the whole chain.

In conclusion, the subgroup has the view that Rec #5 has been fully implemented, while the effectiveness of implementation needs to be further assessed.

3.2 Implementation review of 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".

To address Rec #6, ICANN initiated the [Accuracy Reporting System](https://whois.icann.org/en/whoisars) (ARS) project, with the aim to "proactively identify inaccurate gTLD registration data, explore the use of automated tools, forward potentially inaccurate records to registrars for action, and publicly report on the resulting actions to encourage improvement."

The ARS was designed to be implemented through three Phases based on the types of validations described in the [SAC058 Report](https://www.icann.org/en/system/files/files/sac-058-en.pdf) (syntax, operability, and identity).

(1) [Phase 1](https://whois.icann.org/en/whois-ars-phase-1-reporting): Syntax Accuracy

(2) [Phase 2](https://whois.icann.org/en/whois-ars-phase-2-reporting): Syntax + Operability Accuracy

(3) Phase 3: Syntax + Operability + Identity Accuracy

Phase 1 was completed in August 2015 and assessed the format of a Whois record (i.e. Is the record correctly formatted? Is there an "@" symbol in the email address? Is there a country code in the telephone number?)

Phase 2 reviews both the syntax and operability accuracy of WHOIS records by assessing the functionality of the information in a record (e.g. Does the email go through? Does the phone ring? Will the mail be delivered?). Phase 2 is ongoing with a new report published every 6 months, detailing the leading types of nonconformance, trends and comparisons of WHOIS accuracy across regions, Registrar Accreditation Agreement (RAA) versions and gTLD types. The newest Phase 2 Cycle 5 report was published in December 2017.

Phase 3 has not started yet. 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](https://community.icann.org/display/AFAV) work.

It is worth mentioning that only a sample of WHOIS records is used for accuracy testing. A two-stage sampling method is used to provide a sample to reliably estimate subgroups of interest, such as ICANN region, New gTLD or Prior gTLD, and RAA type. Two samples are prepared at the beginning of each report cycle:

(1) An initial sample of 100,000-200,000 WHOIS records

(2) A sub-sample of the initial sample of 10,000-12,000 WHOIS records, which is used for accuracy testing

Since the sub-sample records falls in both 2009 RAA and 2013 RAA, while the Registrant email address and telephone number are not required for 2009 RAA, the 2013 RAA requires the contact data in a WHOIS record to be more syntactically complete and to be formatted per more specific requirements than that of the 2009 RAA, the accuracy tests were designed in such a way that all records in the analyzed subsample were evaluated against a set of baseline requirements derived from the requirements of the 2009 RAA.

The subgroup mainly focused on the assessment of Phase 2. All WHOIS ARS Phase 2 reporting could be accessed [here](https://whois.icann.org/en/whois-ars-phase-2-reporting). The comparison of all modes accuracy rates between different cycles in Phase 2 are showed in below table:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Phase 2Cycle 1  | Phase 2Cycle 2  | Phase 2Cycle 3  | Phase 2Cycle 4  | Phase 2Cycle 5 |
| Syntax Accuracy | 67.2% | 67.2% | 66.6% | 79.3% | 81.5% |
| Operability Accuracy | 64.7% | 70.2% | 65.1% | 65.4% | 63.4% |

The below table shows syntax and operability accuracy from December 2015 through December 2017 by ICANN region.



ICANN’s Contractual Compliance team supports the WHOIS ARS effort by receiving reports of identified syntax and operational failures and following up with contracted parties to resolve areas of noncompliance. However, 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. As such, above 50% of the tickets were closed before 1st notice, due to either WHOIS data when ticket processed different from sampled WHOIS data, or domain not registered when ticket processed, or Domain already suspended or canceled, or WHOIS format issue identified for 2013 Grandfathered Domain, or Known Privacy/Proxy service. For the left less than 50% of the tickets went to a 1st or further notice, above 60% tickets led to domain suspension or cancellation.

