WHOIS Accuracy Reporting System (ARS):

Phase 2 Cycle 4 Webinar

ICANN GDD Operations
NORC at the University of Chicago

20 July 2017

Report available here: https://whois.icann.org/en/whoisars-reporting

Agenda

1. WHOIS ARS Background
2. Phase 2 Cycle 4: Timeline and Process
3. Phase 2 Cycle 4: Testing Criteria, Sample Design, and Population Information
4. Phase 2 Cycle 4: Results and Major Findings
5. Contractual Compliance Follow-up
6. Demonstration of upcoming WHOIS ARS Data Display Tool

Phase 2 Cycle 4:

WHOIS ARS Implementation Background

WHOIS ARS Implementation

0

Pilot

"Proof of Concept": Tested processes for data collection and validation

Report: **Published 23 December 2014**

Public Comment Report: **Published 3 April 2015**

1

Phase 1: *Syntax Accuracy only*

Is the record correctly formatted?

Report: Published 24 August 2015

2

Phase 2: *Syntax + Operability Accuracy*

Does the email go through, phone ring, mail deliver?

Cycle 1 Report: Published 23 December 2015

Cycle 2 Report: Published 8 June 2016

Cycle 3 Report: Published 12 December 2016

Cycle 4 Report: Published 12 June 2017

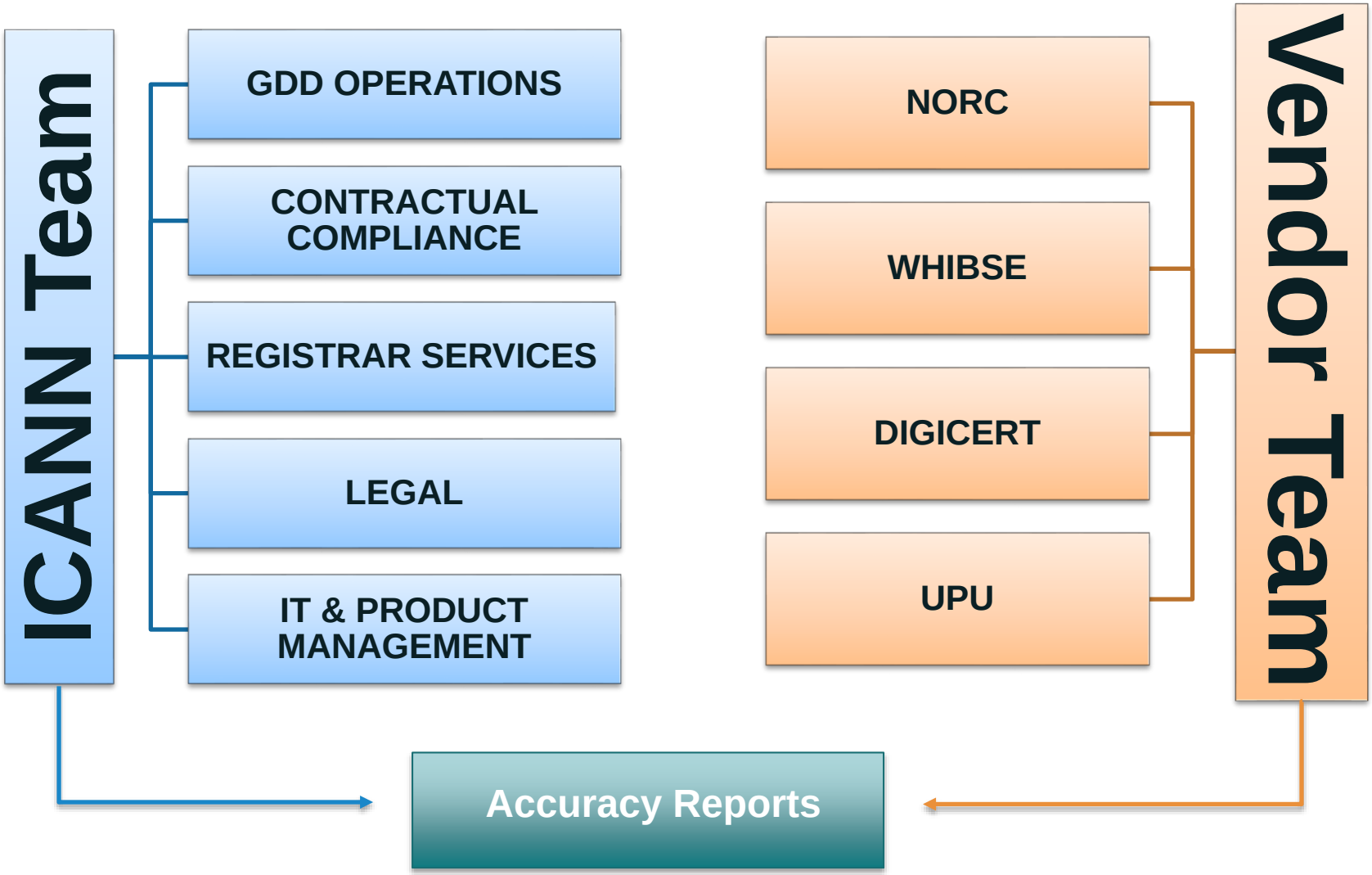
Cycle 5 Report: Expected December 2017

WHOIS ARS Information and Reports available here: https://whois.icann.org/en/whoisars

Phase 2 Cycle 4:

Process and Timeline

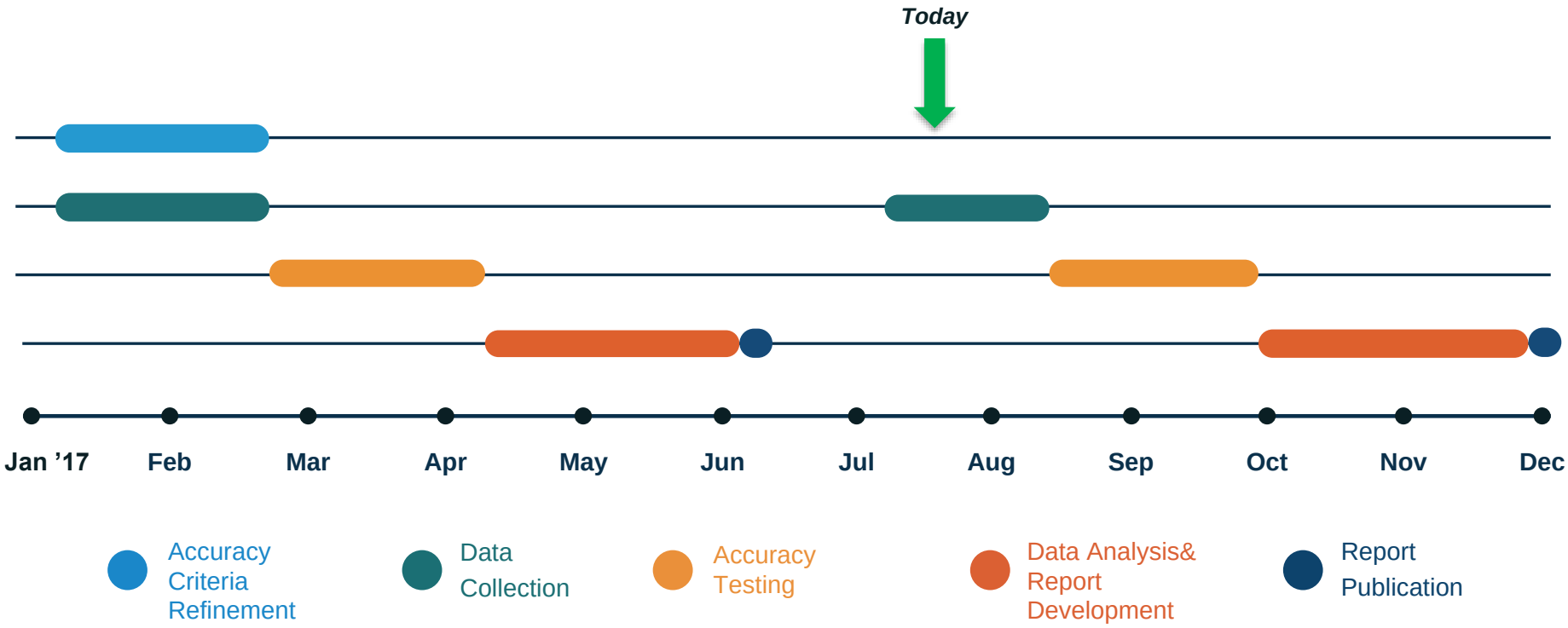
Phase 2 Cross-Functional Team



FY17 Cyclical Timeline – Cycles 4 & 5

Cycle 4: Complete 12 June 2017

Cycle 5: Started July 2017



Cycle 4 Report published 12 June 2017
Cycle 5 has already begun; records will be pulled over the next several weeks

Phase 2 Cycle 4 – Report Highlights

Accuracy Statistics by Subgroup

- Report provides both syntax and operability accuracy rates for:
  - The gTLD space, by region and in total
  - New gTLDs compared to Prior (legacy) gTLDs
  - RAA Type (2009, 2013GF, 2013NGF)
- Data within 95% confidence intervals, ≤+/- 5% margin of error

Report identifies reasons for error

- All domains evaluated against 2009 RAA requirements for both syntax and operability
- Detailed testing results in data that demonstrates in what way a record is inaccurate
- Contains information on regional differences in accuracy.

