WHOIS Accuracy Reporting System (ARS):
Phase 2 Cycle 3 Webinar | 12 January 2017

ICANN GDD Operations
NORC at the University of Chicago

ICANN

Phase 2 Cycle 3 – Webinar Agenda

1 WHOIS ARS Background

2 Phase 2 Cycle 3: Timeline and Process

3 Phase 2 Cycle 3: Population and Sample Information

4 Phase 2 Cycle 3: Results and Major Findings

5 Contractual Compliance Follow-up

6 Summary and Question & Answer Session

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: Expected June 2017
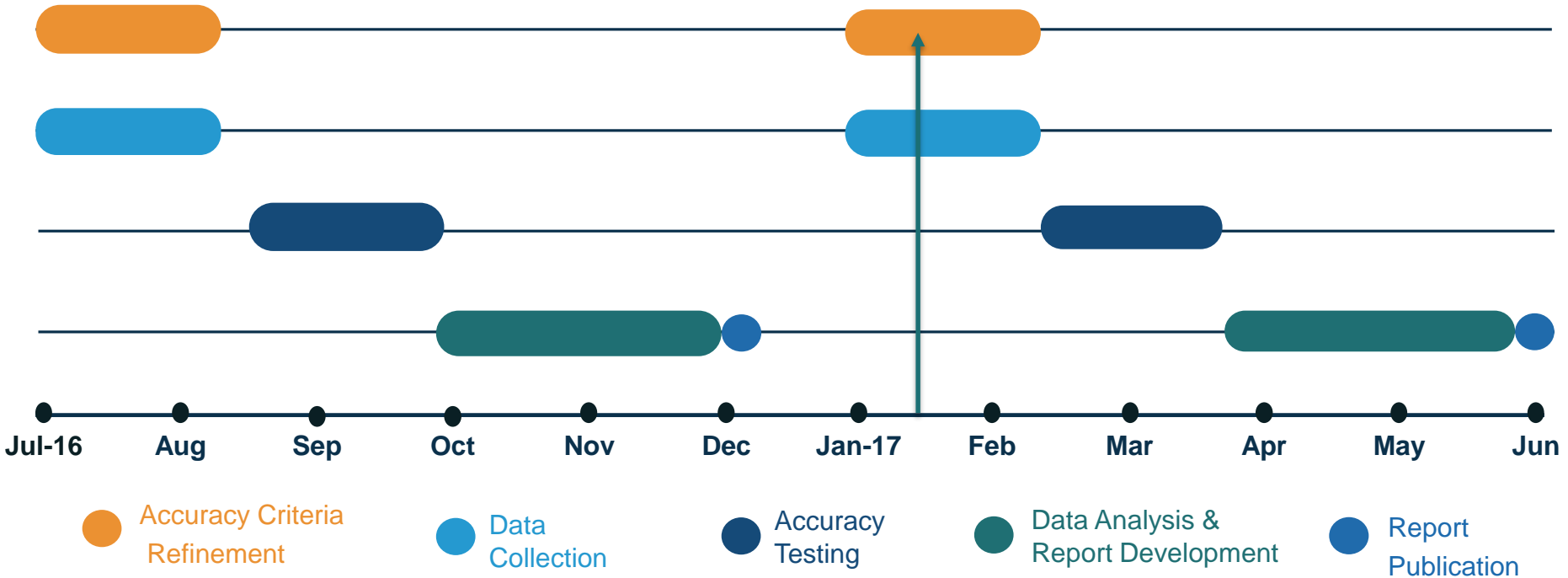
WHOIS ARS Information and Reports available here: https://whois.icann.org/en/whoisars

Phase 2 Cycle 3
Process and Timeline

Phase 2 Cycle 3 – Cyclical Timeline of Phase 2

Cycle 3: Started July 2016

Cycle 4: Jan 2017

Jul-16 Aug Sep Oct Nov Dec Jan-17 Feb Mar Apr May Jun

- ● Accuracy Criteria Refinement
- ● Data Collection
- ● Accuracy Testing
- ● Data Analysis & Report Development
- ● Report Publication

- Cycle 3 Report published 12 December 2016
- Cycle 4 has already started; records will start to be pulled in the next couple of weeks
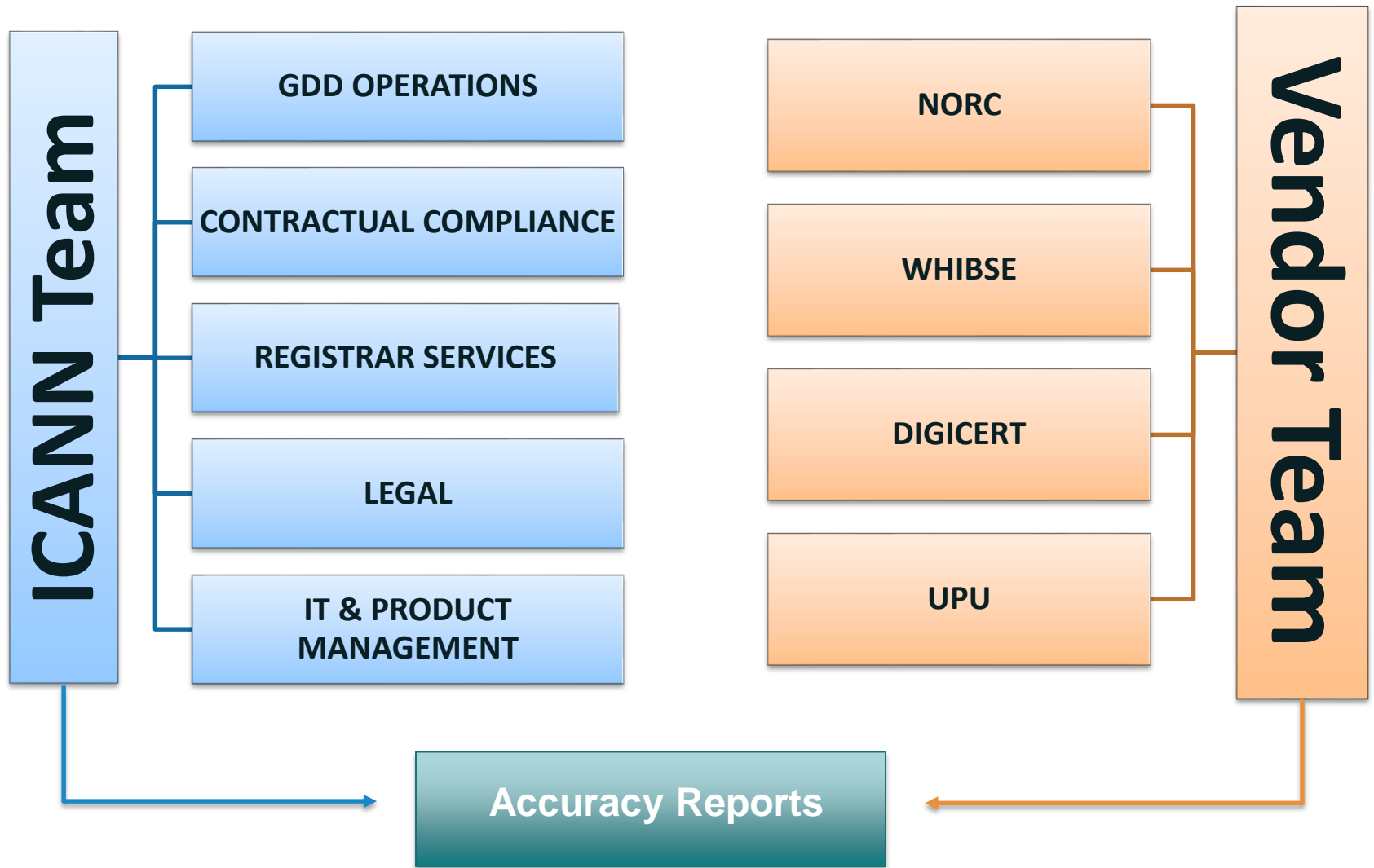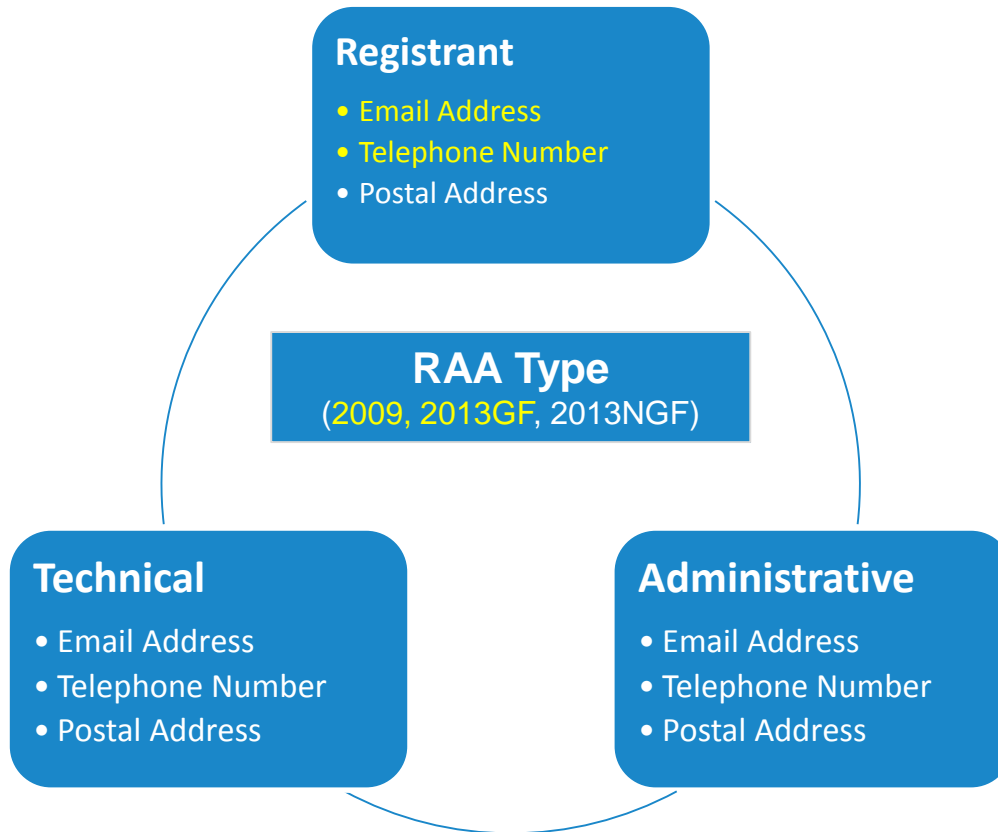
Phase 2 Cycle 3
Testing Criteria and Sample Design

