

WHOIS ACCURACY REPORTING SYSTEM

PILOT REPORT

PRESENTED TO:
ICANN

PUBLISHED BY:
NORC at the
University of Chicago

DECEMBER 23, 2014

Contents

- Executive Summary3**
 - Sample Design5
 - Findings.....6
 - Implications of the Pilot Study7

- Introduction.....8**

- ARS Pilot Study Design10**
 - ARS Report Requirements..... 10
 - ARS Pilot Study 11
 - Sample Sizes of gTLDs..... 13
 - Countries in the Sampling Frame..... 13*
 - Accuracy Definitions 14
 - Automated and Manual Approaches 15

- ARS Pilot Study Methods and Approach17**
 - Validation Perspectives..... 17
 - Rating Scale Categories 18
 - Email Address Validation 18
 - Telephone Number Validation..... 19
 - Postal Address Validation 19
 - Sample Design20

- Main Findings.....24**
 - Accuracy by Validation Perspective 24
 - Accuracy by gTLD.....26
 - Accuracy by Registrar28

- Implications of Findings30**
 - Examining Definitions30
 - Determining an Accuracy Score.....31
 - Sample Size and Confidence Levels.....32
 - Not a Compliance Report.....33
 - Exploring Identity Validation33

NEXT STEPS	34
APPENDIX A. Sample Sizes by gTLD and by Country.	35
Appendix B: Next Steps for the Development of the WHOIS Accuracy Reporting System (ARS).....	45
Public Comment and Consultations in Singapore	45
Process for Correcting Inaccurate Records: ARS Implementation Advisory Group	46
Appendix C: DETAILED DESCRIPTIONS OF VENDOR PROCESSES	48
Telephone Validation: DigiCert.....	49
<i>Syntactical Telephone Validation</i>	<i>49</i>
<i>Operational Telephone Validation.....</i>	<i>52</i>
Email Validation: Strikelron	53
<i>Syntactical Email Validation</i>	<i>54</i>
<i>Operational Email Validation.....</i>	<i>56</i>
Postal Validation: Universal Postal Union (UPU).....	59
<i>Syntactical Postal Validation</i>	<i>59</i>
<i>Operational Postal Validation.....</i>	<i>64</i>

Executive Summary

On 8 November 2012, in response to the [Recommendations of the WHOIS Review Team](#) convened under the [Affirmation of Commitments \(AoC\)](#), the [ICANN Board approved](#) a series of improvements on the manner in which ICANN carries out its oversight of the WHOIS program. As part of these improvements, [ICANN committed](#) to proactively identifying potentially inaccurate gTLD WHOIS contact data and forwarding potentially inaccurate records to gTLD registrars for investigation and follow-up. To accomplish these tasks ICANN initiated the development of the WHOIS Accuracy Reporting System (ARS) – a framework for conducting repeatable assessments of WHOIS accuracy over time and publicly reporting the findings.

The Pilot Study is a precursor to the ARS, designed with the intention of isolating and testing key components of the system. Results from the Pilot Study will be used to inform the design of the ARS.

The Pilot Study was an international collaborative effort led by ICANN, with support from [NORC at the University of Chicago](#), [DigiCert, Inc.](#), [StrikeIron, Inc.](#), [The Universal Postal Union](#) (UPU), and [Whibse, Inc.](#) [Following a Request for Proposals](#), initial planning for the Pilot Study began on 27 August 2014, and data collection began on 4 September 2014. Accuracy validation of the WHOIS records selected for review was completed on a compressed schedule between 12 September 2014 and 19 September 2014. Analysis of the validation results concluded on 8 October 2014.

In the Pilot Study, from a sample of WHOIS records, these organizations validated the telephone, email, and postal address fields respectively. The validation was completed by examining the sample WHOIS records for various factors of syntactic and operational accuracy, which are explained below.

Syntactic perspective: The syntactic perspective examines WHOIS data with the intent of determining which data exists or is missing from within WHOIS data fields and assesses whether data meets specified standards of formatting and structure.

Operational perspective: The operational perspective examines WHOIS data with the intent of determining if contact information can be practically used to achieve the intended function of the data type and involves gathering and utilizing data external to the WHOIS record to inform an accuracy assessment.

An accuracy assessment of WHOIS data, guided by any of the perspectives should result in the data being categorized into a rating scale based on the degree to which it can be considered accurate. The rating scale categories established for the WHOIS accuracy studies describe accuracy within an inclusive range between “No Failure,” the highest degree of accuracy, and “Full Failure,” which represents the lowest degree of accuracy. For the Pilot Study, records were grouped into the superset categories of “Accurate” or “Inaccurate” which encompass the ARS accuracy rating scale categories as follows:

Accurate Categorization

- **No Failure:** Indicates that data are missing no significant information, and that the data provided are accurate using all three accuracy perspectives;
- **Minimal Failure:** Indicates that data are missing no significant information. The record may need clarification or might benefit from additional information, but the data provided are accurate; and
- **Limited Failure:** Indicates that the data are missing at least some significant information, but that the data present are to some degree considered useful.

Inaccurate Categorization

- **Substantial Failure:** Indicates that the data are missing significant information, and that the data provided are mostly inaccurate.
- **Full Failure:** Indicates that significant data elements are missing, or that data provided are overwhelmingly inaccurate.

In the Pilot Study, ICANN contractors applied predetermined validation criteria to the email, telephone, and postal data of WHOIS records to assess accuracy from the syntactic and operational perspectives. The validation results were compiled into datasets, separated by data

type and by perspective, and these datasets were then analyzed to produce accuracy statistics (for New and Prior gTLDs) stratified by gTLD, Geographic Region, and registrar.

Sample Design

The Pilot Study began with the selection of a sample of 100,000 records from gTLD Zone Files. Because New gTLDs are a primary focus of the study, a high percentage of these domains were sampled, accounting for 25 percent of the domains represented in the sample (at the time of sampling, only 1.4 percent of domains were within the New gTLDs). In order to provide some information on all gTLDs, smaller gTLDs were oversampled. Specifically, all New gTLDs had at least 10 records selected (or all if fewer than 10 records were available) and all other (“Prior”) gTLDs had at least 30 records selected (or all if fewer than 30 records were available).

Once the records were selected for the sample, they were then queried via WHOIS lookups to produce data for the Pilot Study. The raw WHOIS data were parsed so that the information in the records could be consistently interpreted during the validation phase of the study. After parsing, all email records moved forward to the syntactical and operational validation phase. The number of successful queries (and parsed records) was 98,821. A subsample of 10,000 records was selected for syntactical validation of postal addresses and telephone numbers. In order to provide estimates at the Geographic Region level, NORC at the University of Chicago determined the country (and Geographic Region) for as many of the domain names as possible, and oversampled the two smallest Regions: Africa and Asia/Australia/Pacific.

A smaller subsample of 1,000 records was drawn (again, Africa and Asia/Australia/Pacific were oversampled) for operational validation of postal addresses and telephone numbers. The same set of sort variables were used to sort the 10,000 records in a serpentine way before selecting a systematic sample of 1,000 records.

A total of 200 countries were represented in the full sample of 98,821 domain names. Of these 200 countries, 91 countries were represented in the subsamples of 10,000 and 1,000 as well as the full sample (African and Latin American/Caribbean countries were oversampled in the subsamples).

The full sample was used for email validation. Syntactical verification for telephone numbers and postal addresses used the 10,000 subsample, while the Operational verification for telephone numbers and postal addresses used the 1,000 subsample. The analysis reports on subgroups with at least 20 records, and only the top seven gTLDs met this criterion in the 1,000 subsample.

It is important to note that the different sample sizes result in a very different uncertainty in the estimates below. Specifically, for the email verification, small differences are statistically significant while for the operational validation of telephone numbers and postal addresses, differences of several percentage points might not be statistically significant.

Findings

The accuracy of email, postal address, and telephone information was assessed from the syntactic and operational perspectives, and group statistics were computed for all gTLDs, 2009 and 2013 RAA, and for New and Prior gTLDs. Some notable findings are:

- The differences in email and telephone syntactical accuracy are not statistically significant for Prior and New gTLDs while syntactical accuracy is better for Prior gTLDs on postal addresses.
- The operational numbers for email addresses indicate that New gTLDs have slightly better accuracy than Prior gTLDs. Prior gTLDs do have more operational accuracy on telephone numbers, but the two groups are equal on operational postal address accuracy.
- The differences in syntactical accuracy between registrars under the 2013 RAA and registrars under the 2009 RAA for email, telephone numbers and postal addresses are not statistically significant.
- Operationally, registrars under the 2013 RAA have more accuracy for email addresses than registrars under the 2009 RAA. Registrars under the 2009 RAA have more operational accuracy in telephone numbers while the two groups are similar on postal addresses.

The Pilot Study was a highly informative exercise, and provided valuable lessons and caveats for the methods and procedures of the ARS. Some of the key lessons learned from the pilot are:

- **Slow Process:** Selecting samples of WHOIS records, turning the raw WHOIS information into parsed fields suitable for evaluation, and evaluating postal addresses, email addresses and telephone numbers takes time, and should not be rushed. The timetable for the Pilot Study was greatly condensed and should be lengthened when the WHOIS Accuracy Reporting System is launched.
- **Postal Address Validation:** Unlike email address and telephone number, a postal address is comprised of a number of fields. This makes it much more difficult to validate postal addresses and the outcomes may be quite different than those of email and telephone. The five category evaluation scale (No Failures, Minimal Failure, Limited Failure, Substantial Failure, and Complete Failure) may not necessarily be useful for postal address validation, while it may be useful for email and telephone validation.
- **Sample Size Issues:** It takes more time to validate postal addresses and telephone numbers than it does email addresses. Furthermore, it takes longer to perform operational validation than syntactical validation for both postal addresses and telephone numbers. Because of this, different sample sizes were used. This means that the accuracy measures based on smaller samples are less reliable than those used on larger samples. In particular, operational validation for postal address and telephone number used a sample of 1,000 WHOIS records. This may be too small to provide good estimates for comparing Geographic Regions, registrars and gTLDs. As such, a larger sample should be considered in the future.

Introduction

On 8 November 2012, the ICANN Board approved a series of improvements to the manner in which ICANN carries out its oversight of the WHOIS Program (the current gTLD registration data directory service offered by registrars and registries), in response to recommendations compiled and delivered by the WHOIS Review Team, under the [Affirmation of Commitments](#) (AoC).

As part of these improvements, ICANN committed to proactively identifying potentially inaccurate gTLD WHOIS contact data and forwarding potentially inaccurate records to gTLD registrars for investigation and follow-up. To accomplish these tasks ICANN initiated the development of the WHOIS Accuracy Reporting System (ARS) – a framework for conducting repeatable assessments of WHOIS accuracy over time and publicly reporting the findings.

The Pilot Study was a precursor to the ARS, designed with the intention of isolating and testing key components of the system. It was also designed to help ensure that ICANN and its efforts are aligned with other activities related to WHOIS, including, examining the effectiveness of the new validation and verification requirements under the Registrar Accreditation Agreement (RAA), and the various GAC Communiqué's that address WHOIS accuracy.