Phase 2 [WHOIS ARS Contractual Compliance Metrics](https://whois.icann.org/en/whoisars-contractual-compliance-metrics) are summarized as below:

Cycle 1: Among 10,000 subsample records, 2,688 tickets were created. 1,324 tickets were closed before 1st notice. For the 1,362 tickets went to 1st or further notice, 60.1% of the related domains were suspended or canceled, 28.2% of the tickets led to changing or updating of WHOIS data by registrar. Four registrars received a Notice of Breach for tickets created. Of the four, one registrar was suspended then terminated.

Cycle 2: Among 12,000 subsample records, 4,001 tickets were created. 2,481 tickets were closed before 1st notice. For the 1,524 tickets went to 1st or further notice, 60.6% of the tickets related domains were suspended or canceled. 25.4% of the tickets led to changing or updating of WHOIS data by registrar. There were no registrars received a Notice of Breach for tickets created.

Cycle 3: Among 12,000 subsample records, 4,552 tickets were created. 2,662 tickets were closed before 1st notice. For the 1,897 tickets went to 1st or further notice, 65% of the tickets related domains were suspended or canceled. 21.5% of the tickets led to changing or updating of WHOIS data by registrar. There were no registrars received a Notice of Breach for tickets created

Cycle 4: 12,000 subsample records. 4,681 tickets were created. 2,498 tickets were closed before 1st notice. For the 1,668 tickets went to 1st or further notice, 72.6% of the tickets related domains were suspended or canceled. 14.9% of the tickets led to changing or updating of WHOIS data by registrar. There were no registrars received a Notice of Breach for tickets created.

Cycle 5: 12,000 subsample records. No ICANN Contractual Compliance Follow-Up Statistics has been published for Cycle 5 yet.

Looking at the WHOIS ARS reporting and Contractual Compliance Metrics above, several observations could be concluded:

(1) Even without identity accuracy check, WHOIS ARS is an effective way to push registrants and registrars to improve Whois data.

(2) For those tickets went to 1st or further notice during Phase 2, above 60% of the tickets related domains were suspended or canceled, and around 20%  of the tickets led to changing or updating of WHOIS data by registrar. Based on the fact, it is suspected that registrars haven't validated and verified Whois data upon registration.

(3) Considering those suspended or canceled domains behind tickets went to 1st or further notice during Phase 2, and those domains had an updated Whois data after tickets issued, the Whois data inaccurate rate across the domain space is still very high (80~90%).

(4) There were seldom Notices of Breach issued by ICANN to registrars for tickets created, while validation and verification of Whois data is already a contractual obligation of registrar as outlined in 2013 RAA.

In consideration that WHOIS ARS is still ongoing, the subgroup has the view that Rec#6 has been partially implemented.

3.3 Implementation review of Recommendation 7 - Annual WHOIS accuracy reports

Instead of an annual WHOIS accuracy report focused on measured reduction in substantial and full failed WHOIS registrations, ICANN has produced and published Annual Report on WHOIS Improvements for [2013](https://whois.icann.org/en/file/improvements-annual-report-04nov13-en), [2014](https://whois.icann.org/en/file/improvements-annual-report-12dec14-en), [2015](https://whois.icann.org/en/file/2015-annual-report-whois-improvements) and [2016](https://whois.icann.org/en/file/2016-annual-report-whois-improvements) separately, which outlined the progress of all WHOIS policy related working streams.

In 2013 Annual Report on WHOIS Improvements, the conclusion of 2013 RAA, the establishment of WHOIS Informational Website, Contractual Compliance Function's enhancement were highlighted as the first year of progress towards fulfilling ICANN’s commitment to improve WHOIS.

In 2014 Annual Report on WHOIS Improvements, besides the progress of other parallel action streams, ARS was firstly introduced. The pilot study results revealed that Registrars under the 2013 RAA experienced better accuracy rates for email addresses than Registrars under prior versions of the RAA. This improvement may be directly related to the introduction of the new validation and verification requirements from the 2013 RAA. It was indicated in the report that as more registrars and more gTLD registrations transition to the new requirements, accuracy rates should continue to be improved.