Phase 2 Cycle 3 – Contact types, modes, and testing criteria

Registrant
- Email Address
- Telephone Number
- Postal Address

RAA Type
(2009, 2013GF, 2013NGF)

Technical
- Email Address
- Telephone Number
- Postal Address

Administrative
- Email Address
- Telephone Number
- Postal Address

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
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 3 – Report Highlights

Accuracy Statistics by Subgroup

⊙ Phase 2 reports provide syntax and operability accuracy rates for:
- The gTLD space, with further breakdown by geographic region
- New gTLDs and Prior (legacy) gTLDs
- RAA Type (2009, 2013GF, 2013NGF)
- Privacy/Proxy status is not evaluated

⊙ 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

⊙ Cycle 3 Report, like Cycle 2, contains section on regional differences in accuracy and reasons for error

⊙ Cycle 3 Report also contains additional information on domain name population demographics

Change in Distribution across Sample Dates
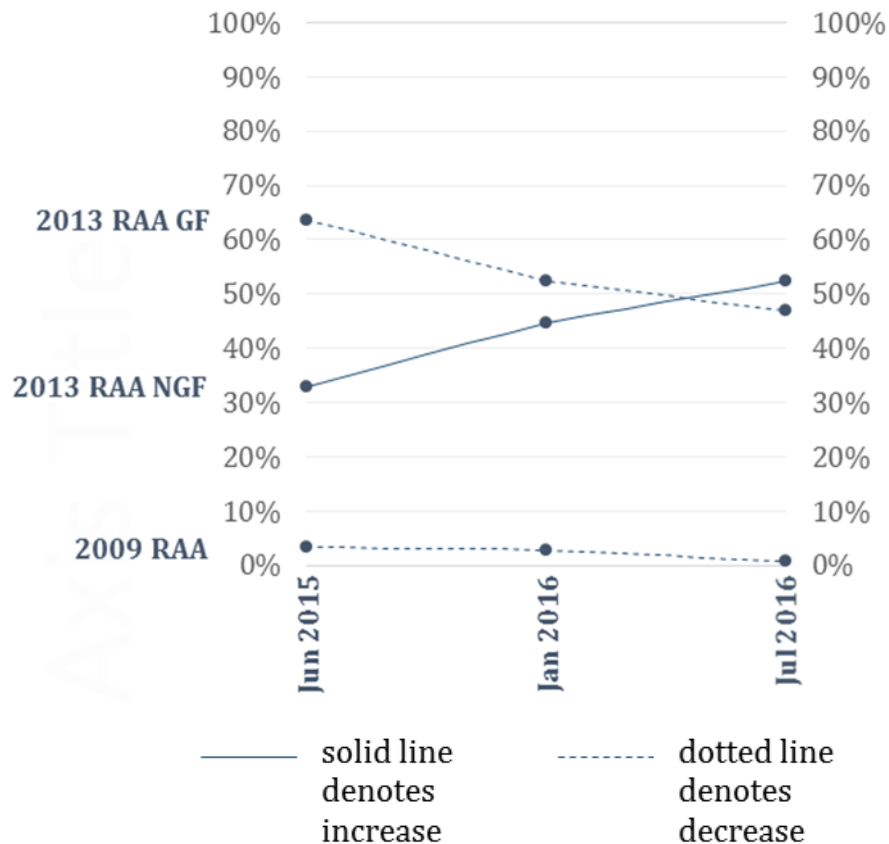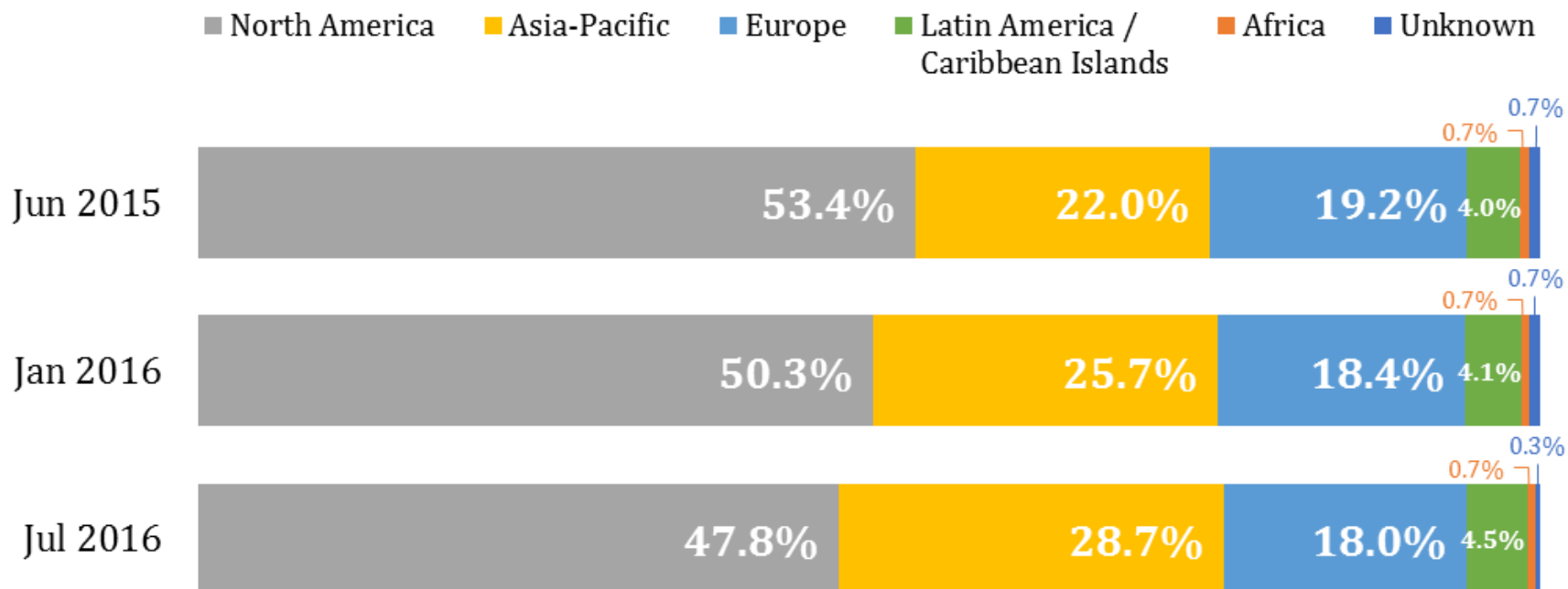

The 2009 RAA share is shrinking;

The share of non-grandfathered 2013 RAA domains are growing rapidly (from ~30% of distribution in June 2015 to ~50% in July 2016).