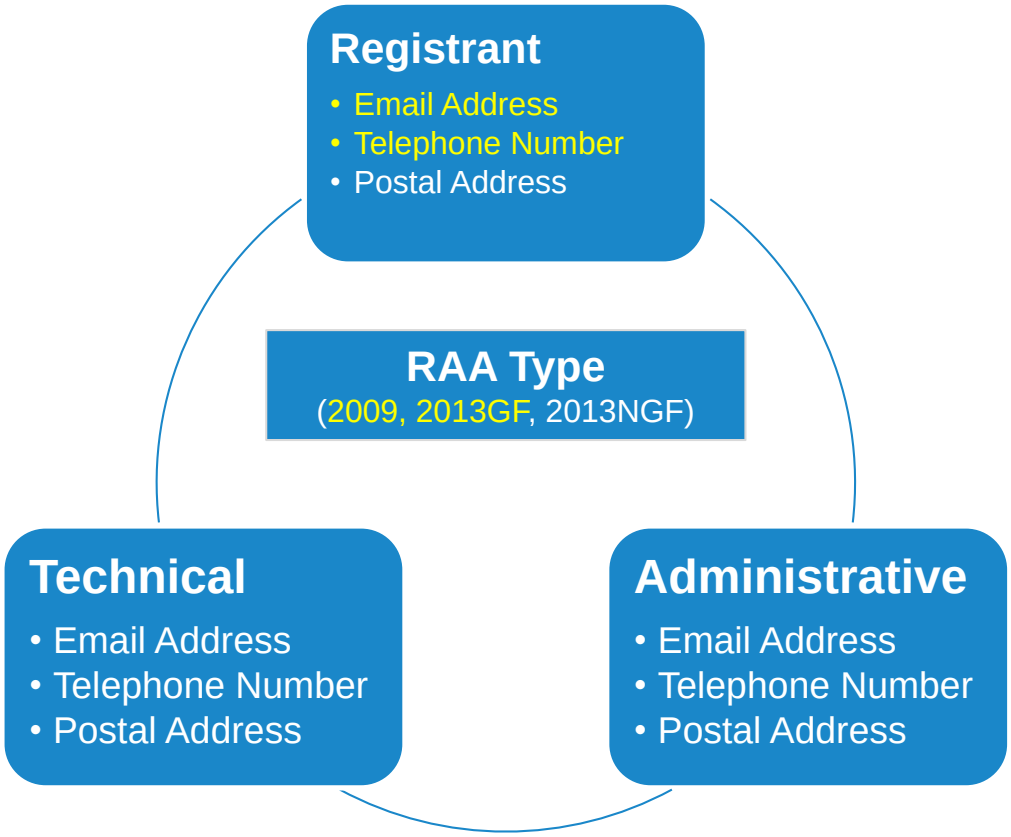
Report & ARS Website now contain Compliance follow-up information

- Provide information in response to community questions

Phase 2 Cycle 4:

Testing Criteria

Phase 2 Cycle 4 – Contact types, modes, and testing criteria



Criteria Examples	**Syntax**: Does the email address contain an “@”? **Operability**: Did the email bounce back?
	Syntax: Does the telephone number have a country code? **Operability**: Did the number ring when dialed?
	Syntax: Does the postal address include an identifiable country?* **Operability**: Can mail be delivered to the address?

Detailed criteria listed at
www.whois.icann.org/en/whoisars-validation

GF = Grandfathered. A domain registered before a registrar changed to the 2013 RAA. Obligated to 2009 RAA requirements.
NGF = Non-grandfathered. Obligated to 2013 RAA requirements.

Phase 2 Cycle 4:

Sample Design and Population Information

Phase 2 Cycle 4 – Demographics

gTLD Population At Time of Sample (January 2017)

Records in gTLDs	Total gTLDs	2009 RAA*	2013GF RAA*	2013 NGF RAA*	New gTLDs	Prior gTLDs
185.7m	1,231	395k	79.7m	103m	1,213	18

200k Sample

AFR	LAC	EUR	APAC	N.A.	2009 RAA	2013GF RAA	2013 NGF RAA	New gTLDs	Prior gTLDs
1.3k	10.1k	34.8k	65.0k	85.8k	370	74.3k	122.8k	718	18

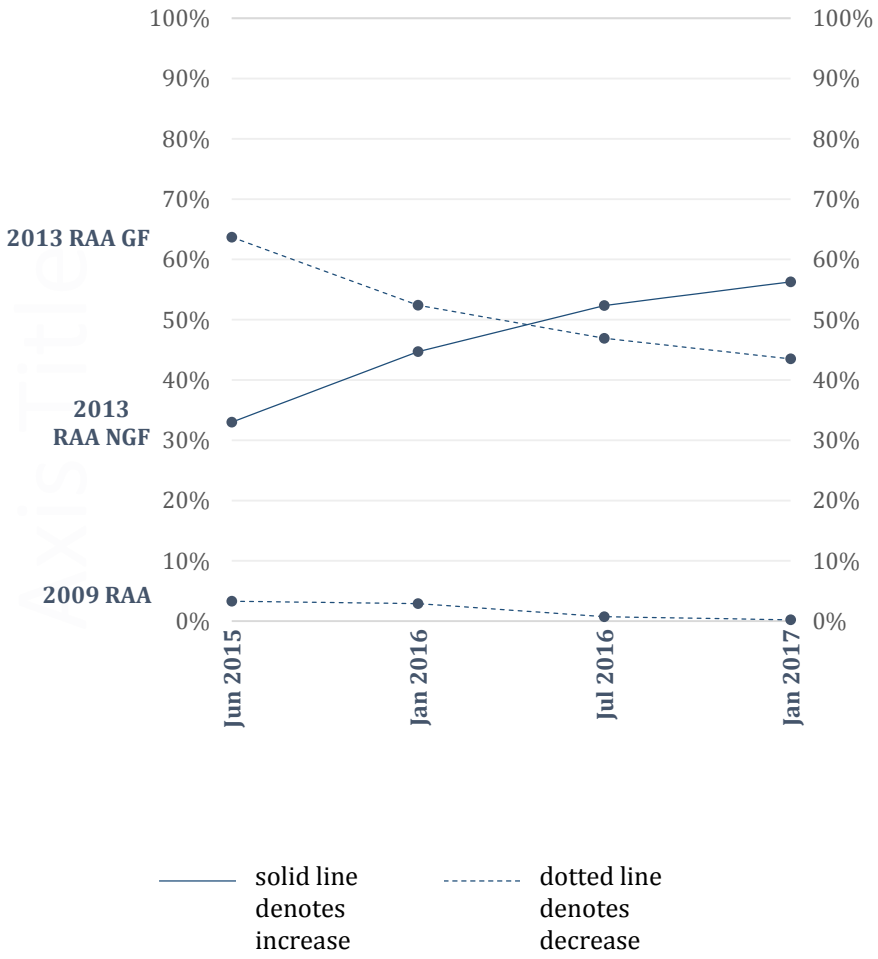
12k Sub-sample

AFR	LAC	EUR	APAC	N.A.	2009 RAA	2013GF RAA	2013 NGF RAA	New gTLDs	Prior gTLDs
1.2k	1.9k	2.3k	2.9k	3.1 k	370	4.9k	6.2k	718	18