In 2015 Annual Report on WHOIS Improvements, a WHOIS quality review process referred to as “WHOIS QR” was introduced. In 2014, ICANN launched the WHOIS QR with the objective is to determine if registrars continue to comply with the WHOIS Accuracy obligations as specified in the 2009 and 2013 Registrar Accreditation Agreements (RAA), with an emphasis on previously closed WHOIS inaccuracy complaints because the domain name was suspended. Staff conducts internal monitoring on regular basis to ensure that registrars are complying with their obligations when removing domain name suspension. Hereafter is a brief summary of WHOIS QR in 2015 from [Contractual Compliance Reports 2015](https://www.icann.org/resources/pages/compliance-reports-2015-04-15-en).



In 2016 Annual Report on WHOIS Improvements, the shift from the Affirmation of Commitments (AoC) to new ICANN Bylaws was highlighted. The WHOIS obligations originally established by AoC were replaced by ICANN Bylaws. Those Bylaws require ICANN to periodically conduct review of the effectiveness of WHOIS (RDS in Bylaws), and use commercially reasonable efforts to enforce relating policies. According to the report, WHOIS complaints on accuracy and record format were still the most common registrar compliance issue addressed by ICANN in 2016.

The Annual Report on WHOIS Improvements presented a big picture of the improvements on WHOIS policy development, several working streams had a positive impact on WHOIS accuracy. However, the measured reduction in WHOIS registrations that fall into the accuracy groups Substantial Failure and Full Failure were missing from the reports. Thus, the subgroup has the view that Rec #7 has been partially implemented

3.4 Implementation review of Recommendation 8 - ICANN should ensure that there is a clear, unambiguous and enforceable chain of contractual agreements.

ICANN-accredited registrars have several WHOIS obligations, including:

(1) Provision of free public WHOIS service on Port 43 and via web with output appearing in the required format and according to certain service level requirements;

(2) Submitting all required data elements to the registries;

(3) Obtaining, retaining and updating data elements in a timely manner;

(4) Escrowing data elements;

(5) Providing for bulk access to WHOIS data in accordance with the required bulk access agreement;

(6) Taking reasonable steps to investigate, and where appropriate, correct inaccuracies upon discovery of information or notification suggesting an inaccuracy exists; and

(7) Providing annual WHOIS data reminders to registrants.

Similarly, registry operators also have contractual obligations related to WHOIS, including:

Provision of free public WHOIS service on Port 43 and via web with output appearing in the required format and according to certain service level requirements.

ICANN has an enforceable chain of contractual agreements with registries, registrars. 2013 RAA includes additional enforcement provisions and sanctions applicable to registrars, registrants, and resellers with regards to WHOIS, including de-accreditation if a registrar fails to respond to reports of inaccurate WHOIS information. New gTLD Registry agreements include enhanced WHOIS obligations Renewals of existing gTLDs to include enhanced WHOIS obligations. Thus, the subgroup has the view that Rec #8 has been fully implemented.

3.5 Implementation review of Recommendation 9 - Compliance Metrics: Impact of the annual

WHOIS Data Reminder Policy (WDRP)

The Whois Data Reminder Policy (WDRP), adopted by ICANN as a consensus policy on 27 March 2003, requires a registrar to present to the registrant the current Whois information at an annual basis, and remind the registrant that provision of false Whois information can be grounds for cancellation of their domain name registration. Thus registrants must review their Whois data, and make any corrections. WDRP is intended to be an additional step towards improving Whois data accuracy. All ICANN-accredited registrars must comply with the WDRP with respect to registrations they sponsor in all top-level domains for which they are accredited.