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%

Regional Distribution of Domains, by Sample Date

Legend: ■ North America ■ Asia-Pacific ■ Europe ■ Latin America / Caribbean Islands ■ Africa ■ Unknown

Jun 2015: North America 53.4%, Asia-Pacific 22.0%, Europe 19.2%, Latin America/Caribbean 4.0%, Africa 0.7%, Unknown 0.7%

Jan 2016: North America 50.3%, Asia-Pacific 25.7%, Europe 18.4%, Latin America/Caribbean 4.1%, Africa 0.7%, Unknown 0.7%

Jul 2016: North America 47.8%, Asia-Pacific 28.7%, Europe 18.0%, Latin America/Caribbean 4.5%, Africa 0.7%, Unknown 0.3%

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**

Phase 2 Cycle 3 – Demographics

gTLD Population At Time of Sample (July 2016)

Records in gTLDs	Total gTLDs	2009 RAA*	2013GF RAA*	2013 NGF RAA*	New gTLDs	Prior gTLDs
181.2m	1,074	1.5m	85.1m	94.7m	1,056	18

200k Sample

AFR	LAC	EUR	APAC	N.A.	2009 RAA	2013GF RAA	2013 NGF RAA	New gTLDs	Prior gTLDs
1.3k	9.9k	36.2k	65.2k	84.2k	1.5k	78.6k	117.1k	646	18

12k Sub-sample

AFR	LAC	EUR	APAC	N.A.	2009 RAA	2013GF RAA	2013 NGF RAA	New gTLDs	Prior gTLDs
1.3k	1.9k	2.2k	2.9k	3.6k	1.4k	4.8k	5.7k	646	18

* Weighted estimates from 200k sample

Phase 2 Cycle 3

Changes in Accuracy from Cycle 2 to Cycle 3

ICANN

Cycle 2 to Cycle 3 Changes

Overall
- Overall Syntax accuracy decreased slightly from 67.4% to 66.6%%.
- Operability accuracy decreased from 70.2% to 65.1%

Email
- Syntax accuracy increased slightly from 99.2% to 99.6%%.
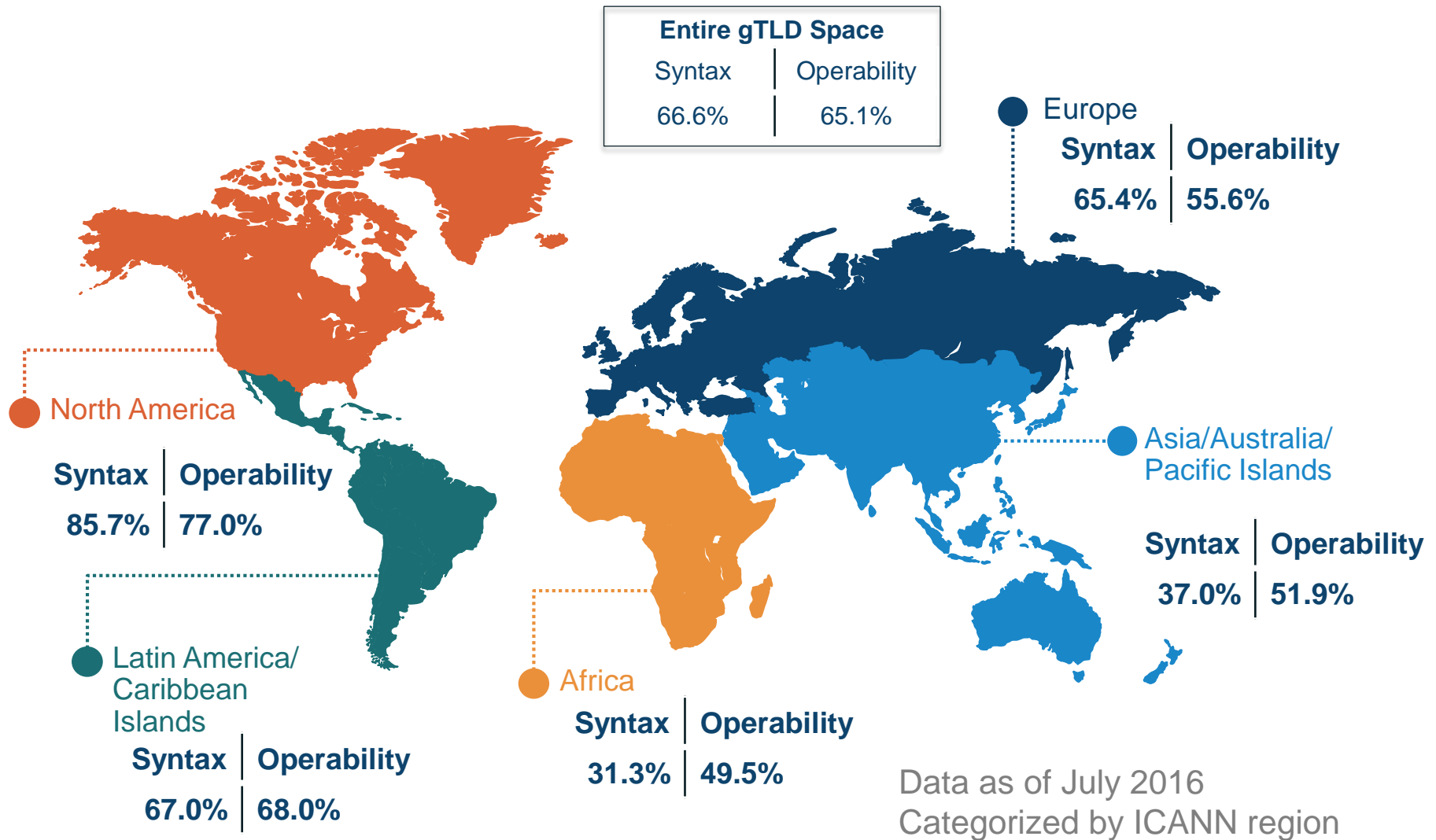- Operability accuracy decreased slightly from 91.4% to 90.1%