The design of the ARS has been developed in consultation with the ICANN community over the last year. Beginning in March, 2014, ICANN posted a [Draft Implementation Plan](#) for the ARS, for [public comment](#), to solicit feedback on the proposed [model](#) that described the methodology and approach. Following the closure of the public comment, ICANN released a [Request for Proposal](#) in May, 2014, to seek vendors to provide services to support the ARS development using an updated methodology that took into account the [feedback](#) received from the public.

ICANN conducted the Pilot Study in August – September 2014 to test the proposed methodology using real data and to determine the feasibility of deploying the technology and services of the vendors identified during the RFP process. [NORC at the University of Chicago](#) was selected to design the Pilot Study and to work with the other vendors to produce an accuracy report based upon the data collected during the pilot. A Preliminary Findings Paper was [published](#) prior to the [ICANN Los Angeles Meeting](#) to facilitate consultations with the ICANN community regarding

the methodology, findings, and next steps for the development of the ARS. ICANN received input from the [GAC](#), Registrar Stakeholder Group, the Intellectual Property Constituency and others on the Preliminary Findings Paper.

This Report summarizes the results of the Pilot Study, addresses issues raised by the community at the ICANN Los Angeles meeting, and provides specific recommendations for finalizing the design of the ARS.

ARS Pilot Study Design

The ARS is being designed to examine a WHOIS record for each of three validation perspectives: syntactic, operational, and identity validation. Applying a scoring methodology, the ARS will enable the publication of statistics based on certain accuracy labels, which will be reported on a regular basis and tracked over time. The ARS will be limited to WHOIS records in generic top level domains (gTLDs), and specifically excludes registrations in country code top level domains (ccTLDs), such as .jp, .cn or .uk.

The proposed process for conducting the ARS would assess the accuracy of email address, telephone number, and postal address information from a sample of WHOIS records, according to assessment protocol established for each of the accuracy perspectives to be considered (syntactic, operational, and possibly identity). The ARS Pilot Study explores and tests techniques for syntactic and operational validation, and will inform the mix of validation perspectives and tools that are utilized in the ARS. The results of validation will then be analyzed, and records will be categorized based on degree of accuracy.

ARS Report Requirements

In addition, the Accuracy Reporting System will support the examination of a statistically significant number of gTLD WHOIS records to enable a series of comparisons and reports to be published, including, for example:

- Accuracy levels associated with new gTLDs¹ and prior gTLDs (e.g., .com., .net, .org);
- Accuracy levels associated with registrations under the 2013 Registrar Accreditation Agreement (RAA) versus the 2009 version of the RAA;
- Compliance levels with regards to the validation and verification of certain WHOIS contact data requirements, as listed in the WHOIS Accuracy Program Specification to the RAA;
- Accuracy levels for registrars, registries and registrants located in each of the five ICANN Geographic Regions; and

¹ New gTLDs are top level domains introduced following the application round conducted by ICANN in 2012.

- Ranking of each ICANN accredited registrar and gTLD registry, based on reported accuracy levels.

ARS Pilot Study

The Pilot Study is a precursor to the ARS, designed with the intention of isolating and testing key components of the proposed ARS. The Pilot Study focused on the syntactic and operational perspectives of data validation, and due to complexity, cost and time constraints, does not employ the identity perspective. The chief differences between the Pilot Study and the ARS are the timeline for conducting the study, the sample sizes of WHOIS data, and the accuracy perspectives by which the data is assessed.

To summarize, the main goals for the Pilot Study are to:

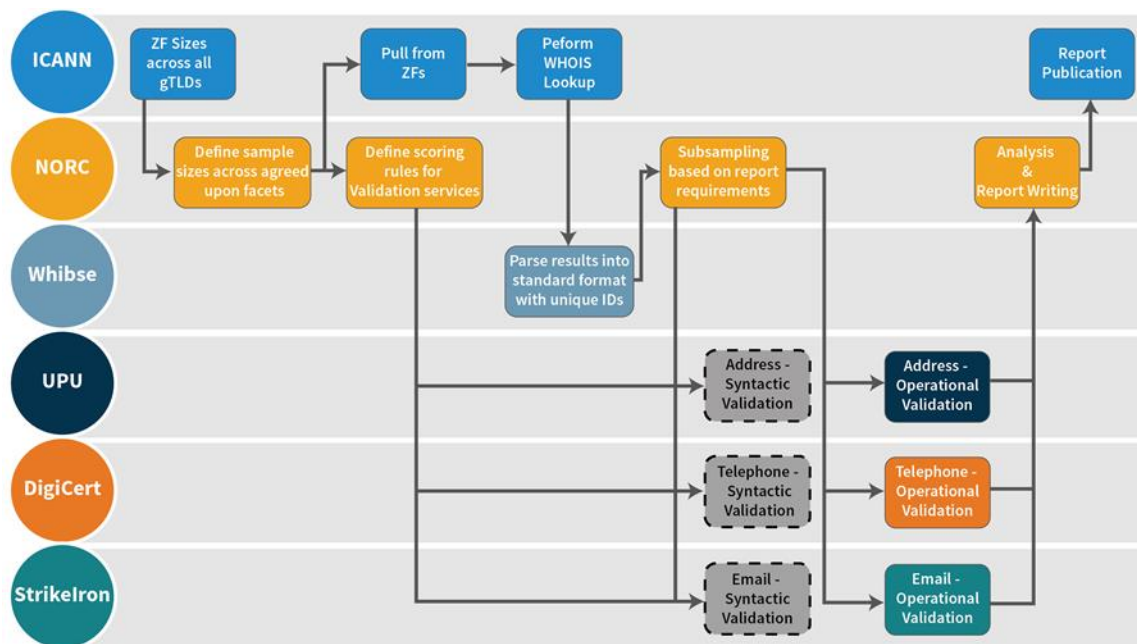
- i. To inform the framework for creating and conducting the ARS; and
- ii. To collect data related to the accuracy rates of WHOIS records.

The Pilot Study was a collaborative effort led by ICANN, with support from [NORC at the University of Chicago](#), [DigiCert, Inc.](#), [StrikeIron, Inc.](#), The [Universal Postal Union](#) (UPU), and [Whibse, Inc.](#) Initial planning for the Pilot began on 27 August 2014, and data collection began on 4 September 2014. Accuracy validation of the WHOIS records selected for review was completed on a compressed schedule between September 12 and 19. Analysis of the validation results concluded on 10 October 2014.

The Pilot Study design began with determining the parameters of the sample, which took into account subsampling considerations and the required sample size for each gTLD included in the study. A representative sample was determined using data from gTLD zone files (ZF's). Records selected from ZF's then underwent WHOIS lookups. The data collected from WHOIS lookups were parsed so that the data could be consistently interpreted by vendors. Representative subsamples of 10,000 and 1,000 were selected for validation from the entire sample of 98,821, with different vendors receiving differing numbers of domain names to validate. The vendors applied predetermined validation criteria to the email address, telephone number, and postal address data in order to assess accuracy from the syntactic perspective and from the operational

perspective. The validation results were compiled into datasets, separated by data type and by perspective, and these datasets were then analyzed to produce accuracy statistics (for new and prior gTLDs) stratified by gTLD, region, and registrar. The design of the Pilot Study flow is depicted below in *Figure A*.

Figure A: ARS Pilot Study Design Flow Chart



The work approach for the Pilot Study was divided into the following tasks:

1. Determine sample size, subsampling, and sample size for each gTLD
2. Sample selection
3. Record parsing
4. Selection of subsamples and delivery of analysis datasets
5. Define scoring rules for validation
6. Validate email address, telephone number, and postal address
7. Analyze datasets

While tasks 1-4 were required to be sequential, tasks 5-7 could be conducted concurrently with other tasks.

Sample Sizes of gTLDs

Table A1 in [Appendix A](#) lists the sample sizes for the top twenty-five gTLDs for the samples of 98,821, 10,000, and 1,000. The full sample was used for Syntactical and Operational email address validation. The 10,000 subsample was used for the Syntactical telephone number and postal address validation. The 1,000 subsample was used for the Operational telephone number and postal address validation.

Countries in the Sampling Frame

Table 1, below, shows the distribution by ICANN Geographic Region for each of the three sample sizes.

Table 1: Sample Sizes by Geographic Region, with Subsampling Rates

Geographic Region	Full Sample	10,000 Subsample	Selection Probability (%)	1,000 Subsample	Selection Probability (%)
Africa (AF)	617	617	100.0	100	16.2
Latin America/Caribbean (LAC)	3,380	1,000	29.6	150	15.0
Asia/Australia/Pacific (AP)	20,176	1,784	8.9	160	8.9
Europe (EUR)	23,417	2,071	8.9	186	8.9
North America (NA)	50,768	4,490	8.9	401	8.9
Missing (.)	463	38	8.9	3	8.9
TOTAL/OVERALL	98,821	10,000	10.1	1,000	10.0

Tables A2a, A2b, and A2c in [Appendix A](#) below show the distribution by country. A total of 200 countries were represented in the full sample of 98,821 domain names. Of these 200 countries, 91 countries were represented in the subsamples of 10,000 and 1,000 as well as the full sample (African and Latin American/Caribbean countries were oversampled in the subsamples). These 91 countries are in Table A2a. An additional 63 countries were represented in the 10,000 subsample, but not the 1,000 subsample. These 63 countries are in Table A2b. The remaining 46 countries are small enough to only be in the full sample, but not in the subsamples. These 46 countries are in Table A2c.

Accuracy Definitions

For the Pilot Study, each vendor applied its technology and expertise (as applicable) to analyze the postal address, email address, and telephone number associated with the WHOIS records in the sample from the different perspectives. These records were grouped into the superset categories of “Accurate” or “Inaccurate” which encompass the ARS accuracy rating scale categories as follows:

Accurate Categorization

- No Failure: Indicates that data are missing no significant information, and that the data provided are accurate using all three accuracy perspectives;
- Minimal Failure: Indicates that data are missing no significant information. The record may need clarification or might benefit from additional information, but the data provided are accurate; and
- Limited Failure: Indicates that the data are missing at least some significant information, but that the data present are to some degree considered useful.

Inaccurate Categorization

- Substantial Failure: Indicates that the data are missing significant information, and that the data provided are mostly inaccurate.
- Full Failure: Indicates that significant data elements are missing, or that data provided are overwhelmingly inaccurate.

Each vendor then determined how to apply its technology to fit into this rating scale. Not all sub-categories were used by the vendors in their analysis. For more information on how each vendor applied this methodology, please refer to [Appendix C](#).

Automated and Manual Approaches

The methods employed by vendors during the ARS Pilot Study included both automated and manual processes. While the ARS seeks to automate as much of the accuracy assessment process as possible, ICANN recognizes that automated approaches have limitations which should be supplemented by manual oversight and intervention.

For the vendor NORC at the University of Chicago, the sampling involved manual and automated processes, while the analysis was mostly automated. The initial sampling process took as input zone file sizes provided by ICANN. Sample sizes per zone for the full sample were determined manually. Instructions for selecting rows in which zone files were created with an automated tool. The strategies for selecting the subsamples of 10,000 and 1,000 were manually created, but the subsampling was done in an automated process. The analysis was automated, but some manual preparation was necessary (data cleaning, for example).

The email analysis conducted by StrikeIron was largely automated, using their proprietary technology to assess over 98,000 records within days. Given the ease by which this work was done, it may be worthwhile to continue to use a large sample size (i.e., 100,000 to analyze email addresses in the final design of the ARS).

