

圣胡安 — 跨社群会议：ICANN 开放数据和信息透明度：未来发展途径
星期三，2018 年 3 月 14 日 — 大西洋标准时间 10:30 至 12:00
ICANN61 | 波多黎各圣胡安

迈克尔·卡拉尼古拉斯

(MICHAEL KARANICOLAS): 大家好。现在我们将开始关于 ICANN 开放数据和信息透明度的会议。我是迈克尔·卡拉尼古拉斯。我来自非商业用户选区 (NCUC)，我也是透明度工作阶段 2 的报告员。

透明度也是我日常工作的一部分，所以能在 ICANN 进行这次讨论对我来说非常棒。

本次会议将首先介绍问题、开放数据及其影响和好处、我们正在讨论什么内容，然后是 ICANN 开放数据倡议和信息透明度倡议的最新信息，之后是几个商业 — 来自 NPOC 和企业选区的一些观点，他们将对 ICANN 的进展以及社群应该寻求什么发表自己的观点。

感谢今天来参加会议的所有人。遗憾的是，这次会议的远程参与会受到限制，但往好的方面想，我猜 Adobe Connect 出问题了，可能意味着没人会用笔记本电脑。不管怎样...

感谢大家前来参加会议，我们可以 — 哦，我可以自己提出这一点。

我们今天在这里讨论透明度，透明度是问责制的命脉。它是关于我们如何执行监督。我们作为社群所做的一切 — 理想情况

注意：本文是一份由音频文件转录而成的 Word/文本文档。虽然转录内容大部分准确无误，但有时可能因无法听清段落内容和纠正语法错误而导致转录不完整或不准确。本文档旨在帮助理解原始音频文件，不应视为权威性的会议记录。

下，我们作为一个社群所做的一切应该是从知情的角度出发，并且从 ICANN 获得及时、准确、全面和可访问的信息是能否参与组织面临的问题的核心。

现在，当我们讨论透明度时，从根本上来说，这意味着从传统模式进行文化转型，在传统模式下，信息被默认分类，并且你看到的信息是有限的，是由组织选择的，而我们要转换到默认开放的模式，所有信息都可以访问。

我们并不是说完全透明。总会有信息需要保密 — 出于合理原因，例如个人隐私或有些信息披露之后会破坏保护互联网安全和稳定措施。

但是，从根本上说，我们正在寻找一个能够最大程度地确保可访问性的系统。我们应该广泛地理解可访问性和开放性，并且应尽可能限制信息披露的例外情况。

今天我们要讨论两个不同的披露途径，即通过开放数据倡议和信息透明度倡议。我们将详细了解这两个计划的内容。但是现在我想简单介绍一下这个主题，举几个例子说明开放数据如何有效，以及它对采纳透明度的机构的影响。

我首先要谈谈透明度的经济效益，因为最近进行了很多关于削减成本的讨论。这是一个自然而然的问题，开放数据对组织净收益可能产生何种影响。经验表明，除了从 ICANN 可提供的信息中获得的外部商业效益之外，至少以采购透明度的形式开放数据实际上可能具有节省组织资金的巨大潜力。

大家现在看到的是巴拉圭的开放合同网站的截图，这个网站是在 2010 年首次推出的。遗憾的是，它只是一个屏幕截图，所以我不能向大家展示各种很棒的功能和花里胡哨的东西。但是如果大家有兴趣进一步了解它，比如说要查找巴拉圭 DNCP，可以到这个网站上看看。这个网站具有很强的互动性。它可以让你跟踪谁花了多少钱，花了多少钱在不同的成本上。它甚至可以让您深入了解政府与不同采购实体签订的各项具体合同。

我再给你们看另一个网站，同样非常好。这个网站来自乌克兰。它被称为 ProZorro。同样，你看到的所有数值都是可点击的。你可以深入挖掘。你可以找到各项合同。您可以看到不同实体在不同时间段内所花的费用。

2015 年开发这个网站作为试点项目的成本，加上截至 2017 年底的运营成本，约为 469 万欧元。但是，由于开放数据 — 开放数据系统，在这两年期间，比较一下实际支付的成本与开始采购流程之前进行的成本估算，可以计算出采购成本节省了约 11.9 亿美元。乌克兰以 469 万美元建立和运营该网站。而在同一时期节省的成本为 11.9 亿美元。这些只是成本估算。

实际支出怎么样？最近，我开展了一项关于开放政府合作伙伴关系的研究，研究巴拉圭在采用其开放合同系统后的采购成本。国防和教育等方面的支出是可变的，所以我们研究了为相对不受波动影响的预算项目支付的价格。大家看看身后的图

表，在 2010 年到 2016 年之间，政府在引入开放合同系统后，单单办公用品一项就节省了 1.5 亿美元。

这些只是一些例子。采购和合同的开放是否以及如何成为这些倡议的一部分是不确定的。这是一个可能成为开放数据方法一部分的数据集，希望它能得到考虑。

在我讲完巴拉圭的例子之前，我还要提一下，透明度对揭示浪费或管理不善是非常有用的。在巴拉圭采纳透明度之后，记者和公民开始挖掘帐簿，发现一次教育部会议上花了 3.5 万的茶点费用。这引发了大规模的抗议活动。

大家可以看到，这次抗议活动被称为 **Cocido de Oro** 黄金和茶抗议，因为一壶绿茶就要 14 美元。由于这些抗议活动，最终导致了教育部长的辞职。

再者，这种监督和公共问责制不仅可以促使对管理不善的资源进行适当的问责，而且还可以促进 — 也可以促进负责任的支出文化。

在继续之前再举几个例子。

这是加拿大的开放议会网站。你可以搞清楚谁在什么情况下说了什么。信息的可访问性是在座的所有人面临的一大挑战。特别是，当我与新成员接触时，我经常听他们说跟上他们想要参与的问题的进度有多困难，这一点让他们有挫败感。这对于老成员来说也是个问题，如果你离开了几个星期，回来之后突然

发现讨论到了个新的阶段，很难弄明白发生了什么，很难获得关于正在进行的讨论的足够信息。所以开放的信息、开放的数据可能是传递这些信息的关键。

例如，通过该网站，您可以搜索任何术语，并了解它是如何在加拿大议会的辩论中被提出的。因此，例如，如果了解在宣布即将对加拿大的钢铁和铝材征收关税后，谁在讨论钢铁和铝材，那么你可以马上找到相关信息。因此，我认为，信息可访问性将成为我们即将听到的有关讨论的重要组成部分。

