# WHOIS Accuracy Reporting System (ARS): Phase 1 Results Webinar | 26 August 2015

ICANN GDD Operations
NORC at the University of Chicago

# Webinar Agenda

**1** — WHOIS ARS Background

**2** — Phase 1: Timeline and Process

**3** — Phase 1: Sample Design and Information
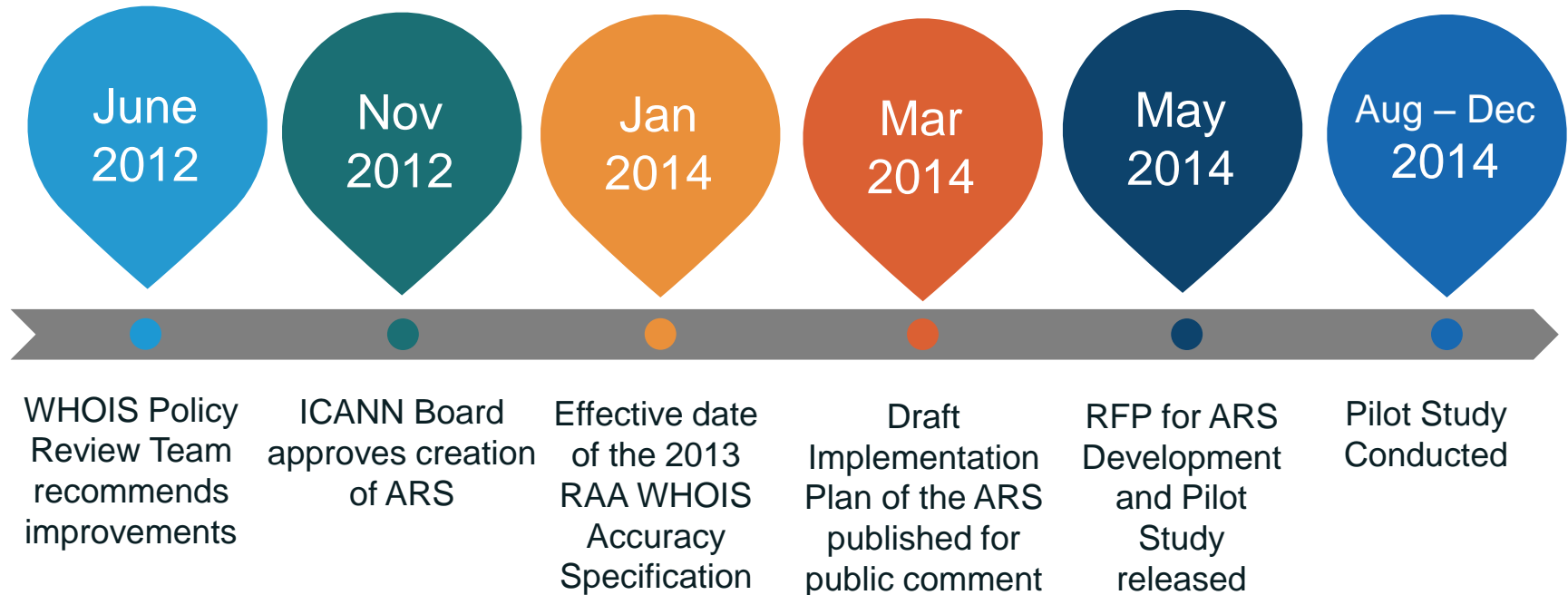
**4** — Phase 1: Results and Major Findings

**5** — Next Steps: Phase 1 and Phase 2

**6** — Question & Answer Session

# **WHOIS ARS Background**

# WHOIS ARS Background

**June 2012**
WHOIS Policy Review Team recommends improvements

**Nov 2012**
ICANN Board approves creation of ARS

**Jan 2014**
Effective date of the 2013 RAA WHOIS Accuracy Specification

**Mar 2014**
Draft Implementation Plan of the ARS published for public comment

**May 2014**
RFP for ARS Development and Pilot Study released

**Aug – Dec 2014**
Pilot Study Conducted

# WHOIS ARS Implementation

**0** — **Pilot (Public Comment Report Completed March 2015)**
"Proof of Concept": Tested processes for data collection and validation

**1** — **Phase 1:** *Syntax Validation*
Is the record correctly formatted?
**Report: 24 August 2015**

**2** — **Phase 2:** *Operational Validation*
Does the email go through, phone ring, mail deliver?
**1st Report: December 2015; 2nd Report: June 2016**

– – – – – – – – – – – – – – – – – – – – – – – – – – – – – – – –

**3** — **Phase 3 TBD, if at all:** *Identity Validation*
Is the contacted individual responsible for the domain?

# Opportunities for Improvement from the Pilot

**Pilot was successful as a proof of concept.** Two major areas of learning:
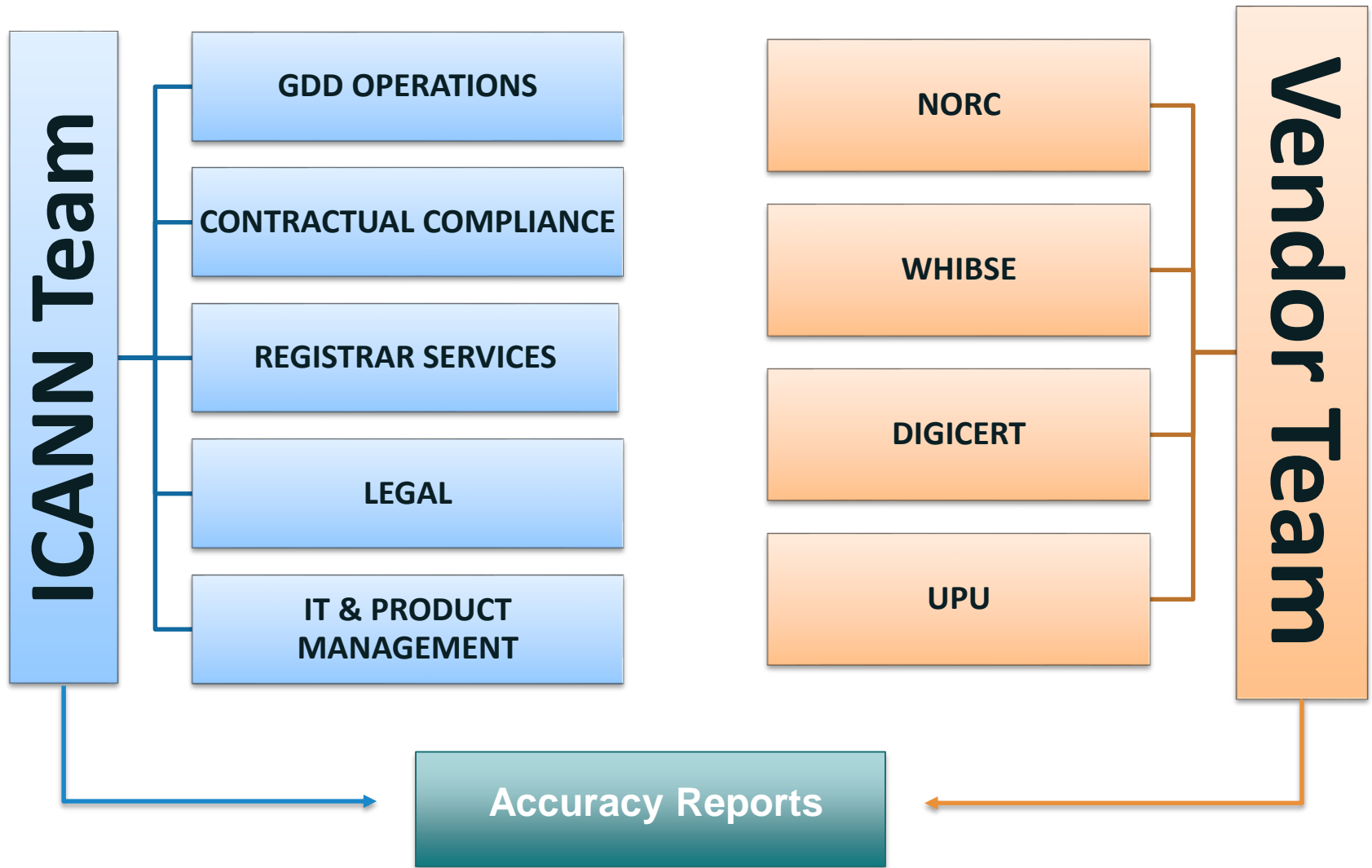
## Criteria

- Align to RAA
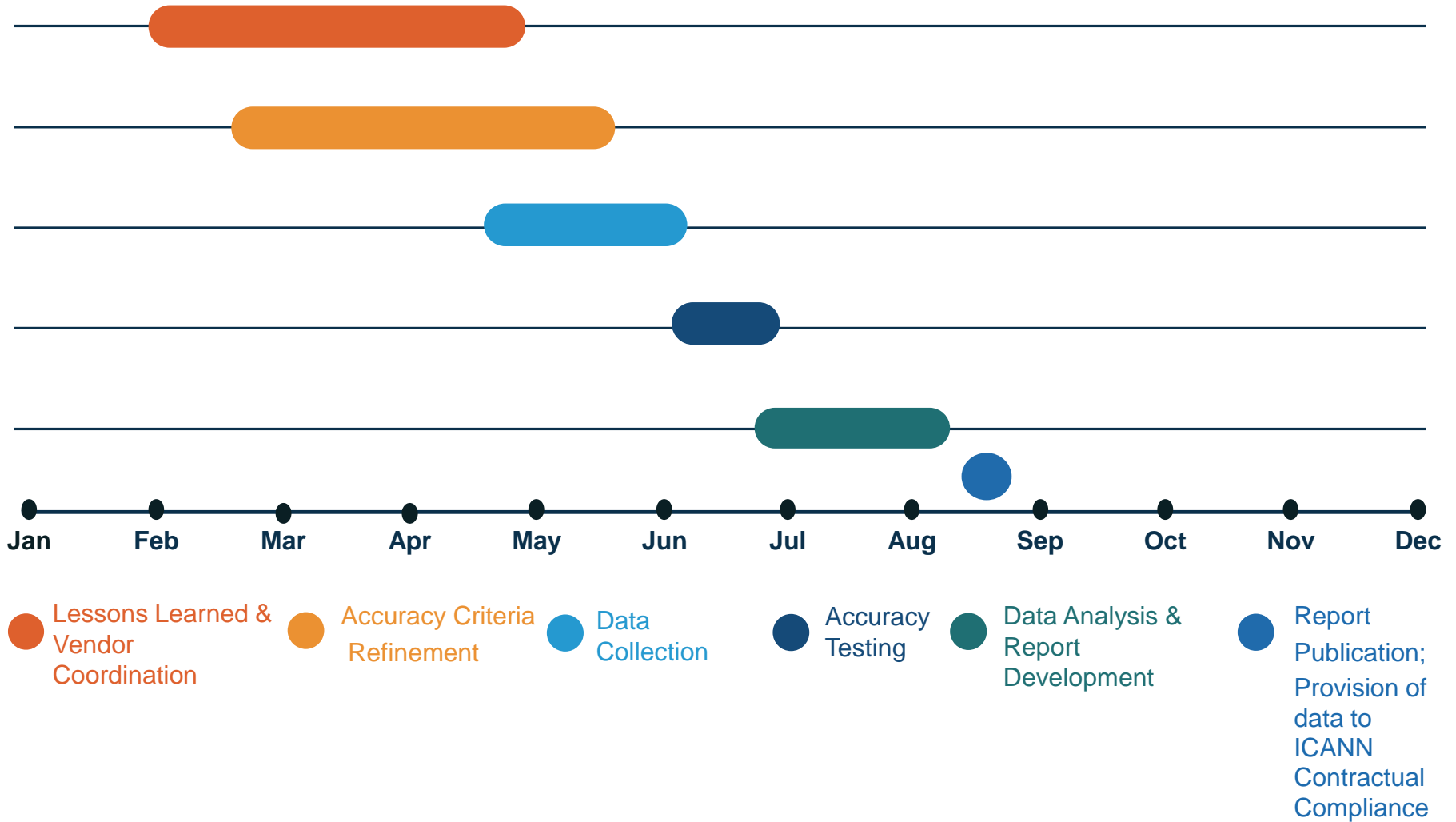- Respect effective date
- Registrant, Tech & Admin contacts

## Reporting

- Statistical significance
- Reasons for inaccuracy
- Actionable data

# Phase 1
Process and Timeline

# Phase 1 Cross-Functional Team

**ICANN Team**

- GDD OPERATIONS
- CONTRACTUAL COMPLIANCE
- REGISTRAR SERVICES
- LEGAL
- IT & PRODUCT MANAGEMENT

**Vendor Team**

- NORC
- WHIBSE
- DIGICERT
- UPU

**Accuracy Reports**

# Phase 1 Timeline



Legend:
- Lessons Learned & Vendor Coordination
- Accuracy Criteria Refinement
- Data Collection
- Accuracy Testing
- Data Analysis & Report Development
- Report Publication; Provision of data to ICANN Contractual Compliance

# Phase 1
## Sample Design and Testing Criteria

# WHOIS ARS Phase 1 Demographics

## gTLD Population At Time of Sample (April 2015)

| Records in gTLDs | Total gTLDs | 2009 RAA* | 2013GF RAA* | 2013 NGF RAA* | New gTLDs | Prior gTLDs |
|---|---|---|---|---|---|---|
| 157m | 610 | 3.5% | 65.7% | 30.8% | 3.15% | 96.85% |

## 100k Sample

| AFR | LAC | EUR | APAC | N.A. | 2009 RAA | 2013GF RAA | 2013 NGF RAA | New gTLDs | Prior gTLDs |
|---|---|---|---|---|---|---|---|---|---|
| .07% | 3% | 23% | 24% | 49% | 3% | 51% | 46% | 25% | 75% |

## 10k Sub-sample

| AFR | LAC | EUR | APAC | N.A. | 2009 RAA | 2013GF RAA | 2013 NGF RAA | New gTLDs | Prior gTLDs |
|---|---|---|---|---|---|---|---|---|---|
| 7% | 18% | 19% | 21% | 35% | 18% | 44% | 38% | 20% | 80% |

* Weighted estimates from 100k sample

# Phase 1 – 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
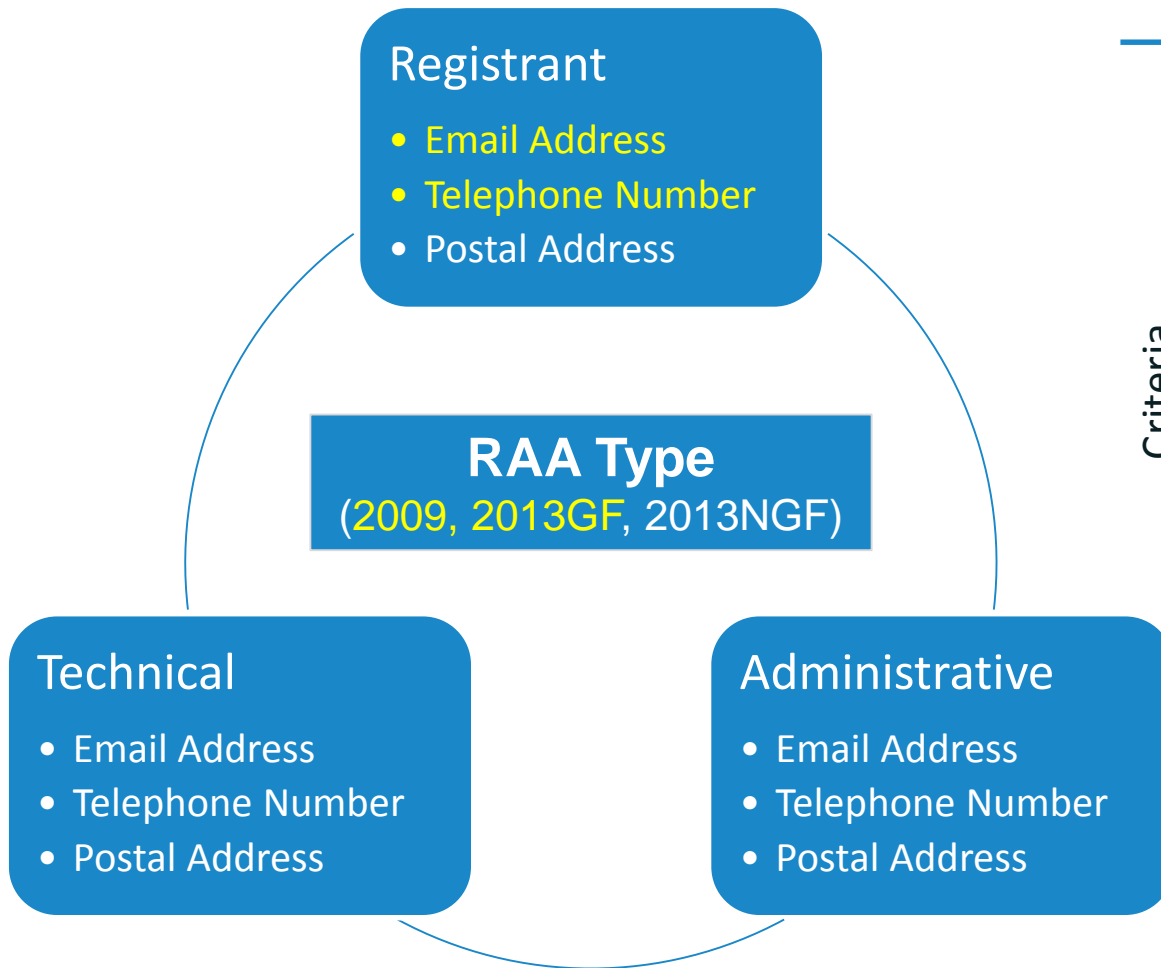
⊙ Criteria listed in Appendix A

**Criteria Examples**

Does the email address contain an "@"?

Does the telephone number have a country code?

Does the postal address include an identifiable country?

# Phase 1 – Report Highlights

**Accuracy Statistics by Subgroup**

⊙ Phase 1 Report provides 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 describes <u>why</u> records are inaccurate**

⊙ All domains evaluated against 2009 RAA requirements

⊙ Appendix C provides results on 2013NGF domains

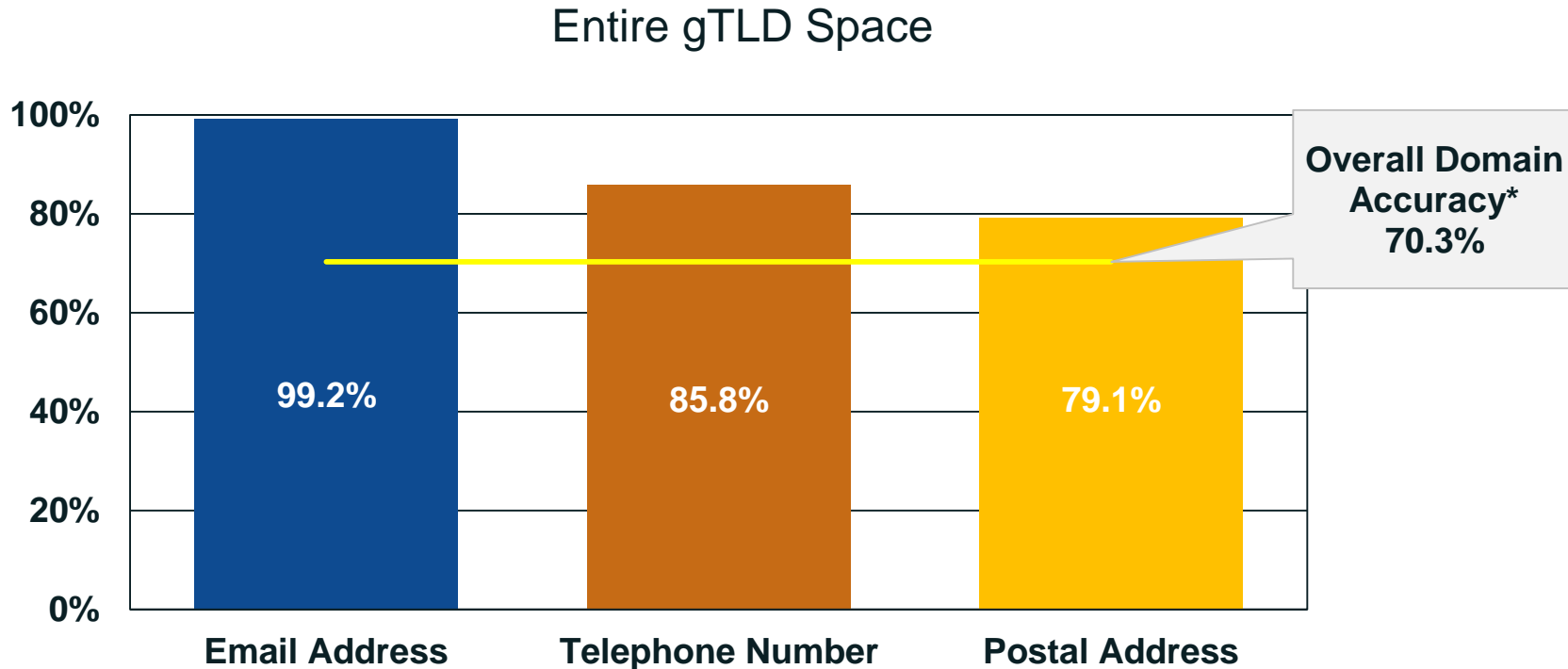⊙ Detailed tests enable us to know why a record is inaccurate

# **Phase 1**
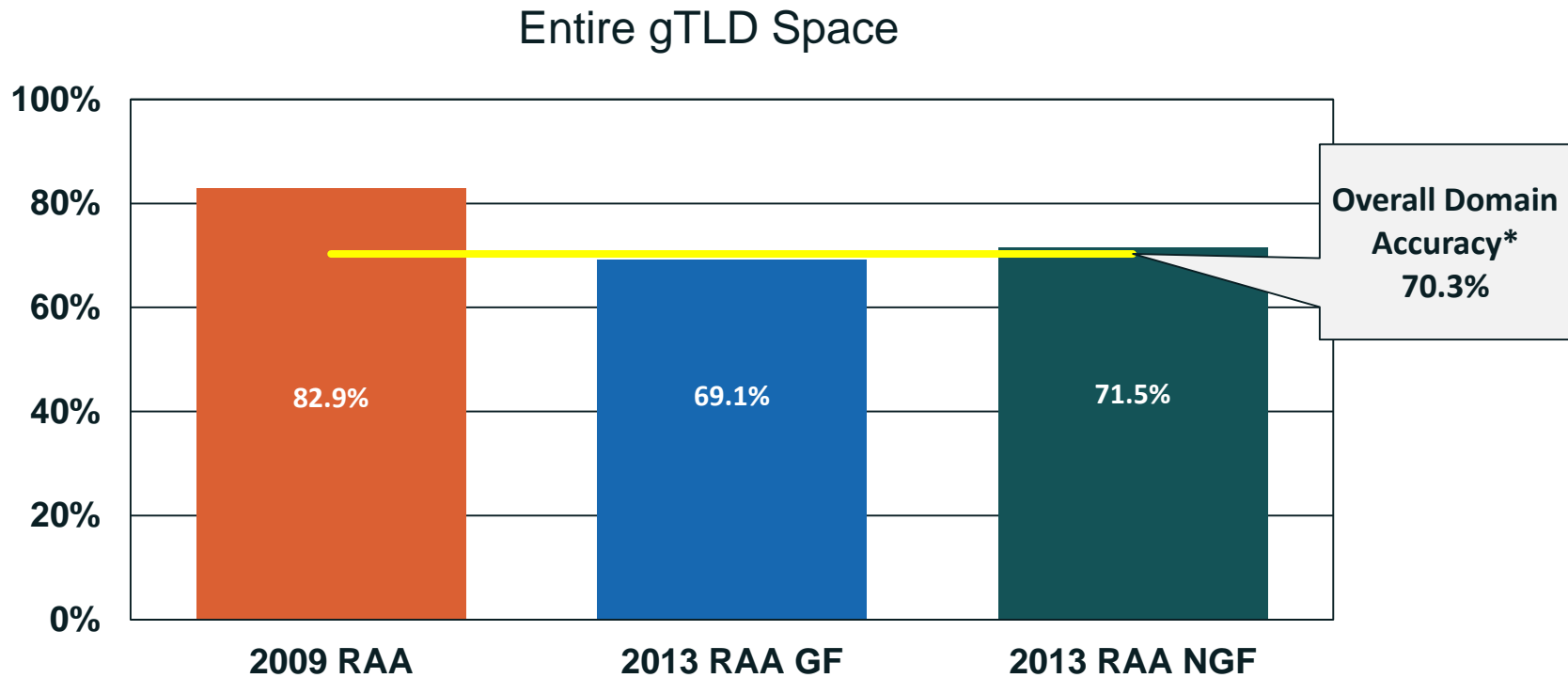# Results and Major Findings

ICANN

# Phase 1 – Summary of Findings

- Accuracy very similar for Registrant, Administrative, and Technical

- Email accuracy very high; postal address lowest (most complex)
  - If an email address is provided, it passed all the tests
  - Two-thirds of telephone number errors due to wrong # of digits
  - Postal address errors: typically missing 1+ required field

- 70.3% of all domains meet all 2009 RAA syntax requirements
  - For all three contact types (Registrant, Administrative, Technical)
  - For all three contact modes (email, telephone, postal address)

- Following findings are estimates applicable to the entire gTLD space or respective subgroup

# Phase 1 - Conformance to 2009 RAA*, by Contact Mode

## Entire gTLD Space



Bar chart showing: Email Address 99.2%, Telephone Number 85.8%, Postal Address 79.1%. Overall Domain Accuracy* 70.3%
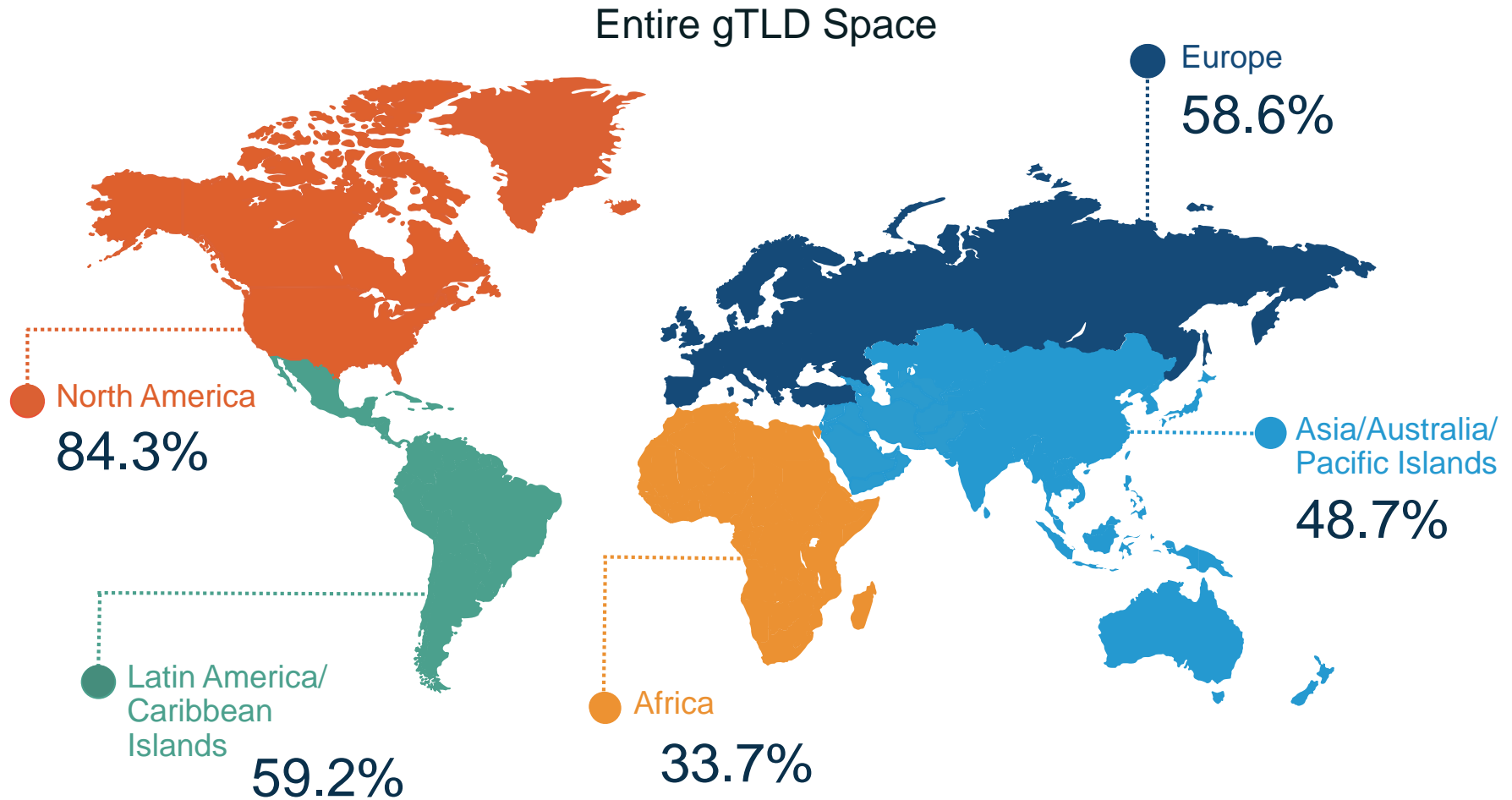
- Contact mode accuracy requires accuracy on all 3 contact types
  - Registrant, Administrative, Technical Contacts
- Overall accuracy requires accuracy on all 3 contact modes and all 3 contact types
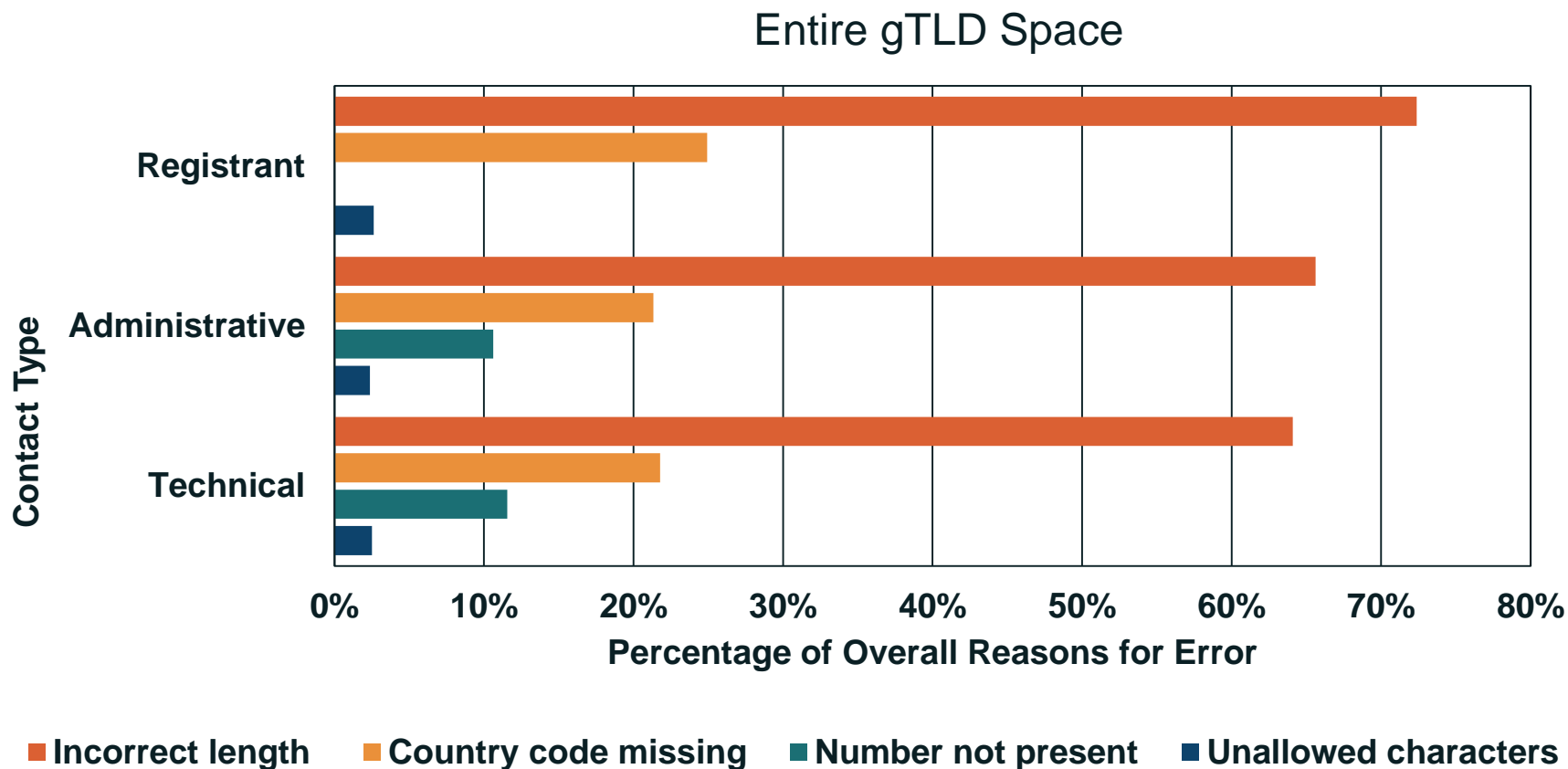
# Phase 1 - Conformance to 2009 RAA*, by RAA Type

## Entire gTLD Space



Bar chart showing:
- 2009 RAA: 82.9%
- 2013 RAA GF: 69.1%
- 2013 RAA NGF: 71.5%
- Overall Domain Accuracy* 70.3%

- Accuracy here also requires accuracy on all 3 contact modes and all 3 contact types

# Phase 1 – Conformance to 2009 RAA by Region

Entire gTLD Space



**Europe**
58.6%

**North America**
84.3%

**Asia/Australia/Pacific Islands**
48.7%

**Latin America/Caribbean Islands**
59.2%

**Africa**
33.7%

- Accuracy here also requires accuracy on all 3 contact modes and all 3 contact types

Entire gTLD Space

Note: "Number not present" error is omitted from the Registrant contact type because it is not a requirement for the Registrant contact type. The "Extension Character" error type has been omitted from this graph because it represents less than 0.1% of errors in each contact type. As a result, the sum of errors shown is slightly less than 100% for each contact type.

## Entire gTLD Space
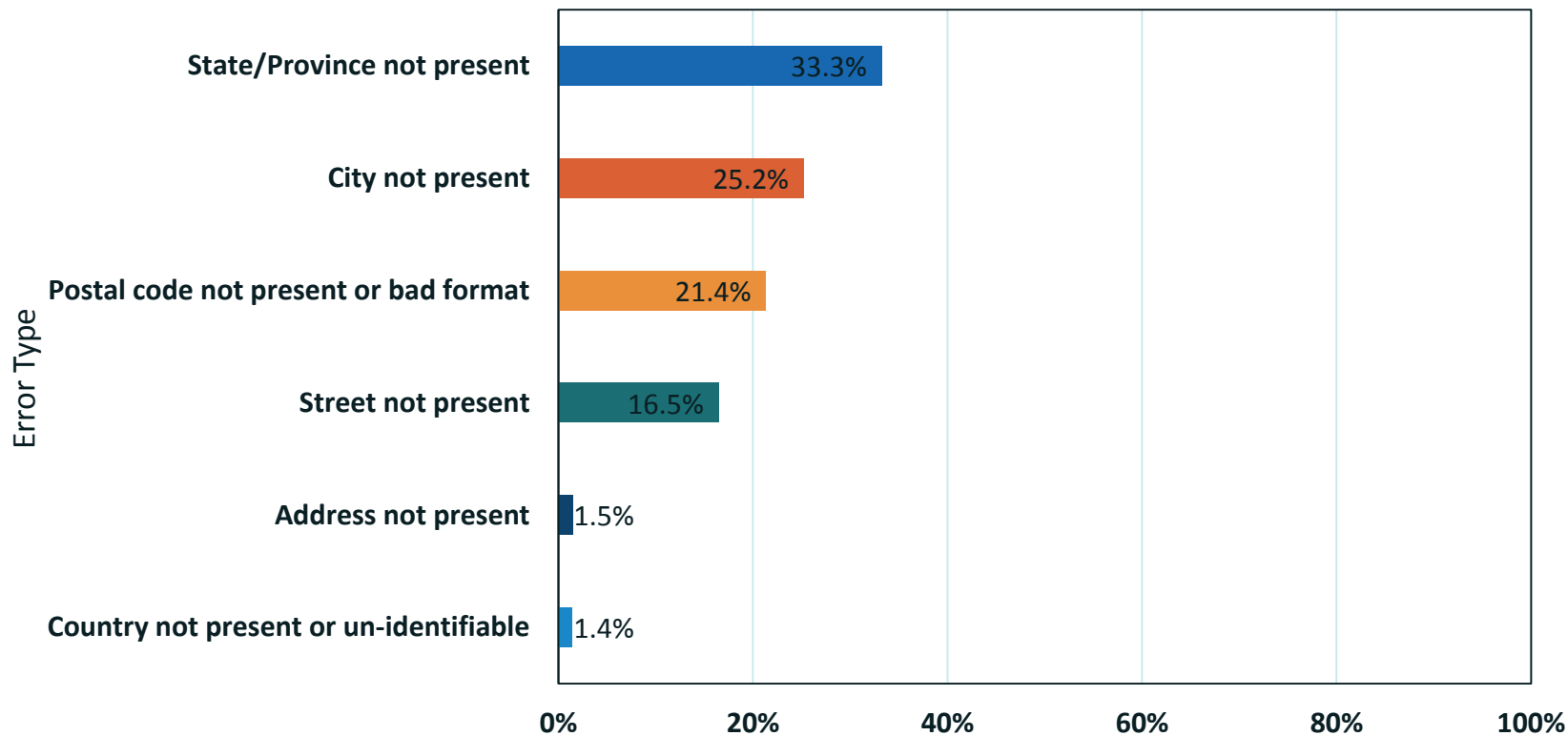


| Error Type | Percent of Overall errors for contact type |
|---|---|
| State/Province not present | 33.3% |
| City not present | 25.2% |
| Postal code not present or bad format | 21.4% |
| Street not present | 16.5% |
| Address not present | 1.5% |
| Country not present or un-identifiable | 1.4% |

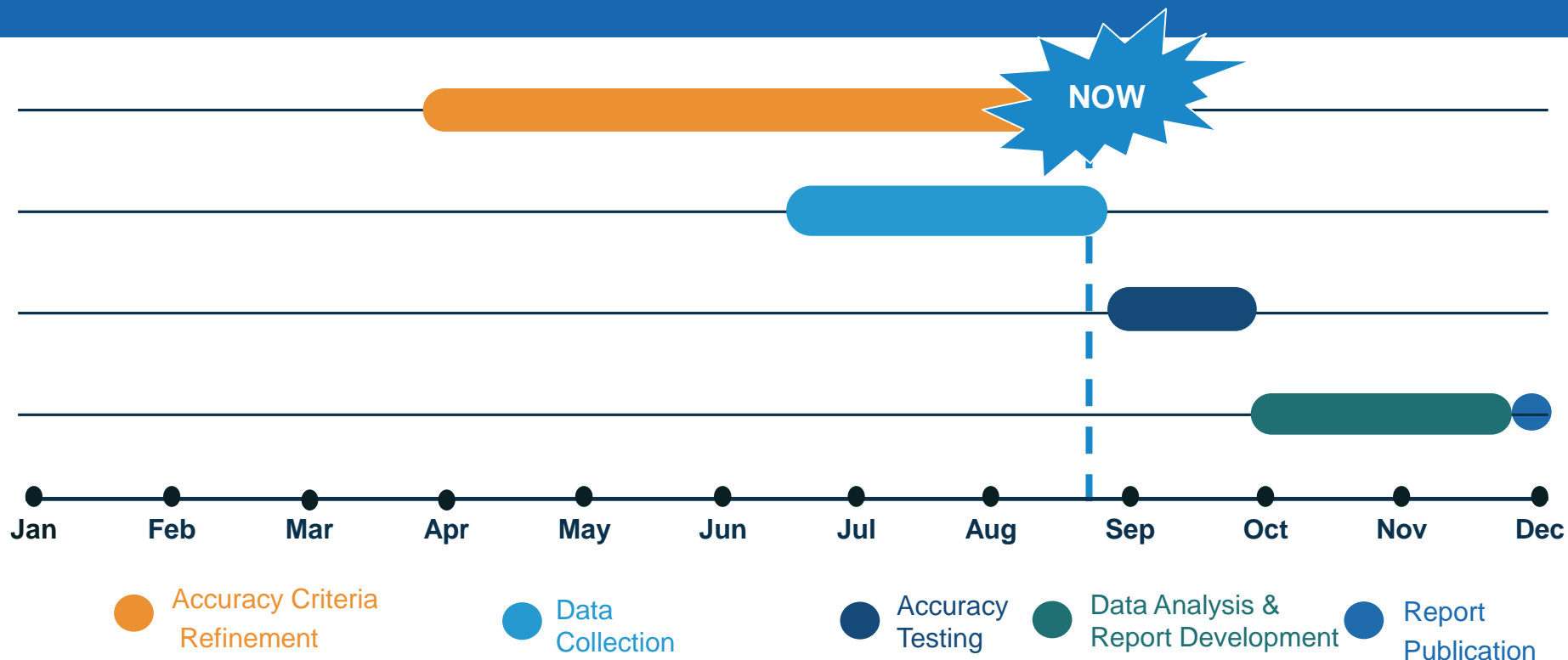Among 9,789 Registrant postal address errors found in the 10,000 record subsample

Next Steps: ICANN Contractual Compliance Follow-Up

# Next Steps – ICANN Contractual Compliance

- ⊙ Potentially inaccurate records have been provided to ICANN Contractual Compliance
    - ⊙ Due to sample weighting and 2013 requirements, total number of records identified as nonconforming exceeds 30% of 10k subsample

- ⊙ 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

- ⊙ Registrars under 2013 RAA must use WHOIS format and layout required by Registration Data Directory Service Specification

- ⊙ 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

# Next Steps: Phase 2

# Phase 2, Cycle 1 – Implementation Underway

**NOW**

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

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

## Phase 2 = Syntax + Operability

*Examples of operability tests:*

- ⊙ *Email*: Does the email go through / not bounced back?
- ⊙ *Phone*: Does the number ring when dialed?
- ⊙ *Postal*: Is the address deliverable?

# Phase 1 Summary

**Phase 1 Syntax Accuracy Report published 24 August 2015**

**Subsample of 10k records; Accounted for regions and RAA type**

**70% accuracy rate on all 2009 RAA requirements**

**Compliance will conduct follow up on potentially inaccurate records**

**Community Feedback in 4th Quarter 2015**

**Phase 2 underway; 1st Report in December 2015**

# Questions
# &
# Answers