The telephone analysis conducted by DigiCert included a mix of manual and automated processes. In cases where the country code for the telephone number was missing, DigiCert manually determined if it could make an assumption of the proper code based upon the remaining WHOIS record, prior to examining whether the appropriate number of digits were present. There was also a need to strip out unnecessary characters of a telephone number, such as a “+”, “()”, “. ” or extra spaces. To test the operational validity of the telephone number, DigiCert used an automated calling system as well as live persons with multiple language skills to decipher the responses received. The complexity of this process suggests deploying a smaller sample size than what may be feasible for emails, in the final design of the ARS.

The postal address validation services conducted by the UPU were also a mix of manual and automated processes. Because postal addresses have multiple fields, with multiple standards, this analysis proved to be relatively complex. Moreover, it is possible to deliver mail to

addresses where some elements are missing. The postal databases are sophisticated enough to determine whether an address is “serviceable” in some locations i.e., whether mail can be delivered to the address even if certain fields are missing. For example, if an address includes a zip code in the United States, but is missing the City, the UPU database extrapolates the correct information to deliver the package. However, the quality of available postal databases varies from country to country, making some assessments more difficult than others. The complexity of this process suggests selecting a smaller sample size for postal addresses than what may be feasible for emails in the final design of the ARS.

ARS Pilot Study Methods and Approach

Validation Perspectives

In the Pilot Study, researchers validated the email address, telephone number, and postal address fields from a subsample of records, by examining the data for factors of syntactic and operational accuracy. As mentioned in Section 1 of this report, the ARS proposes three perspectives by which accuracy can be assessed: syntactic, operational, and identity. For the ARS Pilot Study, researchers focused on just the syntactic and operational perspectives. These perspectives are defined below.

Syntactic Perspective. The syntactic perspective examines WHOIS data with the intent of determining which data exists or is missing from within WHOIS data fields, and assesses whether data meets specified standards of formatting and structure. Syntactic validation offers the highest degree of automation and therefore volume-efficiency in assessing accuracy. A limitation of Syntactic assessment is that it does not provide a conclusive assessment of accuracy, and should be supplemented by other forms of validation.

Operational Perspective. The Operational perspective examines WHOIS data with the intent of determining if the information can be practically used to achieve an intended function. Operational validation combines automated and manual processes, and involves gathering and utilizing external data to inform an accuracy assessment. As a tradeoff for a lower degree of automation, operational validation provides a more tangible notion of accuracy. A limitation of Operational assessment is that while it may confirm practical applicability of the data, it does not determine if the identifying data belongs to the actual domain registrant.

Identity Perspective. A third possible perspective, which was not conducted during the Pilot Study, but for which ICANN is exploring as potentially part of a later iteration of the ARS, would examine WHOIS records with the intent of confirming that WHOIS data corresponds to the person or entity being represented as the registrant by attempting to contact the registrant or entity using the information listed in the WHOIS record. ICANN intends to work with the ICANNCommunity in the coming months to explore the feasibility, costs, and benefits of examining the accuracy of WHOIS from the identity perspective.

Rating Scale Categories

An accuracy assessment of WHOIS data, guided by any of the three perspectives, should result in the data being categorized into a rating scale based on the degree to which it can be considered accurate. The rating scale categories initially established for WHOIS accuracy studies describe accuracy as an inclusive range between No Failure, the highest degree of accuracy, and Full Failure, which represents the lowest degree of accuracy. There are five rating scale categories in total, which are defined as:

- **No Failure:** Indicates that data are missing no significant information, and that the data provided are accurate.
- **Minimal Failure:** Indicates that data are missing no significant information. The record may need clarification or might benefit from additional information, but the data provided are accurate.
- **Limited Failure:** Indicates that the data are missing at least some significant information, but that the data present are to some degree considered useful.
- **Substantial Failure:** Indicates that the data are missing significant information, and that the data provided are mostly inaccurate.
- **Full Failure:** Indicates that significant data elements are missing, or that data provided are overwhelmingly inaccurate.

Email Address Validation

Syntactic validation for email addresses could be described as examining the information contained within the syntax of the address. If an email address was determined to be present in the parsed WHOIS data, the syntax of the email was checked for conformance with RFC specifications.

Operational validation of email addresses was performed by checking the domain within the address for validity and responsiveness and attempting communication with mail servers to determine whether an email address was accurate and functional. The communication attempts returned information on whether the domain, mail server, and email address could be reached, or

provided a reason why they could not be reached, and this data was used to code email addresses into rating scale categories.

See [Appendix C](#) for a full description of the email address validation process.

Telephone Number Validation

Syntactic validation of telephone numbers was performed by examining WHOIS records to determine if a telephone number was provided and, if so, evaluating the format of the number provided for compliance with expected telephone number requirements. Proper format was based on the country in which the number is located and, as a result of the assessment, telephone numbers were coded into a rating scale category.

Operational validation of telephone numbers was performed by evaluating each phone number to determine whether a connection could be established and whether additional information was needed to make a connection possible. Phone numbers were coded into rating scale categories based on whether the phone number as it was provided in the WHOIS record could connect successfully, whether the number required additional information in order to connect, or whether a connection was not possible.

See [Appendix C](#) for a full description of the telephone number validation process.

Postal Address Validation

Syntactic validation of postal address information was performed by examining WHOIS records to determine if a postal address was provided and, if so, whether the format of the address was consistent with expected norms and syntax requirements for the country in which it is located. When components of an address were missing or did not match syntax requirements, additional checks were made to see if the address could be considered serviceable. The expected norms and syntax requirements for these checks were based on guidance from the UPU July 2013 Postal Addressing Systems report and compliance with the UPU S42 Templates. Postal addresses were coded to a rating scale category based on the outcome of the assessment.

Operational validation of postal addresses was performed by determining whether the address provided in a record was an existing address according to UPU’s database of worldwide postal address information. Addresses that were syntactically accurate, whether they were accurate as provided or otherwise serviceable during the syntactic assessment, were included in the assessment and each record was subsequently coded to a rating scale category. It is important to note that validation of postal addresses is the most difficult of the three validations because multiple fields (number, street, city, and so on) must be examined and cross-validated.

See [Appendix C](#) for a full description of the postal address validation process.

Sample Design

The Pilot Study began with the selection of a sample of 100,000 records from gTLD Zone Files. Because New gTLDs² are a primary focus of the study, a high percentage of these domains were sampled, accounting for 25 percent of the domains represented in the sample (at the time of sampling, only 1.4 percent of domains were within the New gTLDs). In order to provide some information on all gTLDs, smaller gTLDs were oversampled. Specifically, all New gTLDs had at least 10 records selected (or all if fewer than 10 records were available) and all other (“Prior”) gTLDs had at least 30 records selected (or all if fewer than 30 records were available). Once these records had been selected, the remaining records were selected proportional to the number of records in each gTLD (separately for New and Prior gTLDs), with the exception that the prevalence of the top four New gTLDs and the Prior gTLD “.com” be allowed selection of less than the full proportion of these records to allow slight increases in the proportion of smaller gTLDs. Within each gTLD, a systematic sample was specified (with a starting record and skip interval) to implicitly stratify by any partial or complete sorting within the gTLD Zone File (newer domains at the top or bottom, for example). All of the data provided in this document is unweighted, and while smaller gTLDs and all New gTLDs are overrepresented, the numbers provided here are roughly representative. Weights will be calculated as the study progresses as part of the full ARS operation.

² New gTLDs are generic top level domains that have been launched as a result of the application round conducted by ICANN in 2012.

Once the records were selected for the sample, they were then queried via WHOIS lookups to produce data for the Pilot Study. A critical next step was to parse the raw WHOIS data so that the information in the records could be consistently interpreted during the validation phase of the study. After parsing, all email address records moved forward to the syntactical and operational validation phase. The number of successful queries (and parsed records) was 98,821. A subsample of 10,000 records was selected for syntactical validation of postal addresses and telephone numbers. In order to provide estimates at the Geographic Region level, NORC at the University of Chicago determined the country (and Geographic Region) for as many of the domain names as possible, and oversampled the two smallest Regions: Africa and Asia/Australia/Pacific.

Within each Geographic Region, the file was serpentine sorted for systematic sampling to guarantee coverage of subgroups. The sorting variables were:

- 1) New versus Prior gTLDs;
- 2) RAA 2009 registrars versus RAA 2013 registrars;
- 3) gTLD;
- 4) Registrar;
- 5) Country; and
- 6) Place in Zone File.

Serpentine sorting makes neighbor records more similar than regular sorting (within new gTLDs, for example, RAA 2009 registrars are last, but within Prior gTLDs, RAA 2009 registrars are first to keep all RAA 2009 registrars together in the sorted list), and results in implicit stratification by the sort variables. A smaller subsample of 1,000 records was drawn (again, Africa and Asia/Australia/Pacific were oversampled) for operational validation of postal addresses and telephone numbers. The same set of sort variables were used to sort the 10,000 records in a serpentine way before selecting a systematic sample of 1,000 records. Table 1, above, shows the distribution by ICANN Geographic Region for each of the three sample sizes. [Appendix A](#) provides the sample sizes for the twenty-five largest gTLDs (Table A1) and for all the countries with selected domains (Tables A2a, A2b, and A2c).

The full sample was used for email address validation. Syntactical verification for telephone numbers and postal addresses used the 10,000 subsample, while the Operational verification for telephone numbers and postal addresses used the 1,000 subsample. The analysis reports on subgroups with at least 20 records, and only the top seven gTLDs met this criterion in the 1,000 subsample.

It is important to note that the different sample sizes result in a very different uncertainty in the estimates below. Specifically, for the email address verification, small differences are statistically significant while for the operational validation of telephone numbers and postal addresses, differences of several percentage points might not be statistically significant. Table 2 gives standard errors and 95 percent confidence intervals for our overall estimates of accuracy:

Table 2: Impact of Sample Size on Standard Errors

	Accuracy Percentage: All gTLDs	Sample Size	Standard Error (%)	95% Confidence Interval
Syntactical:				
Email	99.9	98,821	0.01	(99.7, 100.0)
Telephone	88.7	10,000	0.32	(88.1, 89.3)
Postal	81.2	10,000	0.39	(80.4, 82.0)
Operational:				
Email	89.2	98,821	0.10	(89.0, 89.4)
Telephone	72.4	1,000	1.41	(69.6, 75.2)
Postal	82.4	1,000	1.20	(80.0, 84.8)

Table 2 shows that the standard errors for Email address accuracy are very small, while the standard errors for Operational accuracy for telephone number and postal are much larger than the standard errors for Syntactical accuracy. The standard errors for subgroups within these three sample sizes are larger.

Table 2 shows that the standard errors for Email address accuracy are very small, while the standard errors for Operational accuracy for telephone number and postal are much larger than the standard errors for Syntactical accuracy. The standard errors for subgroups within these three sample sizes are larger.

Main Findings

The ARS Pilot Study analysis results are summarized below to show top-level accuracy statistics by New and Prior gTLDs, by 2009 and 2013 RAAs, and by Geographic Region. These results are further broken down to show accuracy rates from syntactic and operational perspectives for email address, postal address, and telephone number information. Overall accuracy rates for the 10 gTLDs and registrars that had the largest sample sizes in the study are also reported towards the end of this summary of findings.