我将结合来自英国的其他几个例子。关于开放数据，很重要的一点是，一旦信息发布，就不可能预测社群将要采取的行动。所以英国建立了一个名为 **data.gov.uk** 的网站，基本上只是发布人们可以自由使用的数据集。

你们会发现它被用于各种创新方式。这是一幅伦敦地图，里面标出了交通事故的地点，以便查看哪些交叉路口最危险。你可以特别注意看看，如果自己骑自行车或者步行，最危险的地方在哪里，或者规划你的通勤路线，计划在哪里放慢速度。里面有很多非常有用的东西。

这些统计数据显示了各足球俱乐部的逮捕和禁令。所以你可以追踪哪些地方的流氓行为更糟，哪些粉丝做的事情比其他粉丝多。没有什么奇特的图表，但我只想告诉你，纽卡斯尔联仍然遥遥领先，但是切尔西也有很强的表现。

回到波多黎各，这是一个无障碍设施的地图，整合了来自公共数据库的信息，从中可以看出哪些设施可供不同的残障人士使用。你可以将其绘制到我们现在的地图上，并获得对于想要逛逛的人非常有用的所有信息。

我简要介绍了一下开放数据，现在我要将话筒交给杰·戴利 (Jay Daley)，他是为开放数据倡议工作的签约社群成员，长期以来一直积极致力于开放数据问题。他可以介绍一下开放数据倡议并讨论他所取得的进展。

杰伊·戴利：

大家好。我是杰伊·戴利。我们正在准备好我要用的另一组幻灯片。

所以我来补充一下。迈克尔对开放数据倡议做了一个很好的介绍 — 抱歉，他在总体上介绍了开放数据的好处。我不打算讲得太多，但我想解释一下什么是开放数据倡议，以及关于它的具体细节，特别是因为目前有两个倡议正在进行，另一个是信息透明度倡议，稍后会有相关演讲，我们要弄清楚这两个倡议之间的区别。

开放数据倡议是关于增强原始数据的可访问性。我们所说的“原始数据”通常指列表数据，也就是你可能在电子表格或数据库中找到的数据类型，这些数据通常是机读格式，数据分析师或市场营销人员或其他人可能希望将这些数据输入工具，然后在工具中使用。

其目的是让数据移动、让数据流动、使其得到重复利用，因为它被重复利用得越多，就能从中为整个社群和组织获得更多的价值。

这样做会在 ICANN 社群中产生信任。正如迈克尔所说，这是一些旨在提高透明度的项目的一个整体组成部分。

而 ICANN 已经确定的其中一个关键点是促成基于证据的决策制定。正如在座的一些人在之前对此进行的讨论中看到的一样，这其中存在多种误解，比如注册人更喜欢非常短的域名，但所有数据都显示，人们更偏爱中等长度的域名，也就是大约 10 个字符的域名。

所以要明确的是，开放数据倡议与信息透明度倡议不同。信息透明度倡议会晚一些，但不是重复工作。

那么流程是，马跃然 (Goran Marby) 在这方面有着非常强烈的个人兴趣，并且计划在 7 月 1 日之前交付一些工作成果。这是对开放数据的强有力的公开承诺。

我们现在也已开始招标，进行提案征询，以采购一个软件即服务开放数据平台。这是周一启动的。所以 — 这是一个专门的工具，旨在使重复利用和传播开放数据，特别是通过开放标准，以便任何人都可以通过任何开放工具访问它。

整个流程需要几个月的时间，市场中有一些成熟的供应商，还有许多其他人也正在尝试进入市场。所以这应该是一个竞争流程。

现在，在实现这一点之前，我们使用工具整合了一些开放数据试点项目，关于这一点，我会在最后一张幻灯片上详细解释。

这给那些熟悉这方面的社群成员提供了一个机会，使他们能够使用这些工具下载数据集并查看其工作原理。现在这项工作已经完成，并且已经将其中的反馈纳入了 RFP。

在内部，一直在对“数据资产库存”进行这项工作。这是 ICANN 内所有数据集以及及其相关各种属性的列表，例如它们是否包含个人身份信息、如何编辑、如何存储等等详细信息。现在已经 — 我们已经建立了框架，并且已经通过访谈流程和其他研究填补了这个框架。

其目的是提供公共数据目录，然后发布给社群，以便我们可以获得你们对于其中的数据集优先级和重要性的反馈。

发布流程比仅优先级排序更复杂，因为它还必须考虑获取特定数据集的困难程度。因此需要在两者之间进行权衡，但社群意见对此非常重要。

我们希望 7 月 1 日之前能够发布。而且，这些数据集将会进入开放数据门户网站。

最后，在内部，我们正在围绕这些数据集，围绕这些数据资产库存建立一个内部数据治理模型。

那么，最后，我想回顾一下开放数据倡议的试点项目，其中有四个：CKAN，这是许多政府使用的开源工具。它是一个非常强大的编目工具，因此许多政府在使用它。有三个商业提供商：OpenDataSoft、Enigma 和 Socrata。每个都有许多数据集。其中两个仍然可供你们查看。

我要讲的就是这些。现在把话筒交给迈克尔。谢谢。

迈克尔·卡拉尼古拉斯： 非常感谢你介绍的这些最新信息。

现在由 ICANN 工作人员亚娜·尤金诺维克 (Jana Juginovic) 来给我们介绍 ITI 的最新情况。

亚娜·尤金诺维克： 谢谢，迈克尔。我们今天想谈谈信息透明度倡议。

我知道，你们中的一些人可能已经在我们在 icann.org 上发布的一些博客中阅读了相关内容，我们会简单介绍一下自今年 1 月份我们启动项目以来所取得的进展，以及自 2017 年 9 月董事会批准项目以来的进展。