* Weighted estimates from 200k sample

Phase 2 Cycle 4 – Change in Distribution of RAA type

Change in Distribution across Sample Dates

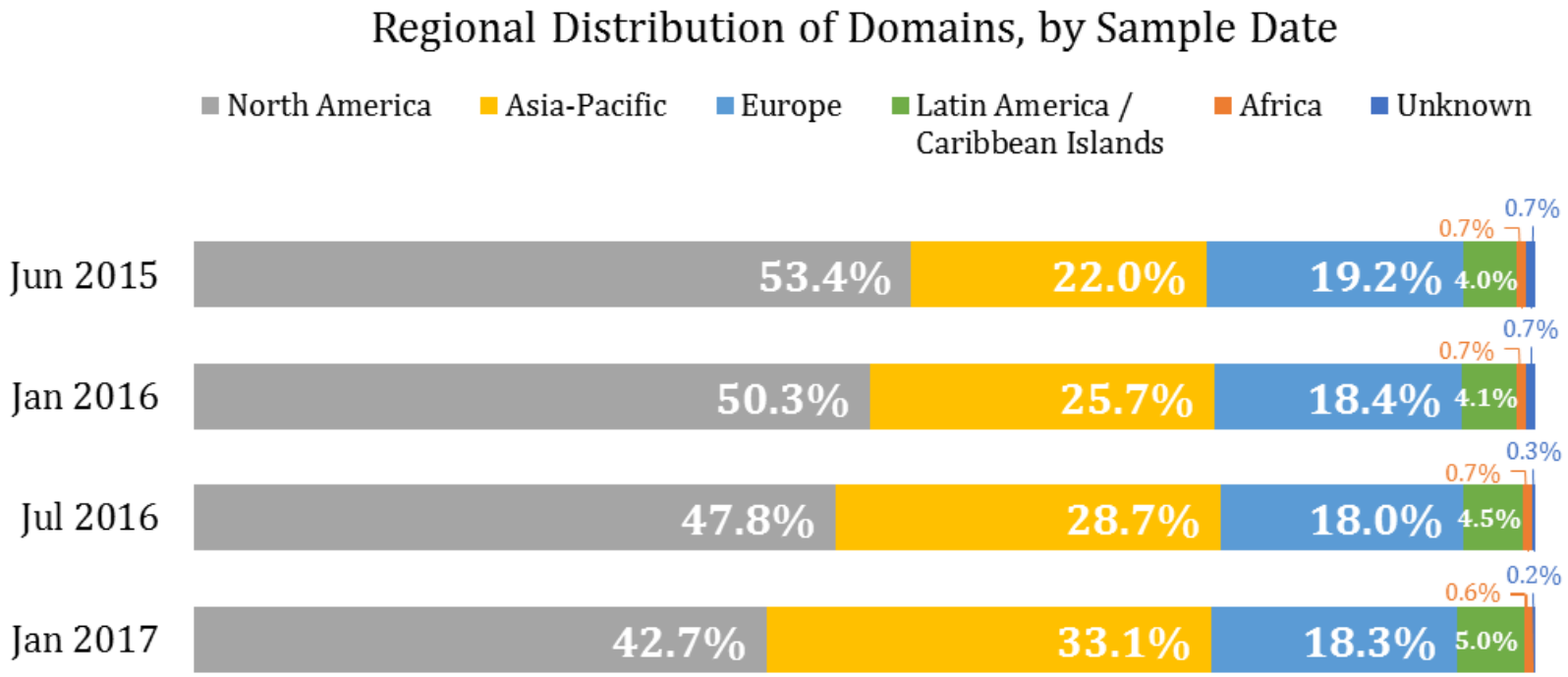


The 2009 RAA share is shrinking;

The share of non-grandfathered 2013 RAA domains are growing rapidly (from ~33% of distribution in June 2015 to ~55% in January 2017).

Sample Date	2009 RAA	2013 RAA GF	2013 RAA NGF
June 2015	3.3%	63.7%	33.0%
January 2016	2.9%	52.4%	44.7%
July 2016	0.7%	46.9%	52.3%
January 2017	0.2%	43.5%	56.3%

Phase 2 Cycle 4 – Domains by Region



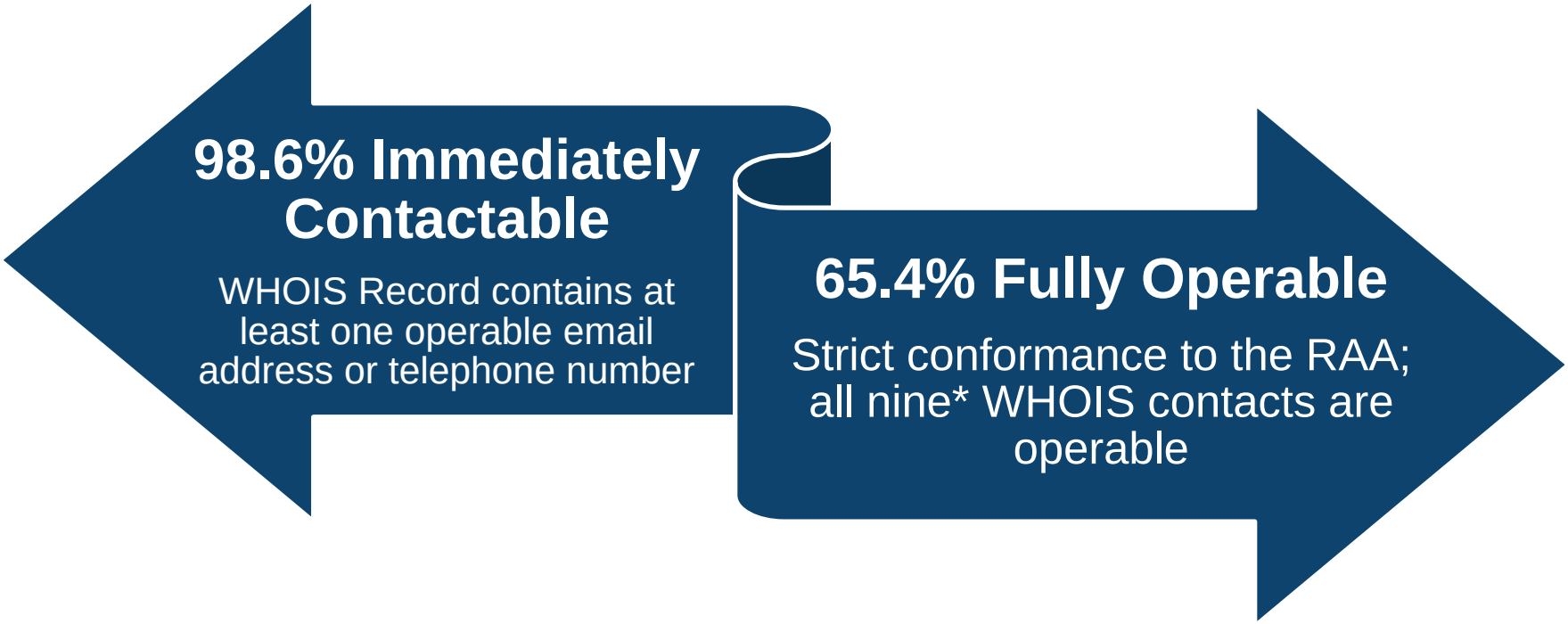
Number of Domains per Region, by Sample Date
(in millions)

Sample Date	NA	AP	EUR	LAC	AF	Unknown	TOTAL
June 2015	84.40	34.70	30.30	6.25	1.12	1.15	**157.92**
January 2016	85.50	43.70	31.30	7.05	1.16	1.26	**169.97**
July 2016	88.00	52.80	33.20	8.31	1.27	0.49	**184.07**
January 2017	79.31	61.43	34.06	9.29	1.19	0.42	**185.70**

Cycle 2 Phase 4:

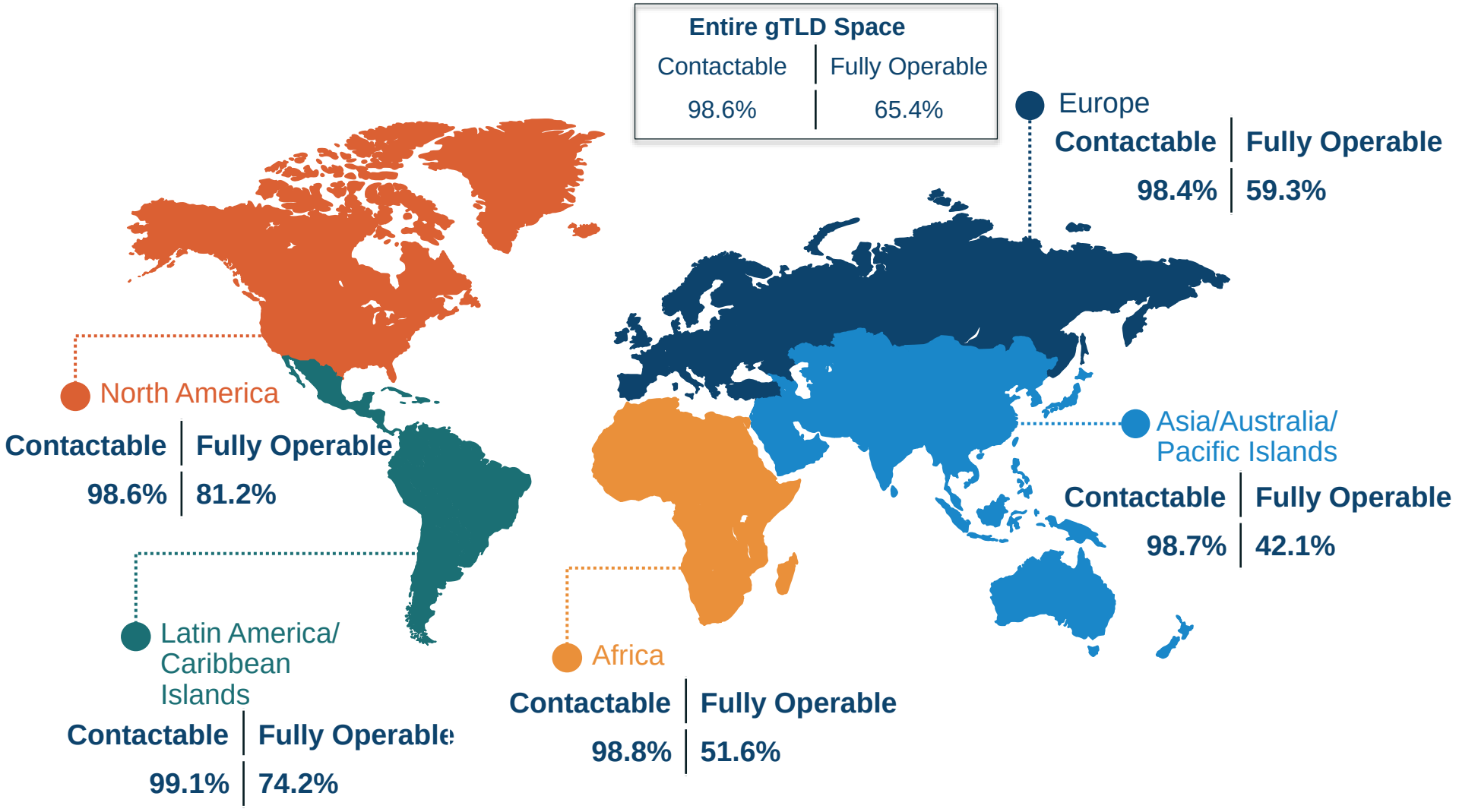
Overall Summary

Phase 2 Cycle 4 – How Contactable are the WHOIS Records?



*9 entries per WHOIS record:
3 Contact Types: Registrant, Administrative, and Technical
3 Contact Modes: Telephone, Email, Address

Phase 2 Cycle 4 – Contactable Domains, by Region



Phase 2 – Cycle 3 to Cycle 4 Changes

Email

- Syntax accuracy decreased slightly from 99.6% to 99.5%%.
- Operability accuracy increased from 90.1% to 94.5%

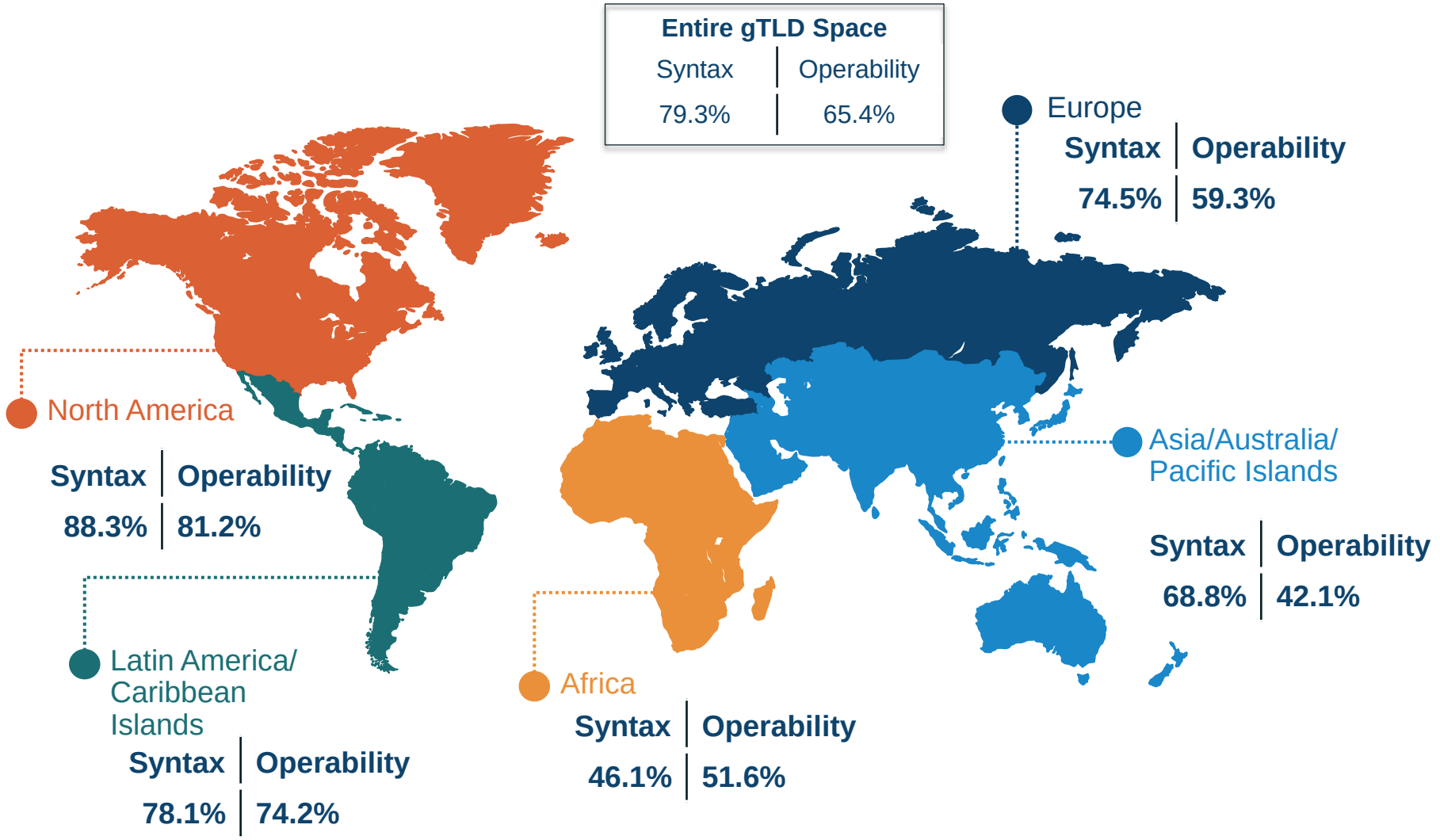
Telephone

- Syntax accuracy increased from 88.5% to 89.5%
- Operability accuracy decreased from 72.4% to 68.9%

Postal

- Syntax accuracy increased from 87.0% to 87.4%
- Operability accuracy increased slightly from 96.8% to 97.2%

Phase 2 Cycle 4 - Syntax and Operability Accuracy by Region



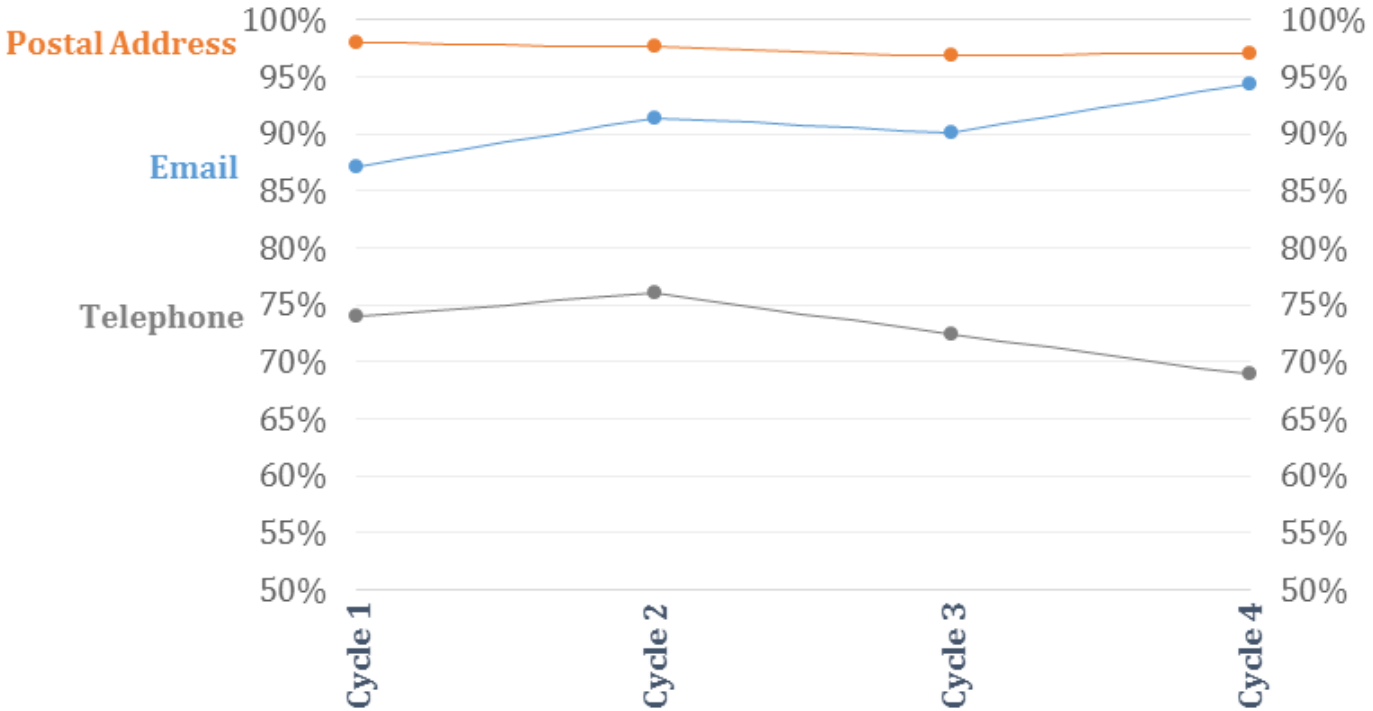
Cycle 2 Phase 4:

Results and Findings: Operability, 2009 RAA

Phase 2 Cycle 4 – Changes Over Time: Operability Accuracy by Contact Mode

Overall Operability Accuracy			
Cycle 1	Cycle 2	Cycle 3	Cycle 4
64.7%	70.2%	65.1%	65.4%
	Δ 5.5%	Δ -5.1%	Δ -0.1%

Entire gTLD Space
Cycle 1 through Cycle 4

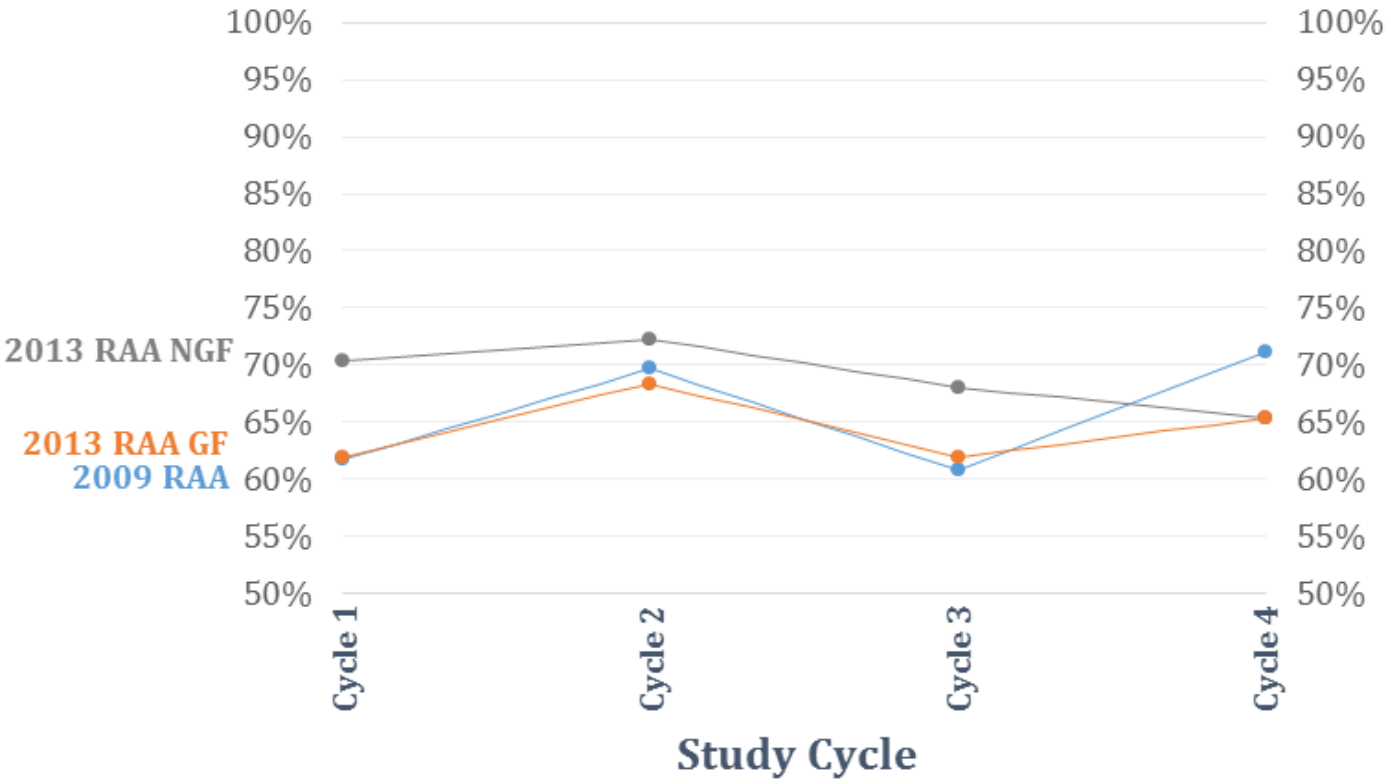


Accuracy of all 3 contact modes and all 3 contact types

Phase 2 Cycle 4 – Changes Over Time: Operability Accuracy by RAA Type

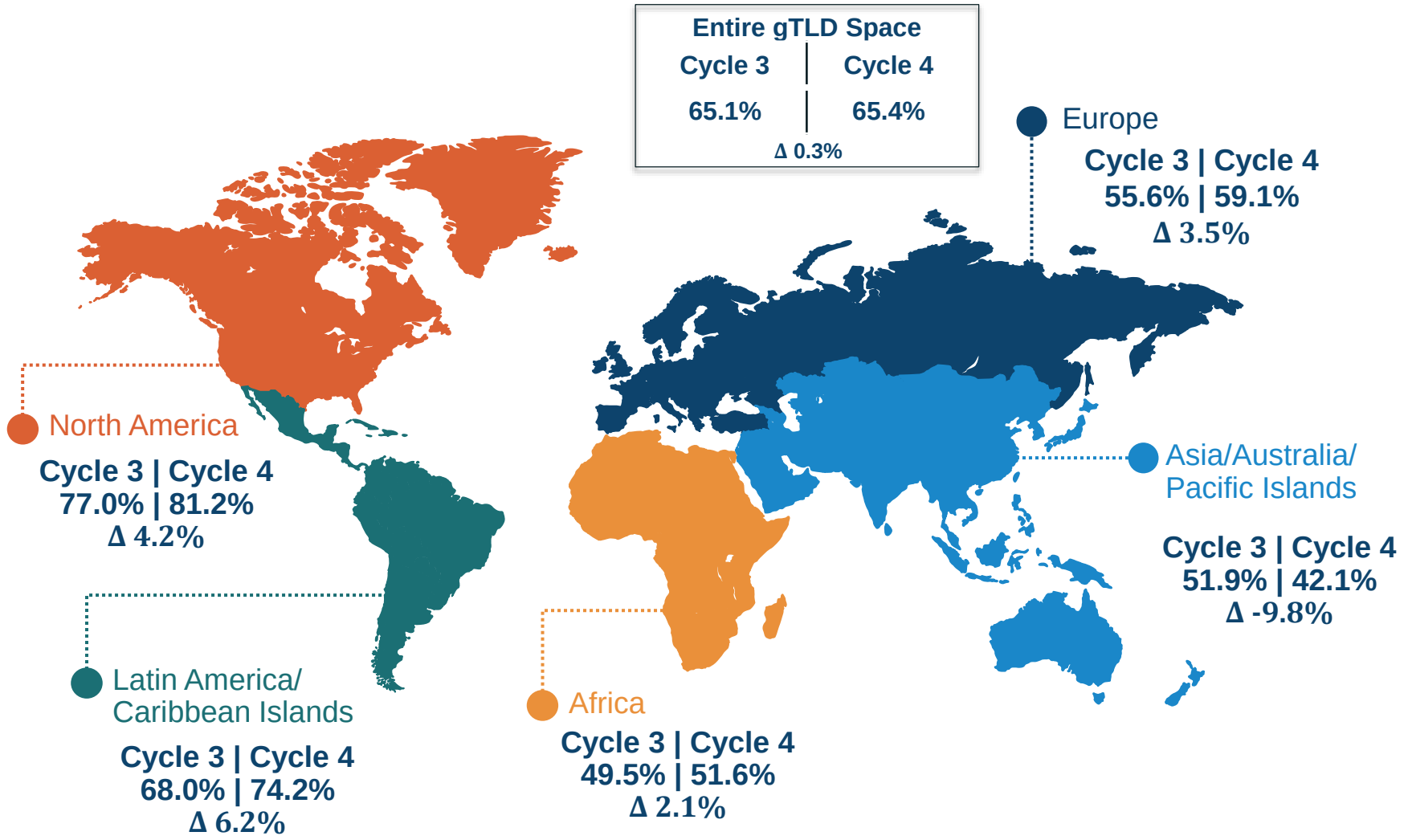
Overall Operability Accuracy			
Cycle 1	Cycle 2	Cycle 3	Cycle 4
64.7%	70.2%	65.1%	65.4%
	Δ 5.5%	Δ -5.1%	Δ -0.1%