The scope of the Pilot Study does not separately address domains that were registered using privacy or proxy services (PP). The existence of these types of registrations is believed to have an impact on the data collected during this study, and thus ICANN would like to engage with community members to discuss and identify potential PP considerations and impacts in preparation for future iterations of the ARS.

True of any study that involves data analysis, the methodological context by which data is analyzed will heavily impact the findings. For example, during the ARS Pilot Study, ICANN found that accuracy assessments of a physical address proved to be significantly more difficult to perform than accuracy assessments of an email address. This was due to the wide variety of postal address norms and expectations that exist across different countries. While email address information may not necessarily be more accurate than postal address information, it is certainly easier to assess in terms of the effort required. ICANN acknowledges that the methodology used in the ARS Pilot Study was designed to explore possible frameworks for assessing WHOIS accuracy, rather than to provide a comprehensive solution to assessing accuracy.

Accuracy by Validation Perspective

As seen in Table 3, below, the accuracy of email address, telephone number, and postal address information was assessed from the syntactic and operational perspectives, and group statistics were computed for all gTLDs, 2009 and 2013 RAA, and for New and Prior gTLDs. Table 3 shows the rates of records which were assessed as accurate, which includes all records categorized as No Failure, Minimal Failure, and Limited Failure.

Table 3: Accuracy Percentages by Perspective, across RAAs and New vs. Prior gTLDs

	All gTLDs	2009 RAA	2013 RAA	Prior gTLDs	New gTLDs
Syntactical:					
Email	99.9	99.9	99.9	99.9	99.9
Telephone	88.7	90.1	89.9	88.0	90.9
Postal	81.2	85.9	81.0	81.8	79.1
Operational:					
Email	89.2	81.7	89.8	88.9	90.2
Telephone	72.4	79.6	73.5	74.3	66.5
Postal	82.4	85.7	82.4	82.4	82.4

Table 3 shows that email address and telephone number syntactical accuracy is not statistically significantly different for Prior and New gTLDs while syntactical accuracy is better for Prior gTLDs on postal addresses. It should be noted that although a registrar may have been under the 2013 RAA at the time of sampling, the sampled domain may have been registered when the registrar was under the 2009 RAA. The additional validation and verification requirements of the 2013 RAA are not retroactive to domain names registered through the registrar, and are only triggered under the situations enumerated in the Whois Accuracy Program Specification of the 2013 RAA. The operational numbers for email addresses indicate that New gTLDs have slightly better accuracy than Prior gTLDs. Prior gTLDs do have more operational accuracy on telephone numbers, but the two groups are equal on operational postal address accuracy. One possible explanation for the increased accuracy for New gTLDs on email addresses and not the postal addresses is that registrations in New gTLDs are subject to the requirements of the 2013 RAA, which requires verification of either the email address or telephone number, but not the postal address.

Table 3 also shows that the differences in syntactical accuracy between registrars under the 2013 RAA and registrars under the 2009 RAA for email address, telephone numbers and postal addresses are not statistically significant. Operationally, registrars under the 2013 RAA have more accuracy for email addresses than registrars under the 2009 RAA. Registrars under the 2009 RAA have more operational accuracy in telephone numbers while the two groups are similar on postal addresses. Further analysis is needed to determine the reason for why registrars

under the 2013 RAA experience less operational accuracy on telephone numbers, since the data collected in the Pilot is not sufficient to explore possible reasons.

Overall, it is not surprising that lower syntactical accuracy occurs for postal addresses because more than one field is involved. Operationally, the accuracy of telephone numbers is lowest while the accuracy of email addresses is highest.

Similar to Table 3, Table 4 reports on accuracy percentages of email address, telephone number, and postal address information by ICANN Geographic Region.

Table 4: Accuracy Percentages by Perspective, across ICANN Geographic Regions

	All Geographic Regions	Africa	Asia Australia Pacific	Europe	Latin America Caribbean	North America
Syntactical:						
Email	99.9	99.5	99.9	99.9	99.9	99.9
Telephone	88.7	67.4	82.6	85.9	88.9	96.1
Postal	81.2	67.4	74.9	91.2	92.0	79.2
Operational:						
Email	89.2	90.9	83.2	90.0	89.8	91.1
Telephone	72.4	52.0	57.5	65.6	76.7	85.5
Postal	82.4	69.0	71.9	93.6	92.0	81.8

Table 4 shows that domain names in Africa have the lowest syntactical accuracy in postal addresses and telephone numbers, but syntactical accuracy in email addresses is similar across the Geographic Regions. Operationally, domains in Africa and Asia/Australia/Pacific have the lowest accuracy for both postal addresses and telephone numbers, but Africa and North America have the highest accuracy for email address.

Accuracy by gTLD

In contrast to the tables above, Tables 5 and 6 report on email address validation for the gTLDs with the ten largest sample sizes in the Pilot Study. The names of these gTLDs are not reported because this is a pilot study, and the accuracy percentages for telephone numbers and postal

addresses are not reported due to sample size concerns. Table 5 reports the Syntactical accuracy percentages for email addresses.

Table 5: Syntactical Email Address Accuracy Percentages for gTLDs with Largest Sample Sizes

TLD	Accuracy
gTLD A	100.00
gTLD B	100.00
gTLD C	100.00
gTLD D	100.00
gTLD E	99.99
gTLD F	99.95
gTLD G	99.94
gTLD H	99.92
gTLD I	99.89
gTLD J	99.88

Table 6 reports the Operational accuracy percentages for email addresses for the gTLDs with the largest sample sizes. Note that the ordering of these gTLDs is different from than that based on Syntactical accuracy.

Table 6: Operational Email Address Accuracy Percentages for gTLDs with Largest Sample Sizes

TLD	Accuracy
gTLD A	91.62
gTLD D	96.53
gTLD B	93.63
gTLD J	92.54
gTLD F	91.46
gTLD I	91.19
gTLD E	90.88
gTLD C	90.65
gTLD H	88.93
gTLD G	88.02

Accuracy by Registrar

Tables 7 and 8 report on the accuracy percentages for email addresses by registrar. Table 7 reports the Syntactical accuracy percentages for email addresses for the registrars with the largest sample sizes.

Table 7: Syntactical Email Address Accuracy Percentages for Registrars with Largest Sample Sizes

Registrar	Accuracy
Registrar A	100.00
Registrar B	100.00
Registrar C	100.00
Registrar D	100.00
Registrar E	100.00
Registrar F	99.97
Registrar G	99.95
Registrar H	99.95
Registrar I	99.93
Registrar J	99.92

Table 8 reports the Operational accuracy percentages for email addresses for the registrars with the largest sample sizes. As with the gTLDs with the largest sample sizes, the Operational accuracy order for registrars changes compared to the Syntactical accuracy order.

Table 8: Operational Email Address Accuracy Percentages for Registrars with Largest Sample Sizes

Registrar	Accuracy
Registrar F	94.6
Registrar D	94.4
Registrar I	94.0
Registrar A	94.0
Registrar E	93.9
Registrar C	92.6
Registrar G	92.6
Registrar H	92.5
Registrar B	92.3
Registrar J	89.6

Implications of Findings

The Pilot Study was a highly informative exercise, and provided valuable lessons and caveats for the methods and procedures of the ARS. Some of the key lessons learned from the pilot are detailed below.

Examining Definitions

In the coming months, ICANN will refine elements of the study that will be critical for establishing a replicable framework for implementing the ARS. The rules for determining accurate and inaccurate WHOIS records need to be considered further. The vendors that performed the validation steps provided reasonable assessments of postal address, email address and telephone number accuracy. However, more thought needs to be given as to whether the assessments fit the definitions of the accuracy categories (No Failures, Minimal Failure, Limited Failure, Substantial Failure, and Complete Failure). An evaluation of the methodologies will be conducted to determine ways to make phone, email and postal address accuracy evaluations comparable.

A key aspect of designing the ARS includes a careful examination of the methodology to be deployed for each perspective, to apply the rating scale. This means examining the error codes that may be generated by the vendor's technology to confirm how it affects accuracy. For example, should an error code that an email address is associated with a malicious website affect the accuracy of the WHOIS record? If a mail server is set up to "catch all" emails sent to the domain, regardless of whether the specific email address was created, should this be considered operationally accurate or inaccurate?

ICANN is interested in receiving input on these aspects of the ARS design during the public comment period on the Pilot Report.

Determining an Accuracy Score

While a lot can be learned from analyzing the accuracy percentages for email address, telephone number, and postal address, a perspective accuracy score and an overall score across perspectives would provide more information about the accuracy state of WHOIS records. In what follows, an outline of a scoring methodology is described that would use the accuracy evaluation of each sampled record to determine a perspective (syntactical, operational) accuracy score and an overall accuracy score. These record scores can be used to analyze the accuracy status of groups of records (by gTLD, country, etc.).

As in the pilot study, the email address, telephone number, and postal address fields for each sampled WHOIS record are evaluated to determine if the field is accurate or inaccurate. Recall that a record's field is considered accurate if its rating is limited failure, minimal failure, or no failure. The record's field is considered inaccurate if its rating is substantial failure or full failure. The accurate/inaccurate determination can be converted to a numerical value by assigning the value 1 to an accurate field and -1 to an inaccurate field. A record's syntactical or operational perspective score is the sum of the email address, telephone number and postal address accuracy value. As shown in Table 9, a record's perspective accuracy score will be -3, -1, 0, 1, or 3.

Table 9: Perspective Accuracy Scoring

Accuracy Value			Record's Perspective Accuracy Score
Email Address	Telephone Number	Postal Address	
-1	-1	-1	-3
-1	-1	1	-1
-1	1	-1	
1	-1	-1	
1	1	-1	1
1	-1	1	
-1	1	1	
1	1	1	3

Perspective accuracy scores can be aggregated across a group of records. The percentage of records within the group that have scores of -3, -1, 1, or 3 can be calculated, and these

percentages can be compared across groups. For example, one could calculate the percentage of records with accuracy scores of 1 or 3 for New gTLDs and compare it to a similar percentage for Prior gTLDs. Additionally, one could average the scores across all records in a group to obtain group perspective composite scores, which will be values between -3 and 3. Group composite scores can also be statistically compared across groups of interest. Furthermore, the perspective composite score can be mapped back to the five accuracy rating scale categories. Table 10 provides a proposed mapping.

Table 10: Proposed Perspective Accuracy Rating Scale to Composite Score Mapping

<u>Accuracy Rating Scale</u>	<u>Perspective Composite Score Range</u>
No Failure	2 – 3
Minimal Failure	0.5 – 2
Limited Failure	-0.5 – 0.5
Substantial Failure	-2 – -0.5
Full Failure	-3 – -2

A record’s overall accuracy score can be determined by averaging the record’s perspective scores. The overall record score can then be averaged across a group of records to provide an accuracy rating for the group of interest. Statistical comparisons between groups can be considered as with perspective scores.

Sample Size and Confidence Levels

It takes more time to validate postal addresses and telephone numbers than it does email addresses. Furthermore, it takes longer to perform operational validation than syntactical validation for both postal addresses and telephone numbers. Because of this, different sample sizes were used. This means that the accuracy measures based on smaller samples are less reliable than those the used on larger samples. In particular, operational validation for postal address and telephone number used a sample of only 1,000 WHOIS records. Table 2 shows the standard errors for the entire 1,000 records as greater than 1 percent. For smaller subgroups such as Geographic Regions, registrars and gTLDs, the standard errors are much larger, and therefore 1,000 is too small to provide good estimates for comparing Geographic Regions, registrars and

gTLDs. As such, a larger sample should be considered in the future. We suggest that an optimal minimum size for any analysis should be 5,000 or 10,000, but even increasing the minimum size from 1,000 to 2,000 would greatly increase the value of the data analysis.

Not a Compliance Report

The Pilot Report was not designed to be a report that assessed compliance with the new 2013 RAA validation and verification requirements. Instead, it examined accuracy from different perspectives, as described above. In the final design of the ARS, ICANN plans to develop and publish additional reports that attempt to more closely align with the 2013 RAA requirements.

Exploring Identity Validation

The Pilot Study did not include an assessment of accuracy from the identity perspective. In the coming months, ICANN plans to explore the possibility of designing the ARS to include an examination of a WHOIS record from the identity perspective.

NEXT STEPS

ICANN plans to publish this Report for public comment for a period that closes on February 27, 2015. During the Public Comment, ICANN is seeking input on:

1. Study Design & Methodology
2. Types of Accuracy Reports to be published through the ARS
3. Whether ICANN should conduct Identity Validation in subsequent phases of the ARS Development
4. Whether the methodology should treat registrations under privacy or proxy services differently, and if so, how
5. Any other aspect of the ARS

This feedback on the Pilot Report during this Public Comment period, as well as via community dialogue at the [ICANN-52 Singapore](#) Meeting in February, 2015, will inform ICANN Staff and help shape the future development the Accuracy Reporting System.

[Appendix B](#) contains additional information from ICANN regarding the next steps for finalizing the development of the ARS.

APPENDIX A. Sample Sizes by gTLD and by Country.

Table A1: Sample Sizes for the Top Twenty-Five gTLDs

gTLD	Full Sample	10,000 Subsample	Selection Probability (%)	1,000 Subsample	Selection Probability (%)
.com	44,399	4,518	10.2	451	10.0
.net	12,539	1,261	10.1	129	10.2
.org	8,726	883	10.1	89	10.1
.info	4,725	456	9.7	44	9.6
.xyz (New)	3,282	293	8.9	26	8.9
.biz	2,215	223	10.1	24	10.8
.berlin (New)	1,504	134	8.9	12	9.0
.club (New)	1,162	141	12.1	17	12.1
.guru (New)	805	82	10.2	10	12.2
.mobi	781	84	10.8	9	10.7
.wang (New)	729	64	8.8	6	9.4
.photography (New)	514	52	10.1	6	11.5
.email (New)	457	49	10.7	7	14.3
.link (New)	439	48	10.9	5	10.4
.xn--3ds443g (New)	420	38	9.0	3	7.9
.xn--55qx5d (New)	420	37	8.8	4	10.8
.today (New)	419	47	11.2	5	10.6
.tips (New)	343	37	10.8	5	13.5
.company (New)	330	36	10.9	3	8.3
.xn--io0a7i (New)	296	27	9.1	2	7.4
.solutions (New)	291	29	10.0	3	10.3
.center (New)	265	28	10.6	4	14.3
.tokyo (New)	255	23	9.0	2	8.7
.asia	253	25	9.9	1	4.0
.expert (New)	247	25	10.1	3	12.0

Table A2a: Sample Sizes for all 91 Countries in the 1,000 Subsample

Country	Full Sample (98,821)	Percentage of Sample	10,000 Subsample	1,000 Subsample
Missing	463	0.47	38	3
United States (US)	46,995	47.65	4,151	367
China (CN)	7,347	7.45	653	57
Germany (DE)	6,402	6.49	570	43
United Kingdom (UK)	4,779	4.85	431	42
Canada (CA)	3,711	3.76	334	34
Japan (JP)	3,349	3.40	298	30
Australia (AU)	3,103	3.15	277	22
France (FR)	2,296	2.33	192	16
Panama (PA)	1,661	1.68	491	77
Spain (ES)	1,511	1.53	147	19
Netherlands (NL)	1,237	1.25	103	8
Turkey (TR)	1,204	1.22	104	8
India (IN)	1,123	1.14	100	5
Italy (IT)	939	0.95	78	7
Cayman Islands (KY)	880	0.89	80	9
Switzerland (CH)	792	0.80	70	4
New Zealand (NZ)	780	0.79	67	7
Russian Federation (RU)	569	0.58	60	3
Austria (AT)	497	0.50	37	2
Brazil (BR)	485	0.49	149	26
Korea, Republic of (KR)	484	0.49	47	4
Sweden (SE)	353	0.36	32	5
Mexico (MX)	338	0.34	99	16
Hong Kong (HK)	310	0.31	25	6

Table A2a: Sample Sizes for all 91 Countries in the 1,000 Subsample

Country	Full Sample (98,821)	Percentage of Sample	10,000 Subsample	1,000 Subsample
Belgium (BE)	295	0.30	24	3
Indonesia (ID)	290	0.29	27	3
Viet Nam (VN)	288	0.29	23	1
Ireland (IE)	283	0.29	26	4
Poland (PL)	281	0.28	27	3
Thailand (TH)	279	0.28	22	3
Norway (NO)	264	0.27	24	2
South Africa (ZA)	223	0.23	223	35
Denmark (DK)	196	0.20	19	1
Finland (FI)	178	0.18	16	2
Singapore (SG)	174	0.18	18	3
Bermuda (BM)	170	0.17	14	1
United Arab Emirates (AE)	166	0.17	15	2
Malaysia (MY)	165	0.17	16	1
Ukraine (UA)	156	0.16	14	1
Portugal (PT)	136	0.14	12	2
Argentina (AR)	132	0.13	42	4
Colombia (CO)	126	0.13	39	4
Bulgaria (BG)	112	0.11	11	3
Greece (GR)	111	0.11	7	1
Saudi Arabia (SA)	110	0.11	10	1
Taiwan (TW)	100	0.10	8	1
Philippines (PH)	97	0.10	6	1
Egypt (EG)	91	0.09	91	16
Pakistan (PK)	87	0.09	6	1
Peru (PE)	80	0.08	24	3

Table A2a: Sample Sizes for all 91 Countries in the 1,000 Subsample

Country	Full Sample (98,821)	Percentage of Sample	10,000 Subsample	1,000 Subsample
Antigua and Barbuda (AG)	78	0.08	22	2
Nigeria (NG)	73	0.07	73	10
Cyprus (CY)	59	0.06	5	1
Luxembourg (LU)	58	0.06	5	1
Costa Rica (CR)	56	0.06	15	2
Venezuela (VE)	55	0.06	12	1
Chile (CL)	52	0.05	16	1
Virgin Islands, British (VG)	48	0.05	5	1
Saint Kitts and Nevis (KN)	46	0.05	13	2
Lebanon (LB)	45	0.05	7	1
Morocco (MA)	38	0.04	38	5
Ecuador (EC)	36	0.04	10	3
Sri Lanka (LK)	25	0.03	1	1
Dominican Republic (DO)	24	0.02	5	2
Belarus (BY)	23	0.02	3	1
Kenya (KE)	22	0.02	22	4
Uruguay (UY)	22	0.02	8	2
Guatemala (GT)	21	0.02	8	1
Lithuania (LT)	17	0.02	2	1
Seychelles (SC)	16	0.02	16	2
Latvia (LV)	16	0.02	1	1
Ghana (GH)	15	0.02	15	4
Namibia (NA)	14	0.01	14	3
Bahamas (BS)	14	0.01	5	3
Uganda (UG)	13	0.01	13	3
Algeria (DZ)	13	0.01	13	1

Table A2a: Sample Sizes for all 91 Countries in the 1,000 Subsample

Country	Full Sample (98,821)	Percentage of Sample	10,000 Subsample	1,000 Subsample
St Vincent & Grenad. (VC)	13	0.01	3	1
Tunisia (TN)	11	0.01	11	4
Swaziland (SZ)	10	0.01	10	1
Uzbekistan (UZ)	7	0.01	1	1
Congo, Demo. Rep. (CD)	6	0.01	6	1
Cote d'Ivoire (CI)	6	0.01	6	1
Western Sahara (EH)	6	0.01	6	1
Zimbabwe (ZW)	5	0.01	5	1
Botswana (BW)	4	0.00	4	2
Burkina Faso (BF)	2	0.00	2	1
Benin (BJ)	2	0.00	2	1
Madagascar (MG)	2	0.00	2	1
Mali (ML)	2	0.00	2	1
Angola (AO)	1	0.00	1	1
Cape Verde (CV)	1	0.00	1	1

Table A2b: Sample Sizes for all other 63 Countries in the 10,000 Subsample

Country	Full Sample (98,821)	Percentage of Sample (%)	10,000 Subsample
Israel (IL)	173	0.18	12
Czech Republic (CZ)	165	0.17	14
Romania (RO)	109	0.11	10
Curaçao (CW)	83	0.08	6
Iran, Islamic Rep. (IR)	79	0.08	7
Hungary (HU)	71	0.07	6
Yug. (Serbia/Montenegro) (YU)	50	0.05	2
Slovakia (SK)	49	0.05	8
Croatia (HR)	46	0.05	4
Afghanistan (AF)	39	0.04	2
Bangladesh (BD)	37	0.04	2
Jordan (JO)	36	0.04	3
Kuwait (KW)	34	0.03	5
Puerto Rico (PR)	31	0.03	3
Barbados (BB)	30	0.03	12
Malta (MT)	30	0.03	1
Slovenia (SI)	30	0.03	1
Trinidad and Tobago (TT)	21	0.02	6
Azerbaijan (AZ)	21	0.02	2
Estonia (EE)	20	0.02	1
Belize (BZ)	18	0.02	5
Bolivia (BO)	18	0.02	3
Gibraltar (GI)	18	0.02	1
Cambodia (KH)	18	0.02	1

Table A2b: Sample Sizes for all other 63 Countries in the 10,000 Subsample

Country	Full Sample (98,821)	Percentage of Sample (%)	10,000 Subsample
Jamaica (JM)	16	0.02	5
Macedonia, FYR (MK)	13	0.01	1
Qatar (QA)	12	0.01	2
El Salvador (SV)	11	0.01	3
Guadeloupe (GP)	11	0.01	1
Nepal (NP)	11	0.01	1
Iceland (IS)	10	0.01	1
Tanzania, United Rep. (TZ)	9	0.01	9
Paraguay (PY)	8	0.01	2
Georgia (GE)	8	0.01	1
Palau (PW)	8	0.01	1
Iraq (IQ)	7	0.01	2
Myanmar (MM)	7	0.01	2
Macao (MO)	7	0.01	2
American Samoa (AS)	7	0.01	1
Brunei Darussalam (BN)	7	0.01	1
Northern Mariana Islands (MP)	7	0.01	1
Mauritius (MU)	6	0.01	6
Andorra (AD)	6	0.01	1
Armenia (AM)	6	0.01	1
St Pierre and Miquelon (PM)	6	0.01	1
Senegal (SN)	5	0.01	5
Nicaragua (NI)	5	0.01	2
Niger (NE)	4	0.00	4
Guernsey (GG)	4	0.00	1