信息透明度倡议的核心是增强我们信息的问责制和透明度，icann.org 已做出这些承诺。你们之前已经提到，和 ODI 一样，ITI 也是跃然的目标之一。

以所有六种联合国语言提供最新的易于查找的信息是满足这些承诺的一个重要部分，也是我们展示透明度和问责制的方式之一。

我们要怎么做？我们需要做的第一件事就是制定一个连贯一致的内容策略。为此，我们将建立一个全面的分类系统和信息架构以及将要应用于我们的外部内容的内容治理。

要做到这一点，我们需要部署一套统一的文件管理系统。这是这个项目的核心。ICANN 将首次实施一套文件管理系统并建立该基础，由此我们将能够实施该分类系统和信息架构。

我们将通过一个新的单一、可扩展、安全的内容管理系统来展示这些内容。这就是我们的公共信息向社群展示的方式。

我们有九个目标，大家可以在 2017 年 9 月 ITI 获得批准时通过的董事会决议中查看这些目标。我之前谈到了其中的一些目标，即构建分类系统和信息架构以及文件管理系统和内容管理系统。但是其中一些目标还包括建立适应未来和内容不可知论的技术格局。这意味着如果我们有一个新的内容管理系统，新技术会发展，那些内容 — 我们不需要改变它，因为我们已经对它应用了分类系统，所以我们可以轻松地将它转移到一个新的内容管理系统。

通过我们正在构建的这个新基础架构，DMS 和 CMS，文件管理系统和内容管理系统，我们正在升级我们的技术基础架构，从而能够通过提高我们内容的可查找性和可访问性来更好地服务我们的全球社群。

我想谈一谈这将为 ICANN 社群和组织带来的好处。大家知道，我们收到很多来自社群的反馈，他们很难在 icann.org 上找到内容。因此，文件管理系统和 ITI 将要解决的第一件事是我们的公共信息。

我们希望为 ICANN 社群带来并展示的第一个好处是提高我们公共内容的可查找性、透明度、安全性和质量。

另外一个好处是 — 在这次 ICANN61 会议上，我们已经就预算进行了很多讨论 — 降低我们的内容的年度维护成本，因为它合并了我们目前管理的 15 个独立的站点。我们为网站提供内容的主要方式是员工自行发布，这样可以在不依赖网络管理团队的情况下降低成本。同样，它将确保我们的内容能够适应未来，因为我们将对其应用分类系统，因此如果我们需要为一个新系统提供内容，分类系统就在那里并且永久存在。

它为 ICANN 的 DMS（文件管理系统）和 CMS（内容管理系统）共享生态系统奠定了基础，一旦奠定了基础，我们便能够将其应用于我们所有的 SO/AC 网站。这将提高采用六种联合国语言的内容的可访问性，并有助于我们履行这些承诺并降低财务风险。

就我们迄今为止所做的工作而言，自 1 月份以来，我们开展了大量的活动。我们 — 我之前提到过，我们是从今年 1 月份开始的，而且我们已经开始标记所有内容。我们在 icann.org 上有很多内容，我之前也提到过，我们就是从这里开始着手的。我们有大约 104,000 条内容。很多内容。我们已经开始为 9,000 条内容做标记，目前已经完成，并且我们正在结合辨别审计来完成这项工作，因此 ICANN 工作人员直接将分类系统应用于该内容，并通过机器学习来实现，该机器学习能够识别某些语言模式，并对内容应用一致的分类系统。

我们已经建立了技术基础架构，DMS 和 CMS，并且完成了两个平台之间的初步集成，这已经完成了。

我们与 ICANN 社群进行了访谈，包括支持组织和咨询委员会的主席、区域领导、注册管理机构和注册服务机构运营商、NextGen 和学员、注册人和其他利益相关方。

我们将推出一个 feedback.icann.org 网站，这非常重要，因为我们需要社群在这里告诉我们他们希望 ITI 是什么样子、他们希望如何展示内容，以及某些功能将会如何推出，例如公众意见。我知道在座的有些人已经讨论过这个问题。因此，我们将在这个反馈网站和内容上启用这些功能，让您直接了解我们正在做的工作，并让您有机会直接提供该反馈，因为我们希望它对你们有帮助。

我们将 — 对于 DMS 和 CMS 之间的这种集成，实际上我们要测试的第一件事是用新的 ICANN 缩略语和术语来替换术语表。我知道有些人在识别 ICANN 缩略语和术语方面有困难，所以我们希望这个功能能够改善这一问题。

这就是我们目前的状况。这就是反馈网站启动后的样子，我们将在 icann.org 的博客上和社交媒体上分享这些信息，并让你们知道何时可以开始访问该反馈网站。

谢谢。

迈克尔·卡拉尼古拉斯： 太棒了。非常感谢。反馈非常重要。这是我们来到这里的最大原因之一。

那么，我们来听听来自社群的几个人的发言，首先是在芬兰开放知识中心工作的 NPOC 成员拉乌尔·普伦默 (Raoul Plommer)。然后，我们会 — 在听完几次演讲之后，我们会从这里直接收集社群的反馈意见。

那么，拉乌尔，请开始。

拉乌尔·普伦默： 谢谢，迈克尔。大家早上好。我是 NPOC 副主席拉乌尔·普伦默，NPOC 是 NCSG 的一个选区，代表 GNSO 下的公民社会。

由于我所属的组织芬兰开放知识中心专门致力于提高世界的透明度和开放度，特别是在芬兰，所以我被要求做这个演讲，经过一番混乱后，我们在演讲稿中发表了对 ICANN ODI 的简单意见。我们为这次演讲做了一些收集信息的工作，演讲稿可以在 okf.fi/ICANN-ODI 找到。也就是 — 后半部分是大写字母。ICANN 连字符 ODI。

请播放下一张幻灯片，或者 — 谢谢。

ODI 国际是一个不断发展的国际工作组织，在 40 多个国家/地区都设有团队。我们欢迎所有人加入。任何人都可以加入这个网络，无论是个人和组织，而且并不需要特殊的技能或背景。我想我就是这样加入的。

这些是 — 提供给我们的工具列表。我今天早上刚刚测试了它们，而 CKAN 实际上并没有解析。花了太长时间。我不知道为什么，可能是我的网络问题，但我表示怀疑。我是否尝试过使用其他浏览器？在 ICANN 的网站上，它表示只有 Enigma 需要注册，但 OpenDataSoft 也需要注册。

最后一个，Socrata，确实可以立即访问，看看有什么。好极了！

现在 Socrata 中有什么？只有两个数据集，是 2017 年 6 月 28 日更新的。我确实在其他地方看过这些内容会被更新，比如拖延三个月，但是，时间比这更长。

就像这个演示文稿一样，如果你真的找到了页面，所有的 — 他们都提供了这些链接。比如，基本上所有加下划线的内容都是到这些页面的链接。

我唯一能够理解的是 1263 个 gTLD 的外部数据集，而实际对我有意义的三列是合同签名日期，gTLD 的名称和注册管理运行机构的名称。现在，我无法 — 我无法真正理解这一点。

现在，衡量开放数据的成功是件好事，我认为这是一个 — 这是开展这个项目的一个非常重要的部分。数据集的数量是一个不好的成功衡量指标。它激起了错误的行为。如果我们每年发布一个不同的数据集，我们的数字会增加，但数据可用性会下降。如果我们发布低价值的数据，我们的结果会上升，但没有人想要看它。

因此，有一些关键绩效指标可用于评估实用性，这些指标有助于我们持续正确实施开放数据倡议。所以我们真的应该用这些指标来评估它。

它们帮助我们识别和传达我们的成功，并指出我们可能需要投入更多的领域。

要衡量这项倡议的成功，有三个主要的方面。准备度，项目承担能力和交付能力；实施、发布和提供的关键数据集是什么；以及影响，比如项目的实际结果。比如新产品和服务以及见解。

此外，采用以需求为中心的方法是非常好的，它实际上将数据用户的需求置于数据发布者的需求之上。

现在，你可能会问，这个项目有什么价值？它不会很便宜，但我确实找到了一个很好的价值清单，比如，在哪里创造了价值。我认为这些都是非常好的事情，是值得争取的有价值的东西。

现在，通过提高公众对数据技能的认识，开放数据为提高员工的数据素养水平创造了机会。开放数据可以鼓励组织内的人员彼此之间以及与外部组织进行沟通。

对于企业来说，开放数据可以改善与消费者的关系并开辟新的客户渠道。

开放数据能够在社会中带来变化。开放政府预算的例子非常有用。

我认为 ICANN 预算的开放数据也是有用的。

开放数据有助于社群相互沟通。这个名为“我的社会”的项目创建了一个“修复我的街道”主题，让居民能够与当地区域分享问题、开放数据并创建共享解决方案。

这些是我提出的建议。我们 — 我们真的需要知道 ICANN 有哪些数据。要找到这些数据，出人意料的困难。我确实听到亚娜说会开放 — 抱歉，是杰伊。杰伊说，我们会在 7 月 1 日收到一份库存报告，至少是初步报告。我非常希望看到这份报告。

我们 — ICANN 网页中也提到的一项功能是让用户能够直接在收件箱中收到新的公共内容，如果他们愿意的话。而且我们也可以不仅仅使用注册管理机构的数据，还应该用它来跟踪政策的有效性。我列举了政府使用的几个衡量标准，即国际开放数据宪章、开放数据指数和开放数据晴雨表。所以我们没有必要白费力气做重复工作。

我们的结论是，是的，为了促进这一倡议，最初我被告知 — 这个项目真的需要社群的推动。现在，我认为，如果我们要这样做，我们真的需要知道那里有什么。所以库存在这方面很重要。

我想我们也应该问问社群，我们应该优先考虑什么。

我就讲到这里。谢谢。

迈克尔·卡拉尼古拉斯： 非常感谢拉乌尔。

那么，现在我把话筒交给马克·达蒂格尔德 (Mark Datysgeld)，他拥有互联网治理的硕士学位，在 ICANN 的企业选区中代表拉丁美洲企业。马克。

马克·达蒂格尔德： 非常感谢，迈克尔。大家早上好。

所以，实际上，从企业选区的角度来看，开放数据倡议已经取得了进展，无论是在预算方面，还是在之前的评论中由企业选区进行代表的需要方面。因此，就域名滥用而言，就域名系统的投资是如何进行的而言，主要的呼吁一直是透明度。但与此同时，我觉得有趣的是，作为一个社群，我们还没有共同理解我们希望从开放数据倡议中获得什么。

对我而言，我认为我们应该研究域名滥用，研究域名行业发展的方式，这是我们的一大要务，因为这是社群中许多人的核心业务，它也有助于制定政策，但同时，我认为有很多机会可以做得更好。因此我以研究员的身份来参加这次会议。

之前我在为我的硕士学位进行关于 ICANN 和决策制定的主题研究时，发现要跟踪 ICANN 政策的制定过程相当难。当然，我们确实有开放数据，因为数据是开放的，但我发现开放数据就只是开放数据。数据是开放的。如果你没有一个系统来正确编目，如果没有什么来通知你使用数据的方式，如果你不明白你想要从数据中获得什么，它就仍然只是一堆数据而已。

所以我们需要聚在一起，开始思考我们想要什么。我认为这是在首次实施 ODI 倡议时所缺乏的。

当然，我们知道它正在进行，我们确实与那些正在执行这一流程的人进行了互动，但与此同时，我们都不知道社群究竟想从中得到什么。我们想要达到什么目的？我们当时的想法是，我

们不知道有哪些数据集可用；因此，我们无法知道我们想要什么。但与此同时，这似乎有点自我限制；对吧？

所以我认为，将来我们可以研究不同的方面，比如我们希望开放数据是什么样的，我们认为开放数据可以是什么样的。

因此，例如，从我的角度来看，在我的研究中，我非常想了解 PDP 中区域参与情况如何。例如，我希望能够以更广泛的方式进行研究。利益相关方的参与也是一样。当我们在拉美地区进行外展活动时，来自其他地区的人们如何以特定方式进行参与？这些都超出了域名行业或 ICANN 组织可以提供的指标。但与此同时，它们是 ICANN 提供的数据，可以很好地用于制定政策。

因此，我认为，参与研究 ICANN 并尝试使用其数据产生有意义的研究结果，我从中获得的关键结论是，我们需要更好地关注我们的需求。我们还需要继续推进这一举措，因为我们已经有预算承诺，我们有员工承诺，我们有一个相关社群正在形成，在我们的选区和我们开始工作之前，开始讨论大家的期望可能对我们最为有利。你们知道什么？因为数据在那里。这是一问题是 — 它太分散了。它在 PDF 和 Excel 电子表格、纯 HTML 和/或文本文件之间，或者在某个地方的含糊不清的 FTP 中。

即使那些看似有点不合适的资源也可以更好地帮助我们制定有意义的政策。我认为，我们作为一个社群可以将它变成更大的

东西，或者我们也可以勉强接受它作为数据分析，因为域名本身是非常有意义的。我们不要减少它的价值。我们需要了解这是如何形成的。但与此同时，我们可以更进一步。

我还有一些 — 还有更多观点。但与此同时，我们延迟太长时间了，我认为现在该为听众提供参与机会了。那么我把话筒还给迈克尔。

迈克尔·卡拉尼古拉斯： 非常感谢。感谢所有发言者都遵守紧张的时间表，这让我们有充足的时间进行社群讨论。

那么我们想邀请大家发表任何评论、任何想法、任何想要分享的问题。请随意发言。

梅森·科尔

(MASON COLE):

大家好。早上好。我是梅森·科尔。我很好奇 — ICANN 尽可能收集了大量数据，对此我表示赞赏。但为了避免你之前提到的问题，也就是有一大堆数据，你们如何在内部整理数据，以便将其应用于 ICANN 职能？例如，你们收集的一些数据是否会应用到 ICANN 组织内部的合规职能或其他职能？

迈克尔·卡拉尼古拉斯： 这是个有趣的问题。我想，这个问题是否与 ITI 更相关？我想这个问题可能是 — 如果我理解正确的话，实际上是问 ICANN 内部是否也有应用程序？

杰伊·戴利： 我所看到的实施开放数据的一个有趣的副作用是，组织内部人员以及组织外部人员都更容易访问数据。因为现在只有一个目录，所有数据都是从一个地方获取的，因为这些数据会保证定期更新。

是的，当然，实际上它并不会导致那个问题。至于是否有具体的计划，还没有。但是，这很好理解。

迈克尔·卡拉尼古拉斯： 好的。那么我们接下来看看数字 3，然后有一个远程问题。

莫娜·爱莎华

(MONA ELSWAH):

大家好。我是来自埃及的莫娜·爱莎华。感谢你们的倡议 — 我想代表所有数据科学家，我们需要为此感谢你们。我想，在一倡议发布之后，我们会看到很多研究 — 关于使用这些数据的研究。

我想知道是否存在 — 数据集是否存在任何限制？例如，有些数据是否会匿名？是否有一些数据不会发布，还是说所有数据都会免费提供？谢谢。

杰伊·戴利：

是的，会有一些限制。当然，我们有对个人身份信息进行编辑的法律义务。还有其他一些数据集可能非常机密，无法发布。

现在，只要有可能，我们仍然会记录这些数据集并让你们知道这些数据集。决策制定流程可以是公开透明的。但我认为有些数据无法做到，这是不可避免的，因为如果允许的话，这样做会适得其反，或者可能会产生法律上的危害。

但是，明确来说，ICANN 内部存在一种非常强大的自上而下的公开透明的文化。我可以这样说，作为一个长期社群成员，现在我能看到它的内部。这不会被用作隐藏数据的借口。

默认情况下数据会发布，但是必须有一些合法的例外来控制它。

迈克尔·卡拉尼古拉斯：

大家是否想要 — 在标记问题上，大家是否想对 ITI 补充什么内容？

亚娜·尤金诺维克： 对于我们现在正在标记的所有信息，这些都是公开信息。但我们正在增加额外的分类 — 以遵循 GDPR。正如你所说，任何数据中都有隐私信息，比如在网站上被标记的电子邮件地址，所以我们遵守这一条例。因此分类系统和标记流程正在考虑这一点。

迈克尔·卡拉尼古拉斯： 这是一个远程问题吗？请讲。

远程发言： 戴夫·阿诺德·蒂拉克辛 (Dev Anand Teelucksingh) 提出了一个远程问题。

我在 Google Sheets 中制作了自己的 ICANN 利益相关方分析工具，尝试按国家和地区或全球范围分析 ICANN 利益相关方社群。该工具背后的理念是，第一，能够找到一个国家或地区的所有 ICANN 社群成员，第二，确定国家和地区的利益相关方差距，以实现有针对性的外展。

到目前为止，该工具拥有按国家搜索的向下钻取能力；第一，一般会员代表，第二，ccNSO 成员；第三，GAC 成员；第四，学员；第五，NextGen；第六，经过认证的注册服务机构，并对一般会员、ccNSO、GAC 和认证注册服务机构进行区域或全球分析。

它尝试从认证注册服务机构的网页、ccNSO 和 GAC 网站上提取实时数据，凸显了一种可能的快速致胜方式，即如果来自此类社群网站的信息以一致的方式显示数据，则可以擦去并重新使用来自当前 ICANN 网站的公开数据。

值得注意的一个挑战是，不同的 AC/SO 使用不同的名称来命名国家/地区，需要在 AC/SO 之间对这些国家/地区的名称进行标准化，以便在工具内进行比较。

最新版利益相关方工具的链接可以通过搜索 Google “ICANN 利益相关方工具” 找到。

迈克尔·卡拉尼古拉斯： 非常感谢。

我想我在某个时候也收到了关于该倡议的电子邮件，我认为至少与部分邮件清单分享了这封邮件。

我确实感觉到这挖掘了巨大的潜力，我确信有一个庞大的研究人员和数据爱好者社群，他们很想访问这些数据，并期待能有不同的机会来利用它们。

马克，你想回答吗？

马克·达蒂格尔德： 好的。其实，我知道这个工具。它已在 ODI 清单中分享。这回到了我刚才所说的问题，也就是我们如何以不同的方式使用开放数据。我们并不知道我们的社群究竟是如何传播的。我们知道谁去参加会议。我们可以有一个大致的概念。但是，这个工具就像是了解 ICANN 当前状况的第一步，不仅是在会议中，因为会议并不能很好地展现社群的面貌。

而且我知道，至少 ICANN 组织中的一些人正在沿着这些方向发展。我可能说错了。但这个项目可能会以某种方式将这些方面整合在一起。

迈克尔·卡拉尼古拉斯： 好的。但是 — 当然，请讲。

杰伊·戴利： 我简单补充一点，有一个众所周知的开放数据五星部署方案，第五颗星与开放数据相链接，这样突出显示的不一致就会被消除，因为有一个中心的、非常清晰的关于一切的数据字典。然后以机读方式链接。

我们还没有打算达到这么远的目标。目前我们只是瞄准三星开放数据，但将来会尽量达到那个水平。达到那个水平需要相当大的努力。

亚娜·尤金诺维克： 我想补充的一点是，关于社群团体来自哪里信息，其中一些信息 — 尤其是对于会议的信息，可在会后报告中获得。这些信息被分解了。人们能够识别他们属于哪个 SO/AC，属于哪个团体。因此，大家可以在每次报告后一个月左右的会后报告中看到这些信息。

迈克尔·卡拉尼古拉斯： 我们回到数字 2。

马克·麦克法登

(MARK McFADDEN)：

大家好，我是来自 ISP 的马克·麦克法登。这个问题是问杰伊的。

我正在查看商业利益相关方团体 2017 年 1 月 5 日发出的一封信。大约在 15 个月前。

其中，这三个选区对访问开放数据和特定数据集提出了非常具体的要求。这不是 — 这不是一个含糊的要求。这实际上是非常具体的要求。

顺便说一下，3 月 — 一年前。3 月，ICANN 首席执行官回答说，事实上，这些数据集将会尽快提供。一年 — 一年多过去了，杰伊。据我所知，ICANN 在提供这些数据集方面没有取得进一步的进展。这些是针对这些选区感兴趣的特定目的非常具体的要求。

而且，我想，首先，我错了吗？没有进展？没有公开进展。例如，没有公开的数据供我们使用。

第二点是，如果我们不能在一年中取得进展，为什么我们要相信下一年，也就是 2018 年会更好？

很抱歉让这个人说话。他的话让人有点恼火。

戴维·康纳德：

我是戴维·康纳德，ICANN 的首席技术官。

我们一直在取得进展。特别是，我们一直在尝试努力缩小我们将发布数据的平台。我们一直在内部努力确定各种数据集。

企业选区提供的三个数据集，可能是 ISPCP，我不太记得了。但是我们知道这些要求，并且一旦我们真正开始将开放数据平台投入运行，就会在部署中优先考虑这些要求。

我们面临的挑战之一是组织内部存在非常大量的数据，其中大部分数据并不适合部署在结构化数据系统中。我们尽量识别这些数据集，并确保这些数据集中指定的数据是已知的，我们有一种字典 — 描述这些数据集的数据字典，并且我们不会越界，违反对于这些数据的隐私或合同义务。

我们对于延迟开放和提供这些数据集表示歉意。但我们显然会继续努力，正如我们所表现的一样，让杰伊担任这个特定项目的顾问，并在内部聘请额外的承包商来推动项目的进展。

迈克尔·卡拉尼古拉斯： 想接着说几句？

马克·麦克法登： 我必须说几句。

谢谢你，戴维，我很欣赏这一点。我想我可以代表社群的许多成员欢迎杰伊加入这个项目。

我现在要提出的要求是 — 我的意思是，对于去年发生的事情，我们已经不能做什么了，只能表达商业利益相关方团体的遗憾，我们在可见数据源方面没有取得很大进展。但是我们理解 ICANN 在收集数据源和减轻其合同义务等方面所面临的复杂性。

但是，我现在请求你和杰伊，当我们在接下来的几个月里继续发展时，社群获得更快的进度报告，让我们对于正在发生的事情有了一定的了解。

我们在本周早些时候看到了 RFP。那很棒。但是，随着杰伊取得进展，随着您的团队取得进展，对我们选区非常有帮助的事情之一是，能够将这一进展反馈给我们选区，表明我们正在取得进展 — 在 ODI 方面确实取得了进展。这是我的一个问题。

戴维·康纳德： 我们将承诺定期提供进度报告。

我们的一个要求是，迄今为止我们收到的来自社群的反馈很少，包括商业利益相关方团体关于他们究竟想要什么样的数据集，他们希望在开放数据平台中看到什么功能。

我们已经选定了两个平台，主要是出于内部观察和内部需求收集的原因。关于社群希望在平台看到什么内容，我们还没有收到很多反馈。所以这是我们要求提供额外帮助的另一方面。

迈克尔·卡拉尼古拉斯： 我不希望这变成一场双人辩论。还有其他人有问题要问。
有请数字 3。

克洛伊·希利

(CHYLOE HEALY):

大家好。我叫克洛伊·希利。我来自加拿大，我是一个新人。很抱歉，我不知道这是不是一个合适的问题，但我不知道 — 因为我不知道数据集。

但我很好奇。我知道一些数据集确实需要语境化，否则会增加误解，这可能是有害的。

我只是好奇，ICANN 是否会扮演一个角色或者采取某种方式来缓解这种情况。

迈克尔·卡拉尼古拉斯： 我来介绍一下背景，克洛伊来自加拿大 First Nations Internet Governance (FNIG) 中心，信息的语境化是围绕加拿大政府所掌握的 First Nations 信息反复进行的有趣辩论。有人一有请。

杰伊·戴利： 根据我的经验，国家统计局非常关心其数据的语境化。他们往往在一定程度上对数据进行注释，有时候进行深入的注释，以便控制语境化。大多数数据科学家通常会立即删除注释并丢弃。

在数据层级语境化是非常困难的。要实现这一目标，通常需要在流程的更高层级进行。

现在，由于 ICANN 是一个开放的组织，坚定承诺开放，因此不会像在国家统计局那样，以那种方式在那样的层级对这些数据进行语境化。除此之外，我不知道还有什么计划，因为我认为我们真正希望看到社群是如何获取这些数据的。

迭戈·堪巴罗

(DIEGO CANABARRO): 大家好。我是迭戈。我为巴西互联网政府工作。能听到我说话吗？

迈克尔·卡拉尼古拉斯： 是 1，麦克风 1。

迭戈·堪巴罗： 好的。谢谢。我是迭戈。我为巴西互联网指导委员会工作，但在这里我以个人身份发言。首先，我要对刚刚介绍的倡议表示祝贺。它们在各方面令人惊叹。但让我感到震惊的是，没有一个人针对网络上的简单数据、结构化数据、创建用于在线显示数据的基础架构提出 W3C 标准。

所以我想知道：W3C 标准在你们的倡议中扮演什么角色？它们有什么作用吗？有还是没有。谢谢。

迈克尔·卡拉尼古拉斯： 杰伊先开始，之后再看看亚娜是否要补充。

杰伊·戴利： 我在演讲中提到了访问数据的开放标准的要求。我只是没有详细说明。

Web 标准中的 CSV 仍然非常新。我认为，距离得到任何供应商的妥善支持，仍然需要一段时间。因此，我们在 RFP 中要求提供有关开放 API 合规、相邻 API 和关于 Web 合规的 CSV 的详细信息。但是可能需要有一个围绕这一点的途径，或者进行一些权衡。我不想说太多，因为 RFP 正在进行中。

正如你们可能知道的那样，我们也是 — Microsoft Odata 是一个重要的行业标准，尽管它也是一个专有标准，并且被嵌入到许多工具中。

几乎所有供应商都通过多种不同的机制让人们获得数据。而且，是的，我同意你的观点，这样做的最终目标应该是完全开放的标准。但这是一个成熟的市场，所以还没有看到这方面的时间标尺。

迈克尔·卡拉尼古拉斯： 回到数字 2。

丹尼斯·米歇尔

(Denise Michel):

大家好，我是来自 Facebook 的丹尼斯·米歇尔。我想 — 首先，我要感谢你们举行这次会议。这是一个非常重要的问题，我们非常感谢戴维·康纳德和他的团队将这个问题放在首位。

我还要指出，Facebook 所属的企业选区最近提交了对 2018 财年预算的评论，强烈支持 ODI 倡议并为其增加资金。

我们已经连续两年为数据集提供了详细反馈。它在某些方面似乎有点像土拨鼠日。如果 ODI 工作人员能够对过去两年中许多个人、公司和选区对数据集提供的可能有用的实质性反馈进行审查，并纳入当前工作中，而不是停下来要求提供全新的意见，这样会比较好。

我们 — 许多公司和组织使用 — 让我们的科学家和专家提供了非常详细的意见，与我们的员工坐在一起数小时，为此提供指导和意见。

我希望你们能利用过去两年的工作成果。而且由于企业选区支持 — 正在为此工作再次寻求财务支持，我认为马克·麦克法登想说的是，我们这次希望看到真正的交付成果和结果以及适当的项目计划和时间表。

谢谢。

迈克尔·卡拉尼古拉斯： 当然，请讲。

杰伊·戴利： 丹尼斯和马克，当然我可以承诺，我们将尽力在我们到你们身边之前实现这一点，你们知道，这对于你们来说不是土拨鼠日。

回到 — 不是。另一个 — 还是数字二。

法夫里西奥·韦拉

(FABRICIO VAYRA):

大家好，我是来自博钦律师事务所的法夫里西奥·韦拉。首先，我想附和一下马克和丹尼斯对时间安排的担忧。我欢迎你们所做的一切，我真的很感激。我的问题是关于 ITHI 的。时

间问题之所以对我来说很敏感，是因为我参加了 2016 年 9 月 7 日的最初的 ITHI 会议。虽然这不是 ITHI，但我确实是在小组中看到有人问这是不是 ITHI 的一部分。

所以除了同样对时间安排表示担忧之外，我的问题是，这两者之间有哪些重叠之处？因为我不完全确定社群清楚这个问题，我在小组中看到了一些疑问。我们如何才能确保我们不会重复工作，并且如果这是一个维恩图，我们可以捕获所有超出重叠范围的内容，以便开放数据，因为尽快开放数据是有用的，而且我们没有存在冲突或带来不同问题的多个工具。

迈克尔·卡拉尼古拉斯： 那么，讨论这两者之间的整合以及重申.....

好的。抱歉。请讲。

杰伊·戴利： 戴维可能需要给出一个全面的回答，但是 — 不，迈特要给出更全面的回答。

迈特·拉森 (Matt Larson)： 嗨，我是首席技术官办公室研究副总裁迈特·拉森。

我想简单谈一谈标识符技术健康指标项目 (ITI [sic]) 和我们今天讨论的两件事之间的区别。因此，ITI 项目正在制定一套衡量标准，以评估 ICANN 帮助协调的唯一标识符系统的健康状

况。RIR 正在研究数字方面，数字的衡量标准方面。在首席技术官办公室里，我们花了很多时间开发名称健康衡量标准，其次是协议参数注册管理机构。因此，这些都是我们将要跟踪的所有数据，其中的理念是，随着时间的推移，能够看到这些衡量标准的变化将反映标识符系统的健康状况。

所以这是我们要跟踪的数据，它是进入 ODI 的完美数据集。

除此之外，除了 ITHI 数据是将在 ODI 平台上发布的数据外，两个项目之间没有其他关系。

迈克尔·卡拉尼古拉斯： 有什么要讲的吗 — 没有？

发言人（姓名不详）： （关掉麦克风）

迈克尔·卡拉尼古拉斯： 当然，请讲。

约翰·莱普瑞斯

(JOHN LAPRISE): 大家好，我是约翰·莱普瑞斯，来自 NARALO 的 ALAC。

我只想恳请小组在 ODI 倡议中反映 UX（用户体验）和 UI（用户界面）的重要性。在我们今天在屏幕上看到的所有内容中，

在设计过程中几乎没有提及 UX。要让最终用户能够有效地使用开放数据，这一点非常重要。

谢谢。

杰伊·戴利：

对于开放数据倡议，这是 — 这主要是关于机器到机器。这是关于数据科学家访问和提取数据的。其中许多工具都包含可视化工具或其他方式，用于理解、表达见解，或者可能是通过数据讲述整个故事，但这不是这个项目的优先事项。它主要关注的是数据科学家能够访问它。

因此，UX 和 UI 主要针对数据科学家以及他们专为 ODI 提取数据的方式。

亚娜·尤金诺维克：

感谢约翰，对于这个问题，我知道我们已经在星期六的一般会员会议上进行了讨论。

显然，UX 对于 ITI 项目非常重要，因为它是公开信息 — 我们从 icann.org 开始。人们难以在 icann.org 上找到信息的部分原因是因为 UX。

在建立这些用户旅程、让人们能够真正快速获取他们需要的内容方面，我们做得不够好。所以 UX 对于 ITI 来说是非常重要的。这是项目计划的一部分。因此，当我们 — 我们将在

feedback.icann.org 上展示这些内容和功能，我们希望你们不仅要评论内容的功能性，还要评论 UX；大家知道，看到一个按钮时，它是否让你更容易访问任何内容，无论是公众意见还是类似的内容。所以绝对重要。

迈克尔·卡拉尼古拉斯： 数字 3，后面那位。

萨蒂什·巴布

(SATISH BABU):

谢谢。我是萨蒂什·巴布，来自亚太地区的一般会员。我们欢迎开放数据倡议，我们认为这将为 ICANN 的运作带来很大的透明度。

我对发布数据的许可条款有个问题。

关于什么许可，你们有详细信息吗？它是开放文档许可还是任何知识共享许可？或者当你们发布数据集时，有什么确切的许可形式？

谢谢。

杰伊·戴利：

这还有待决定，但它将是一个开放许可，所以实际上只有两个适当的可选方案，知识共享或者独立的开放数据许可。后者可能更可取，但仍需要做一些工作来确定用哪一个。

迈克尔·卡拉尼古拉斯： 数字 2。

伊申·梅塔

(ISHAN MEHTA):

大家好。我是来自 NextGen 的伊申。我是 ICANN 的新人。所以 — 但我之前已经处理过与 ICANN 有关的数据，而且听起来好像还有其他一些学术和商业团体已经在 ICANN 工作的外围建立了自己的数据集。我想知道你们是否计划尝试整合第三方已经存在的数据集，或者如果不整合的话，至少将它们编入开放数据倡议的索引中？因为我认为，我们的大学有一个可能对其他人有用的小数据集，但是查询 Georgia Tech ICANN 要比查询 ICANN 开放数据倡议困难。所以这是我的一个建议。谢谢。

杰伊·戴利:

我们已经探讨过这一点，对此还没有做出任何决定。我已经看到许多已经完成并且成功的开放数据平台。

然而，这是一个复杂的环境，关于这些数据的真实性，以及是否会在其他方面进行维护和更新，还有一些重要的问题要询问。

所以我认为遗憾的是，要回答这个特定问题，还有很长的路要走。

迈克尔·卡拉尼古拉斯： 到 1。

法夫里西奥·韦拉： 谢谢。还是来自博钦律师事务所的法夫里西奥·韦拉。感谢你让我问一个后续问题，我很赞同 ITI 解释的区别。ITHI。ITI 和 ITHI。

为了在座各位的利益，既然我们正在讨论 ODI 的时间表，而且我知道我们明天将会举行一次 ITHI 会议，是否有一 我们是否可以提前了解一下我们两年来寻求的数据的时间表 — 什么时候可以提供，目前处于什么阶段？

我所听到的是 ITHI 数据非常适合 ODI 和使用 ODI 平台进行发布。但是，为了让所有人都可以看到这两件事情正在继续，我们能否了解一下 ITHI 数据的进展？

迈特·拉森： 当然。ITHI 项目长期以来一直在咨询社群和制定衡量标准。我可以说，最近这些衡量标准几乎已经固定下来，而且我们正在开始跟踪这些衡量标准所指定的数据，至少在命名方面。数字方面，RIR 正在做这方面的工作；也许稍微落后于时间表，落后于名称方面的工作。所以在这一点上，我会说我们依赖于 ODI 平台，它对于发布 ITHI 衡量标准非常重要。所以我们 — 我想说，大家知道，我们还有几个月，而不是几周。

迈克尔·卡拉尼古拉斯： 好的。我们还有时间回答一个问题 — 哦，拉乌尔。

拉乌尔·普伦默： 我想问一个关于数据管理系统的总体问题。这是免费供所有人使用的吗？它是开源的吗？

亚娜·尤金诺维克： 我纠正一下，它是一个文件管理系统，而不是一个数据管理系统。它称为 **Alfresco**，不，它不会提供给社群。这其实是将内容存储在这个文件管理系统中，但内容将通过内容管理系统呈现给社群。

最终路径是，一旦为 **icann.org** 和其他 15 个公共信息网站设立了基础架构，将为 **SO/AC** 网站推出该基础架构。因此，文件管理系统将被部署并用于 **SO/AC** 网站上当前的内容。

迈克尔·卡拉尼古拉斯： 我在后面看到了数字 3。

凯莱布·奥卢穆伊瓦·

鸥规德勒 (CALEB

OLUMUYIWA OGUNDELE)： 是的，我是来自尼日利亚的凯莱布。

我的问题与文件管理系统和内容管理系统有关，我的意思是与应用程序（听不清）有关，你们是否打算使它成为开源代码，

以便所有的开发人员都可以看到代码？如果将来需要发展，我们可以看看并做出一些贡献。也许我们可以将它用于其他数据流程。

迈克尔·卡拉尼古拉斯： 这可能是 — 这与上一个问题非常相似，但是你们想补充什么吗？

发言人（姓名不详）： 不，这不在计划中。目前，文件管理系统完全是为了内部 — 我们的内部内容，并确保它被恰当地分类，以便它能够更好地在外部呈现。

内容管理系统也是如此。它纯粹是为了呈现内容，没有这方面的计划。还有其他问题吗？

戴维·康纳德： 是的，我想澄清一下。Alfresco 和内容管理系统正在使用 .CMS。两者都有开源版本。我们将要使用的版本实际上来自 Alfresco.com。这是一个仅在内部使用的商业许可版本。它是我们的数据存储库，我们显然不希望外部机构能够进入我们的数据存储库进行涂写。

因此，如果有人对 DMS 或 CMS（内容管理系统）感兴趣，我想告诉他们，我们使用的是知名的开源产品，它们具有开源代码实现，任何人都可以部署，如果他们愿意的话。

迈克尔·卡拉尼古拉斯： 一个优秀的主持人要具备几种特质，但最重要的是不要让人们吃饭迟到。我看到没有其他问题了，现在是中午 12 点，感谢所有嘉宾给我们带来如此精彩的讨论。希望这有助于推动我们一直在寻求的社群反馈，并且我期待在这些倡议发展的过程中对它们进行关注。

[会议记录结束]