



WHOIS 准确度报告体系 (ARS)

第 2 阶段第 2 周期报告：语法与可操作性准确度

全球域名分部 | 2016 年 6 月 8 日

目录

执行摘要.....	4
本报告的主题.....	4
准确度测试方法.....	4
样本设计.....	5
研究结果.....	6
后续步骤.....	8
简介.....	10
本报告的主题.....	10
第 2 阶段第 1 周期概述.....	11
第 2 阶段第 2 周期概述.....	12
研究方案和方法.....	15
简介.....	15
样本设计.....	16
语法与可操作性测试方法.....	21
主要研究结果.....	22
研究结果摘要.....	22
语法准确度 — 2009 RAA 要求.....	23
可操作性准确度 — 2009 RAA 要求.....	30
地区研究结果 — 按地区分析准确度和错误原因.....	39
不同周期之间的对比.....	48
语法和可操作性准确度之间的关系.....	56
挑战和经验教训.....	60
后续步骤.....	62
展望第 2 阶段第 3 周期.....	62
ICANN 合同合规部的后续步骤.....	62
附录 A: 准确度测试标准.....	65
附录 B: 其他分析 — 满足 2009 RAA 要求的准确度.....	66

附录 C：其他分析 — 满足 2013 RAA 要求的准确度.....	73
阶段之间的对比 — 2013 RAA 语法要求.....	77
附录 D：其他分析 — 按地区划分的文字和准确度.....	86

执行摘要

本报告的主题

WHOIS 准确度报告体系 (ARS) 旨在满足 2012 WHOIS 审核小组（依据《义务确认书》[AOC] 组建）的建议。¹根据 2012 年 11 月 8 日的这些建议，ICANN 董事会通过了一系列旨在改进 ICANN 监督 WHOIS 计划方式的措施。这些改进措施包括创建 WHOIS ARS 以解决政府咨询委员会 (GAC) 对于 WHOIS 准确度的担忧。

这份报告是 WHOIS ARS 所生成系列报告中的第四份。WHOIS ARS 设计为分阶段进行，以便 ICANN 社群能够影响其制定过程。试点阶段已于 2015 年 4 月完成，第 1 阶段于 2015 年 8 月完成。第 2 阶段正在进行并且分为不同的周期，第 2 阶段第 1 周期（以下简称“第 1 周期”）于 2015 年 12 月完成，而第 2 阶段第 2 周期（即本报告的主题，以下简称“第 2 周期”）于 2016 年 1 月开始。第 1 阶段只检查语法准确度，第 2 阶段报告同时检查 WHOIS 记录的语法和可操作性准确度。与之前的报告一样，本报告将再次详述主要不合规类型、各地区 WHOIS 准确度的趋势和对比、注册服务机构认证协议 (RAA) 版本和通用顶级域 (gTLD) 类型。

关于 WHOIS ARS 背景的所有详细信息以及结果，都可以在之前的 ARS 报告找到：[试点报告](#)、[第 1 阶段报告](#)、[第 2 阶段第 1 周期报告](#)。

准确度测试方法²

语法和可操作性准确度测试旨在通过将 WHOIS 记录的联系信息与 RAA 适用合同要求对比，从而进行评估。³语法测试评估记录的格式（例如，电子邮件地址是否包含“@”符号？）。可操作性测试评估记录中信息的功能（例如，电子邮件是否发回给发送人？）。记录的全部 9 个联系信息字段都会进行语法和可操作性准确度测试（即注册人联系人、管理联系人和技术联系人的电子邮件地址、电话号码和邮政地址），并将作为完整记录进行编制。所得数据经过分析后，将用于生成各小

¹ 参阅 <https://www.icann.org/resources/pages/aoc-2012-02-25-en>。

² 关于语法和可操作性准确度测试/标准的一般信息可以在[附录 A](#) 中找到。有关更多详细信息，请访问 WHOIS ARS 网页：<https://whois.icann.org/en/whoisars-validation>。

³ 准确度测试/标准就是我们定义的基本要求，满足这些要求的联系数据才会被视为格式正确且可操作。尽管 2009 RAA 没有包含明确的语法要求，但希望提供的联系数据完整且格式正确。

组 WHOIS 联系信息语法和可操作性准确度的统计数据，例如新 gTLD 和早期 gTLD、地区和 RAA 类型（即 2009 RAA 或 2013 RAA⁴）。

有关本研究方案和所进行的准确度测试的详细信息，请参阅[研究方案和方法](#)以及[附录 A: 准确度测试标准](#)。

样本设计

2016 年第一季度初进行初次抽样时，610 个 gTLD⁵ 中分布着近 1.695 亿个域名⁶。在这 1.695 亿个域名中，在 18 个早期 gTLD 中的任意一个进行注册的有近 94%，在 592 个新 gTLD 中的任意一个进行注册的有约 6%。两阶段抽样方法旨在提供足够大的样本，以便对各相关小组进行可靠估计，例如 ICANN 地区、新 gTLD 或早期 gTLD，以及 RAA 类型。初始样本包含 200,000 条记录，而经过分析的子样本则包含 12,000 条代表当时全部活跃 gTLD 的记录。⁷

尽管约 97% 的域名通过根据 2013 RAA 获得认证的注册服务机构注册，但大部分域名可以根据 2009 RAA 的 WHOIS 标准进行运作。⁸因此，本报告使用 2009 RAA 标准作为基准来评估 WHOIS 准确度，但是所有 2013 RAA 非过渡 (NGF) 域名也已根据 [2013 RAA 标准](#)进行测试，测试结果请参阅[附录 C](#)。以下表 Ex1 所示为上述初始样本的细分数据。有关更详细的信息，包括为什么表 Ex1 只有 196,262 个域名，以及如何决定样本尺寸，请参阅[样本设计](#)。

表 Ex1: 初始样本大小（按地区和 RAA 划分）

⁴ 请参阅此处，查看 RAA 的各个版本：<https://www.icann.org/resources/pages/registrars/registrars-en>。

⁵ 抽样时，有 888 个已授权的 gTLD（18 个早期 gTLD 和 870 个新 gTLD），888 个 gTLD 中有 260 个拥有零个域名，以及 40 个刚好拥有一个域名。这 300 个 gTLD 未进行抽样。

⁶ 基于 gTLD 域文件中的信息。

⁷ 552 个新 gTLD 和 18 个早期 gTLD 拥有至少两个域名。另外请注意从第 1 周期开始增加的样品数：150,000 至 200,000 和 10,000 至 12,000。

⁸ 这可能是出于以下两个原因之一：1) 注册服务机构尚未与 ICANN 签署 2013 RAA 并且仅受到 2009 RAA 标准的管辖；或者 2) 注册服务机构与 ICANN 就 2013 RAA 达成了一致，但域名的注册日期在注册服务机构的 2013 RAA 生效日期之前。我们将后一组域名称为 2013 RAA 过渡 (2013 RAA GF) 域名。因此，我们的分析包括三个互不包含的 RAA 小组：2009 RAA、2013 RAA GF 和 2013 RAA 非过渡域名（以下称为 2013 RAA NGF）。

RAA 类型	拉丁美洲						总计
	非洲地区	和加勒比海地区	欧洲地区	亚太地区	北美地区	未知	
2009	30	306	619	450	3,258	19	4,682
2013 GF	457	2,184	18,275	14,379	46,564	438	82,297
2013 NGF	769	6,157	17,571	47,991	36,062	733	109,283
总计	1,256	8,647	36,465	62,820	85,884	1,190	196,262

研究结果

经过分析的子样本中的全部 12,000 条记录都使用 2009 RAA 标准进行评估，该标准可作为基准来评估 gTLD 中 WHOIS 记录的整体准确度。第 2 阶段的重点为在符合 RAA 要求（2009 RAA 或 2013 RAA）方面按联系模式（电子邮件地址、电话号码和邮政地址）划分的语法和可操作性准确率。随后，使用来自经过分析的子样本测试的结果估计整个 gTLD 社群或特定相关小组的结果。这些数据以 95% 的置信区间⁹在本报告中呈现，估计百分比的标准误差约为 $\pm 2\%$ 。根据抽样误差，真参数处于置信区间内的概率为 95%。

建立联系的能力

对于 99% 的记录而言，三个联系人类型中至少有一个联系模式符合 2009 RAA 的所有语法和可操作性要求，这意味着几乎所有记录都包含可用于建立联系的信息。只有 1% 的记录没有符合语法或可操作性要求的任何联系人类型的联系模式。

语法准确度

语法准确度分析发现，约 99% 的电子邮件地址、85% 的电话号码以及 77% 的邮政地址均符合 2009 RAA 中对这三种联系人规定的所有基本语法要求¹⁰。符合 2009 RAA 中规定的语法整体准确度要求的完整 WHOIS 记录（全部三种联系模式的全部三种联系人类型）约占 gTLD 社群的 67%。表 Ex2 按联系模式显示准确度细分，并以 95% 的置信区间呈现。

表 Ex2: gTLD 对 2009 RAA 语法要求的整体¹¹准确度（按模式划分）

	电子邮件地址	电话号码	邮政地址	3 项均准确
3 种联系人类型均准确	99.2% \pm 0.2%	85.3% \pm 0.6%	77.3% \pm 0.7%	67.2% \pm 0.8%

⁹ 这意味着，如果对该社群进行再次抽样，置信区间把小组或参数（例如，按地区划分的准确度）归为一类的概率约为 95%。

有关置信区间的详细信息，请参阅：<http://www.itl.nist.gov/div898/handbook/prc/section1/prc14.htm>。

¹⁰ 请参阅注释 3。

¹¹ 整体准确度指的是全部的 1.695 亿个域名的准确度。有关置信区间和社群的信息，请参阅注释 12。

可操作性准确度

可操作性准确度分析发现，对全部 3 种联系人类型而言，约 91% 的电子邮件地址、76% 的电话号码以及 98% 的邮政地址均可操作。符合可操作性整体准确度要求的完整 WHOIS 记录约占 gTLD 社群的 70%。以下表 Ex3 按联系模式显示准确度细分，并以 95% 的置信区间呈现。

表 Ex3: gTLD 对 2009 RAA 可操作性要求的整体准确度（按模式划分）

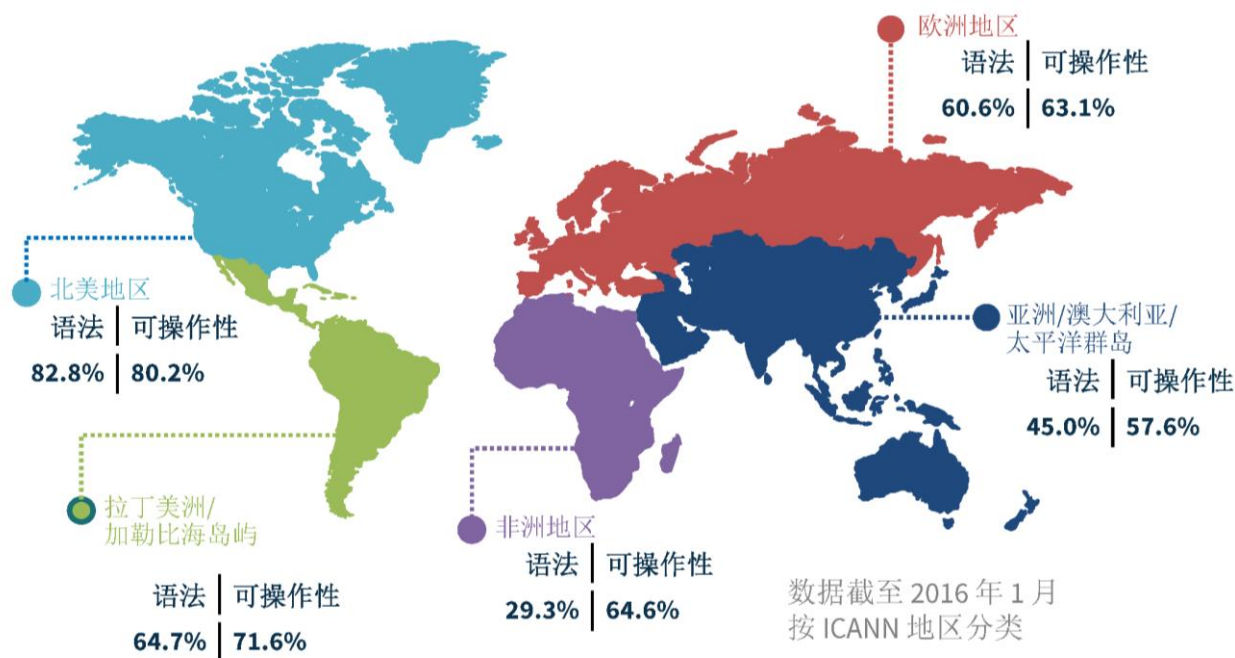
	电子邮件地址	电话号码	邮政地址	3 项均准确
3 种联系人类型均准确	91.4% ± 0.5%	76.0% ± 0.8%	97.7% ± 0.3%	70.2% ± 0.8%

各个小组语法和可操作性不合规的主要原因已进行调查并在[主要研究结果](#)以及[附录 B](#)和[附录 C](#)中予以说明。

区域准确度

在第 2 周期，我们增加了对准确度的区域差异和错误原因的分析。图 Ex1 中的地图基于 ICANN 域名地区显示 WHOIS 记录的整体语法和可操作性准确度，左侧是语法准确度数字，右侧是可操作性。在[其他研究结果](#)中，标题[地区分析](#)下，可以找到其他准确度的地区指标和错误原因。

图 Ex1: 第 2 周期按 ICANN 地区划分的整体语法和可操作性准确度



注：对于每个地区，语法和可操作性准确度数字按以下格式显示：语法 | 可操作性。显示的准确率是所有三种联系人类型、全部三种联系模式的信息均准确的记录的百分比。

报告的正文和附录包括与 [2013 RAA](#) 下准确率相关的其他子分析、[从第 1 周期到第 2 周期的趋势](#)、[新 gTLD](#) 和 [早期 gTLD](#) 之间的差异，以及对用于注册域名的 [文字](#) 进行的分析。

后续步骤

第 2 阶段第 3 周期

WHOIS ARS 旨在成为进行可重复评估的系统；第 2 阶段第 3 周期将重复第 2 阶段第 1 和第 2 周期的语法和可操作性审核，审核将于 2016 年 7 月开始，预计将于 2016 年 12 月初发布报告。

ICANN 合同合规部

本报告发布前，已将第 2 周期的研究结果（即所有可能不准确的记录）提交 ICANN 合同合规部，供其审核和处理。内部审核完成后，ICANN 合同合规部将对发现的错误类型以及需要注册服务机构跟进的类型进行评估。由于第 2 周期包含语法和可操作性研究结果，因此可能通过不同的流程进行合规性跟进和调查，这取决于各记录中发现的不准确问题类型，例如，被判定为“可操作”但格式错误的记录收到的通知类型将不同于被判定为“不可操作”且格式错误的记录。所有的 WHOIS

ARS 投诉单都将根据本报告所述问题类型按照合同合规方法和流程¹²进行处理。如果可能，在咨询注册服务机构后，ICANN 也许能够在处理时合并多个 WHOIS ARS 投诉单。WHOIS ARS 投诉单将与其他投诉一起处理，但 ICANN 将继续优先处理由社群成员提交的投诉。

合同合规部继续在合规季报中提供 WHOIS ARS 的指标（请参阅 <https://www.icann.org/resources/pages/compliance-reports-2016-04-15-en>），并且将在为 2016 年第二季度生成指标时提供其他信息。此外，适当的情况下，指标将在公共 ICANN 会议上提供。

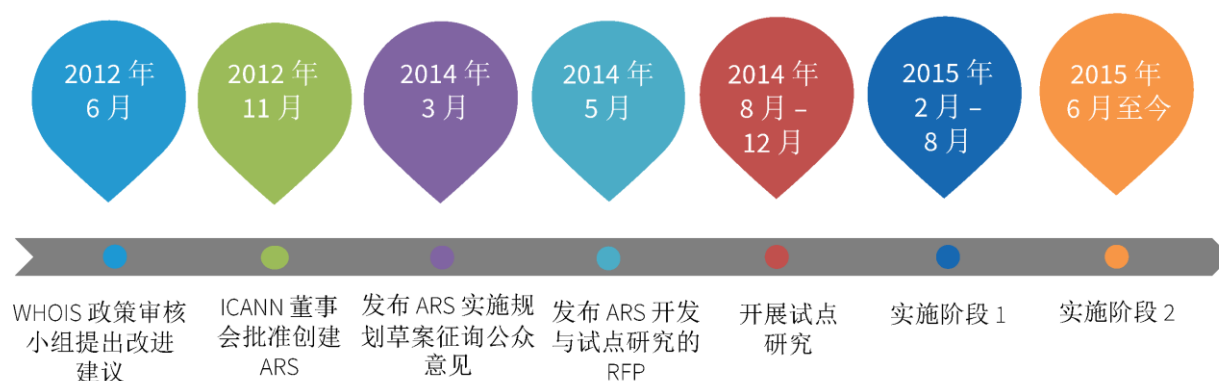
¹² 参阅 ICANN 合同合规方法和流程：<https://www.icann.org/resources/pages/approach-processes-2012-02-25-en>。

简介

本报告的主题

WHOIS 准确度报告体系 (ARS) 旨在满足 2012 WHOIS 审核小组依据《义务确认书》编制和提出的建议。¹³根据 2012 年 11 月 8 日的这些建议，ICANN 董事会通过了一系列旨在改进 ICANN 监督 WHOIS 计划方式的措施。这些改进措施包括创建 WHOIS ARS 以解决政府咨询委员会 (GAC) 对于 WHOIS 准确度的担忧。图 1 围绕 WHOIS ARS 的创建和进度显示事件时间表。

图 1: ARS 背景信息

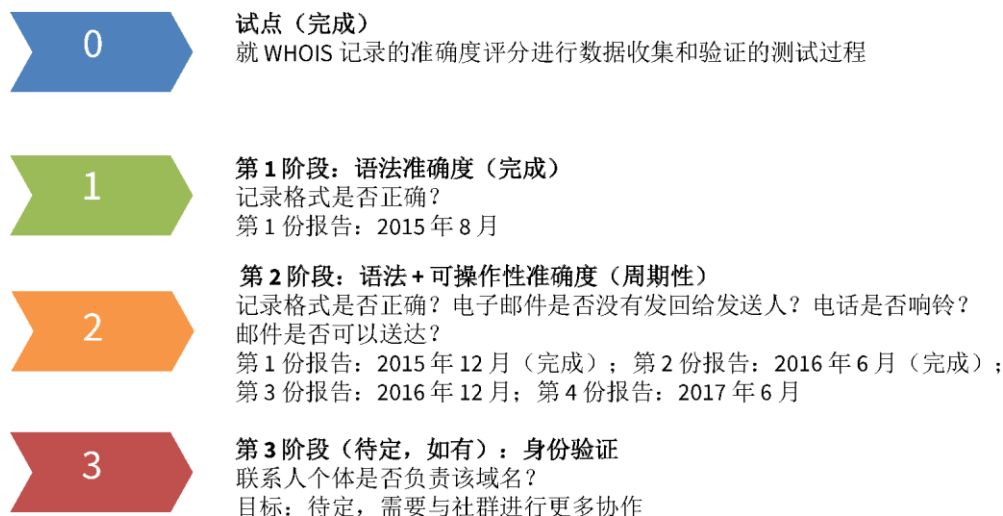


基于 SAC058 报告中所述的验证类型，ARS 设计为三个阶段¹⁴。第 1 阶段分析 WHOIS 联系信息的语法准确度。第 2 阶段正在进行中且具有周期性，将评估记录中联系数据的可操作性，方法是将第 1 阶段的语法测试和可操作性测试相结合。第 3 阶段计划进行身份验证，但实施第 3 阶段的时间表尚未确定。ICANN 仍将与社群继续评估是否会实施第 3 阶段，以及如何实施。图 2 展示了这个按阶段划分的方法。

¹³ 参阅 <https://www.icann.org/resources/pages/aoc-2012-02-25-en>。

¹⁴ 参阅 <https://www.icann.org/en/system/files/files/sac-058-en.pdf>。

图 2: WHOIS ARS 阶段



这份报告是 WHOIS ARS 所生成系列报告中的第四份。试点阶段已于 2015 年 4 月完成，第 1 阶段于 2015 年 8 月完成。第 2 阶段正在进行 — 第 2 阶段第 1 周期（“第 1 周期”）于 2015 年 12 月完成。第 2 阶段第 2 周期（“第 2 周期”），即此报告的主题，于 2016 年 1 月开始。关于 WHOIS ARS 背景的所有详细信息以及结果，都可以在之前的 ARS 报告中找到：[试点报告](#)、[第 1 阶段报告](#)、[第 2 阶段第 1 周期报告](#)。

第 2 阶段第 1 周期概述

WHOIS ARS [第 2 阶段第 1 周期](#)（“第 1 周期”）¹⁵报告发布于 2015 年 12 月，此周期针对 2015 年 4 月至 8 月实施的第 1 阶段研究开展了后续研究。第 1 周期的主要成果包括：

- 就语法准确度而言，[第 1 阶段](#)电话号码的准确度有所下降。电话号码准确度下降的原因似乎是第 1 周期抽取的电话号码样本中国家/地区代码的缺失有所增加。

¹⁵ 在这里，第 2 阶段第 1 周期和第 2 阶段第 2 周期将分别简称为第 1 周期和第 2 周期。从整体上讨论第 1 阶段或第 2 阶段时会明确说明，以避免对周期产生混淆。需要指出的是，由于与荷兰省份相关的一些语法测试结果获得更正，因此一些第 1 周期的计算得到了改进。本文档中使用了经更新的第 1 周期计算，尽管第 1 周期报告尚未更新。

- 87% 的电子邮件地址、74% 的电话号码以及 98% 的邮政地址均符合 2009 RAA 中规定的所有可操作性要求。65% 的域名通过了针对全部联系人类型（注册人联系人、管理联系人、技术联系人）和联系模式（电子邮件地址、电话号码、邮政地址）的所有可操作性测试。
- 与语法准确度测试不同的是，所有可操作性测试通过率最高的联系模式是邮政地址。所有可操作性测试通过率最低的联系模式是电话号码。
- 超过 75% 的域名的注册人联系人、管理联系人和技术联系人的全部三种联系模式的联系信息相同，因此三种联系人类型的准确率类似。

第 2 阶段第 2 周期概述

第 2 阶段第 2 周期目标

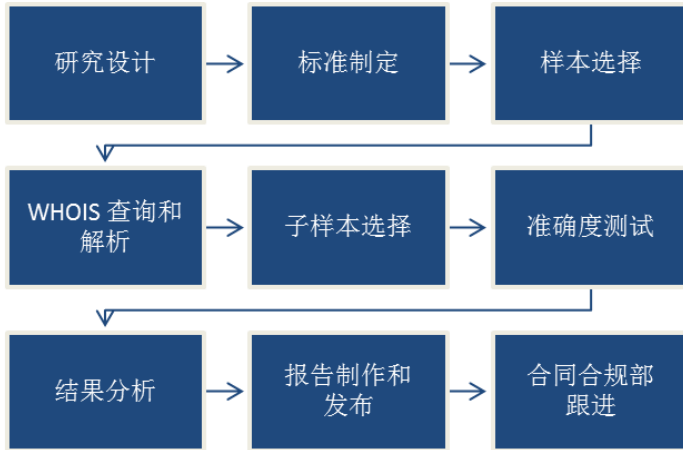
第 2 阶段第 2 周期（“第 2 周期”）的目标与第 1 周期相同：检查 WHOIS 记录的语法和可操作性准确度。ICANN 力求确定 WHOIS 记录是否符合适用 RAA 的格式和内容要求，以及是否可通过所提供的联系数据联系到联系人。第 2 周期报告详细介绍了主要不合规类型、各地区 WHOIS 准确度的趋势和对比、RAA 以及 gTLD 类型。与第 1 周期的一个区别是，第 2 周期报告提供了关于语法和可操作性准确度的区域差异的更多细节。

ICANN 合同合规部可以使用基本数据跟进注册服务机构对可能不准确或不可操作的记录的处理，如有需要，还将进行调查和更正。尽管本报告提供了 ARS 研究之间的对比，但 WHOIS 数据准确度的任何改进都不能直接关联到 ARS。由于数据拉取和 ICANN 将汇总数据提交给 ICANN 社群的时间安排问题，以及合同合规部已经开始跟进注册服务机构，因此 ARS 的潜在影响将被延后。

项目计划、任务和时间表

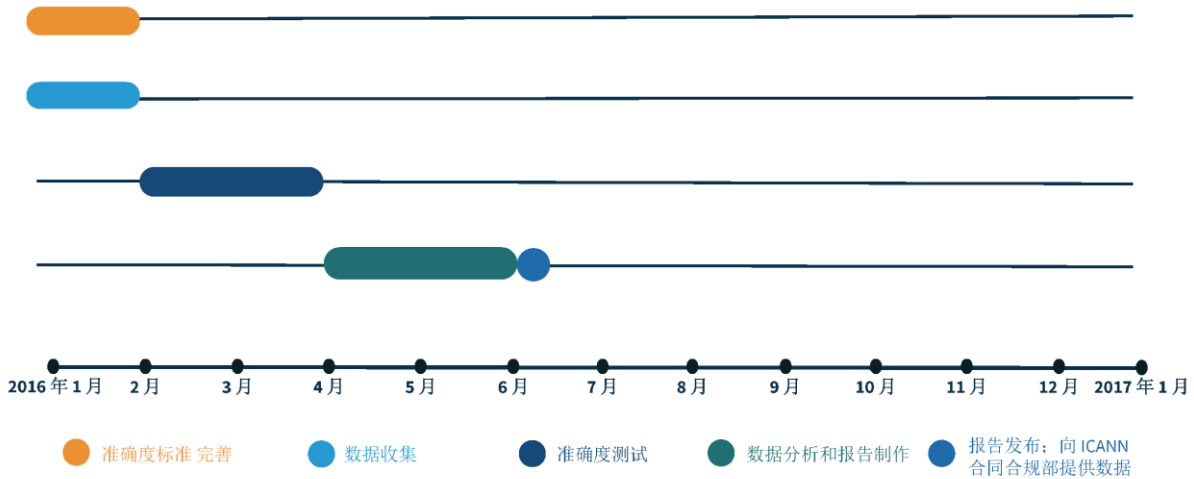
第 2 周期的实施方式与第 1 周期相同。图 3 到图 5 分别图解了流程、时间表和 WHOIS ARS 团队。

图 3：工作流程和任务



小组根据图 4 中所示的时间表执行这些任务¹⁶。

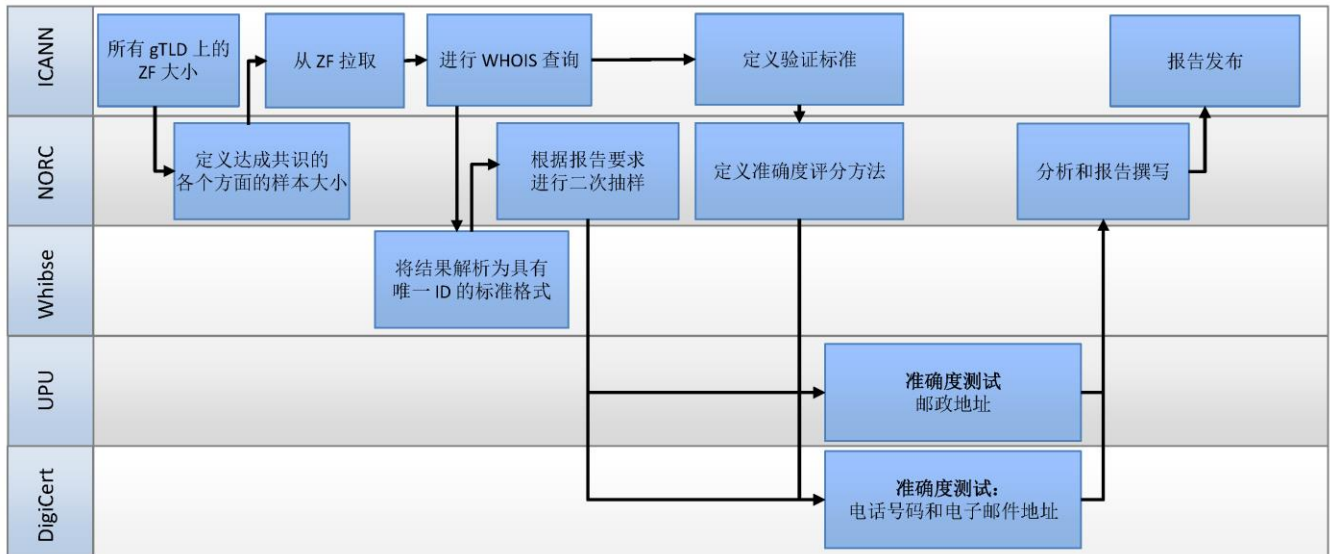
图 4：第 2 周期时间表



WHOIS ARS 团队与之前的阶段相比没有发生变化。图 5 展示了小组成员如何协同制定第 2 周期报告。

¹⁶ 合同合规部跟进除外，该任务将在本报告发布后开始。

图 5: ICANN 供应商协调



研究方案和方法

简介

在第 2 周期中，我们首先从 588 个 gTLD 的域文件中选择约 200,000 条 WHOIS 记录作为样本。根据域名适用的注册服务机构认证协议 (RAA) 中规定的要求，使用系统性测试对 12,000 条记录构成的子样本中的联系信息进行语法标准准确度（例如，值和格式）测试和可操作性标准准确度（即信息可用于建立联系）测试。所得数据经过分析后，将用于生成各小组 WHOIS 联系信息语法和可操作性准确度的统计数据，例如 gTLD 类型（早期或新 gTLD）、ICANN 地区和 RAA 类型。尽管约 98% 的域名通过根据 2013 RAA 运作的注册服务机构进行注册，根据域名的注册时间，大部分在根据 2013 RAA 运作的注册服务机构下注册的域名仍然只需要满足 2009 RAA 中规定的 WHOIS 要求；我们将此类域名称为 2013 RAA 过渡 (2013 RAA GF) 域名。在根据 2013 RAA 运作的注册服务机构下注册并需要满足 2013 RAA 中规定的 WHOIS 要求的域名被称为 2013 RAA 非过渡 (2013 RAA NGF) 域名。因此，分析中考虑了三种 RAA 类型（2009、2013 GF 和 2013 NGF），其分布可以在图表 1 中看到。

图表 1: gTLD 中按 RAA 状态划分的所有注册比例

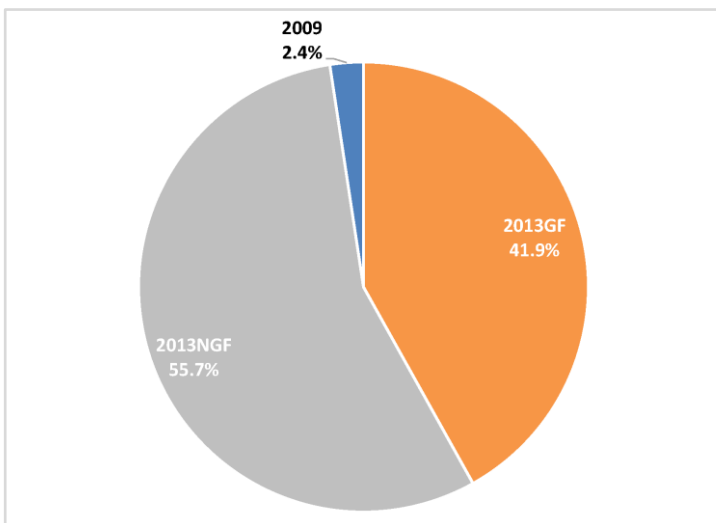
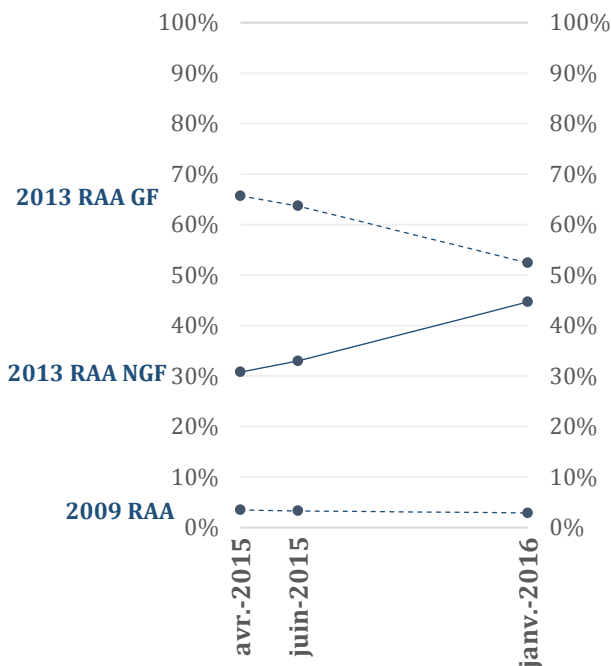


表 1 和统计图 1 显示三种 RRA 类型随时间推移的分布变化。它们显示 2009 RAA 的份额在非常缓慢地缩小，但清楚表明，许多过渡 2013 RAA 域名被转换为非过渡域名。

表 1：按样本日期划分的 RAA 类型分布

	2009 RAA	2013 GF RAA	2013 NGF RAA
2015 年 4 月（第 1 阶段）	3.5%	65.7%	30.8%
2015 年 6 月（第 2 阶段第 1 周期）	3.3%	63.7%	33.0%
2016 年 1 月（第 2 阶段第 2 周期）	2.9%	52.4%	44.7%

统计图 1：按样本日期划分的 RAA 类型分布变化¹⁷



样本设计

研究数据由来自 gTLD 域文件的 200,000 条记录（从第 1 周期的 150,000 增加至此数目）构成的初始样本和 12,000 条记录构成的经过分析的子样本组成。此两阶段抽样旨在提供足够大的样本，以便在收集研究数据存在技术限制的情况下对各相关小组进行可靠估计。gTLD 域文件中的数据是有限的，如果要选择足够大的样本来生成可靠的小组准确度估计，这些数据也并不包含所需的完整 WHOIS 信息（例如，注册人国家/地区、注册服务机构 RAA 版本）。为了获取所需信息，将在初

¹⁷ 这里未包括试点研究样本，因为试点并未从当时所有可用的域名采样。

始抽样中对每条记录执行 WHOIS 查询，所需的其他信息随后将附加到每条记录。通过这些其他信息附加到初始样本的记录，选择包含能够充分代表相关小组的子样本成为可能。初始样本的统计数据摘要和选择子样本的方法如下所述。

初始样本

选择由 200,000 条记录构成的初始样本时，我们审核了域文件摘要数据，其指示每个 gTLD 中有多少个域名。2016 年 1 月初进行第 2 周期初次抽样时，888 个 gTLD 中分布着近 1.7 亿个域名。1.7 亿个域名中的约 94% 都在 18 个早期 gTLD 之一中注册，相比之下，2015 年 6 月在第 1 周期（参阅表 1）收集数据时这个比例为 96%。2016 年 1 月超过 6% 的域名在新 gTLD 中注册，与 2015 年 6 月 4% 的注册比例相比实现了急剧增长。新 gTLD 的整体数据也取得大幅增长，从 2015 年 6 月的 660 增加到 2016 年 1 月的 870。

随着授权 gTLD 的总数出现增长，早期 gTLD 将稳定在 18，而新 gTLD 将持续增长。表 2 显示授权 gTLD 的总数，以及在每个 WHOIS ARS 样本日期，早期和新 gTLD 分别是多少。

表 2：按样本日期划分的授权总数、早期和新 gTLD

	授权 gTLD 总数	早期 gTLD	新 gTLD	gTLD 注册 (单位：百万)
2015 年 4 月（第 1 阶段）	610	18	592	157
2015 年 6 月 （第 2 阶段第 1 周期）	678	18	660	158
2016 年 1 月 （第 2 阶段第 2 周期）	888	18	870	169.5

在 870 个新 gTLD 中，仅 610 个新 gTLD 拥有至少一个域名（有 260 个新 gTLD 尚未拥有任何域名），40 个刚好拥有一个域名（这些通常是 gTLD 的管理域名，不在我们的采样范围内），剩下的 570 个 gTLD 则拥有至少 2 个域名。如上所述，总共增加 18 个早期 gTLD 和 570 个新 gTLD，初始样本代表总计 588 个 gTLD。

与之前的 WHOIS ARS¹⁸ 研究样本类似，我们的第 2 周期样本设计对新 gTLD 采用过抽样，因此，有 25% 的初始样本来自新 gTLD¹⁹。根据在第 1 周期研究中学到的经验，为了减少经过分析的子样本中所需的过抽样（参阅表 3），我们将初始样本大小从 150,000 增加到 200,000。

在 200,000 初始样本中，成功采集和解析的 WHOIS 数据达到 196,262 (98.1%)。剩下的 3,738 个域名中许多不再存在 (1,395)，但由于速率限制一些请求重复超时 (2,134)，并因各种原因出现一些失败的查询 (209)。在 196,262 个域名中，2009 RAA 小组中的记录数量占记录总数的 2.4%，而 2013 过渡 (2013 GF) 和 2013 非过渡 (2013 NGF) 记录的数量分别占域名总数的 41.9% 和 55.7%（请参阅表 3）。

表 3：初始样本大小（按地区和 RAA 划分）

RAA 类型	非洲地区	亚太地区	欧洲地区	拉丁美洲和加勒比海地区	北美地区	未知	总计
2009	30	450	619	306	3,258	19	4,682
2013 GF	457	14,379	18,275	2,184	46,564	438	82,297
2013 NGF	769	47,991	17,571	6,157	36,062	733	109,283
总计	1,256	62,820	36,465	8,647	85,884	1,190	196,262

经过分析的子样本

ICANN 为本报告定义的相关小组是指：使用 2009 RAA 注册服务机构的记录、使用 2013 RAA 注册服务机构的记录、新 gTLD 中的记录、早期 gTLD 中的记录，以及来自全部 5 个 ICANN 地区的记录。相应地，我们选择经过分析的子样本，以便尽可能以 95% ± 5% 的置信区间估计相关小组数据的准确度。这种置信区间需要特定小组根据其在 200,000 域名构成的初始样本中的代表性

¹⁸ 之前的 WHOIS ARS 研究包括[试点](#)、[第 1 阶段研究](#)和[第 2 阶段第 1 周期研究](#)。

¹⁹ 为确保能代表所有拥有至少两个域名的 570 个新 gTLD，我们首先从各 gTLD 中选取了一个样本，其余样本根据大小成比例进行选择（从较大的新 gTLD 中选择更多样本）。类似地，我们从所有早期 gTLD 中至少选择了一个域名，其余样本根据大小成比例进行选择。所有抽样都通过在 gTLD 内进行系统性抽样完成。根据针对每个 gTLD 确定的样本大小确定了抽样间隔（域名总数除以所需的样本大小）。随后，确定零和抽样间隔之间的随机起点。如果此随机起点是 166.2，抽样间隔是 300，则选定的记录将是第 167 条（随机起点向上取整）、第 467 条、第 767 条依此类推。使用此方法可以得到按 gTLD 域文件中的部分或完整排序进行隐形分层的样本（例如，较新的域名排序到域文件中记录列表的顶部或底部）。我们的方法对较小的 gTLD 进行轻微过抽样，并同时让较大的 gTLD 保持非常相似的权重，以确保方差不会按差分权重膨胀。

对其进行过抽样（甚至是直接包含）。虽然抽样不能明确确保包括所有注册服务机构，但根据每个 TLD、RAA 类型和注册人地区进行抽样时，449 个注册服务机构在子样本中得到代表，在经过分析的子样本中实现了注册服务机构多样性。子样本没有考虑 gTLD 类型（早期与新 gTLD），因为初始样本对新 gTLD 进行了过抽样。表 4 显示了按地区和 RAA 划分的经过分析的子样本的大小。²⁰

表 4：经过分析的子样本大小（按地区和 RAA 划分）

RAA 类型	非洲地区	亚太地区	欧洲地区	拉丁美洲 和加勒比海 地区	北美地区	未知	总计
2009	30	450	619	306	800	5	2,210
2013 GF	457	1,000	1,000	800	1,401	13	4,671
2013 NGF	769	1,443	1,000	800	1,085	22	5,119
总计	1,256	2,894	2,619	1,906	3,285	40	12,000

²⁰ 选择由 12,000 个域名构成的子样本进行分析时，目标是让按 RAA 类型划分的每个地区单元格中都有 800 个样本（表 4）。我们选择数字 800 作为目标，以便最小化每个单元格中的置信区间大小。如果某个单元格在初始样本中的样本数量不足 800，应选择所有样本。我们对其他大多数单元格进行过抽样，以便在各单元格中获得 800 个域名，并且如果在初始抽样中某个单元格大于 10,000，则会选择 1,000。只有三个单元格选择的域名超过 1,000，即北美地区 2013 GF、北美地区 2013 NGF 和亚太地区 NGF。对未知地区的单元格抽样的比例与北美地区的相同。

由于少数域名仍在根据 2009 RAA 运作的注册服务机构下注册，经过分析的子样本中包含的此类域名样本比例偏高，因此，与 2009 RAA 域名相关的估计将符合上述可靠性标准。表 5 对比由 196,262 条记录构成的初始样本和由 12,000 条记录构成的经过分析的子样本中按 RAA 类型划分的样本大小。

表 5：样本大小（按 RAA 类型划分）

RAA 类型	占域名总数的百分比	初始样本	经过分析的子样本	占子样本的百分比
2009 RAA	2.4%	4,682	2,210	18.4%
2013 RAA GF	41.9%	82,297	4,671	38.9%
2013 RAA NGF	55.7%	109,283	5,119	42.7%
总计	100.0%	196,262	12,000	100.0%

各联系人类型的相同数据统计

对于全部三种联系模式（电子邮件地址、电话号码和邮政地址），超过 75% 的域名的全部三种联系人类型（注册人联系人、管理联系人和技术联系人）的联系信息相同。表 6 所示为各联系人类型的联系信息相同的频率的完整分布情况。

表 6：各联系人类型和联系模式数据相同的频率²¹

共性	电子邮件地址	电话号码	邮政地址
3 项均相同	77.6% ± 0.7%	80.3% ± 0.7%	78.2% ± 0.7%
两项完全相同，一项不同	20.1% ± 0.7%	18.6% ± 0.7%	19.8% ± 0.7%
3 项均不同	2.3% ± 0.3%	1.0% ± 0.2%	2.0% ± 0.3%

从表 6 的共性数据可以看出，由于注册人联系人、管理联系人和技术联系人所包含的信息通常相同，这三种联系人的准确度差别不大。全部三种联系信息都不同的比例不超过 2.3%。因此，对全部三种联系人类型进行测试和报告时，通常只需查看“全部 3 种”联系人类型均准确的比例。表 6 的扩展版可参阅[附录 B](#) 中的表 B1。

²¹ 表 4 的扩展版可参阅[附录 B](#) 中的表 B1。

语法与可操作性测试方法

语法与可操作性准确度测试设计为，按照根据 2009 RAA 要求制定的一组基本要求对经分析的子样本中的所有记录进行评估²²。记录的全部 9 个联系信息字段都会进行测试（即注册人联系人、管理联系人和技术联系人的电子邮件地址、电话号码和邮政地址这三种联系模式），然后作为完整记录编制结果。关于准确度测试标准的信息和更详细测试信息的链接，请参阅[附录 A](#)。

语法测试方法

语法测试旨在通过将记录的联系信息与 RAA 中规定的合同要求指定的格式对比，从而进行评估。语法测试分两个阶段²³展开：第一阶段测试根据适用 RAA 的要求验证是否存在联系信息，第二阶段进行语法的详细技术测试。语法测试标准在所有先前的 WHOIS ARS 研究中保持一致。

可操作性测试方法

可操作性测试旨在评估记录的联系信息可否实际用于通信。与社群（包括注册服务机构社群的志愿者）磋商后，根据 RAA 要求制定了可操作性验证标准。WHOIS 记录中（例如全部三种联系人类型使用相同的电子邮件地址）和 WHOIS 记录之间（例如多条记录中使用相同的注册人联系数据）的重复性数据仅测试一次（即删除重复数据）。

²² 本报告附录 C 中提供 2013 RAA 要求的其他测试。

²³ 每种联系模式的第一阶段和第二阶段语法测试详细记录在 WHOIS ARS 网页中：
<https://whois.icann.org/en/whoisars-validation>。

主要研究结果

我们在此介绍我们的研究结果，以及与语法和可操作性准确度测试结果相关的统计数据。报告的这一部分包括的重要研究结果的摘要，其次是语法和可操作性测试结果的详细统计分析。这些统计数据分别按联系人类型²⁴（注册人联系人、管理联系人和技术联系人）、联系模式（电子邮件地址、电话号码和邮政地址）、整体和各个小组，如新 gTLD 与早期 gTLD、RAA 类型和 ICANN 地区列出。关于研究结果的进一步详情，包括分析表，请参阅[附录 B](#)。²⁵

由于 2009 和 2013 RAA 版本对有效语法的要求有所不同，我们针对每组要求（2009 和 2013）分别生成了分析表格，并使用 2009 要求作为基准²⁶。由于各 RAA 版本的可操作性结果都类似，因此为各组要求分别制定分析表格在很大程度上是多此一举。关于展示 2013 RAA 要求语法测试结果的分析表，请参阅[附录 C](#)。

研究结果摘要

我们在此列出研究结果的关键要点：

建立联系的能力

- 对于 99% 的记录而言，三个联系人类型中至少有一个联系模式符合 2009 RAA 的所有语法和可操作性要求，这意味着几乎所有记录都包含可用于建立联系的信息。只有 1% 的记录没有符合语法或可操作性要求的任何联系人类型的联系模式。

²⁴ 由于注册人联系人、管理联系人和技术联系人的数据如此相似，我们此处仅介绍“3 项均准确”（即注册人联系人、管理联系人和技术联系人均通过所有准确度测试）的小组准确度。

²⁵ 为了在本部分简洁呈现研究结果，此处讨论的许多分析表格存放在本报告的[附录 B](#) 和 [附录 C](#) 中。

²⁶ 我们选择 2009 RAA 作为基准对全部 10,000 条经过分析的子样本记录进行分析。2013 RAA 的要求比 2009 RAA 的要求更严格，它基于 2009 RAA 的要求构建，因此也涵盖 2009 RAA 的要求。例如，2009 RAA 要求每个联系人都有地址，而 2013 RAA 则要求每个联系人地址遵循特定国家/地区的适用万国邮政联盟 S42 模板的格式。任何满足 2013 RAA 要求的联系人字段也将满足 2009 RAA 的要求。基于此原因，2009 RAA 的要求可用作所有记录的对比基准。

可操作性准确度

- 98% 的邮政地址、76% 的电话号码以及 91% 的电子邮件地址均符合 2009 RAA 中规定的所有可操作性要求。70% 的域名通过了针对全部联系人类型（注册人联系人、管理联系人、技术联系人）和联系模式（电子邮件地址、电话号码、邮政地址）的所有可操作性测试，相比第 1 周期增长了约 6%。
 - 电话号码可操作性准确度的地区差异最明显，准确率范围从 63.7%（亚太地区）到 85.3%（北美地区）。
- 所有可操作性测试通过率最高的联系模式是邮政地址。所有可操作性测试通过率最低的联系模式是电话号码。
 - 对于少数未通过可操作性测试的邮政地址，约四成没有可识别或易于推断的国家/地区名称。
- 关于电子邮件地址的可操作性错误，约 8.5% 的电子邮件地址会将电子邮件退回，而电子邮件地址缺失的比例还不到 1%。

语法准确度：

- 85% 的电话号码符合 2009 RAA 中规定的所有语法要求，比第 1 周期 (83%) 略有上升，并且更加符合第 1 阶段的研究结果 (86%)。语法错误原因的分布与第 1 周期极其相似。
 - 邮政地址可操作性准确度的地区差异最明显，准确率范围从 44.6%（非洲地区）到 96.7%（北美地区）。
 - 大多数地区电话号码语法错误的最常见原因是长度有误，但在北美地区，最常见的错误原因是缺少国家/地区代码。
 - 对于邮政地址，每项研究中的大部分错误始终都是因为缺少必填字段，例如城市、州/省、邮政编码或街道。

语法准确度 — 2009 RAA 要求²⁷

以下章节通过先后论述整体准确度、小组准确度以及错误原因来回顾以 2009 要求为标准的语法准确度测试的结果。

²⁷ 有关 2013 RAA 要求的合规性，请参阅[附录 C](#)。

整体语法准确度

首先，我们来看分析的子样本中全部 12,000 个域名满足 2009 RAA 要求的准确度。从统计图 2 中的黑色虚线可看出，约 67% 的域名语法准确。99% 的电子邮件地址通过了语法准确度测试，而通过测试的电话号码和邮政地址分别为 85% 和 77%。

统计图 2：整体准确度 — 2009 RAA 语法要求

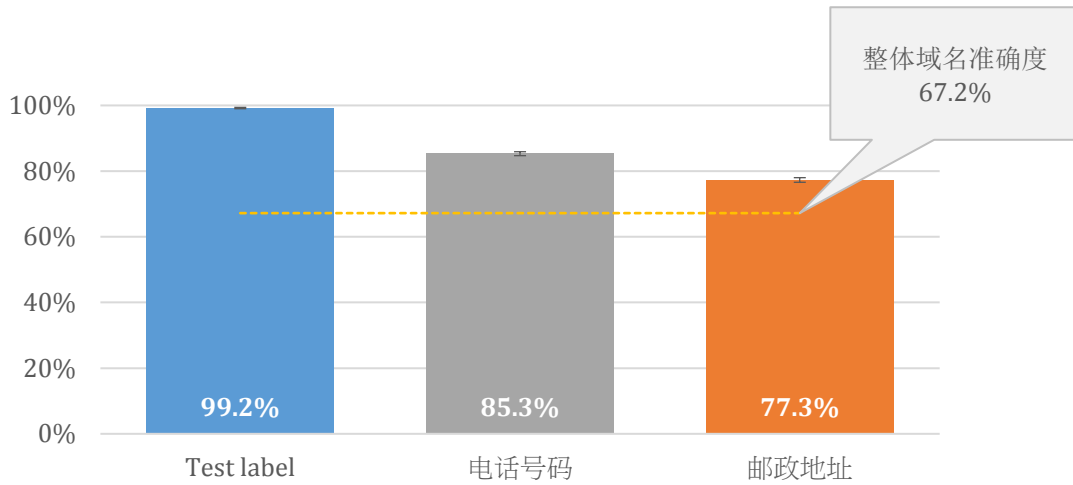


表 7 更详细地显示了各联系人类型的细分数据。这个表格的最后一行表示注册人联系人、管理联系人和技术联系人全都通过特定联系模式（电子邮件地址、电话号码或邮政地址）语法测试的比率。²⁸ 我们将在小组分析中关注全部三种联系模式均通过所有准确度测试（“3 项均准确的行”）的百分比。

表 7：整体准确度（按联系人类型和联系模式划分）— 2009 RAA 语法要求

	电子邮件地址	电话号码	邮政地址	3 项均准确
注册人联系人	100.0% ± 0.0%	88.1% ± 0.6%	79.0% ± 0.7%	69.8% ± 0.8%
管理联系人	99.2% ± 0.2%	86.5% ± 0.6%	78.6% ± 0.7%	69.2% ± 0.8%
技术联系人	99.2% ± 0.2%	86.9% ± 0.6%	80.9% ± 0.7%	71.8% ± 0.8%
总体	99.2% ± 0.2%	85.3% ± 0.6%	77.3% ± 0.7%	67.2% ± 0.8%

²⁸ 由于联系数据之间高频率的共性（如表 4 所示），每种联系模式内的准确度比率在联系人类型之间非常相似，这表明不论信息是包含在注册人联系人、管理联系人或是技术联系人字段中，每种联系模式内的准确度都几乎相同。

语法准确度（早期 gTLD 与新 gTLD）

从统计图 3 和表 8 都可看出，早期 gTLD 的电子邮件地址和电话号码准确度较低²⁹，但邮政地址准确度较高。早期 gTLD 所有三个联系人字段均准确的比例也更高。

统计图 3：按 gTLD 类型划分的准确度 — 2009 RAA 语法要求

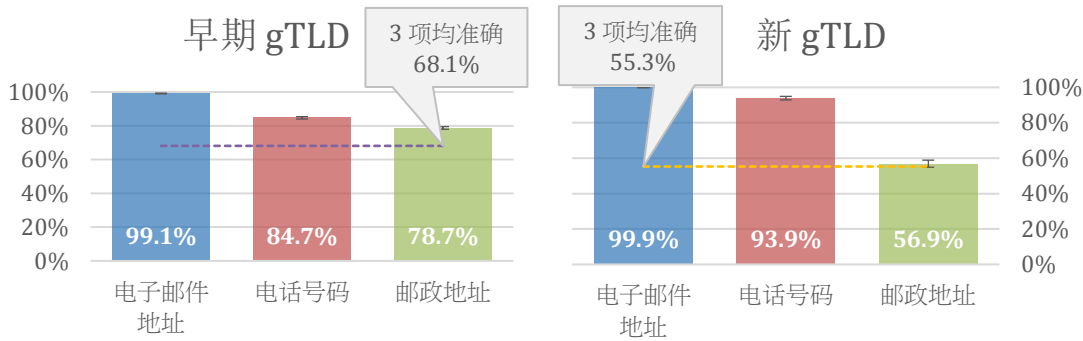


表 8：按 gTLD 类型划分的准确度 — 2009 RAA 语法要求

	电子邮件地址	电话号码	邮政地址	3 项均准确
早期 gTLD	99.1% ± 0.2%	84.7% ± 0.7%	78.7% ± 0.8%	68.1% ± 0.9%
新 gTLD	99.9% ± 0.1%	93.9% ± 1.0%	56.9% ± 2.0%	55.3% ± 2.0%
总体	99.2% ± 0.2%	85.3% ± 0.6%	77.3% ± 0.7%	67.2% ± 0.8%

按 RAA 状态划分的语法准确度

接下来，我们来看按 RAA 状态划分的准确率。统计图 4 和表 9 显示，2013 GF 和 2013 NGF 小组之间电子邮件地址准确度存在的差异虽小，但具有重要的统计学意义，而 2009 RAA 小组拥有与 2013 GF 和 2013 NGF 小组类似的准确度。2013 RAA GF 小组拥有最低的电话号码准确度，而 2009 和 2013 NGF 拥有类似的电话号码准确度。邮政地址准确度最高的是 2009 RAA 小组，最低的是 2013 NGF 小组。尽管 2009 RAA 小组中各个联系模式的准确度并非最高，但该小组全部三种联系模式均准确的记录的百分比最高。

²⁹ 此处的“较高”和“较低”不单指数字，还指统计显著性。为方便您阅读，省略了对后者的大部分描述。

统计图 4：按 RAA 状态划分的准确度 — 2009 RAA 语法要求

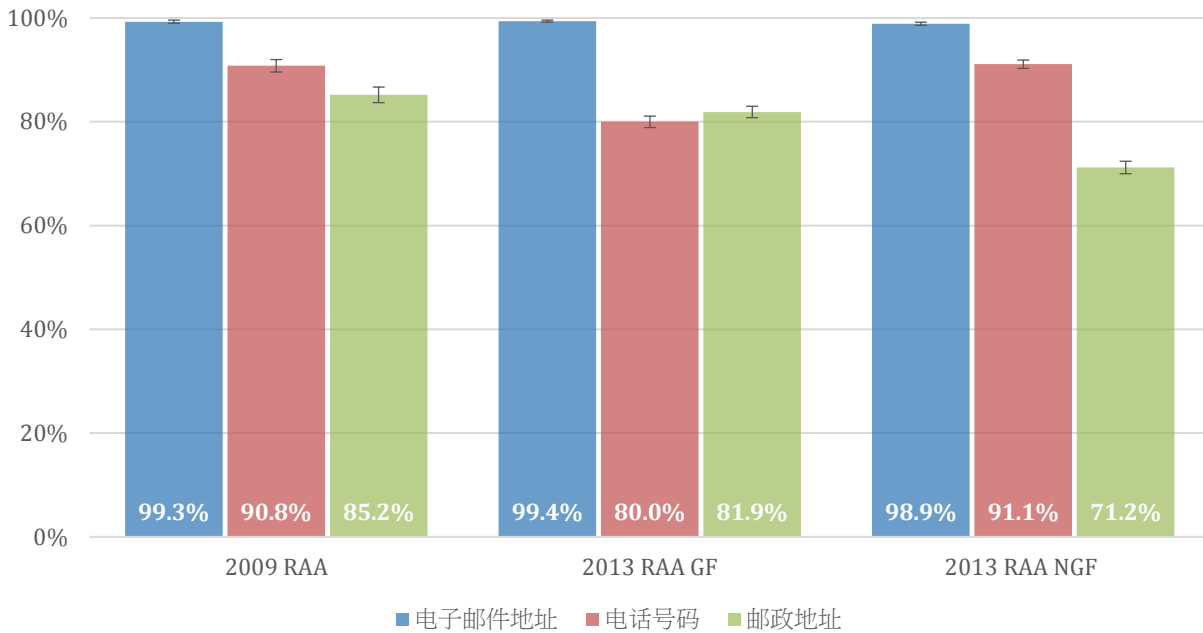


表 9：按 RAA 状态划分的准确度 — 2009 RAA 语法要求

	电子邮件地址	电话号码	邮政地址	3 项均准确
2009 RAA	99.3% ± 0.3%	90.8% ± 1.2%	85.2% ± 1.5%	80.9% ± 1.6%
2013 RAA GF	99.4% ± 0.2%	80.0% ± 1.1%	81.9% ± 1.1%	66.8% ± 1.4%
2013 RAA NGF	98.9% ± 0.3%	91.1% ± 0.8%	71.2% ± 1.2%	66.9% ± 1.3%
总体	99.2% ± 0.2%	85.3% ± 0.6%	77.3% ± 0.7%	67.2% ± 0.8%

按 ICANN 地区划分的语法准确度

最后，我们来看按 ICANN 地区划分的准确度。统计图 5 和表 10 显示，亚太地区的电子邮件地址准确度低于其他四个地区。对于电话号码，除非洲地区外所有地区的语法准确度都在 84.3%（拉丁美洲/加勒比海地区）和 88.9%（亚太地区）之间。与此相反，第 1 周期中仅两个地区（欧洲地区和北美地区）的电话号码准确度达到 80% 以上。邮政地址的结果与第 1 周期类似，北美地区的准确度最高，而非洲和亚太地区的准确度最低。在全部三种联系模式均符合 RAA 2009 标准方面呈现相同的形式，北美地区在“3 项均准确”方面的准确率最高，而非洲地区和亚太地区的准确率最低。如需详细了解按地区划分的地区准确度统计数据 and 错误原因，请参阅[地区研究结果](#)部分。

统计图 5：按 ICANN 地区划分的准确度 — 2009 RAA 语法要求

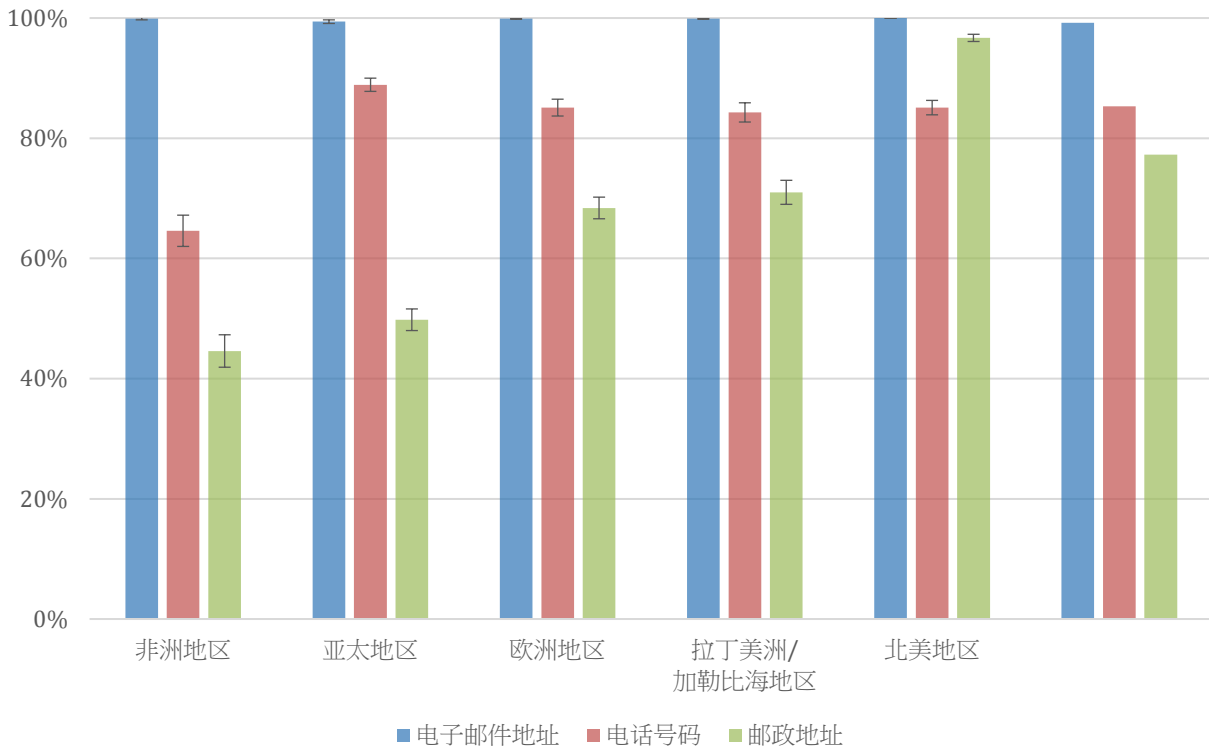


表 10：按 ICANN 地区划分的准确度 — 2009 RAA 语法要求

	电子邮件地址	电话号码	邮政地址	3 项均准确
非洲地区	99.9% ± 0.2%	64.6% ± 2.6%	44.6% ± 2.7%	29.3% ± 2.5%
亚太地区	99.4% ± 0.3%	88.9% ± 1.1%	49.8% ± 1.8%	45.0% ± 1.8%
欧洲地区	99.9% ± 0.1%	85.1% ± 1.4%	68.4% ± 1.8%	60.6% ± 1.9%
拉丁美洲/加勒比海地区	99.9% ± 0.1%	84.3% ± 1.6%	71.0% ± 2.0%	64.7% ± 2.1%
北美地区	100.0% ± 0.0%	85.1% ± 1.2%	96.7% ± 0.6%	82.8% ± 1.3%
总体	99.2% ± 0.2%	85.3% ± 0.6%	77.3% ± 0.7%	67.2% ± 0.8%

错误原因 — 2009 RAA 语法要求

这里我们按联系模式（电子邮件地址、电话号码和邮政地址）对主要错误原因分别进行报告。我们可确定首个未通过测试的电子邮件地址和电话号码。邮政地址需要多个字段，因此可能出现多个语法错误。正如对准确度测试结果所采取的处理方式一样，对于全部 12,000 个经过分析的域名中未

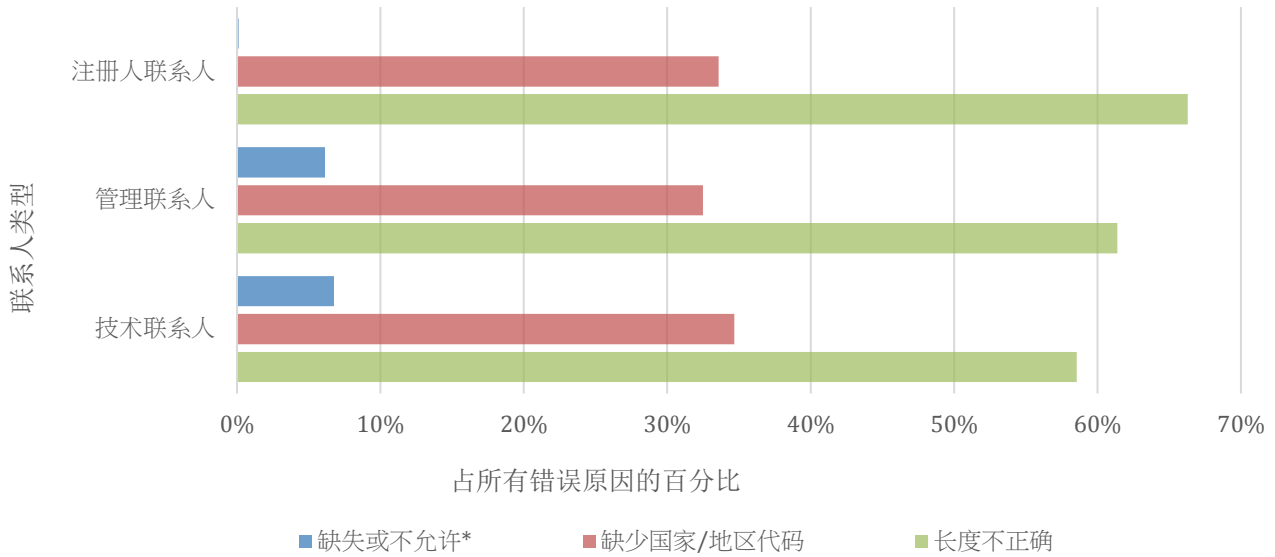
能通过 2009 RAA 要求的主要原因，我们使用单独的表格进行报告。³⁰在第 1 周期，我们按联系模式显示了记录未通过的准确度测试。这里我们也提供第 2 周期的统计图。[附录 B](#) 提供了第 1 周期和第 2 周期的列表数据以进行对比，也包括电子邮件地址的表格（其中的语法错误很少，因而在此忽略）。

电话号码

统计图 6 按联系人类型，以百分比显示所有电话号码错误中电话号码错误的原因。与第 1 周期类似，在电话号码错误中占比最大的原因是适用国家/地区的长度不正确（约占测试的所有电话号码的 8.7%）。第二大常见错误是缺少国家/地区代码（约占测试的所有电话号码的 4.7%）。缺少电话号码的错误不足 1%。

³⁰ 您可在[附录 C](#) 中查看 2013 NGF 小组未能通过 2013 RAA 要求的主要原因。

统计图 6：电话号码错误原因 — 2009 RAA 语法要求

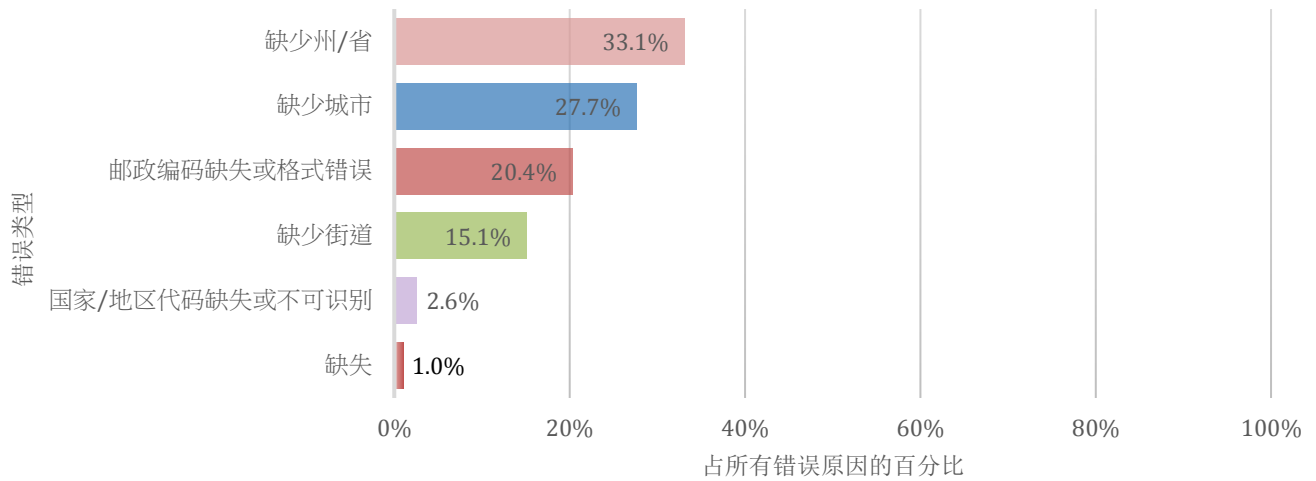


*注：2009 RAA 不要求注册人联系人类型下必须存在电话号码。

邮政地址

统计图 7 以百分比显示所有邮政地址错误中邮政地址错误的原因。与第 1 周期类似，绝大多数的邮政地址语法错误 (96.7%) 是由于缺少地址组成部分，如缺少州/省 (33.1%)、城市 (27.7%)、邮政编码 (20.4%) 和/或街道 (15.1%)。缺少国家/地区代码的错误较少（仅占有所有电话号码语法错误的 2.6%），缺少整个电话号码的错误极少（仅占有所有电话号码语法错误的 1%）。

统计图 7：邮政地址错误的原因（所有联系人类型）— 2009 RAA 语法要求



可操作性准确度 — 2009 RAA 要求³¹

以下章节通过先后论述整体准确度、小组准确度以及错误原因来回顾以 2009 RAA 要求为标准的可操作性准确度测试的结果。有必要在此指出，2013 和 2009 RAA 可操作性要求之间的不同之处仅在于 2009 RAA 要求未规定在注册人电子邮件地址和电话号码字段填写信息，而 2013 RAA 要求则规定了要在这些字段填写信息。

整体可操作性准确度

首先，我们来看全部 12,000 个域名满足 2009 RAA 要求的准确度。在语法测试结果中，我们看到语法准确率最低的是邮政地址。但是，邮政地址的可操作性准确率却是最高的，如以下统计图 8 和表 11 所示。³²其他两种联系模式（电子邮件地址和电话号码）的可操作性准确率低于语法准确率。有关可操作性准确度测试流程的详细信息，请参阅[附录 A](#)。

³¹ 有关 2013 RAA 要求的合规性，请参阅[附录 C](#)。

³² 邮政地址可操作性测试没有邮政地址语法测试那么严格。例如，要符合邮政信件万国邮政联盟标准的语法要求，需缩写州或省的名称（例如，在美国，“DE”符合，但“Delaware”不符合），但是这些语法元素不一定基于包裹能否送达。换言之，语法准确度对电子邮件地址和电话号码而言是可操作性指示因素，对邮政地址而言却不是。语法不准确的电子邮件地址（例如，没有“@”符号）将不可操作，但语法不准确的邮政地址仍可能可操作（即可送达）。另请参阅[挑战和经验教训](#)中的讨论。

统计图 8：整体准确度 — 2009 RAA 可操作性要求

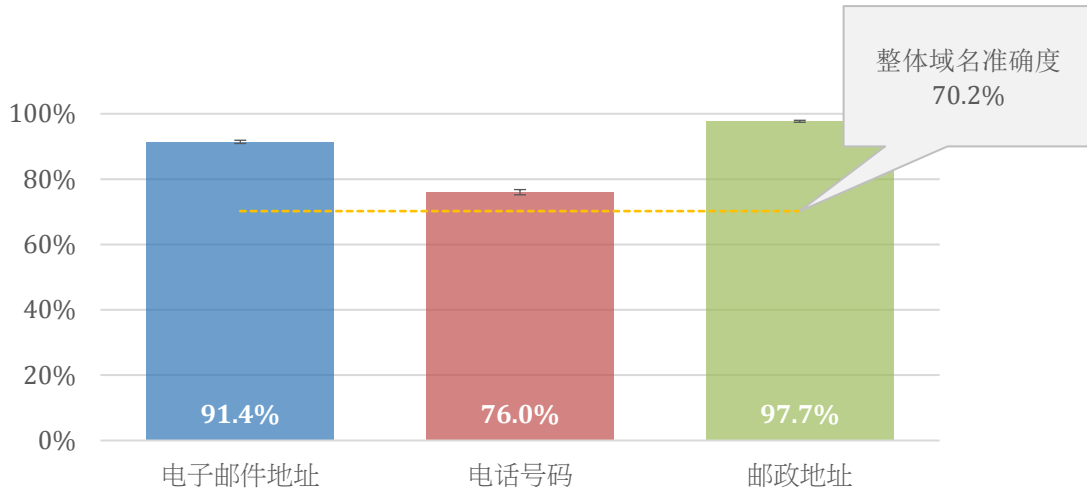


表 11 按联系人类型更详尽地显示细分情况。

表 11：整体准确度（按联系人类型和联系模式划分）— 2009 RAA 可操作性要求

	电子邮件地址	电话号码	邮政地址	3 项均准确
注册人联系人	93.3% ± 0.4%	81.2% ± 0.7%	97.9% ± 0.3%	74.8% ± 0.8%
管理联系人	92.5% ± 0.5%	80.1% ± 0.7%	97.8% ± 0.3%	74.6% ± 0.8%
技术联系人	93.1% ± 0.5%	78.6% ± 0.7%	97.8% ± 0.3%	73.5% ± 0.8%
总体	91.4% ± 0.5%	76.0% ± 0.8%	97.7% ± 0.3%	70.2% ± 0.8%

可操作性准确度（早期 gTLD 与新 gTLD）

从统计图 9 和表 12 可看出，早期 gTLD 的电子邮件地址和邮政地址的可操作性准确度较低，但电话号码的可操作性准确度较高。这些结果与第 1 周期的研究结果类似。

统计图 9：按 gTLD 类型划分的准确度 — 2009 RAA 可操作性要求

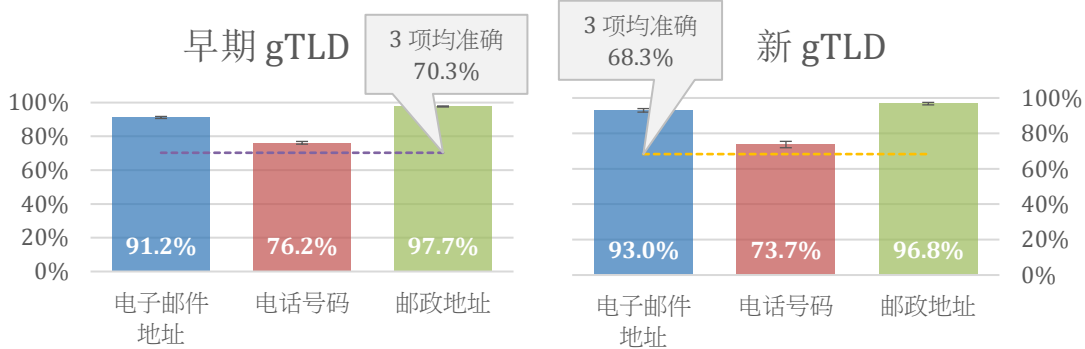


表 12：按 gTLD 类型划分的准确度 — 2009 RAA 可操作性要求

	电子邮件地址	电话号码	邮政地址	3 项均准确
早期 gTLD	91.2% ± 0.6%	76.2% ± 0.8%	97.7% ± 0.3%	70.3% ± 0.9%
新 gTLD	93.0% ± 1.0%	73.7% ± 1.8%	96.8% ± 0.7%	68.3% ± 1.9%
总体	91.4% ± 0.5%	76.0% ± 0.8%	97.7% ± 0.3%	70.2% ± 0.8%

按 RAA 状态划分的可操作性准确度

最后，我们来看按 RAA 状态划分的准确度。从统计图 10 和表 13 可看出，相较于 2009 RAA 小组，2013 RAA GF 和 2013 RAA NGF 小组的电子邮件地址准确度更高，而 2009 RAA 小组的电话号码准确度高于其他两个小组。2013 RAA NGF 小组的电子邮件准确度最高，且其任何联系模式的准确度均不是最低，因此该小组“3 项均准确”的百分比最高。

统计图 10: 按 RAA 状态划分的准确率 — 2009 RAA 可操作性要求

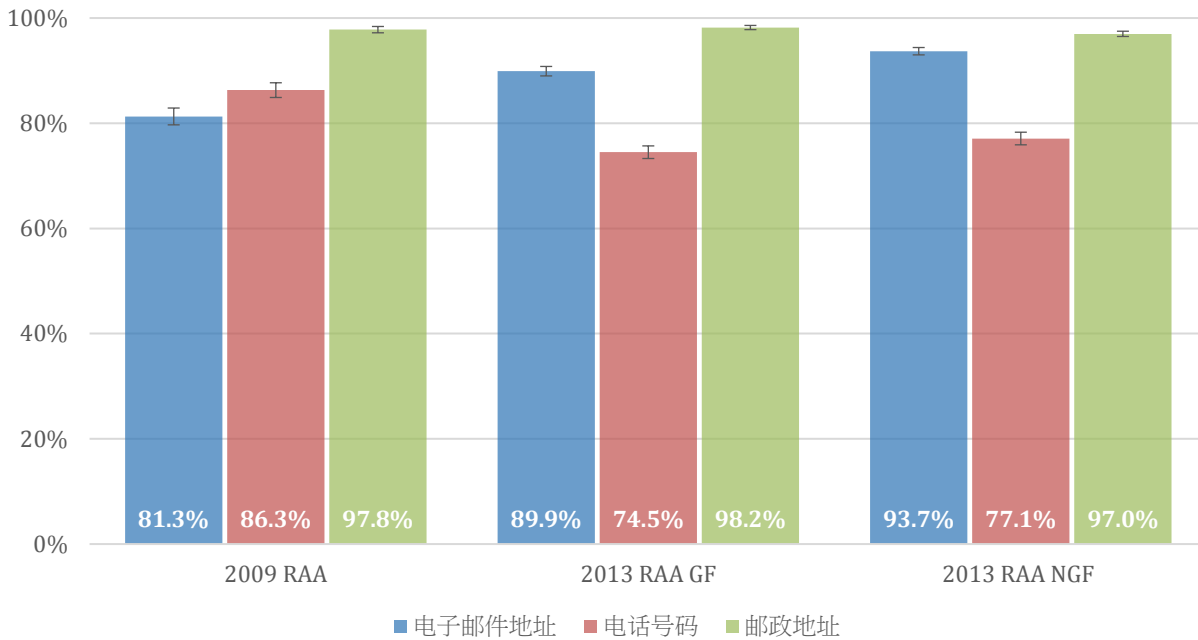


表 13: 按 RAA 状态划分的准确率 — 2009 RAA 可操作性要求

	电子邮件地址	电话号码	邮政地址	3 项均准确
2009 RAA	81.3% ± 1.6%	86.3% ± 1.4%	97.8% ± 0.6%	69.7% ± 1.9%
2013 RAA GF	89.9% ± 0.9%	74.5% ± 1.2%	98.2% ± 0.4%	68.4% ± 1.3%
2013 RAA NGF	93.7% ± 0.7%	77.1% ± 1.2%	97.0% ± 0.5%	72.3% ± 1.2%
总体	91.4% ± 0.5%	76.0% ± 0.8%	97.7% ± 0.3%	70.2% ± 0.8%

按 ICANN 地区划分的可操作性准确度

接下来,我们来看按 ICANN 地区划分的准确度。从统计图 11 和表 14 可看出,非洲地区、亚太地区和北美地区域名的电子邮件地址可操作性准确率更高。拉丁美洲/加勒比海地区和北美地区域名的电话号码可操作性准确率高。非洲地区和亚太地区域名的邮政地址可操作性准确率低于其他地区。对于全部 9 个联系信息字段均通过所有准确度测试而言,拉丁美洲/加勒比海地区和北美地区域名的准确率较高,而亚太地区域名的准确率较低。如需详细了解按地区划分的地区准确度统计数据 and 错误原因, 请参阅[地区研究结果](#)部分。

统计图 11: 按 ICANN 地区划分的准确度 — 2009 RAA 可操作性要求

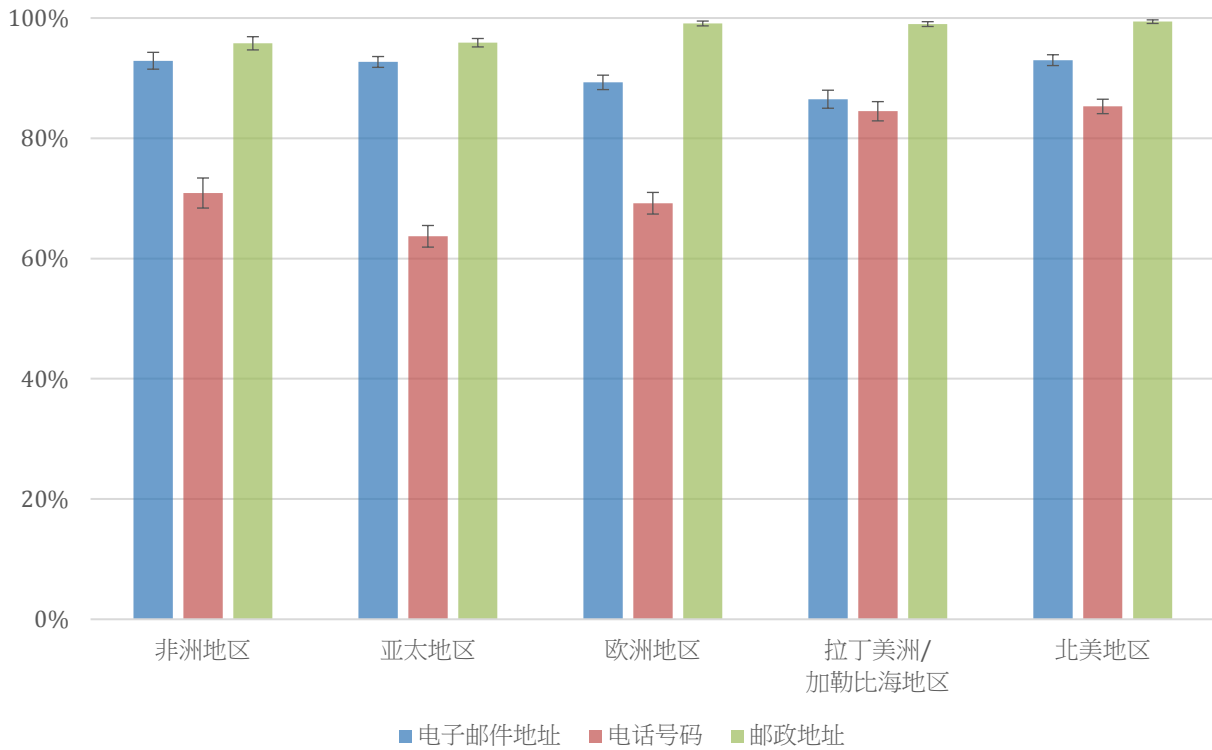


表 14: 按 ICANN 地区划分的准确度 — 2009 RAA 可操作性要求

	电子邮件地址	电话号码	邮政地址	3 项均准确
非洲地区	92.9% ± 1.4%	70.9% ± 2.5%	95.8% ± 1.1%	64.6% ± 2.6%
亚太地区	92.7% ± 0.9%	63.7% ± 1.8%	95.9% ± 0.7%	57.6% ± 1.8%
欧洲地区	89.3% ± 1.2%	69.2% ± 1.8%	99.1% ± 0.4%	63.1% ± 1.8%
拉丁美洲/加勒比海地区	86.5% ± 1.5%	84.5% ± 1.6%	99.0% ± 0.4%	71.6% ± 2.0%
北美地区	93.0% ± 0.9%	85.3% ± 1.2%	99.4% ± 0.3%	80.2% ± 1.4%
总体	91.4% ± 0.5%	76.0% ± 0.8%	97.7% ± 0.3%	70.2% ± 0.8%

错误原因 — 2009 RAA 可操作性要求³³

³³ 要了解关于如何开展测试和错误如何映射到测试的更多信息，请参阅附录 A 或 WHOIS ARS 网页：
<https://whois.icann.org/en/whoisars-validation>。

对于可操作性而言，错误原因比较直接，因为电子邮件地址、电话号码和邮政地址测试是依次进行的。如果一个测试失败，则可操作性不能通过测试。如果测试成功，则联系信息进入下一个测试。

电子邮件地址

从表 15 可看出，约 8.5% 的记录收到了“退回”邮件，这表明电子邮件地址不可操作。再次说明，2009 RAA 不要求提供注册人联系人电子邮件地址。如果提供了可验证的电子邮件地址，则会尝试发送一封电子邮件。如果连接成功，则该电子邮件地址被视为可操作。管理联系人和技术联系人字段中未提供要求的电子邮件地址的仅约 0.3%。

表 15：按联系人类型划分的电子邮件地址错误 — 2009 RAA 可操作性要求

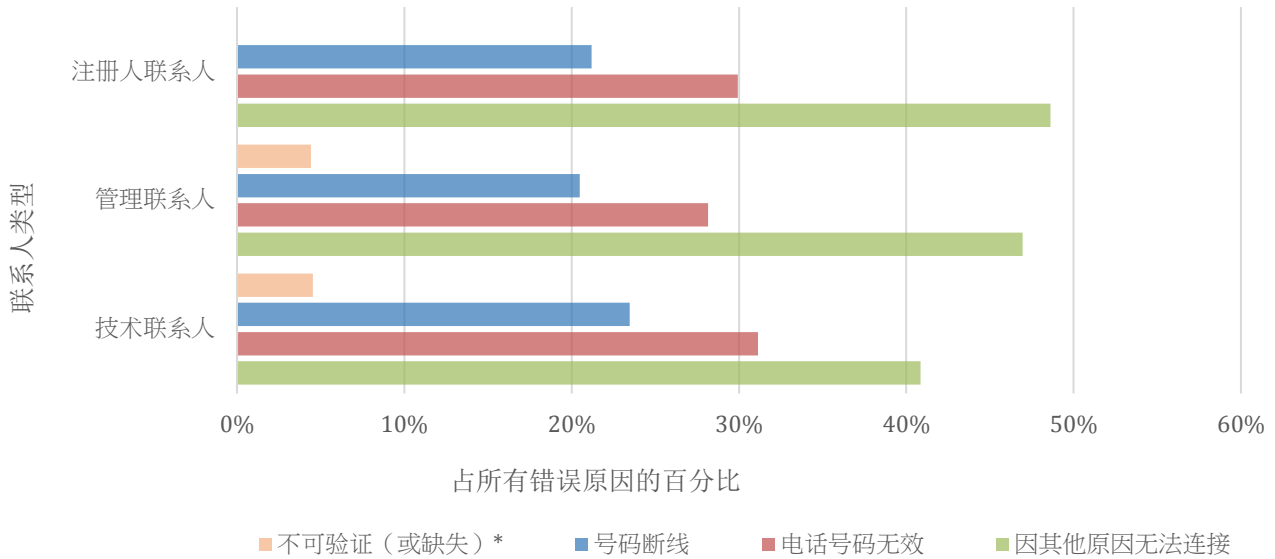
	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	10,995	10,943	10,898	32,836
不可验证（或缺失）	130*	49	51	100
邮件被退回	1,005	1,008	1,051	3,064
总计	12,000	12,000	12,000	36,000

* 2009 RAA 不要求提供注册人联系人电子邮件地址。

电话号码

统计图 12 和表 16 所示为电话号码的可操作性错误。再次说明，2009 RAA 不要求提供注册人联系人电话号码。如果提供了可验证的电话号码，则将进行拨号测试。从表 16 可看出，约 4.5% 的电话号码断线，6.1% 的号码无效，另外 9.4% 的号码无法连接。缺少所要求的电话号码的比例低于 1%。

统计图 12：按联系人类型划分的电话号码错误 — 2009 RAA 可操作性要求



*注：2009 RAA 不要求提供注册人联系人电话号码。

表 16：按联系人类型划分的电话号码错误 — 2009 RAA 可操作性要求

	管理联系人	技术联系人	注册人联系人	总计
通过所有准确度测试	9,618	9,506	9,444	28,568
不可验证（或缺失）	188*	110	116	232
号码断线	505	511	600	1,616
电话号码无效	713	702	796	2,211
因其他原因无法连接	1,158	1,171	1,044	3,373
总计	12,000	12,000	12,000	36,000

* 2009 RAA 不要求提供注册人联系人电话号码。

邮政地址

最后，统计图 13 和表 17 所示为邮政地址可操作性错误。首先由万国邮政联盟的可操作性测试工具自动对邮政地址的可操作性进行编码。³⁴V（按原样验证）、C（修改并验证）和 P3（可能可送达）均被视为可操作的邮政地址。³⁵不可操作的邮政地址会被分类到 P2（可能不可送达）、P1（很

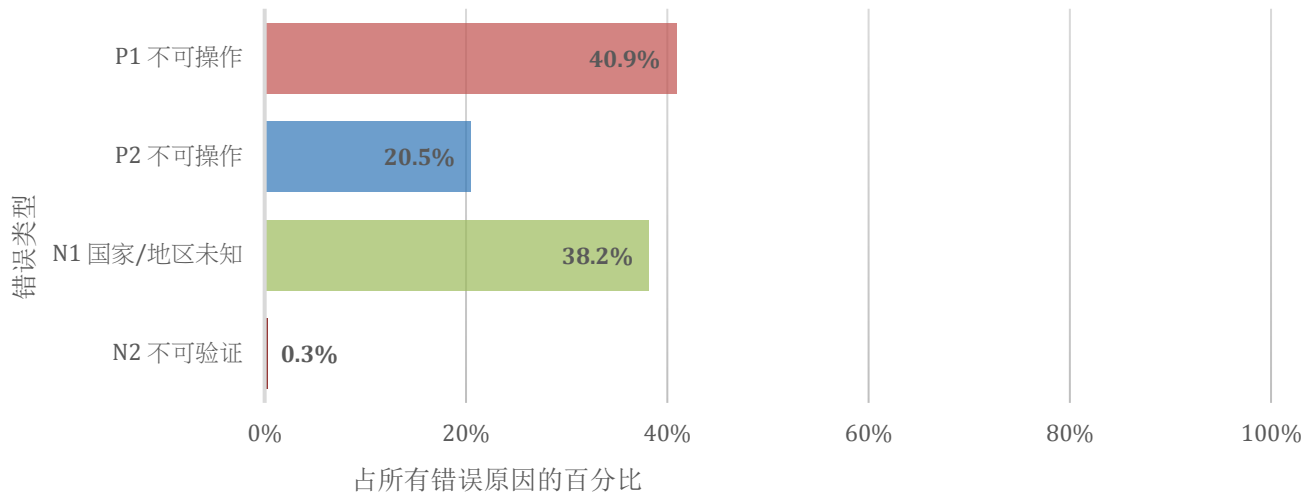
³⁴ 万国邮政联盟（邮政地址准确度测试供应商）模拟邮局协议受理包裹，将包裹发送至记录中的邮政地址，但并未实际将其送至目的地。关于准确度测试的信息，请参阅[附录 A](https://whois.icann.org/en/whoisars-validation) 和 WHOIS ARS 网页：

<https://whois.icann.org/en/whoisars-validation>。

³⁵ 参阅注释 34 以及[挑战和经验教训](#)中的讨论。

可能不可送达)、N1 (国家/地区未知) 和 N2 (由于无可用的国家/地区标准而不可验证的地址)。但是, 我们会对 P1 和 P2 地址实施手动流程³⁶, 该流程可确定大部分邮政地址为可操作。从表 17 可看出, 92.9% 的 P2 地址和 72.0% 的 P1 地址经手动流程确定为可操作。如统计图 13 所示, 在此手动流程后, 仅这些地址仍确定为不可操作。表 17 显示更多细节, 包括各个代码中有多少经手动流程确定为可操作。

统计图 13: 所有联系人类型的邮政地址错误 — 2009 RAA 可操作性要求



³⁶ 此手动流程还旨在针对技术上有误的地址模拟邮局协议, 有时可通过分析地址数据的组成部分来确定可操作的地址 (假定至少提供一些地址数据)。如表中所示, 无法总是确定正确的地址 (即视为可操作)。

表 17：按联系人类型划分的邮政地址错误 — 2009 RAA 可操作性要求

	管理联系人	技术联系人	注册人联系人	总计
可操作	10,432	10,445	10,480	31,357
可操作的 P2	829	809	823	2,461
可操作的 P1	439	447	409	1,295
可操作的总数	11,700	11,701	11,712	35,113
不可操作的 P2	63	61	58	182
不可操作的 P1	122	124	117	363
N1, 国家/地区未知	114	113	112	339
N2, 不可验证	1	1	1	3
不可操作的总数	300	299	288	887
综合总计	12,000	12,000	12,000	36,000

地区研究结果 — 按地区分析准确度和错误原因

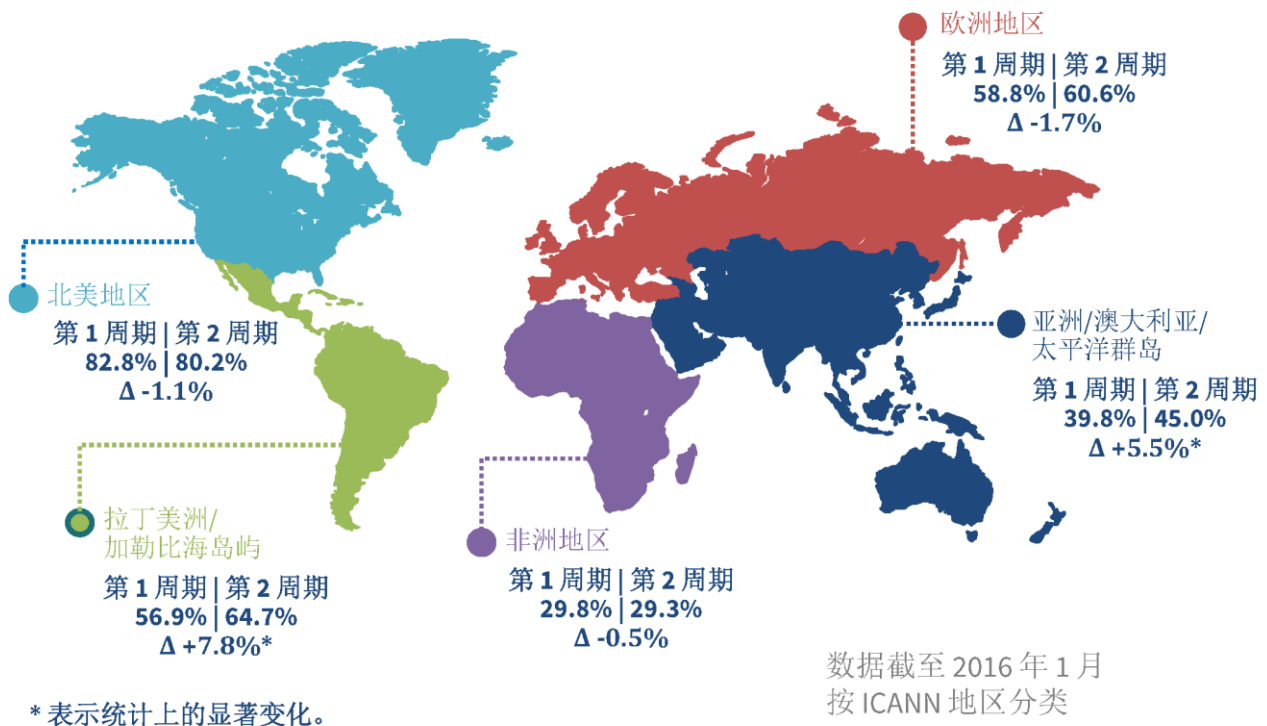
我们在此报告其他地区性统计数据，如第 1 周期和第 2 周期之间各个地区在整体语法和可操作性准确度方面的变化，第 2 周期中各个地区的语法和可操作性错误原因，以及用于注册域名的各类文字语言的语法和可操作性准确度。

按地区划分的整体准确度变化

语法准确度

如图表 2 所示，语法准确度的变化在亚太地区 and 拉丁美洲/加勒比海地区最为明显，分别增长 5.5% 和 7.8%。从第 1 周期到第 2 周期，所有地区的整体语法准确度下降 0.7%，降至 67.2%（参阅[下文](#)的趋势分析）。

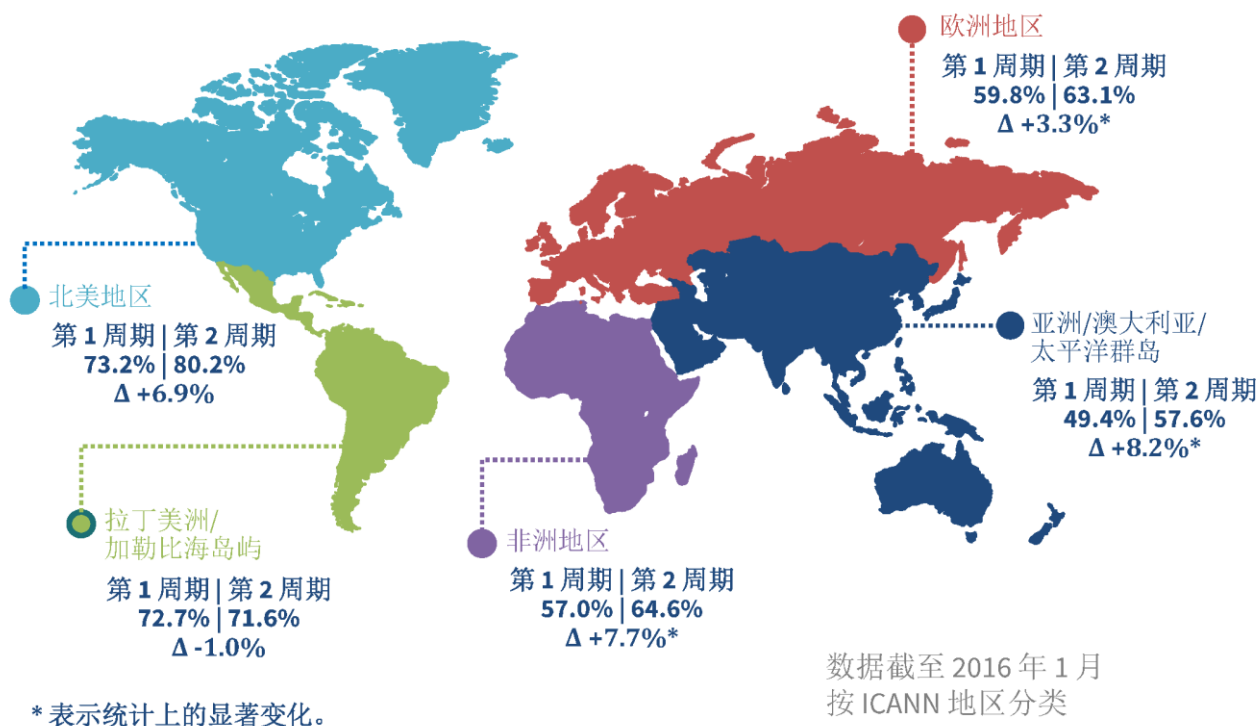
图表 2：按 ARS 周期划分的 ICANN 各地区整体语法准确度变化 — 2009 RAA 要求



可操作性准确度

图表 3 显示，除拉丁美洲/加勒比海地区外各个地区的可操作性准确度均出现增长，增长率为 3.3%（欧洲地区）至 8.2%（亚太地区）。从第 1 周期到第 2 周期，所有地区的整体语法准确度增长约为 5.4%，增至 70.2%（参阅[下文](#)的趋势分析）。

图表 3: 第 1 周期到第 2 周期 ICANN 地区的整体可操作性准确度变化 — 2009 RAA 要求



按地区划分的错误原因

这里我们按联系模式（电子邮件地址、电话号码和邮政地址）对各个地区语法和可操作性测试的主要错误原因分别进行报告。对于电子邮件地址和电话号码，我们报告首个未通过的测试。邮政地址需要多个字段，因此可能出现多个错误。

按地区划分的电子邮件地址语法和可操作性错误原因 — 2009 RAA

表 18 中的语法错误显示不同地区间的显著变化。但是有必要记住，电子邮件地址的实际语法错误数量非常少。大多数错误是由电子邮件地址缺失造成的。

表 18: 按地区划分的电子邮件地址语法错误原因 — 2009 RAA 要求

错误	非洲地区	亚太地区	欧洲地区	拉丁美洲和加勒比海地区	北美地区	未知	所有地区
缺失	0.0%	75.9%	100.0%	0.0%	100.0%	100.0%	86.8%
不允许的字符	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
@ 符号缺失	0.0%	3.4%	0.0%	40.0%	0.0%	0.0%	2.6%
不可解析	100.0%	20.7%	0.0%	60.0%	0.0%	0.0%	10.5%
地区的整体准确度 — 电子邮件地址语法	99.9% ± 0.2%	99.4% ± 0.3%	99.9% ± 0.1%	99.9% ± 0.1%	100.0% ± 0.0%	不适用	99.2% ± 0.2%

注: 此表应按如下方式理解: 在 X 地区的错误中, Y% 的错误是因为 Z (例如, 在非洲地区电子邮件地址的语法错误原因中, 100% 的错误是因为不可解析的地址)。“地区的整体电子邮件地址语法准确度”不是其上方百分比的总和, 但最好列出来, 为错误提供额外的背景信息。这就是说, 非洲地区 100% 的电子邮件地址语法错误是由于不可解析的地址, 但整体上的非洲地区电子邮件地址语法准确度为 99.9%, 这意味着实际的错误数量非常少。

表 19 中的可操作性错误显示, 电子邮件地址具有两大类可操作性错误: 缺失/不可验证, 或会将邮件退回。在每个地区, 电子邮件地址的错误在很大程度上是由于电子邮件被退回, 但当地区未知时, 几乎所有的错误都是由于缺少或不可验证的信息。

表 19: 按地区划分的电子邮件地址可操作性错误原因 — 2009 RAA 要求

错误	非洲地区	亚太地区	欧洲地区	拉丁美洲和加勒比海地区	北美地区	未知	所有地区
不可验证 (或缺失)	0.0%	4.0%	0.2%	0.3%	0.4%	98.6%	3.2%
邮件被退回	100.0%	96.0%	99.8%	99.7%	99.6%	1.4%	96.8%
地区的整体准确度 — 电子邮件地址可操作性	92.9% ± 1.4%	92.7% ± 0.9%	89.3% ± 1.2%	86.5% ± 1.5%	93.0% ± 0.9%	不适用	91.4% ± 0.5%

参阅表 18 中的注释了解如何理解此表格。

按地区划分的电话号码语法和可操作性错误原因 — 2009 RAA

表 20 再次显示, 在地区未知时, 电话号码语法错误的原因是信息缺失或存在不允许的信息。在各地区中, 北美地区在国家/地区代码缺失中所占的百分比最大, 而亚太地区在信息缺失或存在不允许的电话号码中所占的百分比最大。

表 20：按地区划分的电话号码语法错误原因 — 2009 RAA 要求

错误	非洲地区	亚太地区	欧洲地区	拉丁美洲 和加勒比海 地区	北美地区	未知	所有地区
长度不正确	74.4%	58.1%	76.8%	63.6%	44.4%	0.0%	62.0%
缺少国家/ 地区代码	25.2%	32.7%	21.0%	35.5%	54.4%	0.0%	33.6%
缺失或不允许	0.4%	9.2%	2.1%	0.9%	1.2%	100.0%	4.4%
地区的整体准 确度 — 电话号 码语法	64.6% ± 2.6%	88.9% ± 1.1%	85.1% ± 1.4%	84.3% ± 1.6%	85.1% ± 1.2%	不适用	85.3% ± 0.6%

参阅表 18 中的注释了解如何理解此表格。

表 21 显示，各个地区在电话号码可操作性的分布上差距不大。在各个地区的不可操作电话号码中，0.8% 到 3.8% 的号码缺失或不可验证，12.8% 到 28.3% 的号码断线，22.9% 到 34.1% 为无效号码，剩余 33.9% 到 56.7% 的不可操作电话号码因其他原因无法连接。

表 21：按地区划分的电话号码可操作性错误原因 — 2009 RAA 要求

错误	非洲地区	亚太地区	欧洲地区	拉丁美洲 和加勒比海 地区	北美地区	未知	所有地区
不可验证 (或缺失)	0.8%	3.8%	1.1%	1.0%	1.3%	100.0%	3.1%
号码断线	12.8%	28.3%	21.8%	15.6%	19.2%	0.0%	21.7%
电话号码无效	32.2%	34.1%	26.5%	31.9%	22.9%	0.0%	29.7%
因其他原因无法 连接	54.3%	33.9%	50.7%	51.5%	56.7%	0.0%	45.4%
地区的整体准确 度 — 电话号码 可操作性	70.9% ± 2.5%	63.7% ± 1.8%	69.2% ± 1.8%	84.5% ± 1.6%	85.3% ± 1.2%	不适用	76.0% ± 0.8%

参阅表 18 中的注释了解如何理解此表格。

按地区划分的邮政地址错误原因 — 2009 RAA

从表 22 可以看出，所有地区的绝大多数邮政地址语法错误都是因为地址组成部分缺失，如城市或州/省。再次出现这样的情况，即当地区未知时，错误是因为缺少已提供的信息（“缺失”）。

表 22：按地区划分的邮政地址语法错误原因 — 2009 RAA 要求

错误	非洲地区	亚太地区	欧洲地区	拉丁美洲 和加勒比海 地区	北美地区	未知	所有地区
缺失	0.1%	0.4%	0.0%	0.0%	1.0%	100.0%	1.0%
缺少国家/ 地区代码	0.4%	2.3%	8.1%	0.3%	3.4%	0.0%	2.6%
缺少街道	22.6%	14.8%	11.2%	12.1%	13.1%	0.0%	15.1%
邮政编码缺失或 格式错误	23.8%	15.3%	13.7%	39.5%	19.2%	0.0%	20.4%
缺少城市	29.3%	26.6%	33.0%	22.7%	39.1%	0.0%	27.7%
缺少州/省	23.8%	40.7%	34.0%	25.3%	24.1%	0.0%	33.1%
地区的整体准确 度 — 邮政地址 语法	44.6% ± 2.7%	49.8% ± 1.8%	68.4% ± 1.8%	71.0% ± 2.0%	96.7% ± 0.6%	不适用	77.3% ± 0.7%

参阅表 18 中的注释了解如何理解此表格。

表 23 显示，在非洲地区以及拉丁美洲和加勒比海地区，几乎所有的邮政地址可操作性错误的编码都为 P1（很可能不可送达）。在排除“未知”地区的情况下，亚太地区和欧洲地区 N1 “国家/地区未知”错误的百分比最高。

表 23: 按地区划分的邮政地址可操作性错误原因 — 2009 RAA 要求

错误	非洲地区	亚太地区	欧洲地区	拉丁美洲 和加勒比海 地区	北美地区	未知	所有地区
P1 不可操作 ³⁷	81.5%	29.9%	51.9%	79.4%	32.7%	0.0%	40.9%
P2 不可操作	12.1%	28.5%	14.8%	14.3%	44.9%	0.0%	20.5%
N1 国家/ 地区未知	6.4%	40.9%	33.3%	6.3%	22.4%	100.0%	38.2%
N2 不可验证	0.0%	0.7%	0.0%	0.0%	0.0%	0.0%	0.3%
地区的整体准确 度 — 邮政地址 可操作性	95.8% ± 1.1%	95.9% ± 0.7%	99.1% ± 0.4%	99.0% ± 0.4%	99.4% ± 0.3%	不适用	97.7% ± 0.3%

参阅表 18 中的注释了解如何理解此表格。

文字分析

我们按联系人类型和用于注册域名的文字语言在此报告记录的语法和可操作性准确度。所有地区的绝大多数域名都使用拉丁字母注册（参阅表 24），但一些域名使用文字注册，如使用汉字、阿拉伯文或附加符号。采用汉字文字的记录通常在亚太地区注册，而采用附加符号的记录通常在欧洲地区、拉丁美洲/加勒比海地区或亚太地区注册。在[附录 D](#)中可以找到包括与特定地区相关的文字分析的表格。该分析是报告中的新内容，我们尚不能提供周期之间的变化趋势。

³⁷ 参阅第 31 页了解邮政地址可操作性的错误原因解释。

表 24：按联系人类型和文字语言统计的整体语法和可操作性准确度

	文字语言	数量	语法准确度百分比	可操作性准确度百分比
注册人联系人	拉丁文	11,957	65.8% ± 0.9%	75.0% ± 0.8%
	阿拉伯文	2	0.0% ± 0.0%	97.1% ± 23.3%
	中文（汉字）	66	10.7% ± 7.5%	41.6% ± 11.9%
	附加符号	174	57.5% ± 7.3%	66.2% ± 7.0%
管理联系人	拉丁文	11,948	65.7% ± 0.9%	75.2% ± 0.8%
	阿拉伯文	2	0.0% ± 0.0%	97.1% ± 23.3%
	中文（汉字）	65	11.1% ± 7.6%	39.7% ± 11.9%
	附加符号	174	58.3% ± 7.3%	63.9% ± 7.1%
技术联系人	拉丁文	11,943	67.0% ± 0.8%	74.1% ± 0.8%
	阿拉伯文	2	0.0% ± 0.0%	97.1% ± 23.3%
	中文（汉字）	65	6.3% ± 5.9%	34.9% ± 11.6%
	附加符号	141	70.4% ± 7.5%	69.5% ± 7.6%

如我们在[主研究结果](#)中所见，三种联系人类型的准确率均类似。拉丁文文字的可操作性准确率与各个地区的整体可操作性准确度类似，可以预计的是，在各个地区中采用拉丁文注册的记录占据主导地位。与此相反，拉丁文文字的语法准确度高于所有三种联系人类型的整体语法准确率。从地区层面看，拉丁文文字的语法准确度与除亚太地区外的各个地区的整体准确度类似（参阅[附录 D](#)）。由于用其他文字语言注册的记录很少，因此没有必要把重点放在对未占主导地位的文字类型的准确度进行估计，或放在任何文字类型之间的差异上（即拉丁文和阿拉伯文的准确度之间的差异）。

不同周期之间的对比

可以针对第 1 周期和第 2 周期的研究结果进行语法和可操作性准确度的统计对比。我们将比较情况介绍于下，以供参考，并探索可对语法准确度与可操作性准确度之间的关系得出哪些一般性意见。

周期之间的准确度对比 — 2009 RAA 语法要求

在[主要研究结果](#)中，我们根据第 2 周期的 2009 RAA 要求介绍了记录的语法准确度。在这里，我们将第 2 周期语法准确度结果与第 1 周期的结果进行对比。

整体准确度的变化

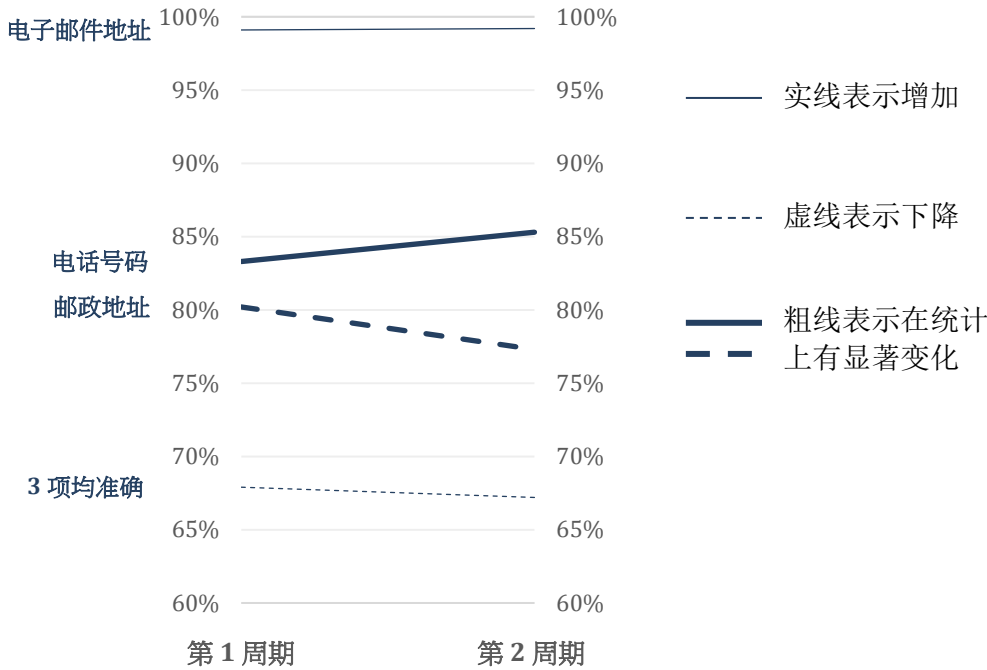
从表 25 和统计图 14 可看出，各阶段的电子邮件准确度非常类似，但电话号码准确度在第 1 周期中较低，邮政地址准确度在第 2 周期中较低。电话号码和邮政地址中展现出的差异最可能的原因是随机变化。³⁸所有模式均准确的比率在各个阶段之间都有所降低。

表 25：按周期划分的整体准确度 — 2009 RAA 语法要求

	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	99.1% ± 0.2%	83.3% ± 0.7%	80.2% ± 0.8%	67.9% ± 0.9%
第 2 周期	99.2% ± 0.2%	85.3% ± 0.6%	77.3% ± 0.7%	67.2% ± 0.8%
变化（第 2 周期 — 第 1 周期）	0.0% ± 0.2%	2.0% ± 1.0%	-3.0% ± 1.1%	-0.7% ± 1.2%

³⁸ 由于我们采用 95% 为置信区间，即使两阶段之间无差异，我们仍有 5% 的可能会呈现明显差异。

统计图 14：按周期划分的整体准确度 — 2009 RAA 语法要求



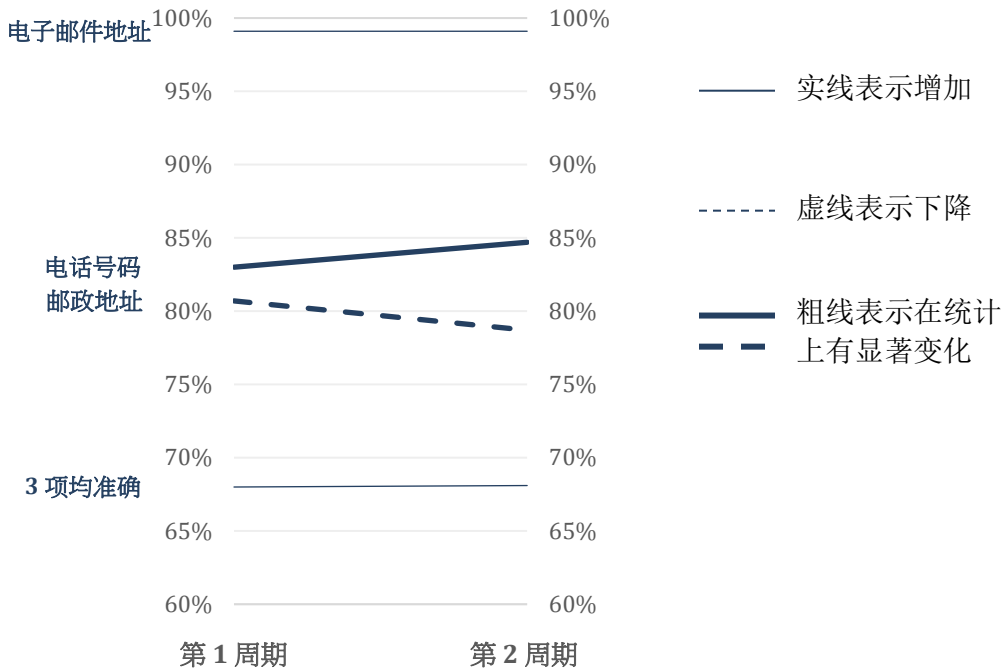
早期 gTLD 的变化

由于域名系统中的大多数域名均来自早期 gTLD，因此表 26 和统计图 15 中的早期 gTLD 的形式与以上表 25 中的整体准确率的形式类似。也就是说，早期 gTLD 的数据显示，第 1 周期电话号码和第 2 周期邮政地址的整体准确度都出现下降。电子邮件地址准确度无变化。

表 26：按周期划分的早期 gTLD 准确度 — 2009 RAA 语法要求

	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	99.1% ± 0.2%	83.0% ± 0.8%	80.7% ± 0.9%	68.0% ± 1.0%
第 2 周期	99.1% ± 0.2%	84.7% ± 0.7%	78.7% ± 0.8%	68.1% ± 0.9%
变化 (第 2 周期 - 第 1 周期)	0.0% ± 0.3%	1.6% ± 1.1%	-2.0% ± 1.2%	0.0% ± 1.4%

统计图 15：按周期划分的早期 gTLD 准确度 — 2009 RAA 要求



新 gTLD 的变化

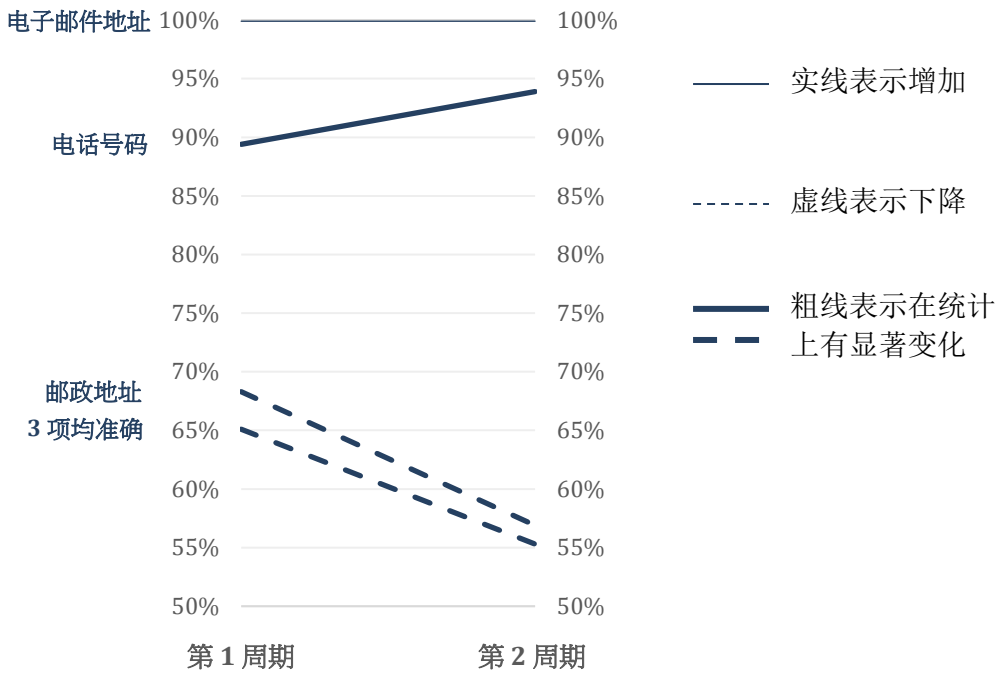
在表 27 和统计图 16 中，新 gTLD 显示与早期 gTLD 相同的形式。此外，周期之间进行对比可发现，电子邮件地址没有变化，而电话号码准确率上升，邮政地址准确率下降。在周期之间，通过对全部 9 个联系信息字段的所有准确度测试的新 gTLD 中域名的百分比有所下降。³⁹

表 27：按周期划分的新 gTLD 准确度 — 2009 RAA 语法要求

	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	99.9% ± 0.1%	89.4% ± 1.4%	68.3% ± 2.1%	65.1% ± 2.1%
第 2 周期	99.9% ± 0.1%	93.9% ± 1.0%	56.9% ± 2.0%	55.3% ± 2.0%
变化（第 2 周期 — 第 1 周期）	0.0% ± 0.2%	4.6% ± 1.7%	-11.4% ± 2.9%	-9.8% ± 3.0%

³⁹ 参阅附录 B 或第 1 周期报告，了解有关结果的更多信息（尤其是按地区划分的）。

统计图 16: 按周期划分的新 gTLD 准确度 — 2009 RAA 语法要求



周期之间的准确度对比 — 2009 RAA 可操作性要求

整体准确度的变化

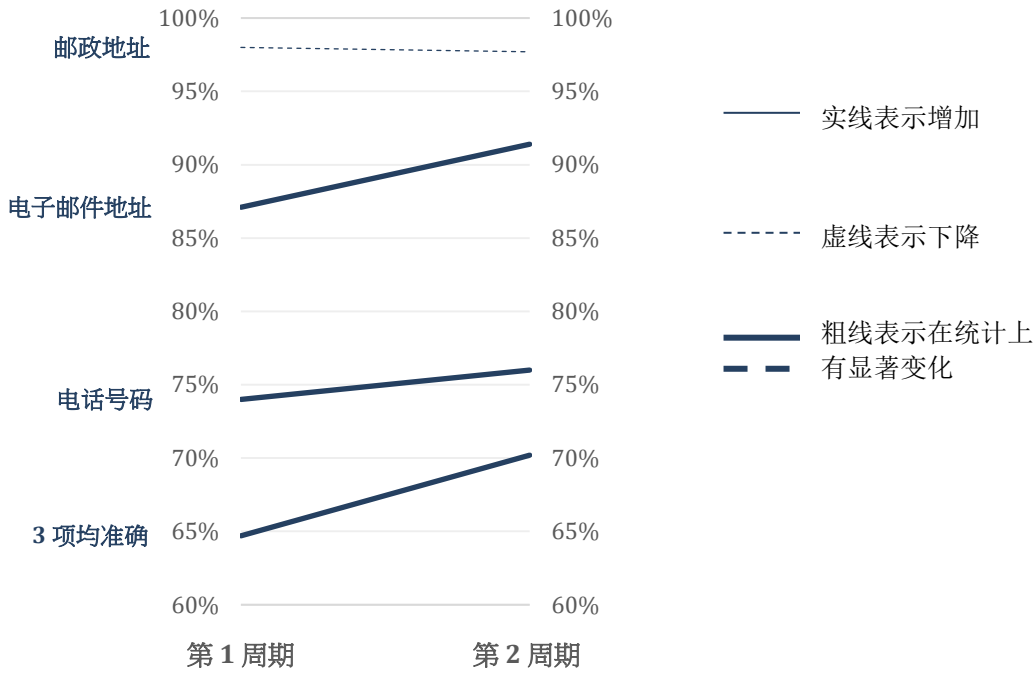
从表 28 和统计图 17 可看出，第 2 周期的电子邮件地址和电话号码整体准确率较高，但周期之间的邮政地址准确率极其相似。第 1 阶段与第 2 阶段之间的差异最可能的原因是随机变化。⁴⁰也有可能分布发生了变化，但是第 1 周期与第 2 周期之间出现此类变化相隔的时间并不久。分布变化即意味着第 1 周期未抽样的最新域名的电子邮件地址和电话号码准确度较高。所有模式均准确的记录比率在第 2 周期出现增长。

⁴⁰ 由于我们采用 95% 为置信区间，即使两阶段之间无差异，我们仍有 5% 的可能会呈现明显差异。

表 28：按周期划分的整体准确度 — 2009 RAA 可操作性要求

	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	87.1% ± 0.7%	74.0% ± 0.9%	98.0% ± 0.3%	64.7% ± 0.9%
第 2 周期	91.4% ± 0.5%	76.0% ± 0.8%	97.7% ± 0.3%	70.2% ± 0.8%
变化 (第 2 周期 — 第 1 周期)	4.3% ± 0.8%	2.1% ± 1.2%	-0.4% ± 0.4%	5.4% ± 1.2%

统计图 17：按周期划分的整体准确度 — 2009 RAA 可操作性要求

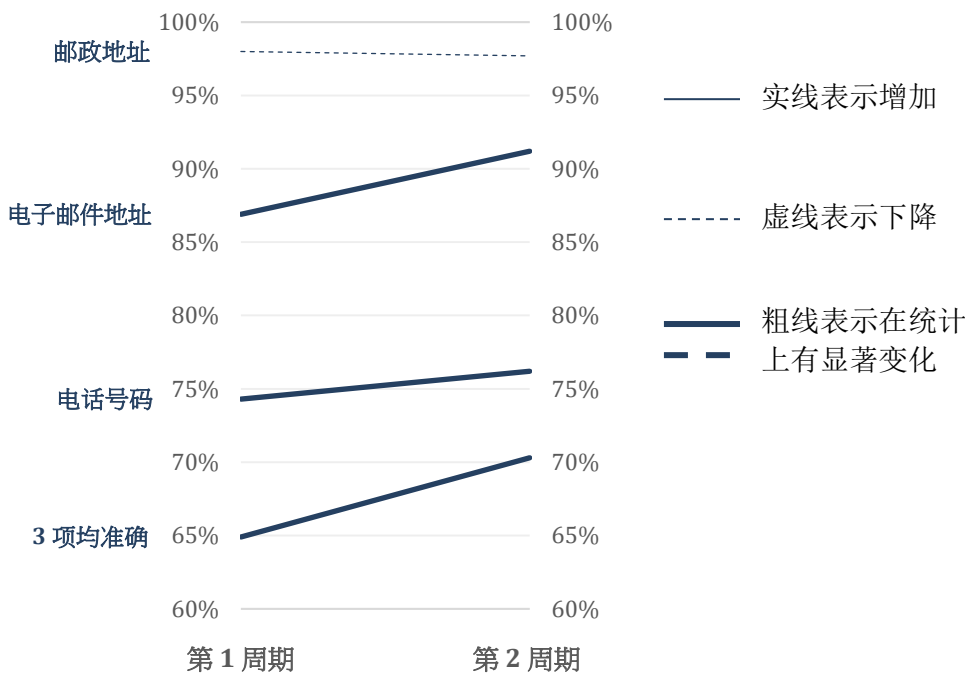


在以下关于早期 gTLD 的数据中，可以看到电子邮件地址、电话号码和所有模式准确度的整体准确度都出现了增长。周期之间的邮政地址准确度没有显示出变化。

表 29：按周期划分的早期 gTLD 准确度 — 2009 RAA 可操作性要求

	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	86.9% ± 0.7%	74.3% ± 1.0%	98.0% ± 0.3%	64.9% ± 1.0%
第 2 周期	91.2% ± 0.6%	76.2% ± 0.8%	97.7% ± 0.3%	70.3% ± 0.9%
变化 (第 2 周期 — 第 1 周期)	4.4% ± 0.9%	1.9% ± 1.3%	-0.3% ± 0.4%	5.4% ± 1.4%

统计图 18：按周期划分的早期 gTLD 准确度 — 2009 RAA 可操作性要求



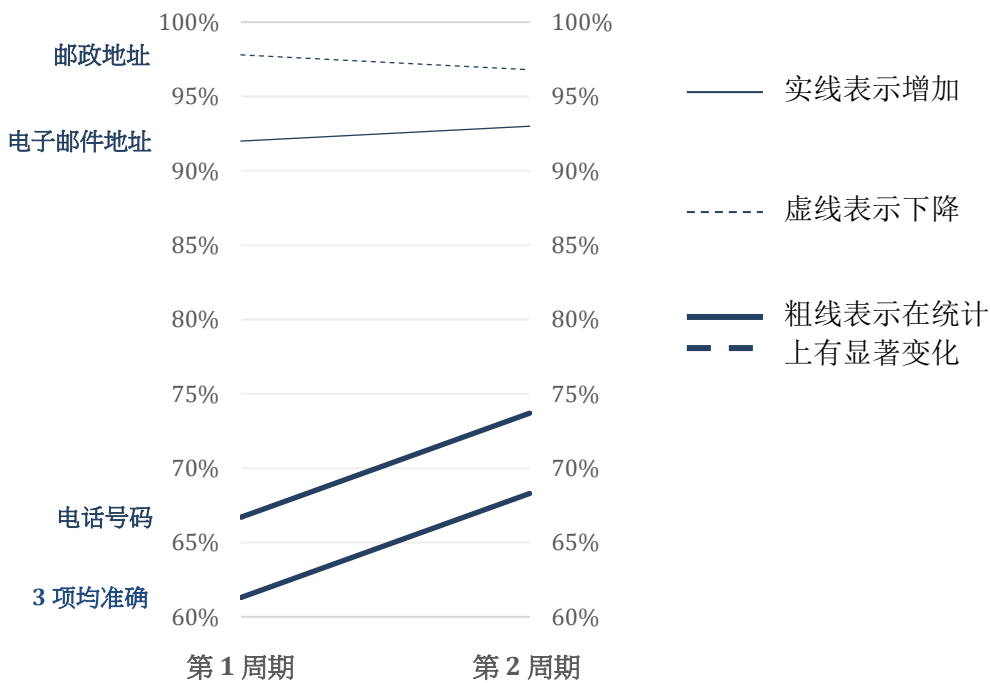
在表 30 和统计图 19 中，新 gTLD 显示出类似的形式。此外，电子邮件地址、电话号码和所有模式准确度在第 2 周期有所增长，而邮政地址准确度变化很小或没有变化。但在新 gTLD 中，电子邮件可操作性的增长并不显著。⁴¹

⁴¹ 参阅附录 B 或第 1 阶段报告，了解有关结果的更多信息（尤其是按地区划分的）。

表 30: 按周期划分的新 gTLD 准确度 — 2009 RAA 可操作性要求

	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	92.0% ± 1.2%	66.7% ± 2.1%	97.8% ± 0.7%	61.3% ± 2.2%
第 2 周期	93.0% ± 1.0%	73.7% ± 1.8%	96.8% ± 0.7%	68.3% ± 1.9%
变化 (第 2 周期 — 第 1 周期)	0.9% ± 1.6%	7.0% ± 2.8%	-1.0% ± 1.0%	6.9% ± 2.9%

统计图 19: 按周期划分的新 gTLD 准确度 — 2009 RAA 可操作性要求



语法和可操作性准确度之间的关系

我们在[主要研究结果](#)中基于 2009 RAA 标准介绍了全部 12,000 个域名的语法准确度和可操作性准确度。表 31 到 33 针对 2009 RAA 标准显示语法准确度和可操作性准确度之间的关系。它们调查了记录中通过两个测试方法中的一个后又通过另一个的比率（例如，通过可操作性测试后又通过语法测试的记录百分比，反之亦然）。

电子邮件地址语法和可操作性准确度

从表 31 可看出，未通过语法准确度测试的电子邮件地址也未通过可操作性准确度测试（即未通过语法准确度测试且通过可操作性准确度测试的百分比为 0）。这种形式是符合逻辑的，因为某些语法错误可表明电子邮件地址不可操作，例如，电子邮件地址缺失“@”符号（或任何其他关键组成部分）。未通过可操作性准确度测试的电子邮件地址（所有域名的 8.6%）的情况恰恰相反：这些电子邮件地址大多数实际上均通过了语法准确度测试（8.6% 的 7.8%）。这种形式也是符合逻辑的，因为特定的可操作性错误即使在语法正确时也可能出现，例如，因电子邮件地址不再使用而导致邮件被退回。从另一种角度来描述表 31 可看出，7.8% 的电子邮件地址通过了语法准确度测试，但不可操作。

表 31：电子邮件地址的语法准确度和可操作性准确度 — 2009 RAA 要求

语法	可操作性		
	通过	未通过	总计
通过	91.4% ± 0.5%	7.8% ± 0.5%	99.2% ± 0.2%
未通过	0.0% ± 0.0%	0.8% ± 0.2%	0.8% ± 0.2%
总计	91.4% ± 0.5%	8.6% ± 0.5%	100%

电话号码语法和可操作性准确度

表 32 显示 14.7% 的电话号码未能通过语法准确度测试，而 24.0% 未能通过可操作性准确度测试。但是这些小组并不完全重合。这与电子邮件地址不同，没有通过语法测试并不代表电话号码将不能通过可操作性测试，因为一些电话号码可能无法通过语法测试，但可以通过可操作性测试。⁴²在未

⁴² 例如，电话号码可能会缺失国家/地区代码，但如果国家/地区在 WHOIS 记录中可用，电话号码则为可操作。

通过语法准确度测试的电话号码中(合计为 14.7%), 过半也未通过可操作性测试(14.7% 的 9.0%, 或 61.2%)。同样, 在未通过可操作性测试的电话号码中(合计为 24.0%), 过半(24.0% 的 15.0%, 或 62.5%) 通过了语法准确度测试。

表 32: 电话号码的语法准确度和可操作性准确度 — 2009 RAA 要求

语法	可操作性			总计
		通过	未通过	
通过		70.3% ± 0.8%	15.0% ± 0.6%	85.3% ± 0.6%
未通过		5.8% ± 0.4%	9.0% ± 0.5%	14.7% ± 0.6%
总计		76.0% ± 0.8%	24.0% ± 0.8%	100%

邮政地址语法和可操作性准确度

从表 33 可看出，未通过可操作性准确度测试的邮政地址也未通过语法测试（即未通过可操作性准确度测试但通过语法准确度测试的百分比为 0）。但是几乎所有未通过语法准确度测试的邮政地址（所有域名的 22.7%）均通过了可操作性准确度测试（22.7% 的 20.4%）。⁴³从另一种角度来看待表 33 可看出，20.4% 的邮政地址未通过语法准确度测试，但通过了可操作性准确度测试。

表 33: 邮政地址语法准确度和可操作性准确度 — 2009 RAA 要求

语法	可操作性			总计
		通过	未通过	
通过		77.3% ± 0.8%	0.0% ± 0.0%	77.3% ± 0.7%
未通过		20.4% ± 0.7%	2.3% ± 0.3%	22.7% ± 0.7%
总计		97.7% ± 0.3%	2.3% ± 0.3%	100%

关于语法和可操作性准确度的要点

这些表格表明语法准确度与可操作性准确度不一样。一方面，语法准确度最高的是电子邮件地址，而最低的是邮政地址。另一方面，可操作性准确度最高的是邮政地址，而最低的是电话号码。对于电子邮件地址而言，语法准确度是可操作性准确度的必要条件，但不是充分条件。对于邮政地址而言，语法准确度不是可操作性准确度的必要条件。但是，语法准确度似乎是可操作性准确度的充分条件。对于电话号码而言，两者之间的关系更复杂，因为 20.8% 的电话号码（15.0% 加上 5.8%）不符合语法或可操作性要求，但不是同时不符合。因此我们可得出结论：语法准确度对电子邮件地址和邮政地址而言是可操作性指示因素，对电话号码而言却不是。也就是说，语法不准确的电子邮件地址（例如，没有“@”符号）将不可操作；而

⁴³ 参阅注释 34 以及[挑战和经验教训](#)中的讨论。

语法准确的邮政地址将可操作（即可送达），但是语法准确的电话号码的可操作性可能准确，也可能不准确。

挑战和经验教训

我们不断寻求方法来改善 ARS，并展望后续的 ARS 报告。本部分提供了在第 2 周期中造成挑战的问题的背景信息，以及如何在后续报告中避免或改善那些问题。

首先，跟进上次报告中提出的问题：

■ 增大样本数将继续提高统计显著性。

根据第 1 周期报告中的讨论，增加初始样本和子样本大小将为最小的小组带来帮助（例如，增加对来自非洲地区的域名进行采样的机率）。对于第 2 周期，我们将样本大小增加到 200,000，将子样本大小增加到 12,000。如[主要研究结果](#)部分中的讨论，这种变化可便于进行与较小小组相关的更强估计。⁴⁴由于 ICANN 已与 WHOIS ARS 供应商达成了一项共识，认为样本和子样本大小中的边际增长对资源要求的影响很小，因此我们再次考虑样本大小中哪种类型的额外增长对于 ARS 研究来说既可行又有益。

■ 分析全部三种联系人类型非常有助于获得数据全貌。

通过查阅各联系人类型之间的共性，我们发现第 1 周期和第 2 周期中逾 75% 的记录注册人联系人、管理联系人和技术联系人信息相同。准确度测试供应商已在测试中去掉了重复记录，以避免效率低下和重复工作，我们的确考虑过仅分析一种联系人类型（例如注册人联系人）是否效率更高。但是我们确定，更为有益的做法是在报告之间保持完全的一致性，以便于对比。同样，仅测试其中一种联系人类型，我们将无法尽可能准确地了解所有 WHOIS ARS 数据，因为有些数据误差可能会缺失。

其次，我们希望解决联系模式（电子邮件地址、电话号码、邮政地址）之间关系的问题，因为它与[可联系性](#)（即通过特定联系模式联系到联系人类型的可能性）相关。

■ 邮政地址可操作性很高，但也许并不是可联系性的最佳指标。

如之前报告中的讨论，邮政地址的语法和可操作性测试提出了独特的挑战（例如，语法上有误的邮政地址可能看上去在可操作性上有误，但事实上，邮政信件可能可以送达该地址）。我们与万国邮政联盟合作，设计了同时使用自动检查和手动检查的可操作性测试，以确保所有可能可送达的地址均标记为可送达。该测试方法在第 1 周期和第 2 周期中均产生了非常高的可操作性准确率。而当供应商的测试寻求模拟实际交付时，它们无法达到 100% 的准确度。在电子邮件地址和电话号码测试中不存在这种怀疑：我们通常不会采用邮寄来检

⁴⁴ 参阅[研究方案和方法](#)部分，了解关于此问题的更多信息。

查送达情况，但我们会通过拨打电话和发送电子邮件来检验其是否可操作。⁴⁵基于这一事实，我们告诫不要过分依赖邮政地址可操作性的准确率来反映可联系性；实际的数字可能会稍低。另外，我们认为试图联系注册人联系人、管理联系人或技术联系人的人不太可能首先尝试邮寄，他们更可能首先选择发送电子邮件或打电话。为此我们认为，记录可联系性的更好指标应采用电子邮件或电话号码的可操作性准确率来表示。如上面的讨论，电子邮件和电话号码的可操作性准确率也非常高，达到约 94%。

最后，我们想要讨论最后一个问题，因为这个问题与邮政地址的准确度测试相关。

■ 邮政地址测试规则继续改进。

如上文和以前的报告所指出，邮政地址测试因许多原因而引起挑战。另外一个原因是，国家/地区的语法准确度规则（即国家/地区的格式要求）可以有例外。在 ARS 项目的整个生命周期，我们收到注册服务机构的反馈称，某些被标记为格式错误或不准确的记录实际上在某个国家/地区是被允许的。在我们收到此类反馈时，我们将此信息报告给万国邮政联盟，然后由其进行备注以供未来测试。我们的目的是在拥有灵活邮政地址规则的国家/地区灵活判断邮政地址的准确度。我们也欢迎在未来提供此类反馈。

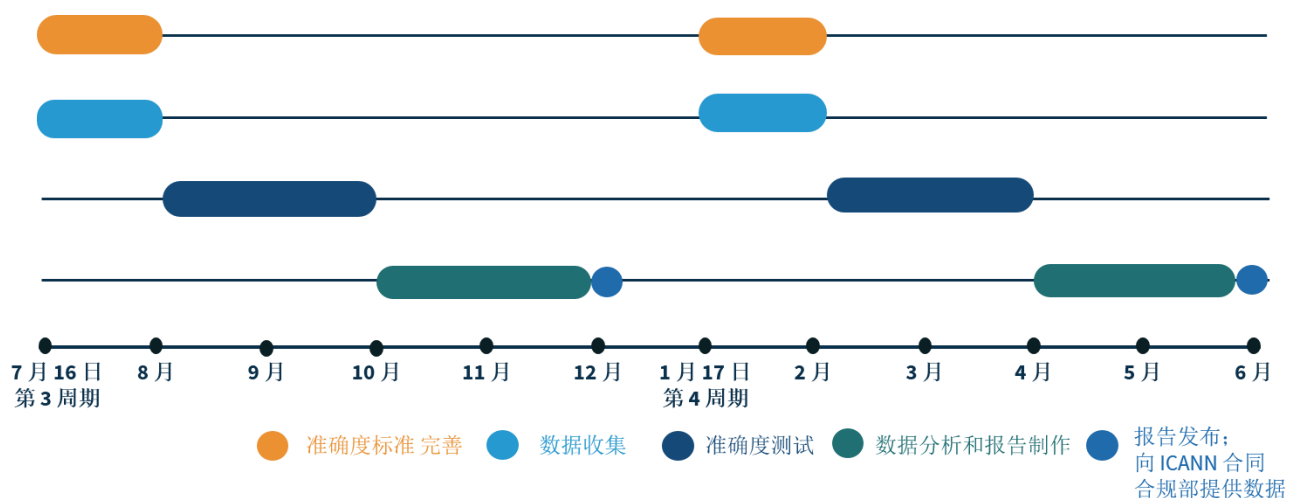
⁴⁵ 但这并不适用于身份验证，身份验证不是本报告的主题，也不由 WHOIS ARS 供应商开展。如之前所述，身份验证的计划尚未确定。

后续步骤

展望第 2 阶段第 3 周期

第 2 阶段第 3 周期和第 4 周期将继续进行 WHOIS 记录的语法和可操作性测试。第 3 周期计划于 2016 年 7 月开始。图 6 中的时间表显示第 2 阶段第 3 周期和第 4 周期直至报告发布的各个阶段。

图 6：第 2 阶段第 3 周期和第 4 周期



ICANN 合同合规部的后续步骤

如上所述，ARS 项目的主要目标之一是能够将可能不准确的问题提交给 ICANN 合同合规部，以便注册服务机构能够调查和跟进。第 2 阶段第 2 周期的研究结果是这些可能不准确的记录已提供给 ICANN 合同合规部，目前正在审核中。根据 WHOIS ARS 结果处理新投诉单的工作将交错进行，以最大程度地减少系统性能问题以及对注册服务机构的影响。WHOIS ARS 投诉单将与其他投诉一起处理，但 ICANN 将继续优先处理由社群成员提交的投诉。

由于第 2 阶段第 2 周期包含语法和可操作性研究结果，因此可能通过不同的流程进行合同合规性审核和跟进，这取决于各记录中发现的不准确问题类型。例如，仅出现格式错误但被判定为“可操作”的记录收到的跟进类型将不同于被判定为“不可操作”的记录。

所有 WHOIS ARS 投诉单都将根据下述问题类型按照合同合规方法和流程⁴⁶进行处理。如果可能，在咨询注册服务机构后，ICANN 也许能够在处理时合并多个 WHOIS ARS 投诉单。

语法错误跟进

如果错误不符合 2013 RAA 中规定的格式要求但信息却是有效且可联系的（例如，位于美国的注册人缺少 +1 国家/地区代码），那么该 WHOIS ARS 投诉将分类为 WHOIS 格式错误。如果错误让联系人变得不可联系（例如，缺少邮政地址），那么该 WHOIS ARS 投诉将作为 WHOIS 信息错误投诉进行处理。WHOIS 格式错误将不会转发给签署 2009 RAA 的注册服务机构。

可操作性错误跟进

可操作性错误生成的 WHOIS ARS 投诉将作为 WHOIS 信息错误投诉进行处理。尽管出现格式问题时可能无需联系注册域名持有者，但出现表明实质性不准确的可操作性错误时则要求注册服务机构采取合理措施调查和纠正（如适用）声称的不符合 2009 和 2013 RAA 的问题。此外，2013 RAA 的 WHOIS 准确度项目规范 (WAPS) 规定了其他要求，包括验证格式要求，以及在出现注册人没有及时响应 WHOIS 信息错误投诉时暂停域名。

合同要求

ICANN 合同合规部就 WHOIS ARS 投诉单向注册服务机构发送通知时，以下合同要求适用：

- 注册服务机构必须根据以下要求调查并更正不准确的 WHOIS 数据：
 - 2009 和 2013 RAA 的第 3.7.8 节以及
 - 2013 RAA 的 WHOIS 准确度项目规范。

⁴⁶ 参阅 ICANN 合同合规方法和流程：<https://www.icann.org/resources/pages/approach-processes-2012-02-25-en>。

- 根据 2013 RAA 运作的注册服务机构必须使用以下文件中要求的 WHOIS 格式和布局：
 - 注册数据目录服务 (WHOIS) 规范⁴⁷和
 - 公告：针对注册管理机构协议以及与适用的注册数据目录服务 (WHOIS) 规范相关的 2013 注册服务机构认证协议 (RAA) 的澄清。⁴⁸

此外，和任何 WHOIS 信息错误投诉或 WHOIS 格式投诉一样，未能在合同合规流程的非正式决议阶段进行回应或展示合规性将导致违规通知（将在 icann.org 上发布）。

第 2 阶段结果

合规部继续在合规部季报中提供 WHOIS ARS 的指标（请参阅 <https://www.icann.org/resources/pages/compliance-reports-2016-04-15-en>），并且将在为 2016 年第二季度生成指标时提供其他信息。此外，适当的情况下，指标将在 ICANN 公共会议上提供。

⁴⁷ 请参阅 <https://www.icann.org/resources/pages/approved-with-specs-2013-09-17-en#whois>。

⁴⁸ 请参阅 <https://www.icann.org/resources/pages/registry-agreement-raa-rdds-2015-04-27-en>。

附录 A：准确度测试标准

ICANN 一直试图使准确度测试标准与注册服务机构认证协议 (RAA) 的合同义务和适用互联网工程任务组 (IETF) 意见征询相一致。当前，gTLD 空间中使用的 RAA 主要有两个版本：2009 版和 2013 版。每个版本的 RAA 均对每个域名的注册人联系人、技术联系人和管理联系人联系信息的特定要素的存在、格式和可操作性提出了要求。每条记录（即域名）将按照创建该域名时的注册服务机构协议中的标准进行评估。ICANN 会将“过渡”记录纳入考虑，“过渡”记录是指在注册服务机构 2013 RAA 生效日期之前创建的记录。例如：

记录创建日期	2013 年 2 月 5 日
注册服务机构 2013 RAA 生效日期	2014 年 1 月 1 日
拟测试的验证标准	2009 RAA 要求

记录创建日期	2014 年 4 月 20 日
注册服务机构 2013 RAA 生效日期	2014 年 1 月 1 日
拟测试的验证标准	2013 RAA 要求

以下地址概述了电子邮件地址、电话号码和邮政地址的语法和可操作性准确度测试标准：<https://whois.icann.org/en/whoisars-validation>。该地址所列标准将由支持 WHOIS ARS 项目的验证供应商使用。

附录 B：其他分析 — 满足 2009 RAA 要求的准确度

联系数据的共性

从表 B1 可看出，当三种联系人类型中的两种信息相同（另一种不同）时，最有可能的是注册人联系人和管理联系人的信息相同，最不可能的是注册人联系人和技术联系人的信息相同。

表 B1：各联系人类型和联系模式的联系信息相同的频率

共性	电子邮件地址	电话号码	邮政地址
3 项均相同	77.6% ± 0.7%	80.3% ± 0.7%	78.2% ± 0.7%
注册人联系人 = 管理联系人	14.0% ± 0.6%	14.0% ± 0.6%	13.2% ± 0.6%
注册人联系人 = 技术联系人	0.4% ± 0.1%	0.3% ± 0.1%	0.4% ± 0.1%
管理联系人 = 技术联系人	5.7% ± 0.4%	4.3% ± 0.4%	6.2% ± 0.4%
3 项均不同	2.3% ± 0.3%	1.0% ± 0.2%	2.0% ± 0.3%

第 1 周期和第 2 周期 2009 RAA 语法错误的原因

[主要研究结果](#) 部分包含第 2 周期的结果，但我们按顺序展示了来自所有三项早期 ARS 研究的结果（第 1 阶段，第 1 周期和第 2 周期）。

表 B2：电子邮件地址错误总数（按联系人类型划分，2009 RAA）— 第 1 阶段

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	10,000	9,950	9,954	29,904
缺失*	[38]*	50	46	96
总计	10,000	10,000	10,000	30,000

* 2009 RAA 不要求提供注册人联系人电子邮件地址。

表 B3: 电子邮件地址错误总数（按联系人类型划分，2009 RAA）— 第 1 周期

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	9,997	9,945	9,933	29,875
缺失*	[124]*	51	62	113
@ 符号缺失	2	2	3	7
不可解析	1	2	2	5
总计	10,000	10,000	10,000	30,000

* 2009 RAA 不要求提供注册人联系人电子邮件地址。

表 B4: 电子邮件地址错误总数（按联系人类型划分，2009 RAA）— 第 2 周期

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	11,994	11,947	11,945	35,886
缺失*	128*	48	51	99
@ 符号缺失	2	1	0	3
不可解析	4	4	4	12
总计	12,000	12,000	12,000	36,000

* 2009 RAA 不要求提供注册人联系人电子邮件地址。

表 B5: 电话号码错误总数（按联系人类型划分，2009 RAA）— 第 1 阶段

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	8,780	8,645	8,719	26,144
不存在*	[234]*	144	148	292
缺少国家/地区代码	304	289	279	872
长度不正确	883	889	821	2,593
不允许的字符	33	33	33	97
总计	10,000	10,000	10,000	30,000

* 2009 RAA 不要求提供注册人联系人电话号码。

表 B6: 电话号码错误总数（按联系人类型划分，2009 RAA）— 第 1 周期

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	8,569	8,409	8,523	25,501
缺失*	[199]*	137	144	281
缺少国家/地区代码	474	499	481	1,454
长度不正确	955	952	849	2,756
不允许的字符	2	3	3	8
总计	10,000	10,000	10,000	30,000

* 2009 RAA 不要求提供注册人联系人电话号码。

表 B7: 电话号码错误总数（按联系人类型划分，2009 RAA）— 第 2 周期

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	10,398	10,224	10,316	30,938
缺失*	182*	107	113	220
缺少国家/地区代码	538	577	584	1,699
长度不正确	1,062	1,090	986	3,138
不允许的字符	2	2	1	5
总计	12,000	12,000	12,000	36,000

* 2009 RAA 不要求提供注册人联系人电话号码。

表 B8: 邮政地址错误总数 (按联系人类型划分, 2009 RAA) — 第 1 阶段

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	7,582	7,570	7,826	22,978
缺失	42	50	56	148
缺少国家/地区	18	22	22	62
国家/地区不可识别	24	26	27	77
缺少邮政编码	691	736	665	2,092
邮政编码格式错误	25	24	20	69
缺少州/省	1,126	1,134	995	3,255
缺少城市	836	858	777	2,471
缺少街道	564	557	494	1,615
总计	10,000	10,000	10,000	30,000
错误总数	3,326	3,407	3,056	9,789
有错误的域名总数	2,418	2,430	2,174	7,022

表 B9: 邮政地址错误总数 (按联系人类型划分, 2009 RAA) — 第 1 周期

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	7,150	7,151	7,511	21,812
缺失	41	54	63	158
缺少国家/地区代码	59	53	52	164
国家/地区不可识别	23	27	30	80
缺少邮政编码	154	144	128	426
邮政编码格式错误	853	901	768	2,522
缺少州/省	720	709	607	2,036
缺少城市	1,125	1,126	1,010	3,261
缺少街道	731	723	637	2,091
总计	10,000	10,000	10,000	30,000
错误总数	3,706	3,737	3,295	10,738
有错误的域名总数	2,850	2,849	2,489	8,188

表 B10: 邮政地址错误总数（按联系人类型划分，2009 RAA）— 第 2 周期

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	8,407	8,383	8,815	25,605
缺失	43	52	57	152
缺少国家/地区代码	71	58	53	182
国家/地区不可识别	65	70	64	199
缺少邮政编码	953	1,039	920	2,912
邮政编码格式错误	23	21	20	64
缺少州/省	1,676	1,699	1,463	4,838
缺少城市	1,398	1,411	1,235	4,044
缺少街道	786	764	662	2,212
总计	12,000	12,000	12,000	36,000
错误总数	5,015	5,114	4,474	14,603
有错误的域名总数	3,593	3,617	3,185	10,395

不同阶段语法准确度的其他比较（按地区和 RAA 小组划分）

表 B11: 非洲地区域名准确度（按周期划分）— 2009 RAA 语法要求

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	99.9% ± 0.2%	66.8% ± 2.9%	42.2% ± 3.1%	29.8% ± 2.9%
第 2 周期	99.9% ± 0.2%	64.6% ± 2.6%	44.6% ± 2.7%	29.3% ± 2.5%
变化（第 2 周期 — 第 1 周期）	0.0% ± 0.3%	-2.1% ± 4.0%	2.4% ± 4.1%	-0.5% ± 3.8%

表 B12: 亚太地区域名准确度（按周期划分）— 2009 RAA 语法要求

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	99.5% ± 0.3%	78.7% ± 1.7%	52.5% ± 2.0%	39.5% ± 2.0%
第 2 周期	99.4% ± 0.3%	88.9% ± 1.1%	49.8% ± 1.8%	45.0% ± 1.8%
变化（第 2 周期 — 第 1 周期）	-0.1% ± 0.4%	10.2% ± 2.0%	-2.7% ± 2.7%	5.5% ± 2.7%

表 B13: 欧洲地区域名准确度 (按周期划分) — 2009 RAA 语法要求

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	99.8% ± 0.2%	85.2% ± 1.5%	72.2% ± 1.9%	62.3% ± 2.1%
第 2 周期	99.9% ± 0.1%	85.1% ± 1.4%	68.4% ± 1.8%	60.6% ± 1.9%
变化 (第 2 周期 — 第 1 周期)	0.1% ± 0.2%	-0.0% ± 2.0%	-3.8% ± 2.6%	-1.7% ± 2.8%

表 B14: 拉丁美洲/加勒比海地区域名准确度 (按周期划分) — 2009 RAA 语法要求

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	99.9% ± 0.1%	79.2% ± 1.9%	67.1% ± 2.1%	56.9% ± 2.3%
第 2 周期	99.9% ± 0.1%	84.3% ± 1.6%	71.0% ± 2.0%	64.7% ± 2.1%
变化 (第 2 周期 — 第 1 周期)	0.0% ± 0.2%	5.1% ± 2.5%	3.8% ± 3.0%	7.8% ± 3.1%

表 B15: 北美地区域名准确度 (按周期划分) — 2009 RAA 语法要求

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	99.9% ± 0.1%	86.1% ± 1.3%	97.1% ± 0.6%	83.9% ± 1.4%
第 2 周期	100.0% ± 0.0%	85.1% ± 1.2%	96.7% ± 0.6%	82.8% ± 1.3%
变化 (第 2 周期 — 第 1 周期)	0.1% ± 0.1%	-1.1% ± 1.8%	-0.4% ± 0.9%	-1.1% ± 1.9%

表 B16: 2009 RAA 域名准确度 (按周期划分) — 2009 RAA 语法要求

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	98.1% ± 0.6%	89.1% ± 1.3%	82.0% ± 1.6%	77.1% ± 1.7%
第 2 周期	99.3% ± 0.3%	90.8% ± 1.2%	85.2% ± 1.5%	80.9% ± 1.6%
变化 (第 2 周期 — 第 1 周期)	1.2% ± 0.7%	1.7% ± 1.7%	3.2% ± 2.2%	3.8% ± 2.4%

表 B17: 2013 RAA GF 域名准确度 (按周期划分) — 2009 RAA 语法要求

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	99.4% ± 0.2%	80.2% ± 1.2%	82.7% ± 1.2%	67.3% ± 1.5%
第 2 周期	99.4% ± 0.2%	80.0% ± 1.1%	81.9% ± 1.1%	66.8% ± 1.4%
变化 (第 2 周期 — 第 1 周期)	-0.0% ± 0.3%	-0.2% ± 1.7%	-0.8% ± 1.6%	-0.6% ± 2.0%

表 B18: 2013 RAA NGF 域名准确度 (按周期划分) — 2009 RAA 语法要求

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	98.8% ± 0.3%	88.6% ± 1.0%	75.3% ± 1.4%	68.2% ± 1.5%
第 2 周期	98.9% ± 0.3%	91.1% ± 0.8%	71.2% ± 1.2%	66.9% ± 1.3%
变化 (第 2 周期 — 第 1 周期)	0.1% ± 0.4%	2.4% ± 1.3%	-4.1% ± 1.9%	-1.3% ± 2.0%

附录 C：其他分析 — 满足 2013 RAA 要求的准确度

在 2013 RAA 中注册的域名目前占有所有域名的近 56%。在本附录中，我们来看看基于 2013 RAA 要求的准确率。如之前在本报告中所述，我们选择 2009 RAA 作为基准对全部 12,000 条经过分析的子样本记录进行分析。2013 RAA 的要求比 2009 RAA 的要求更严格，它基于 2009 RAA 的要求构建，因此也涵盖 2009 RAA 的要求。例如，2009 RAA 要求每个联系人都有地址，而 2013 RAA 则要求每个联系人地址遵循特定国家/地区的适用万国邮政联盟 S42 模板的格式。任何满足 2013 RAA 要求的联系人字段也将满足 2009 RAA 的要求，基于此原因，2009 RAA 的要求可用作所有记录的对比基准。

统计图 C1：整体准确度 — 2013 RAA 语法要求

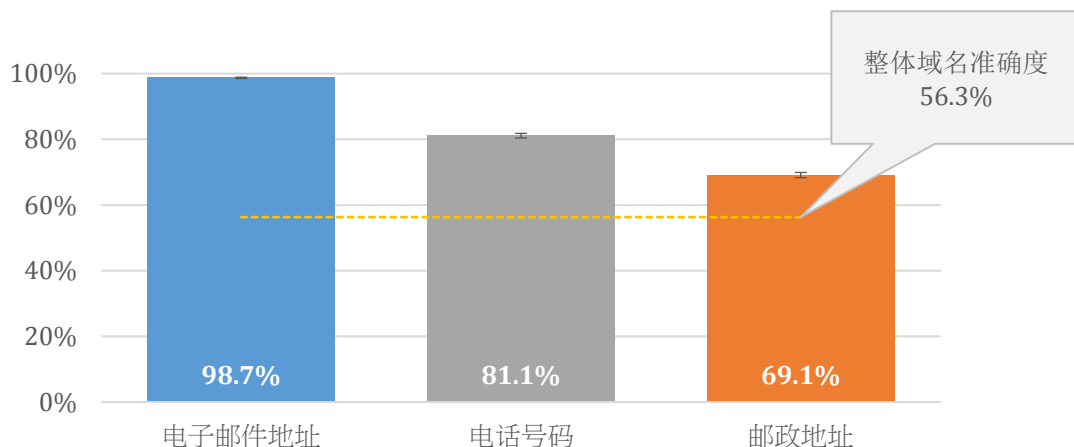


表 C1: 整体准确度 (按联系人类型和联系模式划分) — 2013 RAA 语法要求

	电子邮件地址	电话号码	邮政地址	3 项均准确
注册人联系人	98.8% ± 0.2%	83.3% ± 0.7%	71.0% ± 0.8%	58.9% ± 0.9%
管理联系人	99.2% ± 0.2%	82.9% ± 0.7%	70.8% ± 0.8%	59.0% ± 0.9%
技术联系人	99.2% ± 0.2%	82.7% ± 0.7%	71.1% ± 0.8%	59.3% ± 0.9%
总体	98.7% ± 0.2%	81.1% ± 0.7%	69.1% ± 0.8%	56.3% ± 0.9%

小组准确度 — 2013 RAA 语法要求

接下来, 我们来看第 2 周期的小组, 首先从早期 gTLD 和新 gTLD 的比较开始。由于注册人联系人、管理联系人和技术联系人的数据如此相似 (其信息相同的概率超过四分之三), 我们此处仅介绍注册人联系人、管理联系人和技术联系人均通过准确度测试的小组准确度。

小组 1: 早期 gTLD 与新 gTLD

统计图 C2: 按 gTLD 类型划分的准确度 — 2013 RAA 语法要求

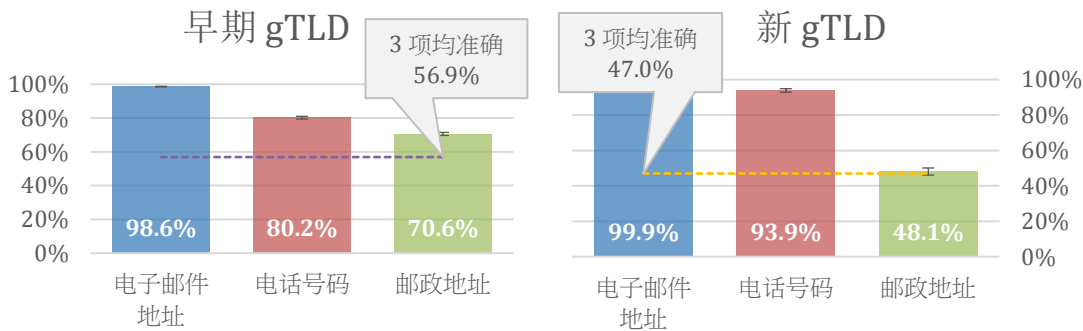


表 C2 显示新 gTLD 的电子邮件地址和电话号码语法准确度较高, 但邮政地址语法的准确度较低。

表 C2: 按 gTLD 类型划分的准确度 — 2013 RAA 语法要求

	电子邮件地址	电话号码	邮政地址	3 项均准确
早期 gTLD	98.6% ± 0.2%	80.2% ± 0.8%	70.6% ± 0.9%	56.9% ± 1.0%
新 gTLD	99.9% ± 0.1%	93.9% ± 1.0%	48.1% ± 2.0%	47.0% ± 2.0%
总体	98.7% ± 0.2%	81.1% ± 0.7%	69.1% ± 0.8%	56.3% ± 0.9%

小组 2: ICANN 地区

接下来, 我们来看按 ICANN 地区划分的准确度。我们再次介绍注册人联系人、管理联系人和技术联系人均通过准确度测试的小组准确度。

统计图 C3: 按 ICANN 地区划分的准确度 — 2013 RAA 语法要求

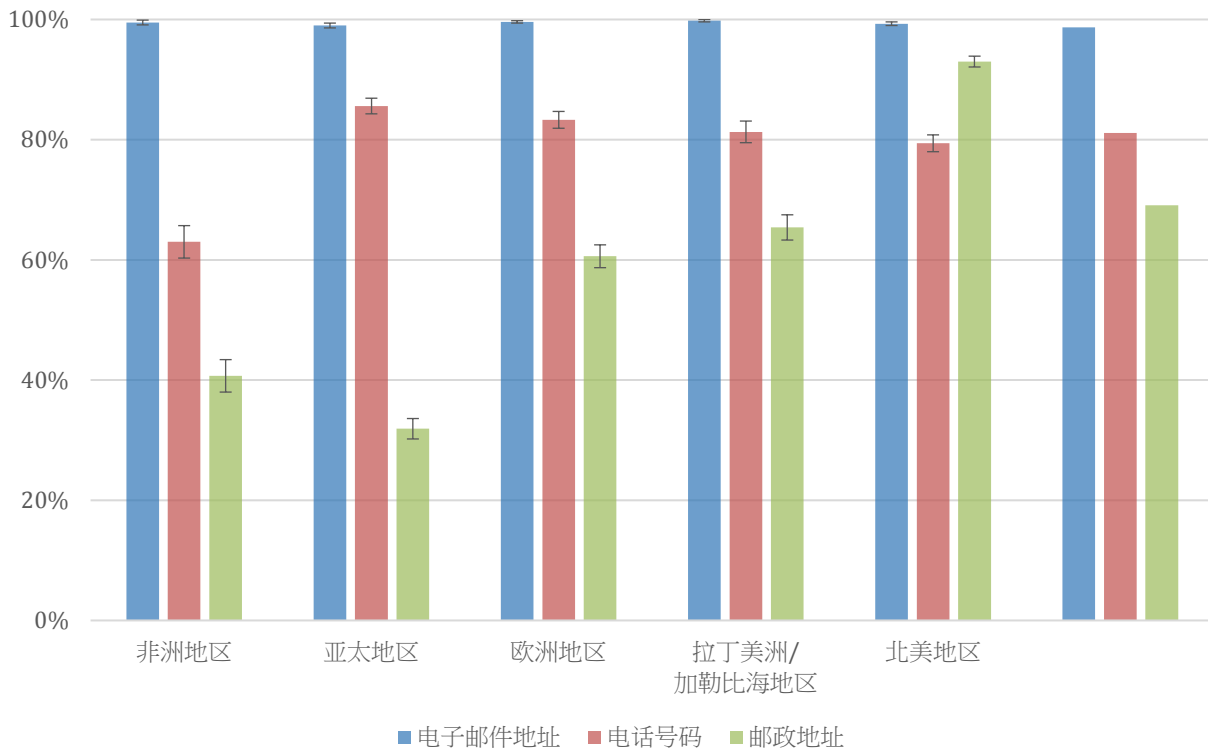


表 C3: 按 ICANN 地区划分的准确度 — 2013 RAA 语法要求

	电子邮件地址	电话号码	邮政地址	3 项均准确
非洲地区	99.5% ± 0.4%	63.0% ± 2.7%	40.7% ± 2.7%	26.3% ± 2.4%
亚太地区	99.0% ± 0.4%	85.6% ± 1.3%	31.9% ± 1.7%	27.3% ± 1.6%
欧洲地区	99.6% ± 0.2%	83.3% ± 1.4%	60.6% ± 1.9%	53.1% ± 1.9%
拉丁美洲/加勒比海地区	99.8% ± 0.2%	81.3% ± 1.8%	65.4% ± 2.1%	59.8% ± 2.2%
北美地区	99.3% ± 0.3%	79.4% ± 1.4%	93.0% ± 0.9%	73.2% ± 1.5%
总体	98.7% ± 0.2%	81.1% ± 0.7%	69.1% ± 0.8%	56.3% ± 0.9%

小组: RAA 状态

最后, 我们来看按 RAA 状态划分的准确度。仅 2013 RAA NGF 小组需要满足 2013 RAA 标准, 因此我们可预计该小组的准确度最高。

统计图 C4: 按 RAA 状态划分的准确度 — 2013 RAA 语法要求

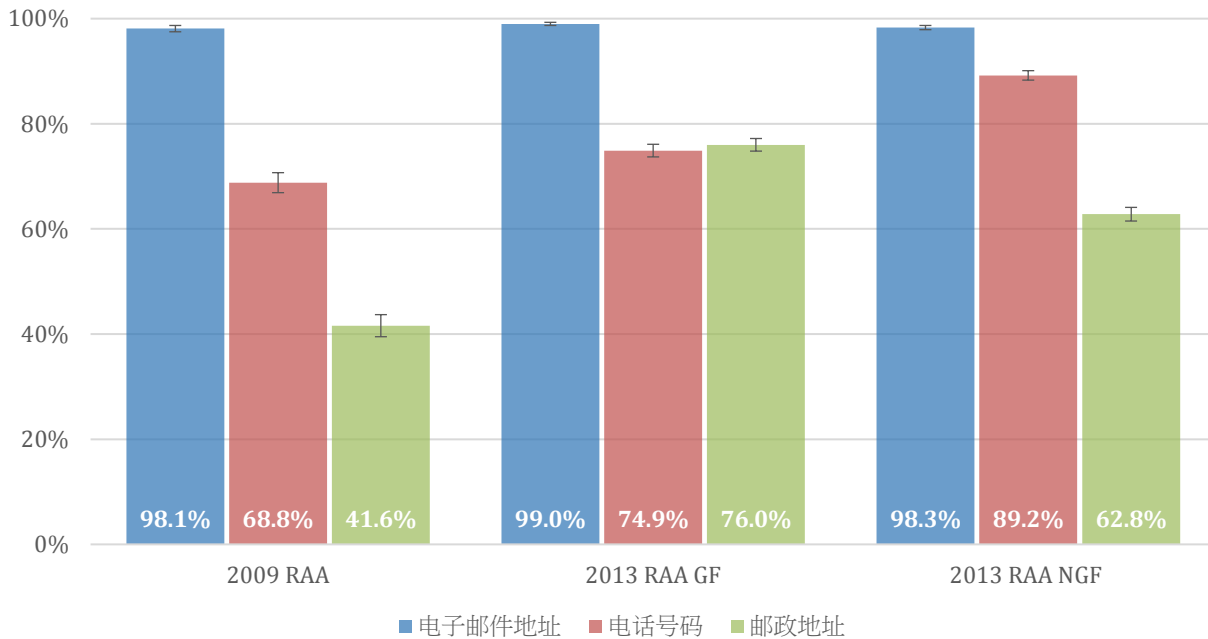


表 C4: 按 RAA 状态划分的准确度 — 2013 RAA 语法要求

	电子邮件地址	电话号码	邮政地址	3 项均准确
2009 RAA	98.1% ± 0.6%	68.8% ± 1.9%	41.6% ± 2.1%	20.2% ± 1.7%
2013 RAA GF	99.0% ± 0.3%	74.9% ± 1.2%	76.0% ± 1.2%	57.6% ± 1.4%
2013 RAA NGF	98.3% ± 0.4%	89.2% ± 0.9%	62.8% ± 1.3%	57.1% ± 1.4%
总体	98.7% ± 0.2%	81.1% ± 0.7%	69.1% ± 0.8%	56.3% ± 0.9%

阶段之间的对比 — 2013 RAA 语法要求

我们已在上文中介绍了第 2 周期符合 2013 RAA 要求的语法准确度。我们在这里针对 5,119 个需要符合这些要求的域名，将第 2 周期的结果与第 1 周期的结果进行对比。

整体准确度

表 C5: 按阶段划分的整体准确度 — 2013 RAA 语法要求

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	97.7% ± 0.5%	85.5% ± 1.1%	41.6% ± 1.6%	36.8% ± 1.5%
第 2 周期	98.3% ± 0.4%	89.2% ± 0.9%	62.8% ± 1.3%	57.1% ± 1.4%
变化 (第 2 周期 — 第 1 周期)	0.5% ± 0.6%	3.7% ± 1.4%	21.2% ± 2.1%	20.3% ± 2.1%

早期 gTLD 与新 gTLD

表 C6: 按阶段划分的早期 gTLD 准确度 — 2013 RAA 语法要求

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	97.5% ± 0.7%	85.0% ± 1.6%	39.8% ± 2.2%	34.6% ± 2.1%
第 2 周期	98.0% ± 0.5%	88.4% ± 1.2%	65.3% ± 1.7%	58.8% ± 1.8%
变化 (第 2 周期 — 第 1 周期)	0.5% ± 0.9%	3.4% ± 2.0%	25.5% ± 2.8%	24.1% ± 2.8%

表 C7: 按阶段划分的新 gTLD 准确度 — 2013 RAA 语法要求

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	99.6% ± 0.3%	89.4% ± 1.4%	56.0% ± 2.3%	53.8% ± 2.3%
第 2 周期	99.9% ± 0.1%	93.9% ± 1.0%	48.1% ± 2.1%	47.0% ± 2.1%
变化 (第 2 周期 — 第 1 周期)	0.3% ± 0.3%	4.5% ± 1.7%	-7.9% ± 3.1%	-6.8% ± 3.1%

ICANN 地区

表 C8: 按阶段划分的非洲地区域名准确度 — 2013 RAA 语法要求

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	100.0% ± 0.0%	70.3% ± 3.9%	32.8% ± 4.0%	24.4% ± 3.7%
第 2 周期	99.8% ± 0.3%	69.2% ± 3.3%	36.3% ± 3.4%	24.8% ± 3.1%
变化 (第 2 周期 — 第 1 周期)	-0.2% ± 0.3%	-1.1% ± 5.1%	3.4% ± 5.2%	0.4% ± 4.8%

表 C9: 亚太地区域名满足 2013 RAA 语法要求的准确度 (按阶段划分)

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	99.7% ± 0.4%	74.3% ± 3.0%	26.9% ± 3.1%	21.8% ± 2.9%
第 2 周期	99.5% ± 0.4%	90.8% ± 1.5%	27.8% ± 2.3%	25.5% ± 2.2%
变化 (第 2 周期 — 第 1 周期)	-0.2% ± 0.5%	16.5% ± 3.4%	1.0% ± 3.8%	3.7% ± 3.6%

表 C10: 欧洲地区域名满足 2013 RAA 语法要求的准确度 (按阶段划分)

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	99.6% ± 0.4%	88.7% ± 2.2%	55.7% ± 3.4%	49.2% ± 3.5%
第 2 周期	100.0% ± 0.0%	90.2% ± 1.8%	59.3% ± 3.0%	56.0% ± 3.1%
变化 (第 2 周期 - 第 1 周期)	0.4% ± 0.4%	1.5% ± 2.9%	3.5% ± 4.6%	6.8% ± 4.6%

表 C11: 拉丁美洲/加勒比海地区域名满足 2013 RAA 语法要求的准确度 (按阶段划分)

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	100.0% ± 0.0%	87.2% ± 2.3%	59.7% ± 3.4%	55.9% ± 3.4%
第 2 周期	100.0% ± 0.1%	91.3% ± 2.0%	74.3% ± 3.0%	72.4% ± 3.1%
变化 (第 2 周期 - 第 1 周期)	-0.0% ± 0.1%	4.0% ± 3.0%	14.6% ± 4.6%	16.4% ± 4.6%

表 C12: 北美地区域名满足 2013 RAA 语法要求的准确度 (按阶段划分)

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	97.6% ± 1.1%	94.8% ± 1.5%	46.6% ± 3.5%	42.3% ± 3.4%
第 2 周期	98.5% ± 0.7%	89.7% ± 1.8%	95.2% ± 1.3%	84.9% ± 2.1%
变化 (第 2 周期 - 第 1 周期)	0.9% ± 1.3%	-5.1% ± 2.4%	48.6% ± 3.7%	42.7% ± 4.0%

RAA 状态

最后, 表 C13-C15 所示为第 1 周期至第 2 周期按联系模式和 RAA 小组划分的准确度变化。

表 C13: 2009 RAA 域名满足 2013 RAA 语法要求的准确度 (按阶段划分)

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	97.4% ± 0.6%	70.8% ± 1.8%	33.7% ± 1.9%	17.7% ± 1.6%
第 2 周期	98.1% ± 0.6%	68.8% ± 1.9%	41.6% ± 2.1%	20.2% ± 1.7%
变化 (第 2 周期 - 第 1 周期)	0.7% ± 0.9%	-2.0% ± 2.7%	7.9% ± 2.8%	2.4% ± 2.3%

表 C14: 2013 RAA GF 域名满足 2013 RAA 语法要求的准确度 (按阶段划分)

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	98.4% ± 0.4%	74.6% ± 1.4%	50.5% ± 1.6%	44.7% ± 1.6%
第 2 周期	99.0% ± 0.3%	74.9% ± 1.2%	76.0% ± 1.2%	57.6% ± 1.4%
变化 (第 2 周期 - 第 1 周期)	0.6% ± 0.5%	0.3% ± 1.8%	25.6% ± 2.0%	13.0% ± 2.1%

表 C15: 2013 RAA NGF 域名满足 2013 RAA 语法要求的准确度 (按阶段划分)

周期	电子邮件地址	电话号码	邮政地址	所有模式均准确
第 1 周期	97.7% ± 0.5%	85.5% ± 1.1%	41.6% ± 1.6%	36.8% ± 1.5%
第 2 周期	98.3% ± 0.4%	89.2% ± 0.9%	62.8% ± 1.3%	57.1% ± 1.4%
变化 (第 2 周期 - 第 1 周期)	0.5% ± 0.6%	3.7% ± 1.4%	21.2% ± 2.1%	20.3% ± 2.1%

2013 RAA 语法错误的原因

在所有早期 WHOIS ARS 研究中, 我们介绍了各联系人类型未通过的准确度测试。下面我们再次呈现第 1 周期的那些表格, 并提供第 2 周期的同等数据。

电子邮件地址

如表 C16 所示, 第 1 阶段中, 除了缺少所要求的电子邮件地址外, 电子邮件地址中未发现其他错误。(2013 RAA 不要求提供注册人联系人电子邮件地址。)

表 C16: 电子邮件地址错误总数 (按联系人类型划分, 2013 RAA) - 第 1 阶段

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	3,802	3,829	3,830	11,461
缺失	46	19	18	83
总计	3,848	3,848	3,848	11,544

表 C17: 电子邮件地址错误总数（按联系人类型划分，2013 RAA）— 第 1 周期

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	3,713	3,725	3,726	11,164
缺失	27	15	14	56
@ 符号缺失	1	1	1	3
不可解析	1	1	1	3
总计	3,742	3,742	3,742	11,226

表 C18: 电子邮件地址错误总数（按联系人类型划分，2013 RAA）— 第 2 周期

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	5,083	5,095	5,095	15,273
缺失	35	23	23	81
不可解析	1	1	1	3
总计	5,119	5,119	5,119	15,357

电话号码

表 C18: 电话号码错误总数（按联系人类型划分，2013 RAA）— 第 1 阶段

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	3,336	3,362	3,389	10,087
缺失	121	102	103	326
缺少国家/地区代码	76	70	65	211
<i>国家/地区代码格式错误</i>	91	90	91	272
长度不正确	223	223	199	645
不允许的字符	1	1	1	3
总计	3,848	3,848	3,848	11,544

注：斜体表示新的 2013 RAA 要求。

表 C19: 电话号码错误总数（按联系人类型划分，2013 RAA）— 第 1 周期

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	3,251	3,267	3,298	9,816
缺失	82	83	83	248
缺少国家/地区代码	82	79	82	243
<i>国家/地区代码格式错误</i>	45	44	42	131
长度不正确	282	269	237	788
不允许的字符	0	0	0	0
总计	3,742	3,742	3,742	11,226

注：斜体表示新的 2013 RAA 要求。

表 C19: 电话号码错误总数（按联系人类型划分，2013 RAA）— 第 2 周期

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	4,606	4,590	4,586	13,782
不存在	60	62	63	185
缺少国家/地区代码	87	86	100	273
<i>国家/地区代码格式错误</i>	53	56	67	176
长度不正确	313	325	303	941
不允许的字符	0	0	0	0
总计	5,119	5,119	5,119	15,357

注：斜体表示新的 2013 RAA 要求。

邮政地址

表 C20: 邮政地址错误总数（按联系人类型划分，2013 RAA）— 第 1 阶段

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	2,213	2,209	2,258	6,680
缺失	16	19	19	54
缺少国家/地区代码	2	3	3	8
国家/地区不可识别	10	10	10	30
<i>国家/地区位于错误字段中</i>	811	812	812	2,435
<i>国家/地区未使用 ISO Alpha 2 格式</i>	0	0	0	0
缺少邮政编码	233	255	250	738
邮政编码格式错误	8	8	8	24
缺少州/省	456	472	440	1,368
<i>州/省位于错误字段中</i>	37	37	34	108
<i>州/省格式错误</i>	47	46	45	138
缺少城市	278	283	255	816
<i>城市位于错误字段中</i>	165	165	186	516
缺少街道	249	242	224	715
<i>街道位于错误字段中</i>	30	31	28	89
总计	3,848	3,848	3,848	11,544
错误总数	2,342	2,383	2,314	7,039
有错误的域名总数	1,635	1,639	1,590	4,864

注：斜体表示新的 2013 RAA 要求。

表 C21: 邮政地址错误总数 (按联系人类型划分, 2013 RAA) — 第 1 周期

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	2,010	2,011	2,051	6,072
缺失	15	17	17	49
缺少国家/地区代码	1	2	2	5
国家/地区不可识别	13	13	14	40
<i>国家/地区位于错误字段中</i>	865	865	864	2,594
缺少邮政编码	270	283	270	823
邮政编码格式错误	7	8	8	23
<i>邮政编码位于错误字段中</i>	0	0	0	0
缺少州/省	459	468	409	1,336
<i>州/省位于错误字段中</i>	33	27	24	84
<i>州/省格式错误</i>	62	62	71	195
缺少城市	366	365	337	1,068
<i>城市位于错误字段中</i>	244	250	304	798
缺少街道	312	305	286	903
<i>街道位于错误字段中</i>	38	38	28	104
总计	3,742	3,742	3,742	11,226
错误总数	2,685	2,703	2,634	8,022
有错误的域名总数	1,732	1,731	1,691	5,154

注: 斜体表示新的 2013 RAA 要求。

表 C21: 邮政地址错误总数 (按联系人类型划分, 2013 RAA) — 第 1 周期

	注册人联系人	管理联系人	技术联系人	总计
通过所有准确度测试	3,041	3,037	3,093	9,171
缺失	23	25	25	73
缺少国家/地区代码	1	1	1	3
国家/地区不可识别	33	33	32	98
<i>国家/地区位于错误字段中</i>	61	61	61	183
国家/地区未使用 ISO Alpha 2 格式	0	0	0	0
缺少邮政编码	362	389	386	1,137
邮政编码格式错误	12	11	11	34
<i>邮政编码位于错误字段中</i>	0	0	0	0
缺少州/省	893	914	805	2,612
<i>州/省位于错误字段中</i>	45	44	42	131
<i>州/省格式错误</i>	96	98	116	310
缺少城市	568	582	538	1,688
<i>城市位于错误字段中</i>	389	386	483	1,258
缺少街道	425	415	386	1,226
<i>街道位于错误字段中</i>	109	108	95	312
总计	5,119	5,119	5,119	15,357
错误总数	3,017	3,067	2,981	9,065

注：斜体表示新的 2013 RAA 要求。

按小组分析：满足 2013 RAA 要求的准确度 — 可操作性

2013 RAA 对可操作性的唯一其他要求是必须提供注册人联系人电子邮件地址和电话号码。满足 2013 RAA 可操作性要求的准确度结果将极其重复，因此在本报告中未重复呈现。

附录 D：其他分析 — 按地区划分的文字和准确度

表 D1：按联系人类型和文字语言划分的非洲地区语法和可操作性准确度

	文字语言	数量	语法准确度百分比	可操作性准确度百分比
注册人联系人	拉丁文	1,256	34.4% ± 2.6%	68.4% ± 2.6%
	阿拉伯文	1	0.0% ± 0.0%	0.0% ± 0.0%
	中文（汉字）	0	NA ± NA	NA ± NA
	附加符号	14	42.9% ± 25.9%	71.4% ± 23.7%
管理联系人	拉丁文	1,255	35.4% ± 2.6%	68.9% ± 2.6%
	阿拉伯文	1	0.0% ± 0.0%	0.0% ± 0.0%
	中文（汉字）	0	NA ± NA	NA ± NA
	附加符号	13	46.2% ± 27.1%	69.2% ± 25.1%
技术联系人	拉丁文	1,254	40.7% ± 2.7%	70.7% ± 2.5%
	阿拉伯文	1	0.0% ± 0.0%	0.0% ± 0.0%
	中文（汉字）	0	NA ± NA	NA ± NA
	附加符号	14	42.9% ± 25.9%	71.4% ± 23.7%

表 D2: 按联系人类型和文字语言划分的亚太地区语法和可操作性准确度

	文字语言	数量	语法准确度百分比	可操作性准确度百分比
注册人联系人	拉丁文	2,891	47.1% ± 1.8%	63.3% ± 1.8%
	阿拉伯文	1	0.0% ± 0.0%	100.0% ± 0.0%
	中文(汉字)	66	15.3% ± 8.7%	73.9% ± 10.6%
	附加符号	7	19.5% ± 29.3%	61.1% ± 36.1%
管理联系人	拉丁文	2,882	45.9% ± 1.8%	62.8% ± 1.8%
	阿拉伯文	1	0.0% ± 0.0%	100.0% ± 0.0%
	中文(汉字)	65	15.8% ± 8.9%	39.7% ± 11.9%
	附加符号	7	19.5% ± 29.3%	61.1% ± 36.1%
技术联系人	拉丁文	2,882	51.8% ± 1.8%	60.3% ± 1.8%
	阿拉伯文	1	0.0% ± 0.0%	100.0% ± 0.0%
	中文(汉字)	65	12.7% ± 8.1%	34.9% ± 11.6%
	附加符号	5	20.0% ± 35.1%	60.0% ± 42.9%

表 D3: 按联系人类型和文字语言划分的欧洲地区语法和可操作性准确度

	文字语言	数量	语法准确度百分比	可操作性准确度百分比
注册人联系人	拉丁文	2,619	65.1% ± 1.8%	74.1% ± 1.7%
	阿拉伯文	0	NA ± NA	NA ± NA
	中文 (汉字)	0	NA ± NA	NA ± NA
	附加符号	123	64.4% ± 8.5%	66.6% ± 8.3%
管理联系人	拉丁文	2,619	65.0% ± 1.8%	74.2% ± 1.7%
	阿拉伯文	0	NA ± NA	NA ± NA
	中文 (汉字)	0	NA ± NA	NA ± NA
	附加符号	125	65.0% ± 8.4%	64.2% ± 8.4%
技术联系人	拉丁文	2,619	71.4% ± 1.7%	71.2% ± 1.7%
	阿拉伯文	0	NA ± NA	NA ± NA
	中文 (汉字)	0	NA ± NA	NA ± NA
	附加符号	99	81.1% ± 7.7%	71.1% ± 8.9%

表 D4：按联系人类型和文字语言划分的拉丁美洲/加勒比海地区语法和可操作性准确度

	文字语言	数量	语法准确度百分比	可操作性准确度百分比
注册人联系人	拉丁文	1,905	66.2% ± 2.1%	73.1% ± 2.0%
	阿拉伯文	0	NA ± NA	NA ± NA
	中文（汉字）	0	NA ± NA	NA ± NA
	附加符号	30	19.7% ± 14.2%	63.7% ± 17.2%
管理联系人	拉丁文	1,906	66.8% ± 2.1%	73.6% ± 2.0%
	阿拉伯文	0	NA ± NA	NA ± NA
	中文（汉字）	0	NA ± NA	NA ± NA
	附加符号	29	20.3% ± 14.6%	61.4% ± 17.7%
技术联系人	拉丁文	1,906	68.5% ± 2.1%	74.4% ± 2.0%
	阿拉伯文	0	NA ± NA	NA ± NA
	中文（汉字）	0	NA ± NA	NA ± NA
	附加符号	23	26.2% ± 18.0%	59.8% ± 20.0%

表 D5: 按联系人类型和文字语言划分的北美地区语法和可操作性准确度

	文字语言	数量	语法准确度百分比	可操作性准确度百分比
注册人联系人	拉丁文	3,286	85.0% ± 1.2%	82.4% ± 1.3%
	阿拉伯文	0	NA ± NA	NA ± NA
	中文 (汉字)	0	NA ± NA	NA ± NA
	附加符号	0	NA ± NA	NA ± NA
管理联系人	拉丁文	3,286	84.4% ± 1.2%	82.1% ± 1.3%
	阿拉伯文	0	NA ± NA	NA ± NA
	中文 (汉字)	0	NA ± NA	NA ± NA
	附加符号	0	NA ± NA	NA ± NA
技术联系人	拉丁文	3,282	84.1% ± 1.3%	82.3% ± 1.3%
	阿拉伯文	0	NA ± NA	NA ± NA
	中文 (汉字)	0	NA ± NA	NA ± NA
	附加符号	0	NA ± NA	NA ± NA