Entire gTLD Space
Cycle 1 through Cycle 4



Accuracy of all 3 contact modes and all 3 contact types

Overall Operability Accuracy by Region – Cycle 3 v. Cycle 4



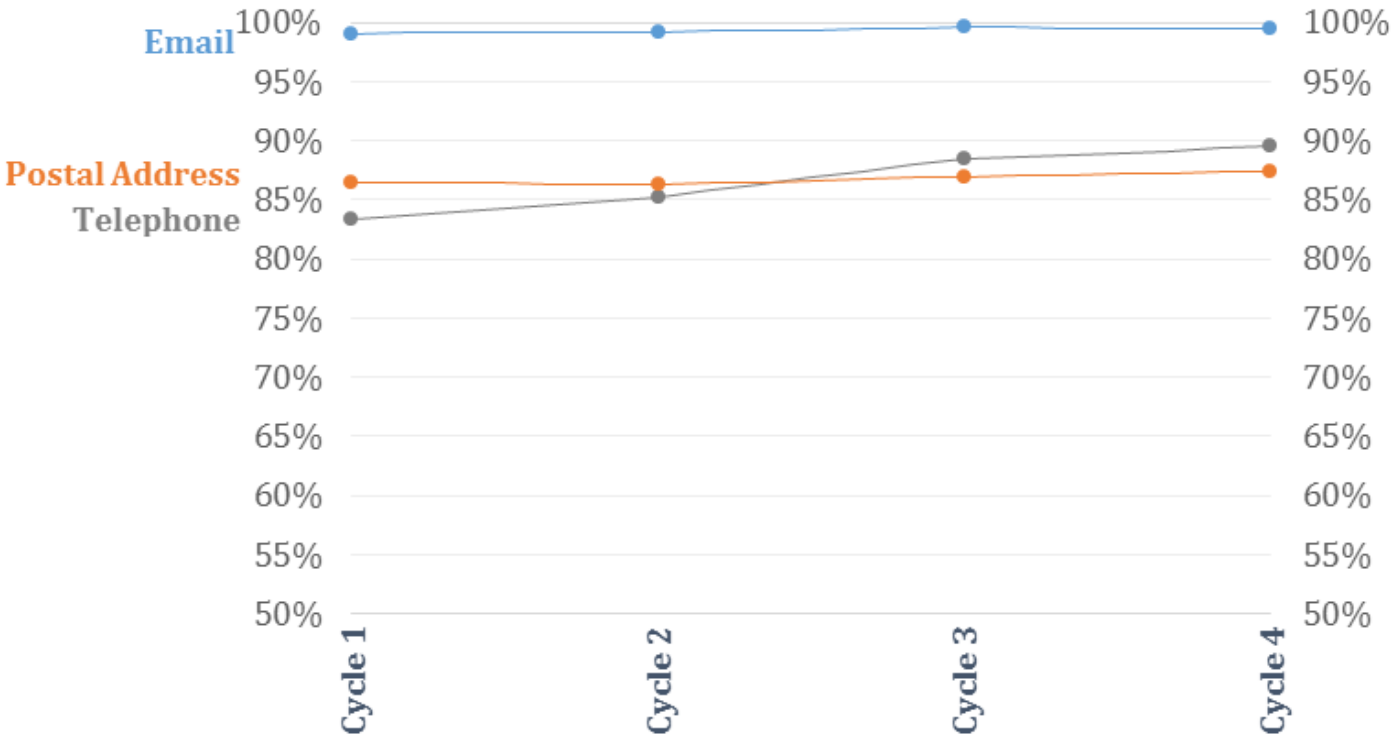
Cycle 2 Phase 4:

Results and Findings: Syntax, 2009 RAA

Phase 2 Cycle 4 – Changes Over Time: Syntax Accuracy by Contact Mode

Overall Syntax Accuracy			
Cycle 1	Cycle 2	Cycle 3	Cycle 4
73.1%	75.3%	78.0%	79.3%
	Δ 2.2%	Δ 2.7%	Δ 1.3%