Table A2b: Sample Sizes for all other 63 Countries in the 10,000 Subsample

Country	Full Sample (98,821)	Percentage of Sample (%)	10,000 Subsample
Ethiopia (ET)	3	0.00	3
Turks and Caicos Islands (TC)	3	0.00	1
Cameroon (CM)	2	0.00	2
Sudan (SD)	2	0.00	2
Somalia (SO)	2	0.00	2
Zambia (ZM)	2	0.00	2
Dominica (DM)	2	0.00	1
Burundi (BI)	1	0.00	1
Bhutan (BT)	1	0.00	1
Congo (CG)	1	0.00	1
Djibouti (DJ)	1	0.00	1
Libyan Arab Jamahiriya (LY)	1	0.00	1
Malawi (MW)	1	0.00	1
Togo (TG)	1	0.00	1

Table A2c: Sample Sizes for all other 46 Countries in the Full Subsample

Country	Full Sample (98,821)	Percentage of Sample (%)	10,000 Subsample
Albania (AL)	16	0.02	0
Liechtenstein (LI)	12	0.01	0
Bosnia and Herzegovina (BA)	11	0.01	0
Bahrain (BH)	10	0.01	0
Kazakhstan (KZ)	10	0.01	0
Monaco (MC)	10	0.01	0
Norfolk Island (NF)	10	0.01	0
Isle of Man (IM)	9	0.01	0
US Minor Islands (UM)	8	0.01	0
Cocos (Keeling) Islands (CC)	7	0.01	0
Moldova, Republic of (MD)	7	0.01	0
San Marino (SM)	7	0.01	0
French Southern Terr.(TF)	6	0.01	0
Virgin Islands, U.S. (VI)	6	0.01	0
Honduras (HN)	5	0.01	0
Reunion (RE)	5	0.01	0
Vanuatu (VU)	5	0.01	0
Maldives (MV)	4	0.00	0
Oman (OM)	4	0.00	0
Guam (GU)	3	0.00	0
Mongolia (MN)	3	0.00	0
French Polynesia (PF)	3	0.00	0
Papua New Guinea (PG)	3	0.00	0
Palestinian Territories (PS)	3	0.00	0
Yemen (YE)	3	0.00	0

Table A2c: Sample Sizes for all other 46 Countries in the Full Subsample

Country	Full Sample (98,821)	Percentage of Sample (%)	10,000 Subsample
Jersey (JE)	2	0.00	0
Kyrgyzstan (KG)	2	0.00	0
Saint Lucia (LC)	2	0.00	0
Marshall Islands (MH)	2	0.00	0
Suriname (SR)	2	0.00	0
Ascension Island (AC)	1	0.00	0
Aruba (AW)	1	0.00	0
Cuba (CU)	1	0.00	0
Christmas Island (CX)	1	0.00	0
Grenada (GD)	1	0.00	0
Greenland (GL)	1	0.00	0
Haiti (HT)	1	0.00	0
Korea, Democratic People (KP)	1	0.00	0
Lao People (LA)	1	0.00	0
New Caledonia (NC)	1	0.00	0
Saint Helena (SH)	1	0.00	0
Syrian Arab Republic (SY)	1	0.00	0
Tajikistan (TJ)	1	0.00	0
Timor-Leste (TL)	1	0.00	0
Tuvalu (TV)	1	0.00	0
Wallis And Futuna (WF)	1	0.00	0

Appendix B: Next Steps for the Development of the WHOIS Accuracy Reporting System (ARS)

This Paper is published by ICANN to accompany the NORC Final Accuracy Pilot Report (NORC Report) to describe the next steps for development of the WHOIS Accuracy Reporting System (ARS).

Public Comment and Consultations in Singapore

As described in the NORC Report, a public comment forum will be open until February 27, 2015. In addition, ICANN plans to engage with affected stakeholders at the [ICANN 52 Meeting in Singapore](#) in February, 2015, to obtain feedback on the proposed design for the ARS as described in the NORC Report. Upon review of these comments and feedback, ICANN will update the design of the ARS and begin development in phases, as described below.

During the Public Comment, ICANN is seeking input on:

1. Study Design & Methodology
2. Types of Accuracy Reports to be published through the ARS
3. Whether ICANN should conduct Identity Validation in subsequent phases of the ARS Development
4. Whether the methodology should treat registrations under privacy or proxy services differently, and if so, how
5. Any other aspect of the ARS

a. **Phase 1 – Syntactic Accuracy**

Phase 1 will focus on reporting accuracy levels that examine the syntactic aspects of the email, telephone number, and postal addresses. As these examinations can be conducted with the use of largely automated processes, it is expected that this portion of the ARS can be launched in early to mid-2015. A **Final Implementation Plan - Phase 1** will be developed to specify the sample size, process, and classification methodology to be deployed for examining the accuracy of WHOIS from the syntactic perspective.

b. **Phase 2 – Operational Accuracy**

Next, ICANN plans to further develop the ARS to report accuracy levels that examine WHOIS records from an operational perspective, with regard to the email, telephone number and post addresses. As this examination involves largely manual processes, smaller sample sizes are likely to be deployed. ICANN will develop a **Final Implementation Plan - Phase 2**, to take into account any lessons learned during Phase 1, for a launch in mid-late 2015.

c. **Phase 3 – Exploring Accuracy from an Identity Perspective**

The final phase examines whether and how to conduct ongoing accuracy studies from the perspective of confirming the identity of the registrant. The Governmental Advisory Committee (GAC) [Los Angeles Communiqué](#) has advised ICANN to take steps to scope and examine the risks, feasibility, costs and benefits of conducting ongoing accuracy studies to validate and verify the identity of the registrant. Staff is in the process of developing a response to the GAC that will provide this additional information in January, in advance of the [ICANN 52 Singapore Meeting](#). Phase 3 of ARS will assess the feasibility and costs to conduct accuracy studies based on the identity of the registrant.

Process for Correcting Inaccurate Records: ARS Implementation Advisory Group

A key function of the ARS will be to forward records identified as potentially inaccurate to registrars for follow-up to confirm their accuracy. The ARS is being designed to track and report on the progress of these records.

Initially, ICANN will kick-off a Compliance Pilot in January 2015, to examine the results of the Pilot to determine if a compliance response is appropriate for the WHOIS records that have been categorized as inaccurate from a syntactical perspective. ICANN's Contractual Compliance Department is in the process of auditing the results of the Pilot Study as part of a compliance pilot to determine whether a compliance follow-up is needed.

In addition, ICANN plans to engage with registrars and other interested stakeholders as necessary to define how to integrate transmitting, reviewing, and updating, as appropriate, the volume of identified WHOIS records that have been identified as potentially inaccurate into the existing Compliance processes and systems. ICANN plans to work with registrars and the broader Community in the months ahead in order to develop and refine this process.

Appendix C: DETAILED DESCRIPTIONS OF VENDOR PROCESSES

For the Pilot Study, each vendor applied their proprietary technology and expertise (as applicable) to analyze the postal address, email address, and telephone number associated with the WHOIS records in the sample from the different perspectives. These records were grouped into the superset categories of “Accurate” or “Inaccurate” which encompass the ARS accuracy rating scale categories as follows:

Accurate Categorization

- No Failure: Indicates that data are missing no significant information, and that the data provided are accurate using all three accuracy perspectives;
- Minimal Failure: Indicates that data are missing no significant information. The record may need clarification or might benefit from additional information, but the data provided are accurate; and
- Limited Failure: Indicates that the data are missing at least some significant information, but that the data present are to some degree considered useful.

Inaccurate Categorization

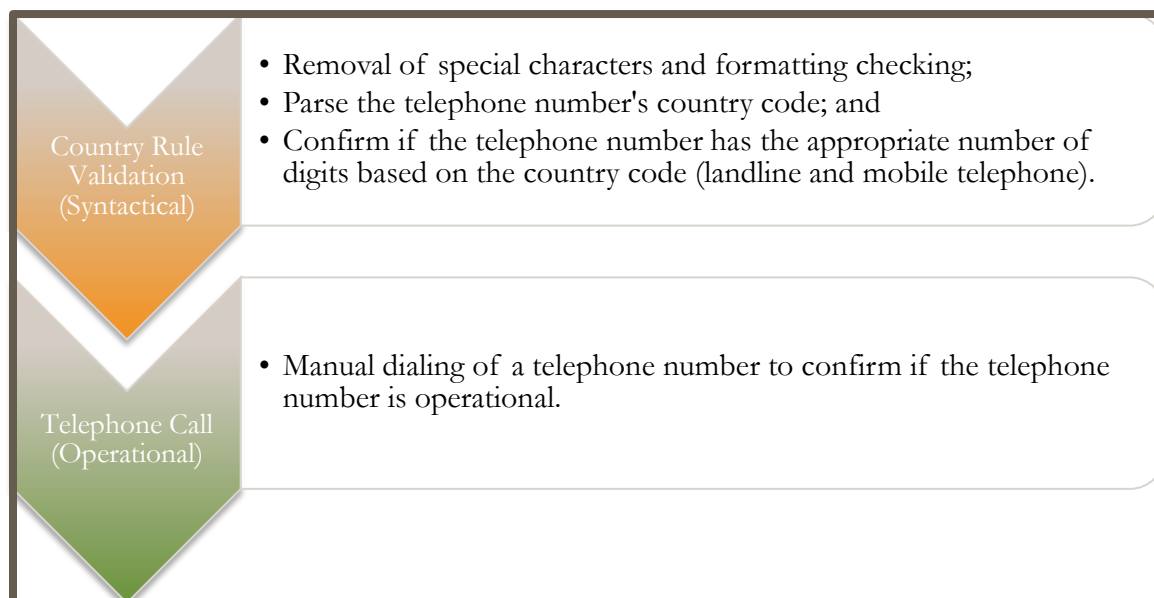
- Substantial Failure: Indicates that the data are missing significant information, and that the data provided are mostly inaccurate.
- Full Failure: Indicates that significant data elements are missing, or that data provided are overwhelmingly inaccurate.

This Appendix provides detailed explanations of the processes followed by each vendor for the specific data elements they analyzed, and an overview of how each vendor applied their analysis to the NORC rating scale.

Telephone Validation: DigiCert

ICANN contracted with DigiCert to assist in the assessment of telephone WHOIS accuracy.

The chart below briefly describes the general steps taken by DigiCert, with a detailed explanation below.



Syntactical Telephone Validation

Once the parsed records were received by DigiCert, the initial step was to “clean” the telephone record data to remove special characters (parenthesis, pluses, periods, etc.) and properly format.

After this initial review was completed, DigiCert checked each record to determine if a country code was located in the phone number. If so, DigiCert would then parse the country code from the records and begin its next step of verification.

If the record did have a country code in the Registrant field, DigiCert would compare the remaining digits in the telephone number to see if it contained the appropriate number of digits, taking into consideration both landline and mobile telephone numbers, for the country. To determine the appropriate number of digits for a country, DigiCert relied on ITU Telecommunication Standardization Sector Recommendation E.164 (ITU-T E.164).

Moreover, DigiCert did not attempt to cross-validate telephone numbers with postal addresses. For example, if a telephone number contained a country code and the acceptable number of digits for the country code, the country code was not cross-referenced to the address. As there is no requirement to match the countries of the telephone numbers with the postal addresses in a WHOIS record, the postal addresses were only referenced if there was a need to “assume” a country code for a telephone record that was missing the country code prefix information.

A record received a scoring of “No Failure (+2)” if the telephone number was present in the WHOIS data and it contained an acceptable number of digits under the respective country code’s applicable telephone number requirements. A country code prefix was required to be present in the record in order to achieve a score of (+2).

A record received a scoring of “Minimal Failure (+2)” if there were formatting errors within the telephone record. However, the telephone number contained the acceptable number of digits for the country code. As such, DigiCert ignored formatting issues and essentially applied a “No Failure (+2)” rating to the record. An example of this might be the inclusion of (“.”) or (“+”) in the data.

A record received a scoring of “Limited Failure (+0)” if the telephone number lacked a country code but the telephone number contained an acceptable number of digits, if the country code from the corresponding country identified in the address field of the specific contact (i.e., the administrative contact) was appended to the record.

A record received a scoring of “Substantial Failure (-1)” if the Administrative or Technical contact telephone number lacked a country code but the Registrant contact address field contained a country. DigiCert then determined if the telephone number, appended with the Registrant country code based on the address in the Registrant contact field, contained an acceptable number of digits. It should be noted that this scoring was not used for Registrant records, only for Technical or Administrative records.

A record received a scoring of “Full Failure (-2)” if the telephone number was missing or if the telephone number contained the wrong number of digits based on its country code (this criteria was applied to records that had a country code listed and to records that had the country code appended based on an address).

Description	Explanation	NORC Scoring if Error Code Applies
No Failure	The telephone number is present, contains a country code and an acceptable number of digits under the country's applicable telephone number requirements.	+2
Minimal Failure	As formatting is not a requirement for telephone numbers under ITU-T E.164, records with formatting issues received the same score as those that received "No Failure (2)" rating.	+2
Limited Failure	The telephone number lacks a country code but the corresponding address field notes the country. The telephone number, appended with the country code based on the corresponding address, contains an acceptable number of digits under the country's applicable telephone number requirements.	+0
Substantial Failure	The Administrative or Technical contact telephone number lacks a country code but the Registrant contact address field contains a country. The telephone number, appended with the Registrant country code based on the address, contains an acceptable number of digits under the country's applicable telephone number requirements. Note that this rating is not used for Registrant Records	-1
Full Failure	The record is missing a telephone number or the telephone number contains the wrong number of digits based on its country code (this includes records where the country code was appended based on the address).	-2

Operational Telephone Validation

For the purposes of the Pilot Study, DigiCert defined a telephone number as operationally valid if the number was not disconnected, a connection was made within 30 seconds of being dialed, and there were no error messages once connected. Moreover, if the number dialed, gave a busy signal, went to an answering service, or responded in a manner that indicated the number functions properly, the telephone number was considered operationally valid.

DigiCert conducted this process by having its live support personal manually dial each number. To ensure a higher rate of accuracy in its efforts, DigiCert employed a multilingual Staff so that error messages in languages other than English could be translated.³

A record received a scoring of “No Failure (+2)” if the telephone number listed in the WHOIS record successfully dialed based on the aforementioned specifications.

No records received a scoring of “Minimal Failure” as the number either connected or it did not. As such, DigiCert did not apply this scoring to any of the records.

A record received a scoring of “Limited Failure (+0)” if the telephone number listed in the WHOIS record failed to dial but the telephone number dialed after appending a country code based on the corresponding country address information.

A record received a scoring of “Substantial Failure (-1)” if an Administrative or Technical contact’s telephone number listed in the WHOIS record failed to dial, but the telephone number dialed after appending the Registrant’s dialing country code based on the Registrant’s country address information.

A record received a scoring of “Full Failure (-2)” if the telephone number was missing or failed to connect after appending country address information.

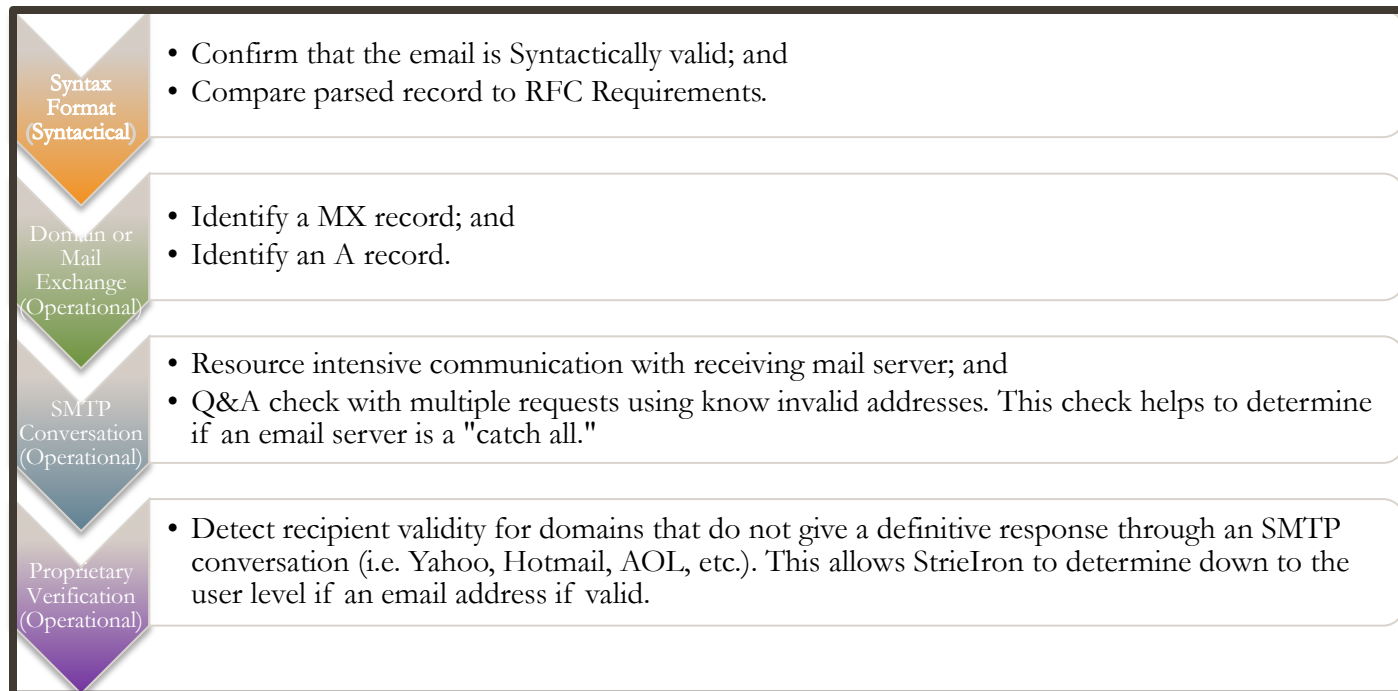
³ DigiCert’s live support personnel included the following languages: Chinese, English, French, German, Japanese, Korean, Portuguese, Russian, and Spanish.

Description	Explanation	NORC Scoring if Error Code Applies
No Failure	The telephone number listed in the WHOIS record successfully dials.	+2
Minimal Failure	For the purpose of the pilot, this rating was not included.	n/a
Limited Failure	The telephone number listed in the WHOIS record fails to dial but the telephone number dials after appending the corresponding dialing country code based on the corresponding address country information.	+0
Substantial Failure	An Administrative or Technical contact's telephone number listed in the WHOIS record fails to dial but the telephone number dials after appending the Registrant's dialing country code based on the Registrant's address country information.	-1
Full Failure	The telephone number is missing or fails to connect after appending address country information.	-2

Email Validation: StrikeIron

ICANN contracted with StrikeIron to perform the assessment of accuracy of email addresses in the sampled WHOIS records.

The chart below briefly describes the general steps taken by StrikeIron, with a detailed explanation below.



Syntactical Email Validation

Once the parsed records were received, StrikeIron undertook its Syntactical Validation for email addresses by verifying that email addresses were in the correct Syntax Format. The standard used for this verification was RFC compliance (e.g. [RFC 5322](#)).

Specifically, StrikeIron checked emails to ensure they had the following four key RFC specifications:

- (1) The email address contained an “@” symbol,
- (2) The domain portion of the email address was valid;
- (3) The local portion of the email address was valid;
- (4) The email address adheres to all other RFC requirements.

Once this step was completed, StrikeIron applied the following scoring to the results.

A record received a scoring of “No Failure (+2)” if it was in compliance with all RFC specifications.

No records received a scoring of “Minimal Failure” as an email was either RFC compliant or it was not.

No records received a scoring of “Limited Failure” as an email was either RFC compliant or it was not.

No records received a scoring of “Substantial Failure” as an email was either RFC compliant or it was not.

A record received a scoring of “Full Failure (-2)” if it was not in compliance with all RFC specifications.

Description	Explanation	NORC Scoring if Error Code Applies
No Failure	<ul style="list-style-type: none"> The parsed record contains the "@" symbol; The parsed record contains a valid domain; The parsed record contains a valid LocalPart before the @ symbol; The parsed record's syntax is valid; and The parsed record adheres to all other RFC specifications. 	+2
Minimal Failure	No records received a scoring of "Minimal Failure" as an email was either RFC compliant or it was not.	n/a
Limited Failure	No records received a scoring of "Limited Failure" as an email was either RFC compliant or it was not.	n/a
Substantial Failure	No records received a scoring of "Substantial Failure" as an email was either RFC compliant or it was not.	n/a
Full Failure	<ul style="list-style-type: none"> The parsed record did not contain the "@" symbol; The parsed record did not contain a valid domain; The parsed record did not contain a valid LocalPart before the @ symbol; The parsed record's syntax was invalid; and The parsed record did not adhere to all other RFC specifications. 	-2

Operational Email Validation

Strikelron took three steps to assess operational validity for email addresses. The first step was Domain or Mail Exchange (MX) verification, the second step was SMTP Conversation, and the third step was Proprietary Verification of the results. In performing this assessment, Strikelron did not attempt to send an email to the addresses in the sample.

Once the parsed records were received, Strikelron took the syntactically valid emails and conducted three checks.

The first check that Strikelron conducted was a Domain or Mail Exchange (MX) verification. In this step, Strikelron checked if the mail server for a domain indicated that the email was valid and then checked.

To validate against “catch all” servers, Strikelron would then test an assuredly invalid email against the same server to see if an error message was received against the invalid email. Once this step was completed, Strikelron applied the following scoring to the results.

A record received a scoring of “No Failure (+2)” rating if there was an MX Record or an A Record for the domain and the email's domain had a DNS entry for the A record and there was a mail exchange server listed and the A record responded as a mail server. A domain could also receive a No Failure rating if it had a MX Record, an A Record, and was responding.

No records received a scoring of “Minimal Failure” as an email was either operational or non-operational.

No records received a scoring of “Limited Failure” as an email was either operational or non-operational.

No records received a scoring of “Substantial Failure” as an email was either operational or non-operational.

A record received a scoring of “Full Failure (-2)” rating if there was no MX Record or A Record for the domain and the email's domain did not have a DNS entry for the A record, there was no mail exchange server listed, and the A record did not respond as a mail server. A domain could also receive a Full Failure rating if it did not have a MX Record, an A Record, and it was not responding.

The second check that Strikelron conducted was a SMTP Conversation with a mail server account to determine if the mail servers returned a "Server Will Accept" message. Once this step was completed, Strikelron applied the following scoring to the results.

A record received a scoring of “No Failure (+2)” if the username or mailbox was accepted as a valid email address, the email was valid according to RFC specifications, and the email was not identified as a spam trap.

No records received a scoring of “Minimal Failure” as the email was either operational or non-operational.

No records received a scoring of “Limited Failure” as the email was either operational or non-operational.

No records received a scoring of “Substantial Failure” as the email was either operational or non-operational.

A record received a scoring of “Full Failure (-2)” if the username or mailbox was rejected as a valid email address, the email was invalid according to RFC specifications, and the email was identified as a spam trap.

Description	Explanation	NORC Scoring if Error Code Applies
No Failure	<ul style="list-style-type: none"> • Domain found: MX record or A record exists for this domain. • Valid mail domain: The email's domain has a DNS entry for the A record, and there is a mail exchange server listed and the A record responds as a mail server. • Mail Domain is Responsive: The email's domain has a MX record and an A record and is responding. • The username or mailbox was accepted as a valid email address. • The email is valid according to RFC specifications. • The email has not been identified as a spam trap. 	+2
Minimal Failure	No records received a scoring of "Minimal Failure" as the email was either operational or non-operational.	n/a
Limited Failure	No records received a scoring of "Limited Failure" as the email was either operational or non-operational.	n/a
Substantial Failure	No records received a scoring of "Substantial Failure" as the email was either operational or non-operational.	n/a
Full Failure	<ul style="list-style-type: none"> • Domain not found: No MX record or A record exists for this domain. • Not a valid mail domain: The email's domain has a DNS entry for the A record, however, there is not a mail exchange server listed and the A record does not respond as a mail server. • Mail Domain is Non-Responsive: The email's domain has a MX record and an A record but is not responding. • The username or mailbox was rejected as a valid email address. • The email is invalid according to RFC specifications. • The email has been identified as a spam trap. 	-2

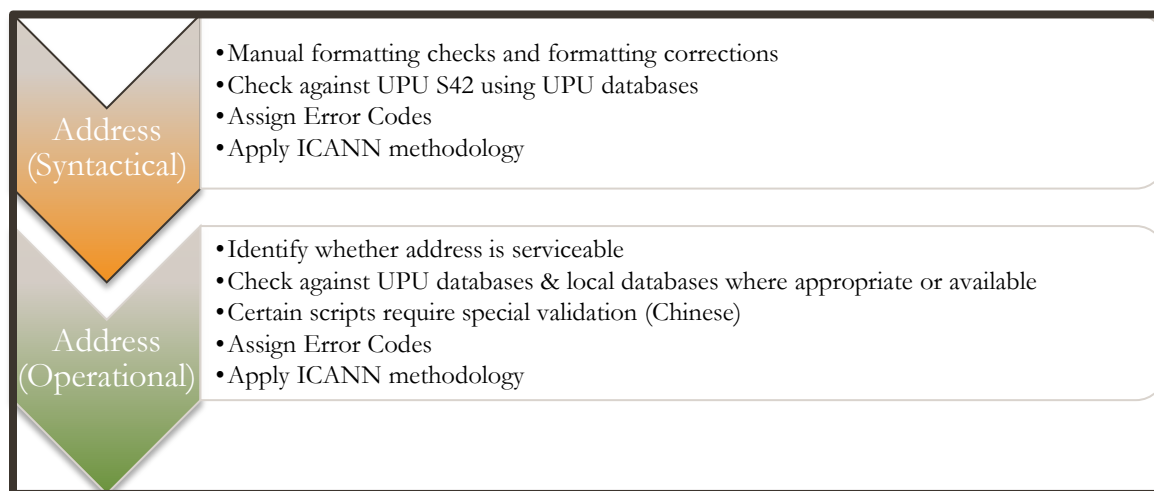
The last check conducted by Strikelron was a propriety verification test to detect recipient validity for domains that do not give definitive responses through SMTP Conversation (i.e. Yahoo, Hotmail, AOL,

etc.). This final form of verification allowed Strikelron to determine, down to the user level, if an email address is valid.

Postal Validation: Universal Postal Union (UPU)

ICANN contracted with the Universal Postal Union (UPU) to perform the assessment of accuracy of postal addresses in the sampled WHOIS records.

The chart below briefly describes the general steps taken by the UPU, with a detailed explanation below.



Syntactical Postal Validation

Once the parsed records were received, the UPU manually verified that each address field was in the correct fields – if errors were found they were corrected. For example, if a parsed record contained all of the address information in the country field rather than in the respective, correct field it was adjusted.

The UPU cross-referenced the manually verified records against UPU Addressing Standard S42, which defines the internationally agreed upon method for classifying international postal address components and templates. Once these steps were completed, the UPU applied the following scoring to the results.

A record received a scoring of “No Failure (+2)” if the address components are clear and the address data is sufficient for Syntactical Validation.

A record received a scoring of “Minimal Failure (+1)” if there was a sufficient amount of correct values in the address fields for further automatic validation.

To achieve a scoring of “Limited Failure (+0), a record would have to have at least one of the following errors:

- (1) The record could not be processed as the country was not recognized or the address field(s) contained values that were not recognized as address components. However, the address data may nevertheless be accurate upon further manual address validation.
- (2) The record could not be processed as the reference data for this country was not valid. However, the address data may nevertheless be accurate upon further manual address validation.
- (3) Some address fields contain multiple address components or more data than is necessary. However, sufficient address component exist to deliver to this address and the address can be corrected.

No records received a scoring of “Substantial Failure (-1)” as addresses with multiple address fields are still likely serviceable.

No records received a scoring of “Full Failure (-2)” as addresses with multiple address fields are still likely serviceable.

Description	Explanation	Detailed Explanation	NORC Scoring if Error Code Applies
No Failure	Syntactically correct	The address components (values in fields) are clear and the address data is sufficient for this syntactical validation.	2
Minimal Failure		Sufficient correct values in address fields for further operational automatic validation.	1

Description	Explanation	Detailed Explanation	NORC Scoring if Error Code Applies
Limited Failure	Not processed	<ul style="list-style-type: none"> Country not recognized; Some address fields may contain values that are not recognized as address components, Registrant address field is empty, special characters, diacritic characters, c/o address, or too many address components are present in a single address field; and Address data may be considered accurate for further manual address validation based on the current address data or when using complementary information (phone, Registrant-name, Registrant organization or Registrant technical and/or administrative) contact data. 	0
Limited Failure	Not processed	<ul style="list-style-type: none"> Reference data for this country not valid; Possible state/province/territory name in the Registrant country field; and Address data may nevertheless be accurate for further manual address validation. 	0
Limited Failure	Syntactic concerns Unlikely to be deliverable	<ul style="list-style-type: none"> Some address fields contain multiple address components or twice the address data; But sufficient address components exist to deliver to this address; and Address can be corrected. 	0

Description	Explanation	Detailed Explanation	NORC Scoring if Error Code Applies
Limited Failure	Syntactic concerns Fairly likely to be deliverable	<ul style="list-style-type: none"> • Some address fields contain multiple address components or twice the address data or the postcode is missing; • But sufficient address components exist to deliver to this address; and • Address can be corrected. 	0
Limited Failure	Syntactic concerns Very likely to be deliverable	<ul style="list-style-type: none"> • Some address fields contain multiple address components or twice the address data or the postcode is missing; • Sufficient address components exist to deliver to this address; and • Address can be corrected. 	0
Limited Failure	Syntactic concerns Address can be corrected	Sufficient correct values in address fields for further operational automatic validation.	1
Substantial Failure	n/a	For the syntactical address validation, scoring -2 or -1 was not assigned as with multiple address fields there is often a good chance to determine the correct address and if it is serviceable.	n/a
Full Failure		For the syntactical address validation, scoring -2 or -1 was not assigned as with multiple address fields there is often a good chance to determine the correct address and if it is serviceable.	n/a

Description	Explanation	Detailed Explanation	NORC Scoring if Error Code Applies
No Failure	Validated Operationally correct	The address components (values in fields) are clear and the address data is sufficient to render a correct address for delivery.	2
Minimal Failure	Validated Address was corrected	Some doubts about some fields but sufficient correct values in address fields to render a correct address for delivery.	1
Limited Failure	Fair deliverability Address was corrected	Some address fields contain multiple address components or twice the address data or the postcode is missing. However, sufficient address components exist to render an address and deliver to this address.	0
Limited Failure	Good deliverability Address was corrected	Some address fields contain multiple address components or twice the address data or the postcode is missing. However, sufficient address components exist to deliver mail correctly to this address.	0
Substantial Failure	Poor deliverability Address was corrected	Some address fields contain multiple address components or twice the address data. However, sufficient address components exist to render a correct address and maybe deliver to this address	-1
Full Failure	Not processed	<p>This record could not be processed for one or more of the following reasons:</p> <ul style="list-style-type: none"> • Country not recognized; • Some address fields may contain values that are not recognized as address components, Registrant address field is empty, special characters, diacritic characters, c/o address, or too many address components are present in a single address field; and/or • Address data may nevertheless be accurate for further manual address validation based on the current address data or when using complementary information (phone, Registrant-name, Registrant_organization or Registrant 	-2

		technical and/or administrative) contact data.	
Full Failure	Not processed	This record could not be processed for one or more of the following reasons: <ul style="list-style-type: none"> • Reference data for this country not valid; • Possible state/province/territory name in the Registrant country field; and/or • Address data may nevertheless be accurate for further manual address validation. 	-2

Operational Postal Validation

UPU conducted an assessment for operational validity for physical addresses was to determine if an address was serviceable by using a combination of manual and automated processes. The UPU defines serviceable as, “capable of being put to use.” Essentially, if the UPU determined, based on the various international standards and norms, that a letter could be sent and received at the address, it was deemed serviceable.

Once this initial check was completed, the UPU then checked the address against its own databases, as well as the databases of partner countries, to further determine if an address was serviceable. In a case where the UPU was unsure of an address’ serviceability, the record was forward to the relevant country’s postal authority to confirm if the address was serviceable. For example, some countries require the UPU to send inquiries regarding serviceability on a case-by-case basis.

Additionally, if an address was in a script not supported by the UPU’s proprietary software, the UPU forwarded the address to the relevant partner country for assistance for confirmation of serviceability. Once these steps were completed, the UPU applied the following scoring to the results.

A record received a scoring of “No Failure (+2)” if the address components were clear and the address data was sufficient to render a correct address for delivery.

A record received a scoring of “Minimal Failure (+1)” if there were some doubts about some fields, but there were sufficient correct values in the address fields to render a correct address for delivery.

A record received a scoring of “Substantial Failure (-1)” if: (1) some address fields contained multiple address components, twice the necessary address data, or were missing the postcode. However, sufficient address components existed to render a correct address and may make it serviceable.

A record received a scoring of “Substantial Failure (-1)” if some address fields contained multiple address components or twice the address data. However, sufficient address components existed to render a correct address and may make it serviceable.

A record received a scoring of “Full Failure (-2)” if the record could not be processed for one or more of the following reasons: the country was not recognized, some address fields contained values that are not recognized as address components, the field was empty, or reference data for this country is not valid. However, the address data may nevertheless be accurate upon further manual address validation.