Telephone
- Syntax accuracy increased from 85.3% to 88.5%
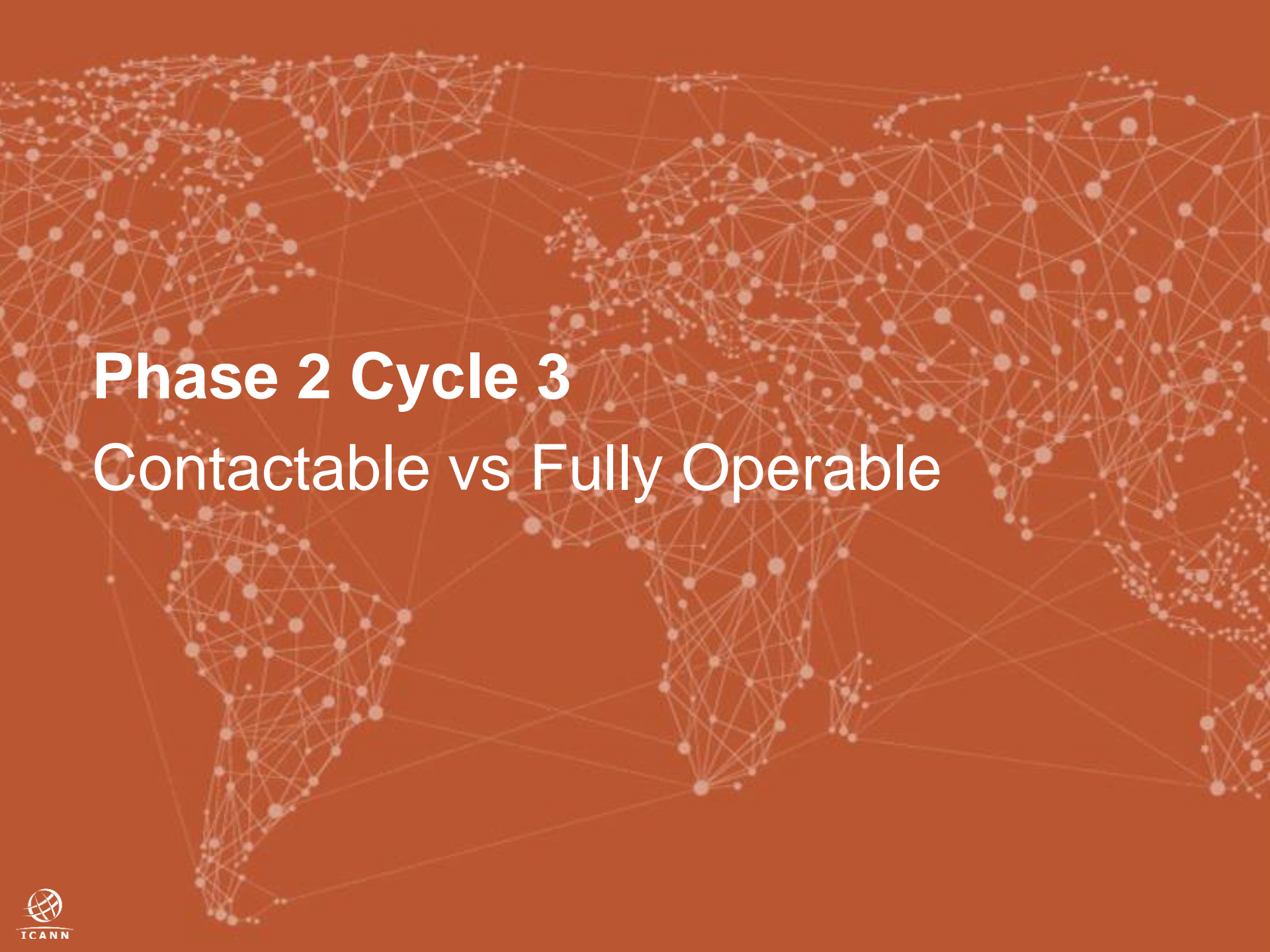- Operability decreased from 76.0% to 72.4%

Postal
- Syntax accuracy decreased from 77.4% to 74.7%
- Operability accuracy decreased slightly from 97.6% to 96.8%

Phase 2 Cycle 3 – Overall Syntax and Operability Accuracy by Region

Entire gTLD Space

Syntax	Operability
66.6%	65.1%

Europe

Syntax	Operability
65.4%	55.6%

North America

Syntax	Operability
85.7%	77.0%

Asia/Australia/Pacific Islands

Syntax	Operability
37.0%	51.9%

Latin America/Caribbean Islands

Syntax	Operability
67.0%	68.0%

Africa

Syntax	Operability
31.3%	49.5%

Data as of July 2016
Categorized by ICANN region

Phase 2 Cycle 3
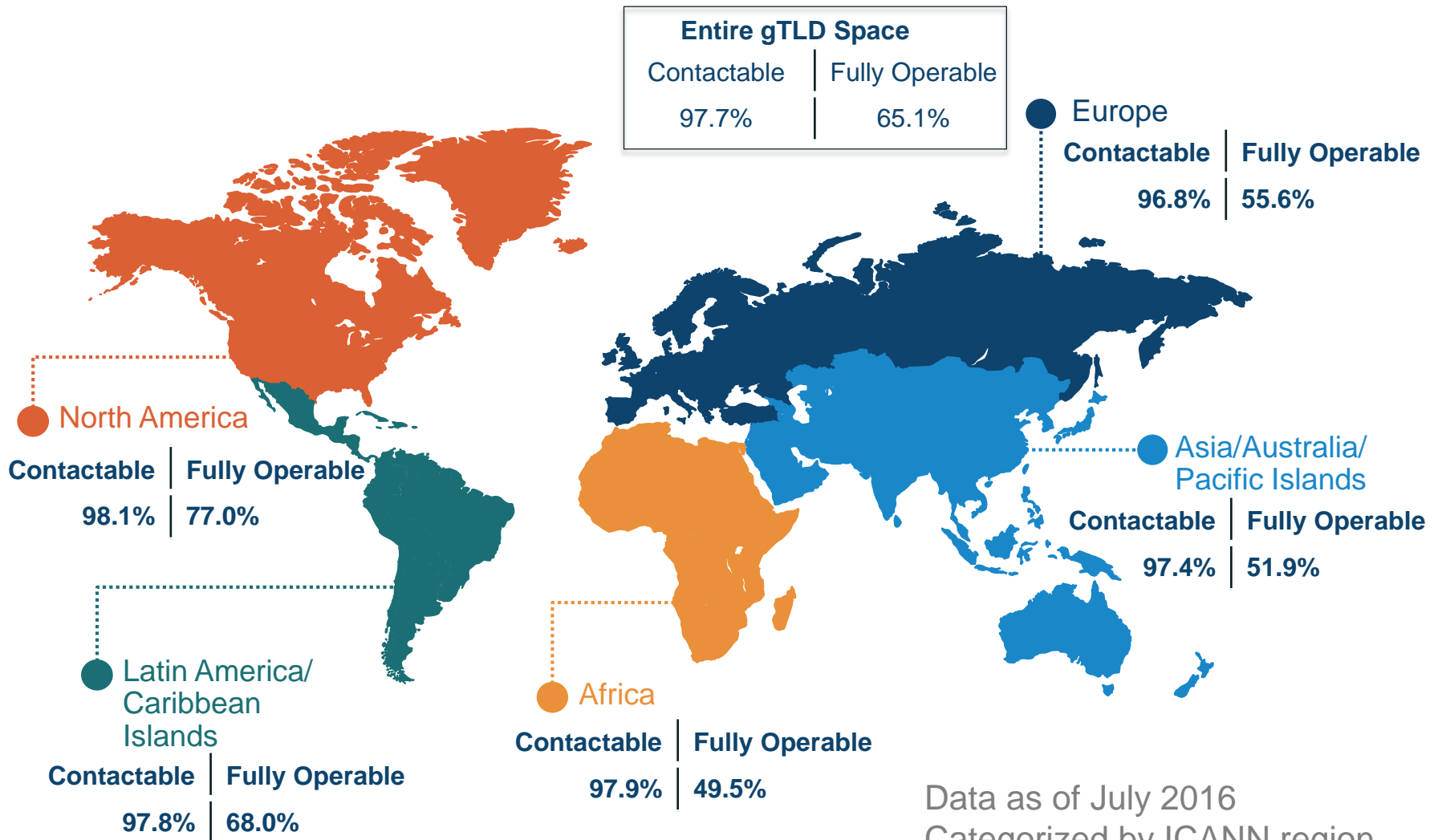Contactable vs Fully Operable

ICANN

97% Contactable

Record contains at least one operable email address or telephone number

65% Fully Operable

Strict conformance to the RAA; all nine contacts are operable

Phase 2 Cycle 3 – Contactable Domains, by Region

Entire gTLD Space

Contactable	Fully Operable
97.7%	65.1%

Europe

Contactable	Fully Operable
96.8%	55.6%

North America

Contactable	Fully Operable
98.1%	77.0%

Asia/Australia/Pacific Islands

Contactable	Fully Operable
97.4%	51.9%

Latin America/Caribbean Islands

Contactable	Fully Operable
97.8%	68.0%

Africa

Contactable	Fully Operable
97.9%	49.5%

Data as of July 2016
Categorized by ICANN region

Phase 2 Cycle 3

Results and Major Findings: Operability, 2009 RAA

ICANN

Entire gTLD Space
Cycle 1 through Cycle 3

Overall Operability Accuracy		
Cycle 1	Cycle 2	Cycle 3
64.7%	70.2%	65.1%
Δ 5.5%		Δ -5.1%
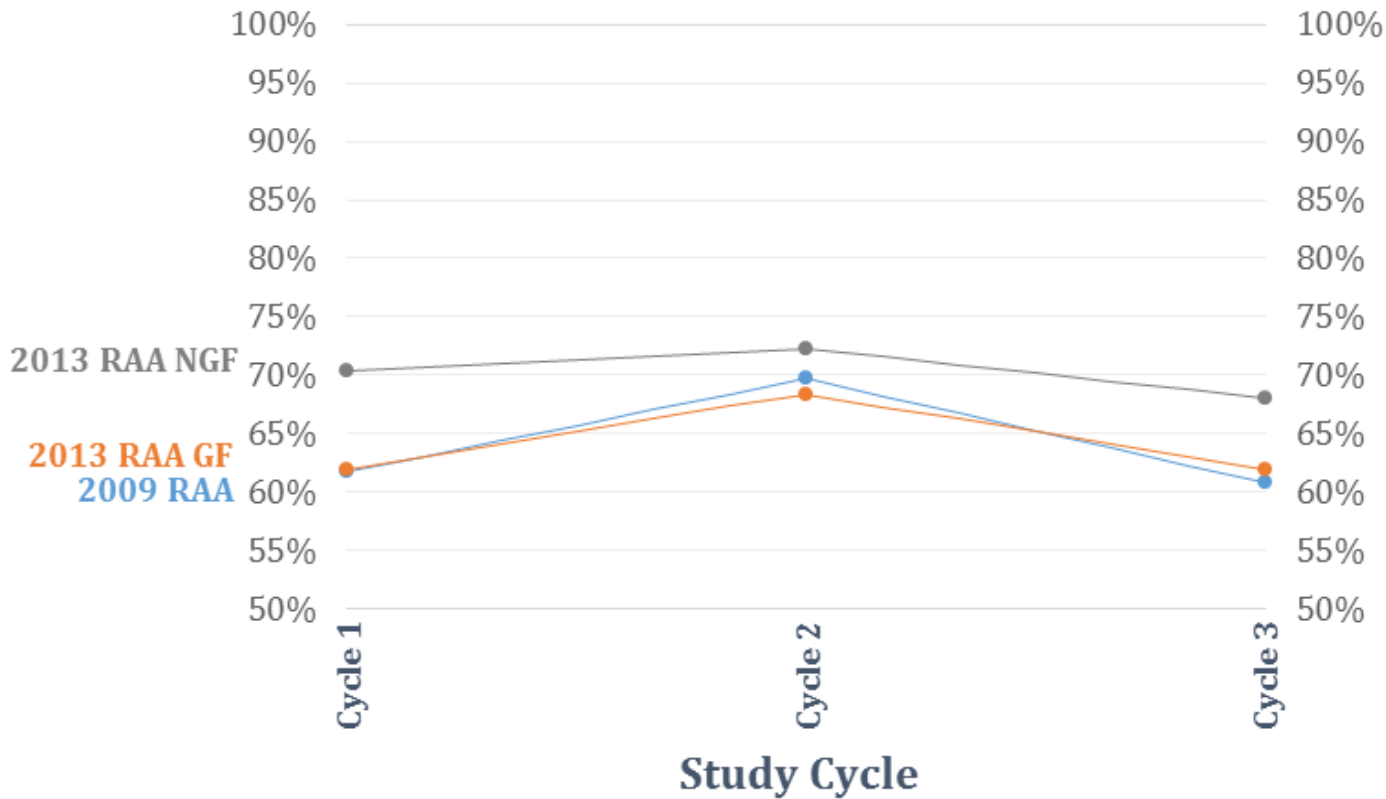
Accuracy of all 3 contact modes and all 3 contact types

Overall Operability Accuracy

Cycle 1	Cycle 2	Cycle 3
64.7%	70.2%	65.1%
	Δ 5.5%	Δ -5.1%

Entire gTLD Space
Cycle 1 through Cycle 3

Accuracy of all 3 contact modes and all 3 contact types

Phase 2 Cycle 3 – Overall Operability Accuracy by Region

Entire gTLD Space

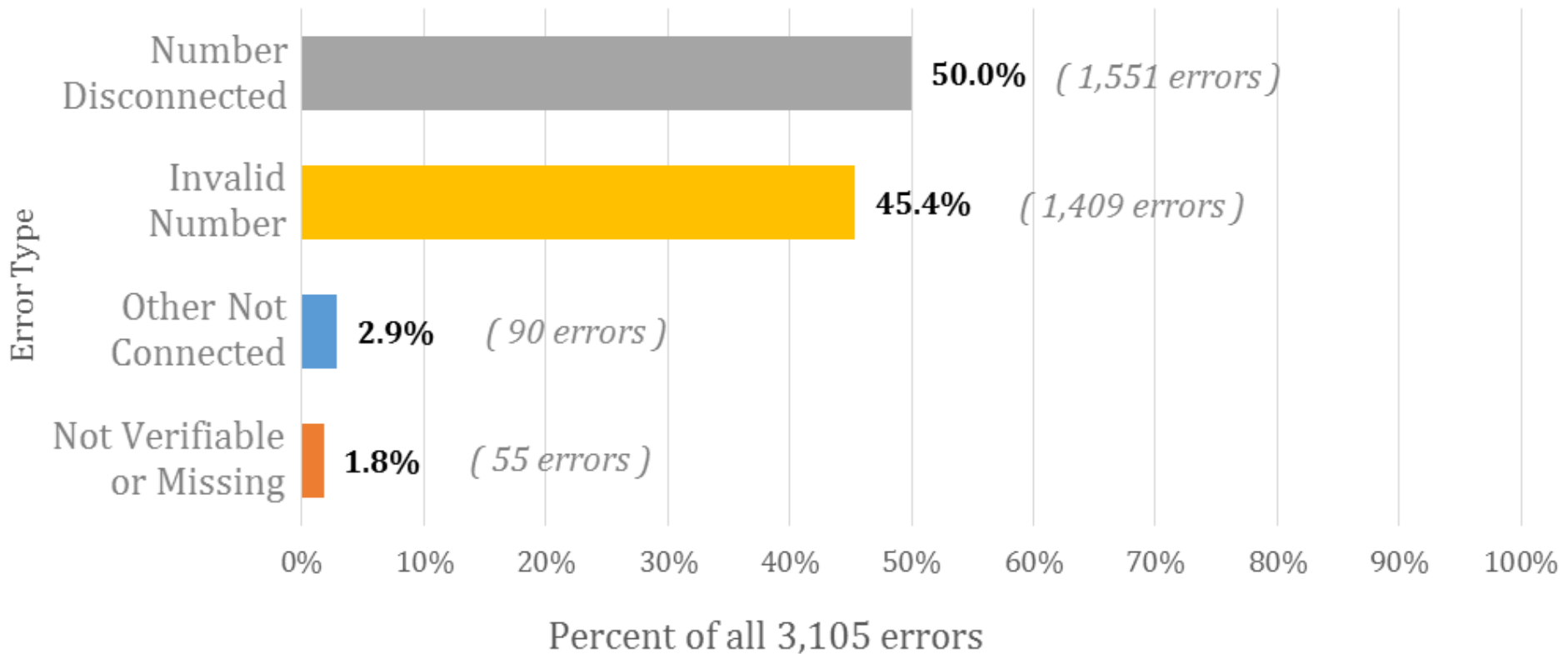| Cycle 2	Cycle 3
70.2%	65.1%

Δ -5.0%*

Europe
Cycle 2 | Cycle 3
63.0% | 55.6%
Δ -7.4%*

North America
Cycle 2 | Cycle 3
80.2% | 77.0%
Δ -3.2%

Asia/Australia/Pacific Islands
Cycle 2 | Cycle 3
57.6% | 51.9%
Δ -5.7%*

Latin America/Caribbean Islands
Cycle 2 | Cycle 3
71.6% | 68.0%
Δ -3.6%

Africa
Cycle 2 | Cycle 3
64.6% | 49.5%
Δ -15.1%*
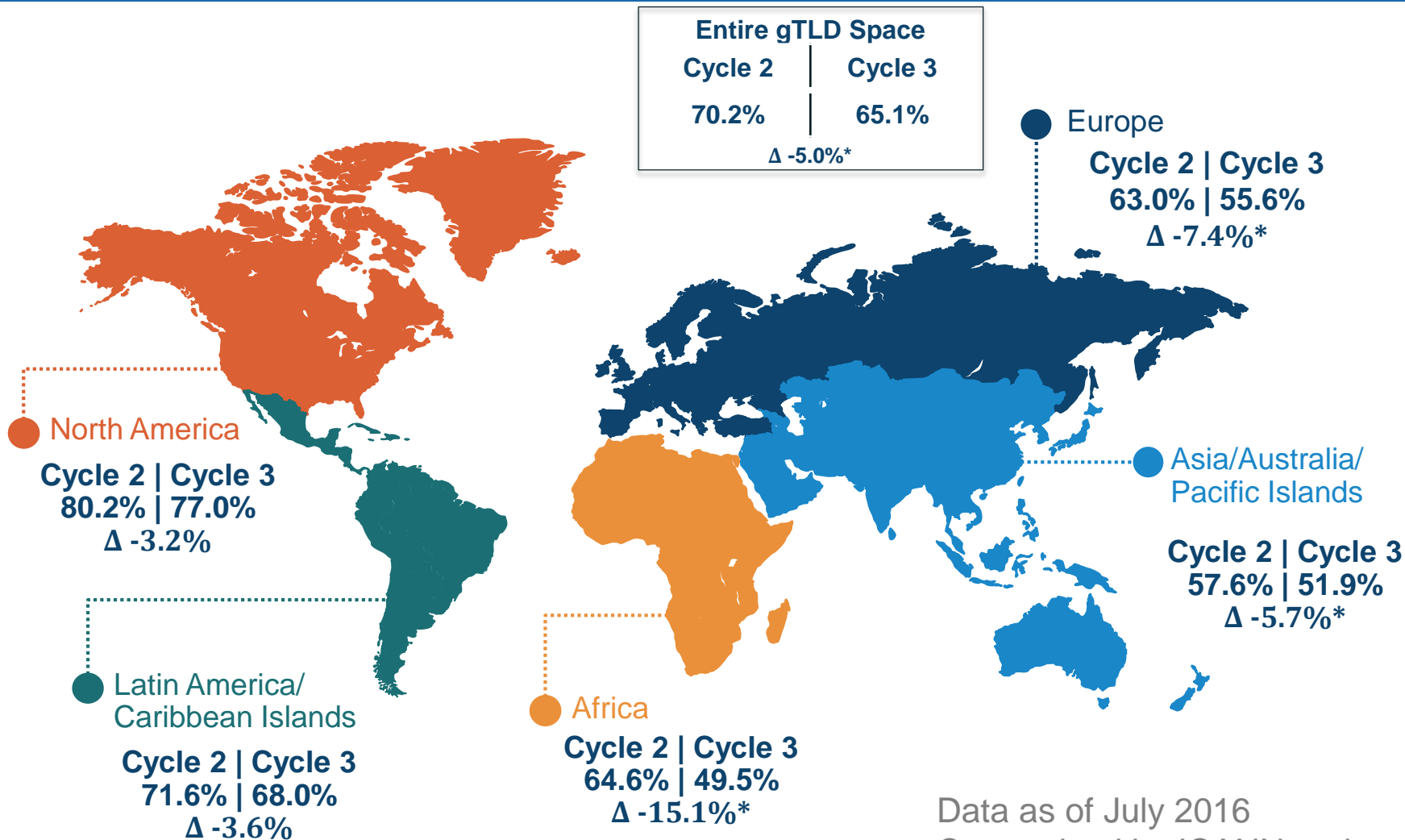
Data as of July 2016
Categorized by ICANN region

* Denotes significant change

Reasons for Telephone Number Operability Error
Administrative Contact Type
3,105 total errors

Percent of all 3,105 errors

Note: A telephone number in the Registrant contact type is not a requirement of the 2009 RAA. This graph shows the percentage of overall error types found in the Administrative contact type.
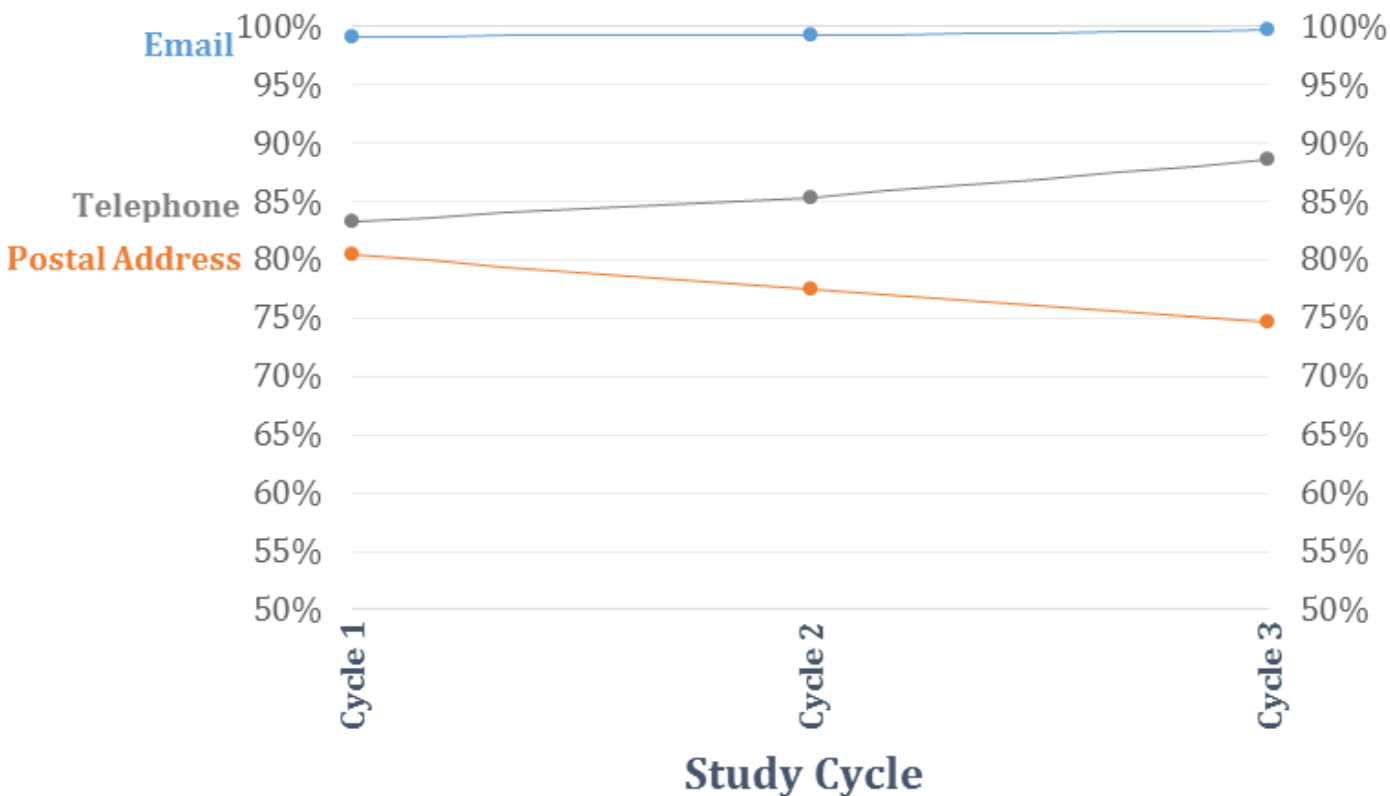
Phase 2 Cycle 3

Results and Major Findings:
Syntax, 2009 RAA

ICANN

Overall Syntax Accuracy		
Cycle 1	Cycle 2	Cycle 3
68.0%	67.4%	66.6%
Δ -0.6%		Δ -0.8%

Entire gTLD Space
Cycle 1 through Cycle 3

Accuracy of all 3 contact modes and all 3 contact types

Overall Syntax Accuracy		
Cycle 1	Cycle 2	Cycle 3
68.0%	67.4%	66.6%
Δ -0.6%		Δ -0.8%

Entire gTLD Space
Cycle 1 through Cycle 3
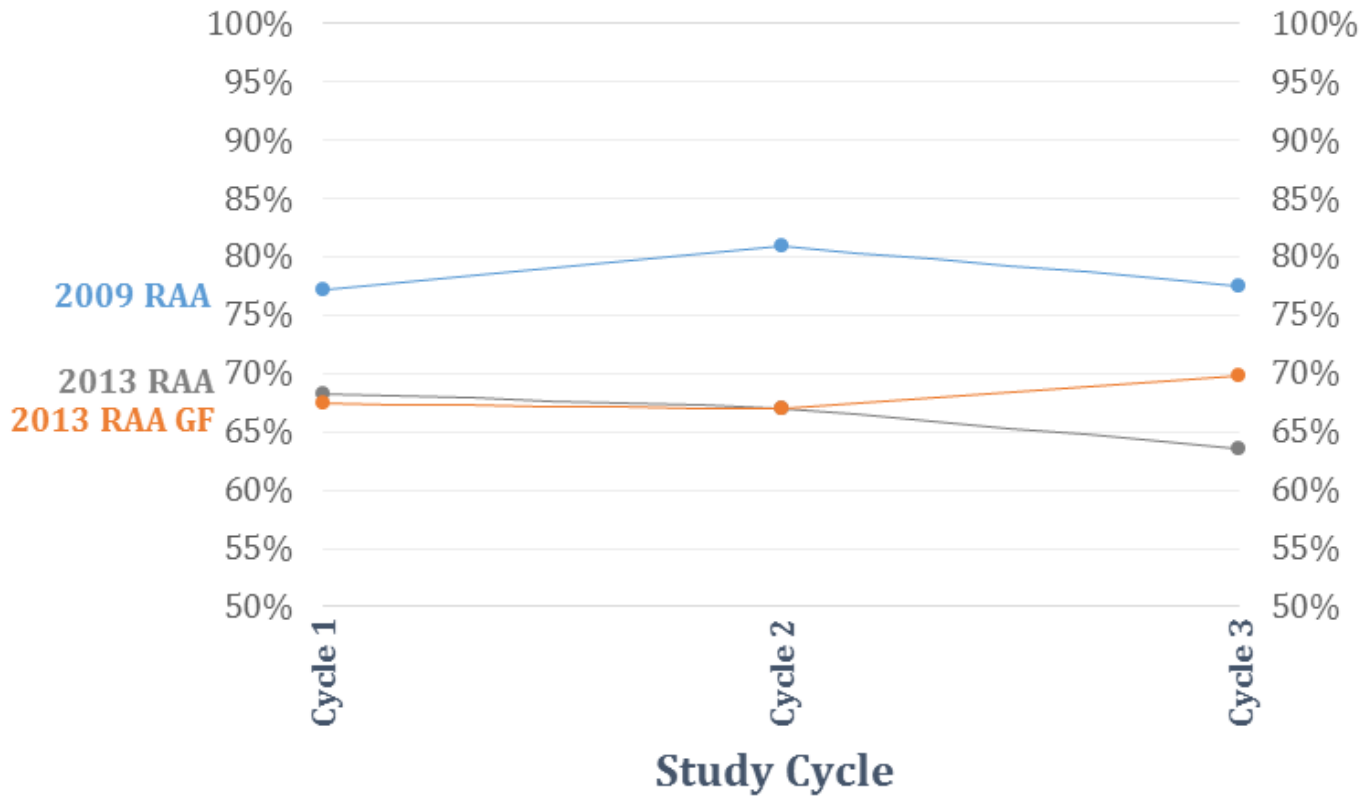

Accuracy of all 3 contact modes and all 3 contact types

Phase 2 Cycle 3 – Overall Syntax Accuracy by Region

Entire gTLD Space

| Cycle 2	Cycle 3
67.4%	66.6%

Δ -0.8%
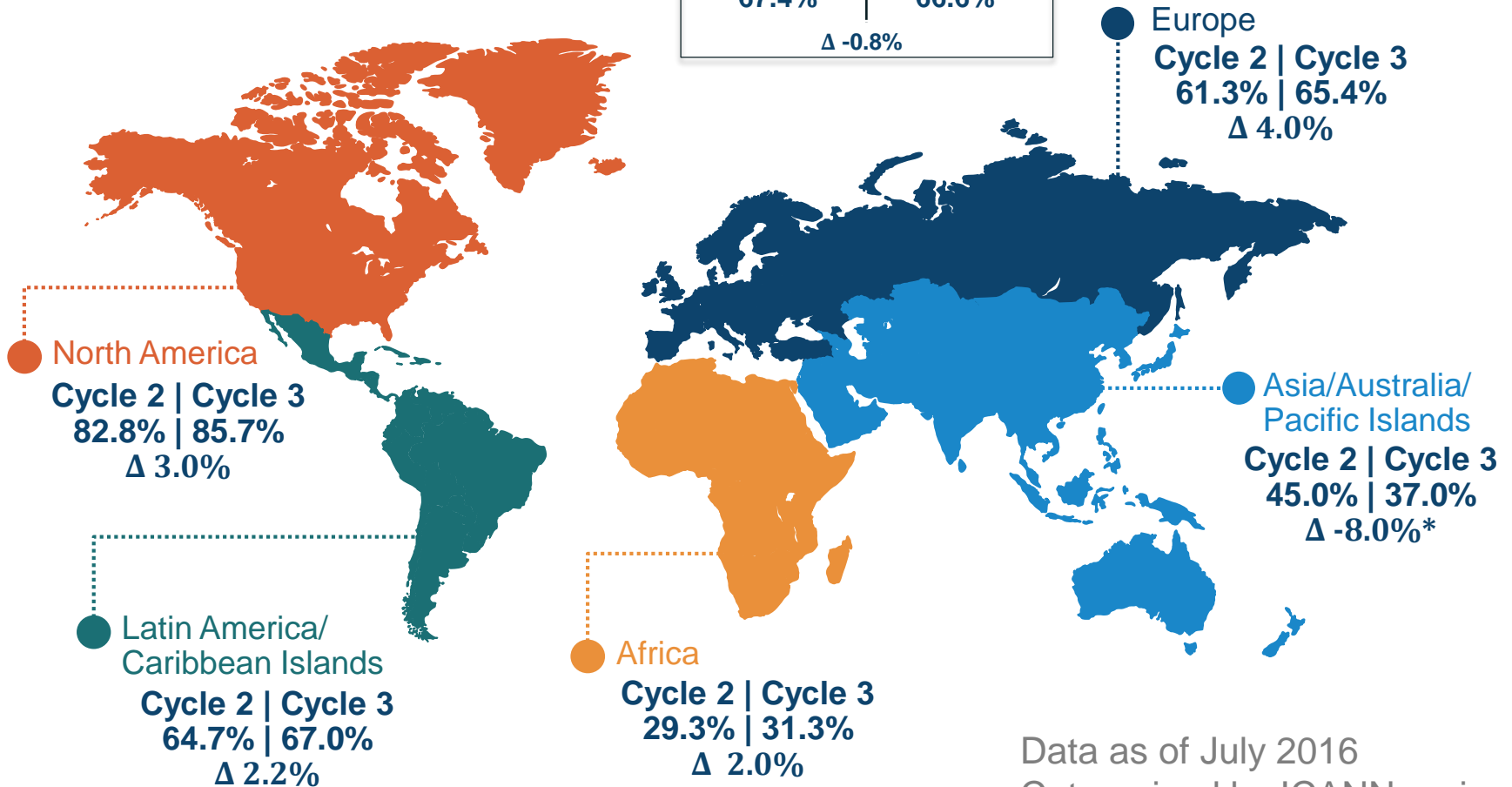
Europe
Cycle 2 | Cycle 3
61.3% | 65.4%
Δ 4.0%

North America
Cycle 2 | Cycle 3
82.8% | 85.7%
Δ 3.0%

**Asia/Australia/
Pacific Islands**
Cycle 2 | Cycle 3
45.0% | 37.0%
Δ -8.0%*

**Latin America/
Caribbean Islands**
Cycle 2 | Cycle 3
64.7% | 67.0%
Δ 2.2%

Africa
Cycle 2 | Cycle 3
29.3% | 31.3%
Δ 2.0%

Data as of July 2016
Categorized by ICANN region
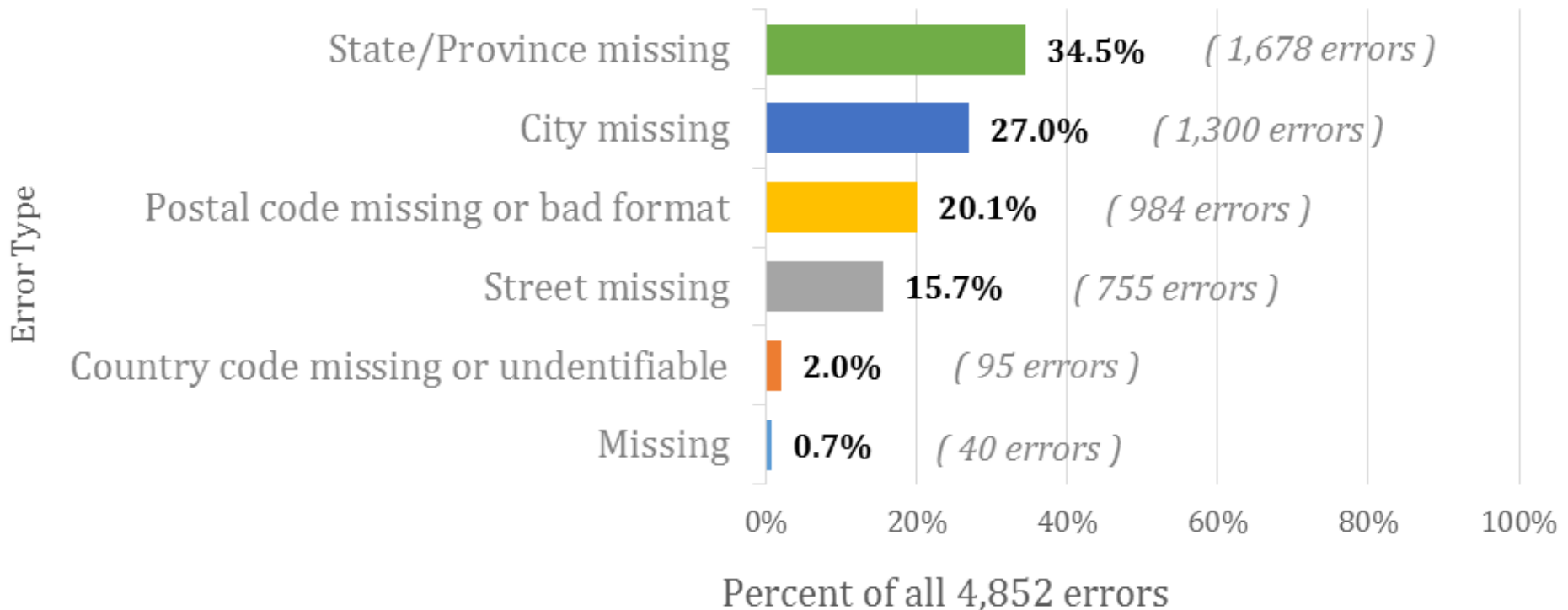
* Denotes significant change

Reasons for Telephone Number Syntax Error
Administrative Contact Type
1,588 total errors
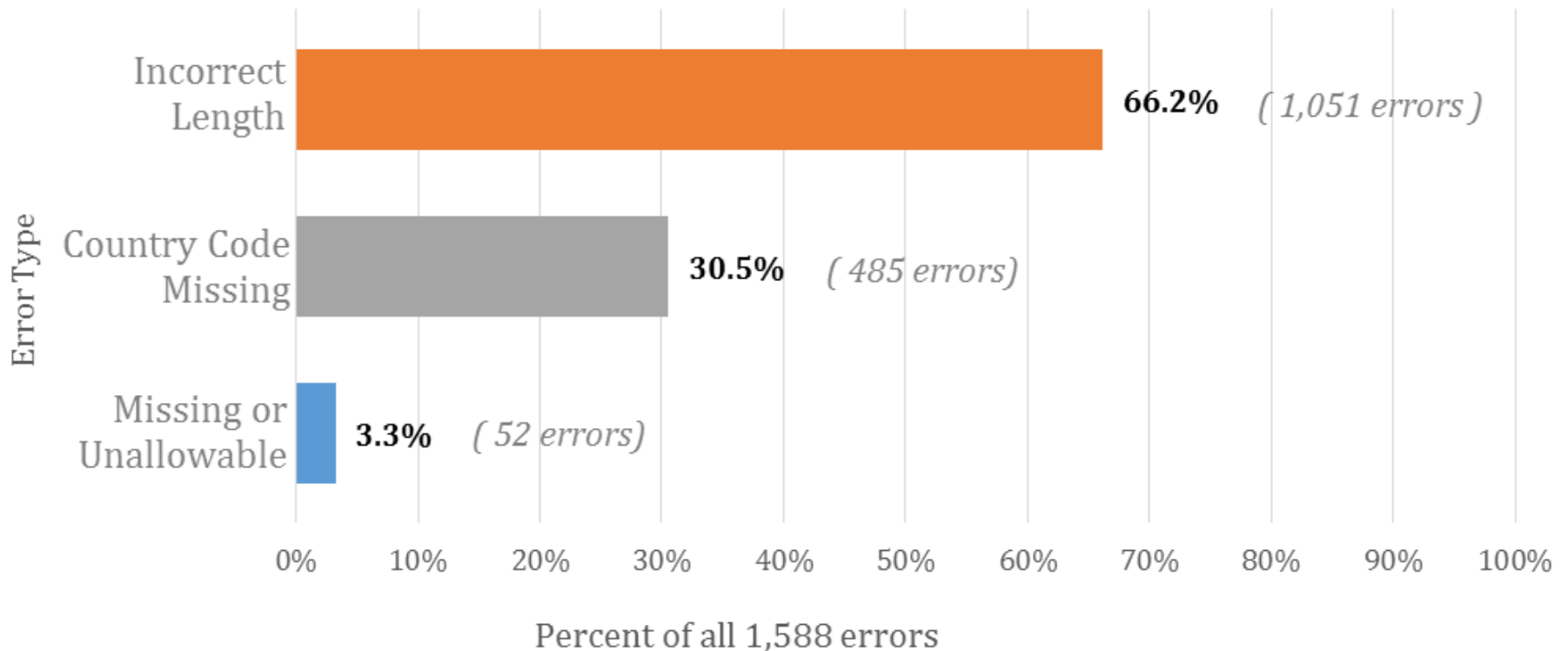
Note: Presence of a telephone number in the Registrant contact type is not a requirement of the 2009 RAA. This graph shows the percentage of overall error types found in the Administrative contact type. The "Unallowable Character" error type has been combined with the "Missing" error type, because unallowable character errors represent less than 0.2% of overall errors.

Reasons for Postal Address Syntax Error
Administrative Contact Type
4,852 Total Errors

Error Type	Percent	Errors
State/Province missing	34.5%	(1,678 errors)
City missing	27.0%	(1,300 errors)
Postal code missing or bad format	20.1%	(984 errors)
Street missing	15.7%	(755 errors)
Country code missing or undentifiable	2.0%	(95 errors)
Missing	0.7%	(40 errors)

Percent of all 4,852 errors

Phase 2 Cycle 3

ICANN Contractual Compliance Follow-Up

Phase 2 Cycle 3 – ICANN Contractual Compliance

- ◎ Potentially inaccurate records have been provided to ICANN Contractual Compliance

- ◎ 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

- ◎ WHOIS inaccuracy and format complaints will follow the Contractual Compliance Approach and Process

- ◎ ICANN will continue to give priority to complaints submitted by community members

- ◎ Whois ARS Compliance Metrics (as of 1 January 2017)
  - ◎ Phase 2, Cycle 2:
    - ◎ 4,006 tickets created
    - ◎ 2,468 closed before first notice (generally because the domain names either had already been suspended or the current WHOIS data was not consistent with the sampled data).
    - ◎ 3,986 closed
  - ◎ Phase 2, Cycle 3:
    - ◎ 4,552 tickets created
    - ◎ 1,171 closed before first notice (generally because the domain names either had already been suspended or the current WHOIS data was not consistent with the sampled data).
    - ◎ 1,471 closed
  - ◎ More detailed metrics are available in the Compliance Quarterly Reports (see https://www.icann.org/resources/pages/compliance-reports-2016-04-15-en).

Summary and Q&A

Phase 2 Cycle 3 Summary

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

Subsample of 12k records; Accounted for regions and RAA type

97% of records immediately contactable; 65% operability full accuracy rate on all 2009 RAA requirements

Decrease in overall Op accuracy; Email Syntax Accuracy and Postal Operability accuracy again very high

Compliance will conduct follow up on potentially inaccurate records

Next Cycle has already begun; Report expected June 2017

Questions
&
Answers