If the Whois information is correct and up-to-date, no further action is needed from registrant side. If the registrant does need to update the Whois information, the registrant will be directed by corresponding registrar to options available for updating. Notably, registrations under privacy and/or proxy service are subject to WDRP as well.

According to [Implementation of the Whois Data Reminder Policy (WDRP) – 30 November 2004](https://www.icann.org/en/system/files/files/wdrp-implementation-30nov04-en.pdf), a total of 254 registrars (70% of all ICANN-accredited registrars at that time) responded to the “Whois Data Reminder Policy Survey and Compliance Audit.”, and only 44% (111 out of 254 ) of the respondant registrars did sent WDRP Notices. Reading from the results of the survey, for those registrars who sent out WDRP Notices, most of them covered 50% or less of all registrations under sponsorship, and there were considerable WDRP Notices undeliverable. Even with all the aspects above, there were still at least several thousands of WDRP Notices led to changes in registrant data. Thus, there is good reason for this subgroup to believe that if the WDRP were fully enforced at annual basis, there would be a quite positive impact on Whois accuracy.

Surprisingly, there have been no further updates from ICANN on the implementation of WDRP ever since then, except a [FAQ webpage](https://www.icann.org/resources/pages/faqs-f0-2012-02-25-en) was online on Feb. 25 2012. It's the subgroup's understanding that WDRP is still in effect till today, however it's not clear to the subgroup whether or not ICANN has enforced this policy afterwards, what impact this policy has placed in improving Whois accuracy.

# Problem/Issue

[What observed fact-based issue is the recommendation intending to solve? What is the “problem statement”?]

WHOIS remains one of the ICANN’s toughest issues over the years, besides ICANN's initiatives and policy development, it needs more proactive efforts from registrants, registrars to fix it. After looking into the all the measures have been taken and Whois improvements have been progressed so far, there are still some gaps to be bridged to meet the prior Whois review recommendations.

4.1 The objective of reliable Whois data has not been achieved

WHOIS data, with the purpose to be able to contact registrants, needs to be accurate. It's explicitly required by ICANN for the registrants to be solely responsible for the registration and use of the domain name registered, and must provide accurate information for Whois data publication, and promptly update this to reflect any changes

How does one determine whether the data displayed in a WHOIS Record is accurate? There may be contact information that appears correct – i.e. that represents a valid and viable name and address (electronic and/or physical) – but is not necessarily accurate, i.e. it does not correspond to the person/entity registering, managing or owning the domain name.

Until the adoption of the 2013 RAA, registrars were not required to verify or validate WHOIS data. The 2013 RAA includes obligations to validate certain WHOIS data fields, and verify either the email address or telephone number displayed. Besides the contractual obligations in the 2013 RAA, ICANN is in the midst of developing a WHOIS ARS with the aim to proactively indentify inaccurate Whois data for improvement.

However, neither the WHOIS ACCURACY PROGRAM SPECIFICATION (WAPS) in 2013 RAA, nor the WHOIS ARS has touched upon the reliability of Whois data yet. Complying with WAPS or WHOIS ARS doesn't necessarily mean the Whois record is not falling into the accuracy groups Substantial Failure and Full Failure (as defined by the NORC Data Accuracy Study, 2009/10). The verification measures have been taken are beating around the bush. In other words, a pragmatic approach for the validation and verification of Whois data is still missing.

4.2 Whois inaccuracy is believed to be largely under-reported

Whois has been a free public lookup service for years. Although ICANN encourages the ICANN community to submit a complaint to ICANN regarding incomplete or incorrect Whois data, however, there are no due resources for general public to judge the accuracy of Whois data, let alone whether the Whois data corresponding to the owner of the domain name. As outlined in the [Whois Inaccuracy Complaint Form](https://forms.icann.org/en/resources/compliance/complaints/whois/inaccuracy-form), the general public is only capable to report on missing information, incorrect address, no such person or entity, etc.