Entire gTLD Space
Cycle 1 through Cycle 4

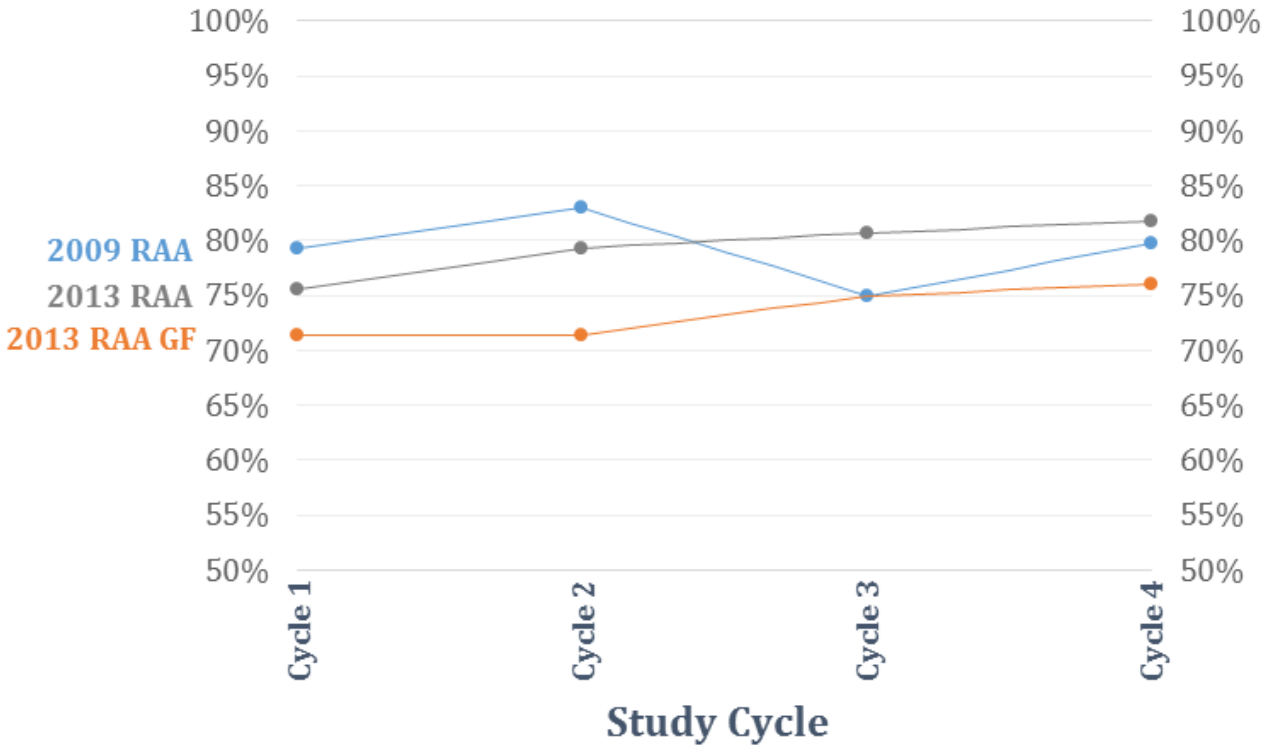


Accuracy of all 3 contact modes and all 3 contact types

Phase 2 Cycle 4 – Changes Over Time: Syntax Accuracy by RAA Type

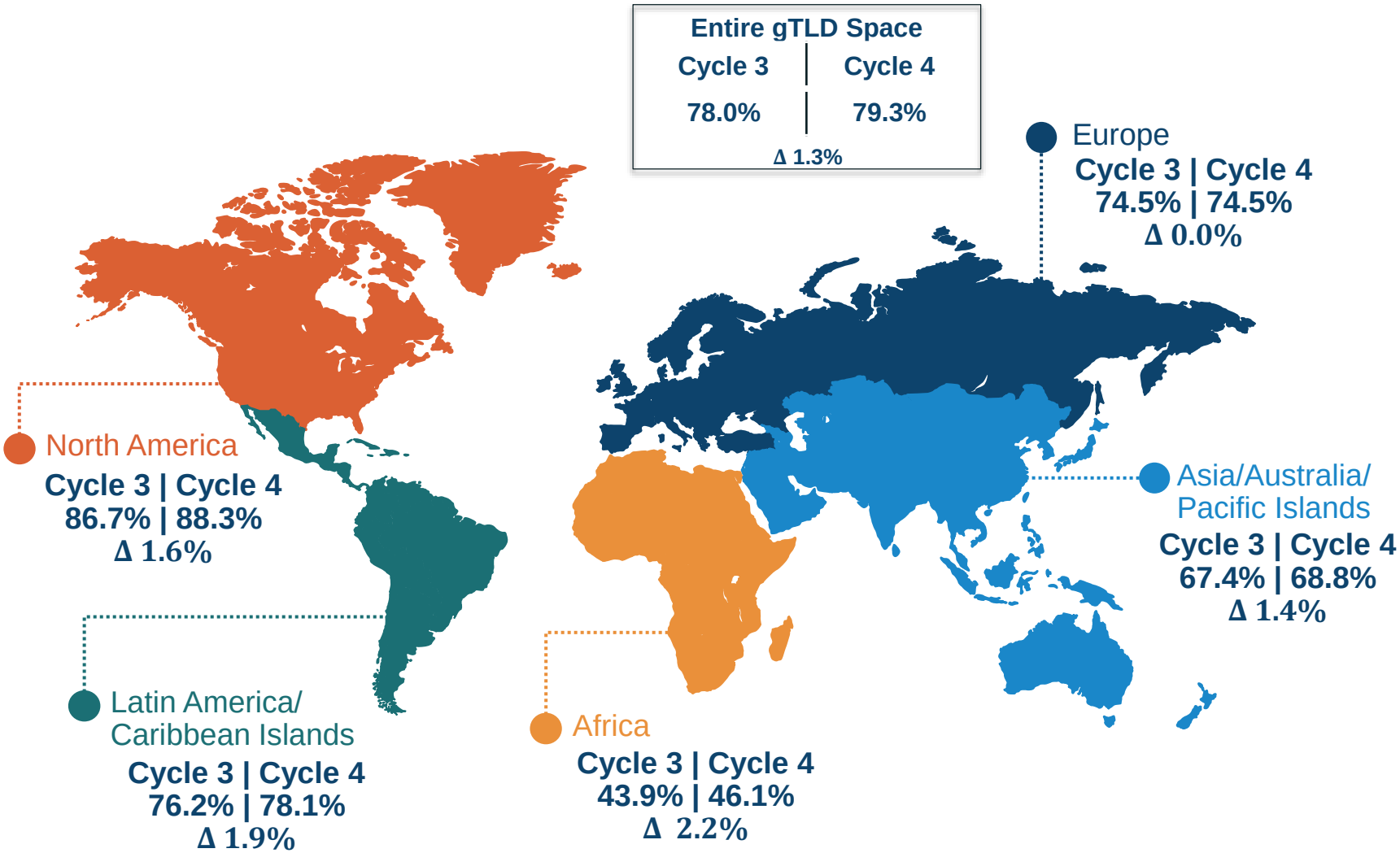
Entire gTLD Space
Cycle 1 through Cycle 4

Overall Syntax Accuracy			
Cycle 1	Cycle 2	Cycle 3	Cycle 4
73.1%	75.3%	78.0%	79.3%
	Δ 2.2%	Δ 2.7%	Δ 1.3%



Accuracy of all 3 contact modes and all 3 contact types

Overall Syntax Accuracy by Region – Cycle 3 v. Cycle 4



Cycle 2 Phase 4:

ICANN Contractual Compliance Follow-Up

Phase 2 Cycle 4 – ICANN Contractual Compliance

- Potentially inaccurate records identified by the ARS are provided to ICANN Contractual Compliance
- WHOIS inaccuracy and format complaints will follow the Contractual Compliance Approach and Process
- Registrars must investigate and correct inaccurate WHOIS data:
  - Section 3.7.8 of 2009 and 2013 RAA (and WHOIS Accuracy Program Specification)
  - Failure to respond or demonstrate compliance during complaint processing will result in a Notice of Breach
- ICANN will continue to give priority to complaints submitted by community members
- The process of reviewing and reporting WHOIS ARS test results is time consuming such that it takes anywhere from four to five months before ICANN Contractual Compliance can begin processing the WHOIS ARS tickets. This lag can result in outdated WHOIS ARS test results. However, with each new WHOIS ARS test cycle, the WHOIS ARS and ICANN Contractual Compliance teams are working to reduce this lag time

Phase 2 Cycle 4 – ICANN Contractual Compliance

WHOIS ARS Compliance Metrics (as of 1 July 2017)

- Phase 2, Cycle 3:
  - 4,552 tickets created, all have been completed.
  - 2,662 were closed prior to 1st notice. Of those, closure reasons as follows:
    - WHOIS data when ticket processed different from sampled WHOIS data: 60.1%
    - WHOIS format issue identified for 2013 Grandfathered Domain: 14.3%
    - Domain suspended or canceled: 7.9%
    - Domain not registered when ticket processed: 7.7%
    - Known Privacy/Proxy service: 6.3%
    - Duplicate WHOIS compliant already pending: 3.8%
    - Other (remaining closure reasons representing less than 0.5% of cases): 0.4%

- Phase 2, Cycle 4 (In Progress):
  - 4,681 tickets created. 1,424 have been closed, 3,256 remaining to be processed.
  - 984 were closed prior to 1st notice. Of those, closure reasons as follows:
    - WHOIS data when ticket processed different from sampled WHOIS data: 45.4%
    - Domain not registered when ticket processed: 26.7%
    - Domain suspended or canceled: 13.1%
    - WHOIS format issue identified for 2013 Grandfathered Domain: 13.0%
    - Known Privacy/Proxy service
    - Other (remaining closure reasons representing less than 0.5% of cases): 0.1%

- More WHOIS ARS Compliance follow-up metrics are now available on the ICANN.org WHOIS ARS page here: https://whois.icann.org/en/whoisars-contractual-compliance-metrics

Cycle 2 Phase 4:

Summary

Phase 2 Cycle 4 – Summary

Report included information on population demographics; Have seen a large growth in NGF domains and domains from AP region

Subsample of 11.5k records; Accounted for regions and RAA type

98.6% of records immediately contactable; 65.4% operability full accuracy rate on all 2009 RAA requirements

Increase in Email Operability, decrease in Telephone Operability

Syntax Accuracy for all 3 modes remains high

Compliance continues to monitor inaccuracies, ICANN working to decrease lag between record pull and ticket creation.

Next Cycle (5) has already begun; Report expected Dec 2017

WHOIS ARS Data Display Tool

- In response to Community feedback, ICANN will launch a Data Display Tool to allow Community members to sort and display WHOIS ARS results data.
- Beginning next month, visitors to ICANN.org will be able to use the Data Display Tool to sift through ARS data as they see fit.
- No personally identifiable data will be available, only summary statistics.
- Ongoing community feedback of the tool encouraged.

Engage with ICANN – Thank You and Questions

One World, One Internet

Visit us at **icann.org**


@icann


facebook.com/icannorg

youtube.com/icannnews


flickr.com/icann

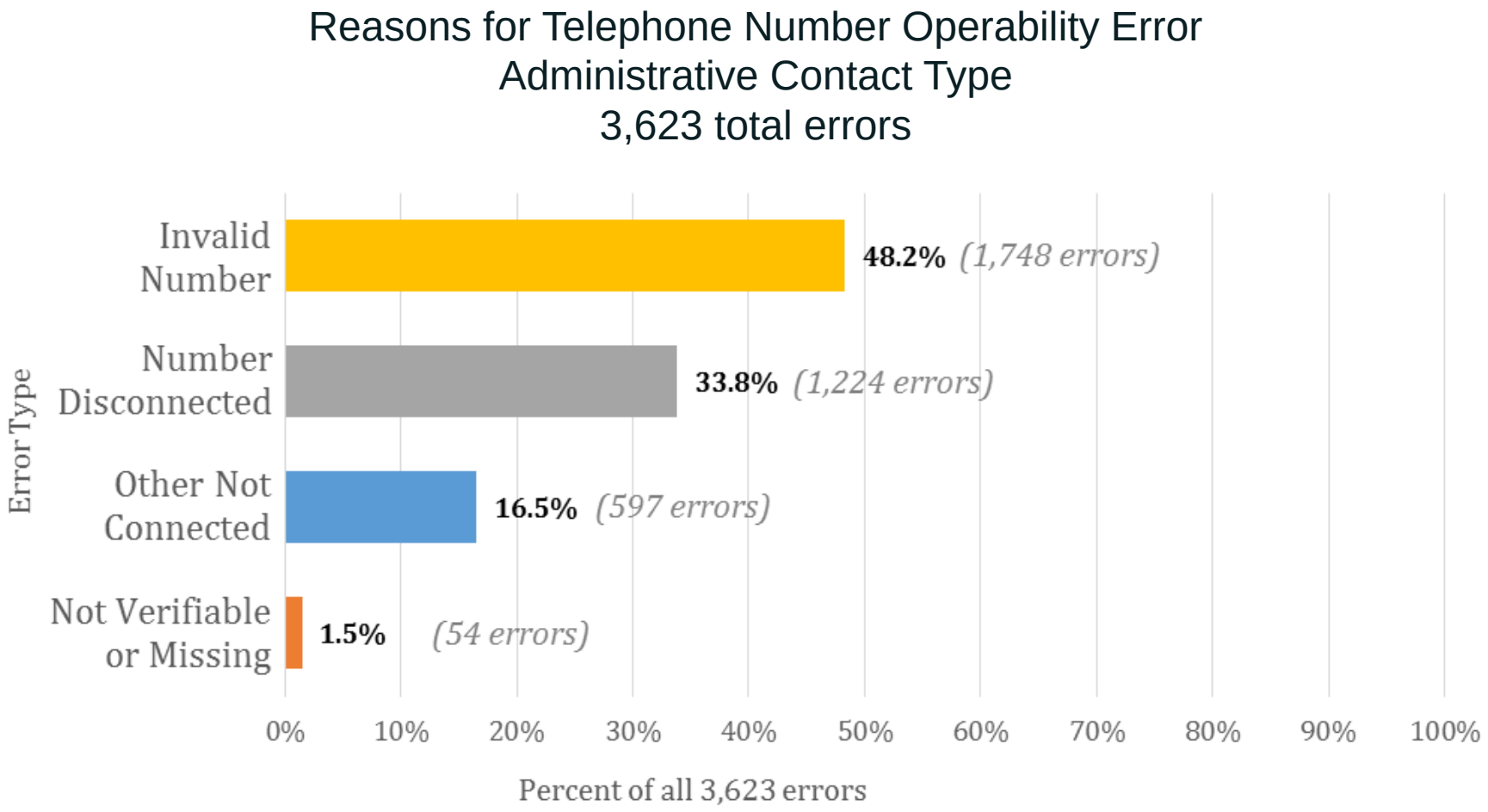

linkedin/company/icann


slideshare/icannpresentations


soundcloud/icann

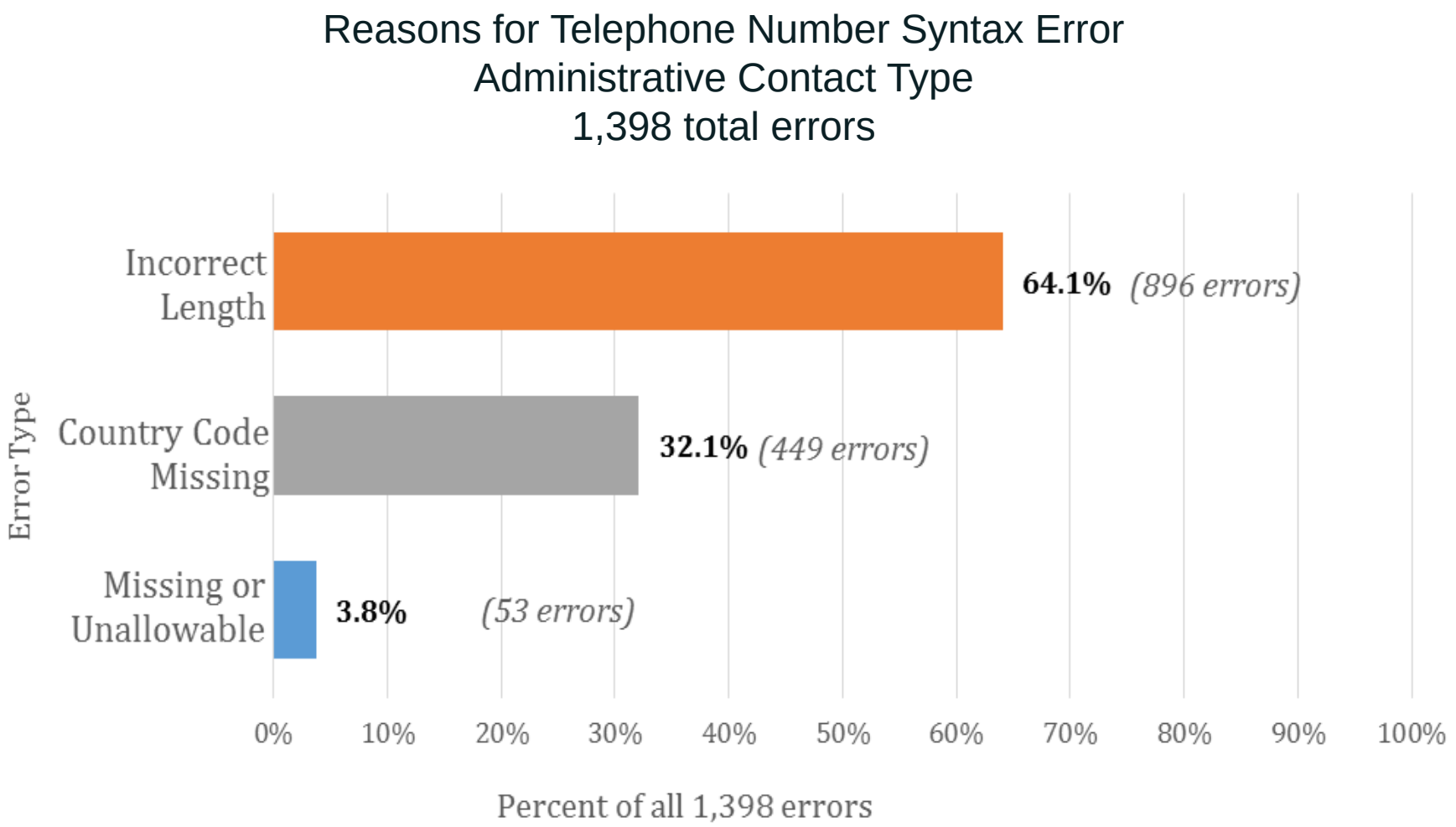
APPENDIX

Phase 2 Cycle 4 – Reasons for Telephone Operability Error



Note: A missing telephone number in the Registrant contact type is not is not a requirement of the 2009 RAA. This graph shows the percentage of overall error types found in the Administrative contact type.

Phase 2 Cycle 4 – Reasons for Telephone Syntax Error



Note: Presence of a telephone number in the Registrant contact type is not is not a requirement of the 2009 RAA. This graph shows the percentage of overall error types found in the Administrative contact type. The "Unallowable Character" error type has been combined with the "Missing" error type, because unallowable character errors represent less than 0.2% of overall errors.

Phase 2 Cycle 4 – Reasons for Address Syntax Error