Furthermore, as analyzed in the implementation of Rec #6, the inaccurate Whois records exposed in the WHOIS ARS are only a tip of the iceberg. Thus, there is reasonable ground to believe that the Whois inaccuracy is largely under-reported.

4.3 Incentives for registrant to provide accurate Whois data and for registrars to validate and verify Whois data are missing

Domain name registrants play a key role in ensuring the accuracy of WHOIS. Registrant's willful breach of Whois accuracy policy is supposed to lead to suspension and/or cancellation of the registered domain name. But as long as the related Whois record is not identified by WHOIS ARS, or complained by community, or verified by registrar, there is no risk to falsify the Whois data while keep it fomatly correct, whether for legitimate needs for else.

Looking at the Contractual Compliance Annual Report in [2016](https://www.icann.org/en/system/files/files/annual-2016-31jan17-en.pdf), [2017](https://www.icann.org/en/system/files/files/annual-2017-30jan18-en.pdf), the most common issues addressed by ICANN with regards to registrar compliance on WHOIS Inaccuracy:



In other words, the involved registrars usually didn't comply with the contractual obligations. Actually, registrars (or resellers) are in best position to validate and verify Whois data. As indicated in the [Implementation of the Whois Data Reminder Policy (WDRP) – 30 November 2004](https://www.icann.org/en/system/files/files/wdrp-implementation-30nov04-en.pdf), One registrar noted that its most accurate contact information is contained in its internal accounting system. It wrote that “[w]e have been fairly successful in keeping this data up to date as registrants who are interested in keeping their domain keep their billing information accurate.” Another registrar also suggested that "the billing contact information" to be showed on any given Whois record (see below).



Some best practices on verification of Whois data has emerged from industry. In the years of fighting Avalanche (phishing group), Interdomain, a Spanish registrar, began [requiring a confirmation code delivered by mobile phone](https://slidex.tips/download/phishing-trends-report) in April 2009 which successfully forced Avalanche to stop registering fraudulent domains with them.

It's also worth mentioning the highly regulated domains where Registry rules require “provide appropriate jurisdictional authorities with the capability at their option and at no cost to make designations in the WHOIS record relevant to the registrant’s organizational status in the registrant’s jurisdiction.” It should provide more accurate data.

4.4 The Whois accuracy of domain names that utilize Privacy and Proxy Services is in misty

The Whois data relating to domain names that utilize P/P services has not been touched upon by WHOIS ARS. According to [WHOIS ARS Contractual Compliance Metrics](https://whois.icann.org/en/whoisars-contractual-compliance-metrics), all tickets were closed before 1st Notice for Known Privacy/Proxy service.

Although ICANN Contractual Compliance receives and processes WHOIS inaccuracy complaints regarding domain names that utilize Privacy and Proxy Services, it does not identify the proportion of complaints this represents. Absent an accreditation system for Privacy and Proxy service providers, it is difficult to automate the accurate identification of domain names subject to Privacy and Proxy services in WHOIS inaccuracy complaints.

# Recommendations

Recommendation: xxx

Findings: [what are the findings that support the recommendation]

Rationale:

[What is Intent of recommendation and envisioned outcome?

How did the finding lead to this recommendation?

How significant would impact be if recommendation not addressed?

Is it aligned with ICANN’s Strategic Plan and Mission?

Is it in compliance with scope Review Team set?]

Impact of Recommendation: [What are the impacted areas, e.g. security, transparency, legitimacy, efficiency, diversity etc. Which group/audience will be impacted by this recommendation]

Feasibility of Recommendation: [Document feasibility of recommendation]

Implementation:

[Who are responsible parties that need to be involved in implementation? Community/ICANN org/combination)

What is the target for a successful implementation?

Is related work already underway and how will that dovetail with recommendation?

What is the envisioned implementation timeline? Within 6 months/12 months/more than 12 months]

Priority: [If only 5 recommendations could be implemented due to community bandwidth and other resource constraints, would this recommendation be one of the top 5? Why or why not?]

Consensus: