

萨尔麦德·侯赛因

(SARMAD HUSSAIN):

好的。会议马上开始了。能把录音打开吗？

感谢大家参加今天的会议，本次会议的主题是，庆祝多利益相关方模型在根区标签生成规则编制中的应用。总的来说，之所以要定义这些规则，是为了在互联网根区中支持顶级域。这是一个从 2011 年就开始做的项目。在本次会议中，我们会获得更多有关这项工作的信息和细节。这项工作由来自世界各地的语言文字社群推动，许许多多的志愿者参与其中，共同编制了用于确定在特定文字及使用这些文字的语言中，哪些是唯一标签和有效标签的规则。

就像我说的，这项工作由许多不同的语言文字社群完成，包括阿拉伯文、亚美尼亚文、孟加拉文、中文、西里尔文、梵文、埃塞俄比亚文、格鲁吉亚文、希腊文、古吉拉特文、果鲁穆奇文、希伯来文、日文、埃纳德文、高棉文、韩文、老挝文、拉丁文、马拉雅拉姆文、缅甸文、奥里亚文、僧伽罗文、泰米尔文、泰卢固文和泰文。正是因为这些社群和他们多年来的出色工作，现在，我们才有了非常具体的规则，可以推进在定义这些文字中的有效域名（特别是顶级域名）方面的工作。

接下来，请允许介绍一下本次会议的议程。本次会议分为两个部分，第一部分是专家组讨论，内容包括概述根区 LGR 项目是什么，为什么需要 RZ-LGR，它是如何编制出来的，社群的工作方式是怎样的，以及在获得通过后的后续步骤是什么，即，如何通过政策制定

流程将社群完成的 RZ-LGR 工作向前推进。然后是问答环节。在专家组成员发言后，我们会接受大家的提问并进行回答。

在会议的第二部分，我们会感谢所有参与这个项目的社群成员，你们为此努力了近十年，编制出了这些规则。我们会邀请 ICANN 总裁兼首席执行官马跃然 (Göran Marby) 和 ICANN 董事会主席马腾·波特曼 (Maarten Botterman) 致辞，感谢社群对这项工作所做的巨大贡献。

我们还会与生成专家组的所有成员以及为这项工作做出贡献的社群成员合影，大家可以看到，他们中的许多人都穿着绿色 T 恤，不过也有许多人是远程参会，他们会在 Zoom 中加入我们的合影行动。在那之后，大家可以与各自地区的语言文字社群互动，相互交流。以上就是本次会议的议程安排。请翻到下一页。

专家组成员方面，我们有一群非常有经验的专家组成员。我们有 ICANN 董事会 IDN-UA 工作组的主席钟宏安 (Edmon Chung)。我们有整合专家组的代表马克·布兰切特 (Marc Blanchet)。我们还有来自不同语言文字社群、一直致力于 RZ-LGR 编制工作的社群成员，包括中文生成专家组的联合主席王伟 (Wang Wei)。他不在这里，所以可能没办法出席。这样的话，同为中文生成专家组联合主席的黄胜雄 (Kenny Huang) 将跟大家分享他们的经历。我们有拉丁文生成专家组的成员迈克尔·巴德兰 (Michael Badland)，以及新婆罗米文生成专家组的联合主席阿嘉·达塔 (Ajay Data) 博士。他们将从社群工作的角度分享他们的经历。

然后，随着 RZ-LGR 编制的推进，技术方面的工作完成了，下一步就是促成政策对它的采纳。我们邀请了 GNSO IDN EPDP (快速政策

制定流程) 工作组的主席唐娜·奥斯汀 (Donna Austin)。但很遗憾, 她今天没能来到现场。所以她的部分将由 ICANN 的爱丽儿 (Ariel) 代为发言。另外, 由于黄胜雄身兼 ccNSO IDN 政策制定流程工作组的主席, 所以将由他跟大家介绍 RZ-LGR 的运作方式, 然后由爱丽儿跟大家介绍如何将 RZ-LGR 的工作纳入到政策中。我叫萨尔麦德·侯赛因。今天我的职责是协助大家。我会与皮提南 (Pitinan) 一起主持这次会议。

好了, 我们进入正题吧。这是一条时间线, 从这个项目的最初阶段开始, 当时它叫做“变体问题项目”, 那是 2010 年到 2012 年, 稍后我们会请钟宏安谈谈那个时候的情况。然后, 在最初阶段的所有工作完成后, 2013 年, 这个项目正式启动。在过去 9 年左右的时间里, 我们的社群一直在做这件事。但在启动之前, 我们设立了一个整合专家组, 他们在这个过程中发挥了自己的作用。然后, 这个过程就是生成专家组和整合专家组的互动协作过程。现在, 项目已经到了这样的阶段, 基本上, 在 2022 年, 在今年早些时候的 5 月份, 我们发布了第 5 版根区 LGR, 它整合了所有活跃生成专家组的工作, 也就是最初计划的 28 种文字中的 26 种。换言之, 至少这些活跃的生成专家组的工作已经完成。今天我们在这里, 就是要感谢他们做出的贡献, 感谢社群在这项非常艰巨但重要的工作中做出的贡献。

接下来, 让我们有请今天的第一位专家组成员, 钟宏安, 从这个项目的策划阶段开始, 他就一直参与其中。请你跟我们大家大致谈谈你对多语言互联网需求的看法, 我们知道, 这是非常贴近你内心的东西, 然后, 或许你还可以分享一些背景信息, 让大家知道当初促成根区 LGR 项目的因素是什么。谢谢。

钟宏安：

谢谢萨尔麦德 (Sarmad)。欢迎大家。我很高兴能谈到这个话题。在座的许多人已经知道，这对我来说绝对是一个充满激情的项目，它可以一直追溯到 1999 年我刚加入 ICANN 的时候。当时我就坐在这里跟马克 (Marc) 讲话，马克坐在我旁边。我第一次见到马克时，我们谈论的就是 IDN，而且是在机场。

他向我提出了一个问题，我认为是一个非常大的问题，为什么我们需要多语言互联网？有意思的是，如果我没记错，到现在，IDN 已经发展了大约 23 年了。但在最初的几年里，要说服人们相信我们确实需要多语言互联网，这并不容易。事实上，我仍然清楚地记得，在一次会议上，有个人走到麦克风前说，“我只知道两种语言：英语和计算机语言 C。”所以，我们花了很多时间，才让人们明白在互联网上支持多语言的重要性。现实情况是，世界上大多数人的母语并非英语，也不会只使用英语的字母数字字符。本地企业会使用本地语言给自己起名。还有人名，很显然，也会使用本地语言。既然如此，为什么人们不能使用母语来创建他们的网络身份呢？我觉得这才是促成 IDN 和部分国际化电子邮件地址的真正因素。如今，也许大家会说，使用国际化域名的人还是不多。这是当然的，肯定会存在挑战。其中的一个挑战是，不同 IDN 和电子邮件地址的普遍适用性。

不过，除此之外，还有另一个方面也是我们今天要讨论的。另一个挑战是政策。什么样的政策？它们只是一些跟英语域名很像的域名而已，我们已经有相应的技术支持其他语言的域名。那么问题是什么呢？没错，确实。有趣的是，早期的时候，我们发现，当你在互联网上添加不同的语言时，特别是唯一标识符系统，特别是域名和电子邮件地址，会存在不同字符拥有不同性质的情况。

其实严格说来，这不是一个很好的例子，不过我还是想用它，因为这是向不懂很多语言的人解释的最简单方法。大家最好这样来看待这个问题，比如说英语域名，它有大写字母和小写字母，二者的用法是一样的。在输入某个域名的时候，不管你输入的是大写字母还是小写字母，跳转的都是同一个网站。国际化域名则不是这样，一个经典的例子就是中文域名。我们有繁体中文和简体中文，从技术上说，它们是域名的不同码点。那么需要什么呢？但是，人们会互用它们，举个例子，特别是在香港，越来越多的人会互用简体中文和繁体中文。

所以，我们需要的是一个能将二者映射到一起的系统。大家不妨想想，当你要注册一个域名的时候，你不得不注册这个域名的所有大写字母和小写字母变体，也就是说，最终你注册的不是一个域名，而是几十个域名。这就是政策。这就是我们所讨论的政策的一部分。另外，这还跟我心目中非常重要的东西有关，那就是减少滥用，提高终端用户对 DNS 的信任。当人们在看到域名或在输入域名时，政策应该确保，他们对用户体验的期望能够尽可能地得到满足。这是我们所说的根区 LGR，即标签生成规则集的重要组成部分。

接下来我谈一点历史吧。首先，我要说的是这一点非常重要。很早的时候，我曾在想，到底是什么问题？如果这些不同的语言 — 这就像是，“颜色”这个词在英式英语的拼写中为“colour”，有“u”，而在美式英语的拼写中为“color”，没有“u”，可那又怎样？你注册两个不同的域名不就行了。但后来我意识到，情况并非如此。不同的语言确实有不同的特点，这就是根区 LGR 政策要解决的问题，我们要确保这些政策在技术上能够同样地实施，让根区和 IDN 能够统一协调地运作。

然后，我想再讲一点背景信息，萨尔麦德刚才提到，根区 LGR 的工作始于 2010 年。事实上，那是 IDN ccTLD 快速通道首次应用的时候。也就是说，第一批 IDN TLD 于 2010 年开始落地。当时也是在 GNSO 和 ccNSO 之间组建联合工作组的时候。我很荣幸能够成为这个工作组的主席，它的名字是 ccNSO 和 GNSO IDN 联合工作组。喜欢了解 ICANN 历史的人可能知道，这代表着我们第一次重新把 ccNSO 和 GNSO 放在了一起。如果大家去看看有关 ICANN 的冷知识，会发现 ccTLD 和 gTLD 最初都在 DNSO 负责的范畴内。但这个工作组把它们重新放在了一起，因为二者有很多共同的问题，包括单字符 IDN TLD、IDN TLD 的普遍适用性，当然还有 IDN 变体 TLD，这在很大程度上都是根区 LGR 解决的问题。那是在 2010 年。

之后，随着时间的推移，我觉得很多问题都通过政策得到了解决。而今天，我们之所以要在这里庆祝 — 看着周围的环境，我觉得下次如果疫情真的结束了，我们应该弄点葡萄酒和啤酒来庆祝一下。到现在的 2022 年，我认为真正令人兴奋的是看到，世界上大多数活跃的语言和人们经常使用的文字现在都能够在根区中使用。

最后我想说的是，在 ICANN，人们对很多不同的事情都有争论，IDN 只是其中之一，就像我说的，从 1999 年开始就一直有争论。另一件自 1999 年以来大家一直讨论和争论的事情可能就是 WHOIS 了。这方面我们不多讲，现在我们要关注的是将 IDN 变为现实的最后阶段。我很高兴能成为 IDN-UA 工作组的主席，亲眼见证它的落地，见证 IDN 得到实际应用。在我看来，这个问题可能代表了最强烈的社群共识之一，让 IDN 得到普遍应用并制定政策来防止滥用和支持 IDN，这完全符合全球公共利益。

我想，以上就是我认为对这个项目而言很重要的一个小故事。从那时起，经过多年的努力，现在我们终于有了非常有用、一致并且在技术上和政策上都非常强大的根区 LGR，关于这个根区 LGR，我想其他专家组成员稍后会谈到。谢谢。

皮提南·古阿莫帕塔纳

(PITINAN

KOOARMORNPATANA):

刚才宏安分享了这个项目的历史，接下来，我们继续往前看，邀请其他专家组成员分享一下他们的经历以及他们开展工作的方式。首先有请来自整合专家组的马克·布兰切特。他要向大家介绍的是，我们是如何组织工作，将这么多语言文字社群的成果整合到单个规则集中的。那么下面交给你，马克。

马克·布兰切特:

谢谢皮提南。在座的各位下午好，还有远程参会的各位，上午好、晚上好、晚安。很荣幸来到这里，跟大家一起回顾这 10 年的工作。这是一项非常艰巨的任务，特别是对我右手边的同事来说，因为大部分工作都是他们做的，他们就是生成专家组。

首先，我想从统一码 (Unicode) 开始，自上而下地谈谈我们都做了些什么。统一码是一套标准，它从开始编写时就规范了对世界上任何文字和语言中使用的所有可能字形的编码。实际上，统一码 — 正如最近编码的 [听不清] 欧洲字形。但这些字形并不适合互联网 [听不清]。其中一个最重要的标准是安全性。我们不希望看到基于相似字形发起的网络钓鱼攻击。

由 IETF 定义的 IDN 协议（叫做 IDNA）根据字符的属性（而非外观或视觉效果），将整个统一码集缩小到了一个更小的集。例如，IETF 不会去查看每一对字符来验证它们的视觉相似性。LGR 工作进一步限制了字符集，并编制了相应的规则，使任何文字中的字符串都有机会成为 DNS 标签或域名。

为了创建各种文字的字符集，我们针对每种文字召集了一群专家，组建了多个生成专家组。他们花了数月甚至数年的时间，来定义各自文字中应该在 TLD 根区中使用但实际上对第二级也相当有用的受限字符和变体字汇以及相关的规则集。在这项工作中，他们实际上排除了一些字符，比如已成为历史或当下没有使用的字符。这项工作的正式交付成果是 LGR，即各种文字的标签生成规则，对技术人员来说，它实际上是一个 XML 文件。

这项工作遵循了一些原则，例如稳健、简单和安全。关于这些原则的运用，举个例子，我们会排除当下人们没有使用的字符。比如说，我们不希望看到垃圾邮件发送人使用一些当下没有使用但与当下正在使用的字符很类似的模糊字符。所以在这个项目中，生成专家组必须记录他们的字符，并确定人们是否真的在使用它们。而我们整合专家组，我们的工作就是验证他们的说法。

我们希望被纳入的字符能长期稳定存在。例如，如果统一码最近编码了一个新字符，那么该字符实际上并非纳入 LGR 的理想字符。为什么这么说呢？这是因为，就像软件开发中的其他人一样，有时候，你会因为字符的实际属性不合适而做出更改。而在标识符方面，我们不希望出现这样的情况，比如说，某个 TLD 字符串包含的字符从有效变为无效。这不是什么好消息。

另一个设计原则是简单。在工作过程中，有时我们从生成专家组那里收到的 LGR 会看起来像语言语法。但这不是放语言语法的地方。它只能包含有关你如何编写文字的内容。

整合专家组的目的是审核这些不同文字的 LGR，确定它们是否符合相应原则，然后将所有 LGR 整合成单个根区 LGR。在整合时，整合专家组还会看看有无跨不同文字的问题，很显然，单个文字生成专家组是没法发现这些问题的，但是整合专家组在综合来看所有这些 LGR 时，就可能会发现问题。整合专家组详细审核了收到的所有文字 LGR，然后开始唱黑脸，要求生成专家组就在提案中纳入的每个字符提供理由。大家可能知道，有些文字的字符真的很多很多。整合专家组由来自不同领域的 5 名专家组成，包括统一码、IDNA、DNS 以及语言和文字领域。我本人就是整合专家组的一员。

对所有 LGR 的整合为整合专家组带来了一些设计选择。例如，跨文字变体有时候会产生雪崩效应，毕竟太多文字了。某个变体可能与另一个 LGR 中的字符相关，而该字符又是另一个的变体，最终一发不可收拾。因此，为了简单起见，有时候我们会做出对整个社群来说似乎是最好的，但可能会或多或少地影响一些 LGR 文件的设计选择。但我们绝不会影响生成专家组的工作，这一点毋庸置疑。核心生成专家组的 LGR 工作 — 字汇和规则从未受到我们工作的影响。

对某些文字来说，究竟要不要在 LGR 中包含某个字符，这之中的界限非常模糊。此外，一些字符之间，特别是不同文字之间的相似性往往非常明显。所以很多时候，生成专家组在最后做出选择以及整合专家组在达成一致意见时其实非常艰难。虽然我们不需要到街上打一架，但一两杯啤酒肯定是需要的。能不能请你翻到上一页？谢谢。大家可以看到，在这页幻灯片的下面部分有一个小箭头，写着

“需要改进”。有时候我们会像高中的教授一样，教授在收到学生的家庭作业时，会说，“不够好，继续努力。”然后这个学生会做一个新的版本，再来找教授，教授仍然说，“不够好”。所以，抱歉了，生成专家组的各位同事，我们不是有意要这么做的。

在这几年里，我们经常会遇到生成专家组为了让他们的字符能被纳入 LGR 而贿赂我们的情况。有一件大事，你们之中只有极少数人知道，但我要在这里说出来。我们被收买了。怎么样，有没有激起大家的兴趣？既然现在这项工作已经完成了，那么为了社群的利益，我需要公开这件事，对吧？我们从某个生成专家组那里收到了一份礼物，给大家看看。有没有激起大家的兴趣？

男性发言人（姓名不详）： 一堆现金吗？

马克·布兰切特： 不是，抱歉。不是现金。

男性发言人（姓名不详）： 支票？

马克·布兰切特： 就是这个礼物。我们从某个生成专家组那里收到了一件印有他们文字的 T 恤。真是太贴心了。我很乐意收到根区 LGR 中每种文字的 T 恤作为贿赂，虽然对于某些文字来说，这可能比较困难，因为它们的字符实在太多了。为避免大家误会，我想说的是，这只是个玩笑。

总而言之，如果没有各个生成工作组的出色工作，这项艰巨的任务根本不可能完成。整合专家组在其中只是扮演了协助和促进的角色。在此，我要感谢整合专家组的各位同事，他们是阿斯穆斯·弗赖塔格 (Asmus Freytag)、米歇尔·叙伊尼亚尔 (Michel Suignard)、威尔·塔恩 (Will Tan) 和尼古拉斯·奥斯勒 (Nicholas Ostler)。我还要感谢所有从一开始就参与进来的 ICANN 工作人员，包括纳尔拉·萨拉斯 (Naela Sarras)、尼科勒塔·蒙蒂努 (Nicoleta Munteanu)、萨尔麦德·侯赛因、阿里礼萨·萨利赫 (Alireza Saleh)，还有皮提南·古阿莫帕塔纳。今天我还请她帮我拼读了她的姓，不过现在对我来说还是很难，希望我的发音没有太糟。另外我还要感谢生成专家组的各位成员，但我就不一一列出来了，因为可能有好几百人。女王万岁 — 抱歉 — 国王万岁 — 抱歉，应该是 LGR 万岁，开放多样的多语言互联网万岁。谢谢。

皮提南·古阿莫帕塔纳： 谢谢马克。现在大家知道了为什么我们花了这么多年，就是因为，我们都忙着开玩笑。好的。我们继续吧。

接下来有请各个生成专家组分享一些他们的经历。实际上，我们有 7 个不同生成专家组的超过 270 名志愿者参与了这项工作。但很显然，我们没办法让所有人都来到这里，很遗憾。今天我们请到了三个生成专家组的代表。我们请到了中文、拉丁文和新婆罗米文生成专家组的成员来到现场。你们每个人只有一分钟的时间，首先是胜雄 (Kenny)，然后是迈克尔 (Michael)，最后是阿嘉 (Ajay) 博士，简单介绍一下你们的生成专家组是什么样的，成员组成如何，以及你们的文字有什么特点？先每个人一分钟，然后我们看看大家有没有什么问题。谢谢。

黄胜雄：

谢谢。我是中文标签生成专家组的联合主席黄胜雄。很高兴能给大家简单介绍一下中文生成专家组。基本上，中文标签生成专家组的成员都来自于 — 实际上，在加入生成专家组之前，我们已经开始一起研究国际化域名了，例如，以说中文的地区来划分，我们的成员有来自台湾、中国大陆、澳门、香港还有新加坡的。也就是说，我们这几个说中文的地区已经开始合作，加入了我们称为中文域名协调联合会 (CDNC) 的联盟。除此之外，我们还与其他使用汉字的地区（例如日本、韩国）建立了合作。所以说，基本上，我们标签生成专家组的成员大多分布在台湾、澳门、香港和新加坡这些地方。但就中文的覆盖范围而言，它是世界上的第二大语言。所以就覆盖范围而言，我们的政策提案中也有写道，基本上，你会发现很多地方，他们都在使用汉字。很多时候，即使是马来文，比如 [听不清]，在树下走的时候，你会看到 [听不清] 使用了很多中文标签，比如在新加坡，到处都是中文标签。所以可以说，汉字几乎覆盖了整个世界。至于字符，我们同时涵盖了繁体中文和简体中文，涵盖了很多语言，很多我们在很久以前就已经弃用的字符。我们希望通过整合这方面的工作，把所有东西整合到单个中文标签生成规则集中，然后纳入我们的词汇。抱歉，超过一分钟了。谢谢。

皮提南·古阿莫帕塔纳：

谢谢。迈克尔，有请。

迈克尔·巴德兰：

谢谢。我是拉丁文生成专家组的成员。可以说，我们有很多兼职成员，因为很多人在加入后又离开了。最终，我们的核心团队由 7 人组成，他们分别是比尔 (Bill)、丹尼斯 (Dennis)、哈齐姆 (Hazem)、

马茨 (Mats)、梅卡尔 (Meikal) 以及我们的主席米里亚娜 (Mirjana)。我们有来自不同领域的专家，例如语言专家，一些统一码专家，还有一些注册管理机构和注册服务机构方面的专家。

拉丁文的覆盖范围是众所周知的，因为我们所有人都会说英语，因此也会使用拉丁文。但实际上，全世界有数百种语言都在使用拉丁文，因此我们所面临的挑战是，如何覆盖它们之中的大部分语言，因为要覆盖所有语言是一件不可能的事。拉丁文的主要特点，当然就是语言的多样性了。谢谢。

皮提南·古阿莫帕塔纳：

谢谢迈克尔·阿嘉博士？我们知道你现在是风云人物，得离开这去参加其他与普遍适用性相关的会议。所以，在介绍完新婆罗米文生成专家组之后，你能不能接着分享一下你们专家组是如何完成这项艰巨的任务的？谢谢。

阿嘉·达塔：

谢谢皮提南为我考虑。大家好，我是新婆罗米文生成专家组的联合主席阿嘉·达塔。我们覆盖了四个国家，分别是印度、尼泊尔、斯里兰卡和孟加拉国。显然，我们都知道，新婆罗米文生成专家组是多利益相关方流程的一个成功例子，人们聚在一起达成共识，实现共同的目标。我们的成员来自于学术界、公民社会团体、政府、技术和语言社群以及工业界，他们共同携手，为印度以及我刚才提到的所有国家的所有 9 种文字和 22 种官方语言编制了规则。

从这个角度来看，我们遇到的挑战很多，因为这不仅仅涉及到语言。有些时候，比如人们想要维持他们自己的规则的时候，你得去处理这带来的政治影响，但最关键的挑战在于，在社群内部展开讨论和达成共识。我们走访了每个国家，在每个国家都举行了线下会议，与他们的有关部门讨论，与会受到影响的利益相关方讨论，看看他们有什么疑虑，或者有什么其他想法，然后消除他们的疑虑。说到底，这是在增强他们国家在自己地区的力量。

简单地说，我们遵循了交通信号灯的处理过程，这点很有意思。我们使用了绿灯、红灯和黄灯。如果有人亮绿灯，这意味着这是一个可区分的标签。你可以把它纳入，没有人会担心它。大家都在说，“亮出你的绿灯、红灯或黄灯”，这使得我们很容易专注于讨论我们需要讨论的东西。在整个讨论过程中，我们都采用了这样的规则，不断将字符从一个颜色移到另一个颜色。通过这样做，我们就可以只关注所讨论的 9 种文字中有问题的方面。正如其他同事和整合专家组的成员所理解的那样，我们面对的不是一种文字。我们面对的是 9 种文字。因此，我们还需要将这些规则标准化。我们从这个角度对 [听不清] 规则进行了标准化，确保不同文字之间不会产生冲突。因此，我们跨不同文字对这些规则进行了标准化。来自所有利益相关方团体的近 70 名志愿者组成了这些专家组。我记得，有两年的时间里，我们同时处理着所有这 9 种文字，我们做到了。对于其中一种文字，我们多花了一年的时间，那就是孟加拉文。现在，我想这方面应该也已经解决了。我就说到这里，抱歉我要去参加另一场会议，10 分钟后我会再回来。非常感谢。

皮提南·古阿莫帕塔纳： 谢谢阿嘉博士。接下来我们一起来看看为胜雄和迈克尔准备的一些问题。我们先从中文生成专家组开始吧，因为你们使用的是汉字，很显然，韩文和日文中也有使用汉字。那么，能否分享一下在这方面，中文生成专家组和其他生成专家组是如何协调的吗？谢谢。

黄胜雄： 好的。谢谢。基本上，就像我刚才说的，在组建生成专家组之前，我们就已经成立了技术工程组。所以，我们与来自日本还有韩国的工程团队非常熟悉，不仅仅是工程团队，与来自日本和韩国的语言专家也都非常熟悉。可以说，在加入生成专家组之前，我们就已经在中文、日文、韩文之间建立了一种协调机制，以此确定我们可以如何合作。因为在组建根区生成专家组之前，所有国际化域名都注册在 ccTLD 下，不存在任何码点重叠的问题。但如果我们想要把汉字放在根区中一个单独的空间，我们就必须相互协调。因此在此之前，我们已经预料到了合作的必要。

另一个方面是，我们已经认识彼此很久了。所以我们很高兴能在这个问题上合作。其他组织呢，他们也愿意投入一定的资源来帮助我们解决问题。顺便说一下，因为，举个例子，日文或韩文中的汉字与中文汉字之间有很多字符重叠，因此实际上有很多码点重叠。为了解决重叠问题，我们举行了大量的会议，不仅仅是在 ICANN 内部，在 ICANN 社群之外也是如此，我们还与韩文汉字社群，以及与日文汉字社群举行了多次独立会议。

我记得有一次，我们甚至来了个首尔一日游，就因为另一间会议室被预订了。最终，我们那天在首尔的一个火车站进行了汉字讨论。会议结束后已经差不多是半夜了，然后我们连夜飞回了台北。可以

说，我们进行了非常非常全面的讨论，包括很多工程方面的讨论，很多汉字方面的讨论。我觉得，我们与日本和韩国的社群成员之间产生了很好的化学反应。这种和谐的化学反应，我觉得它是社群合作的最佳状态，使我们最终能够交付可预见的成果，生成我们所期望的字汇。谢谢。

皮提南·古阿莫帕塔纳：

谢谢胜雄。今天，日本和韩国的生成专家组代表也来到了现场。在这个结束后，我们应该一起拍个照。好的。我们继续。接下来，我想问问拉丁文生成专家组的成员迈克尔，因为你们的文字在很多很多语言中都有使用。那么，你们生成专家组是如何最终得出结论的呢？谢谢。

迈克尔·巴德兰：

是的，的确如此。拉丁文在世界各地都有使用，不仅仅是在欧洲或美洲，在非洲、亚洲、澳大利亚以及其他地方，都有语言在使用拉丁文。所以我们面临的第一个困难就是选择我们能够分析的语言，为此我们使用了所谓的 EGIDS 分数，它会为每种语言分配一个分数，这个分数表明了该语言的使用程度。例如，英语的得分为零，因为它几乎在世界各地都有使用。然后比如说降到 5 分，这个分数意味着，该语言在文学作品中以一种标准化的形式被大力使用，但使用范围不太普遍，或者不可持续，再比如 6 分，它意味着该语言被所有年龄段的人用于面对面交流，并且这种情况是可持续的。但由于我们做的事情与域名有关，很显然，只有那些有稳定书写系统的语言对我们才有用。因此，我们决定只考虑 EGIDS 分数不超过 4 分的语

言，外加分数达到 5 分但拥有至少一百万用户的语言。当然了，这个决定有点专横，但我们必须就要纳入哪些东西做出决定。

即使是这些语言，仍然有超过 200 种，然后，我们在网上搜索了这些语言中所使用的字符的来源。我们将所有这些字符都纳入了字汇，但有一些例外。例如，如果某个字符同时用于表示哇哒声，或者如果它看起来像感叹号，那我们就不能把它包含在内了。尽管我们只考虑这些语言，但这并不意味着，EGIDS 分数低于 5 分的所有其他语言会被完全排除在外。这只是意味着我们不会去关注它们的字符。最有可能的情况是，这些语言中 99% 的字符已经包含进来，因为我们发现，我们包含的大多数字符并不是只在一种语言中有使用，而是在许多语言中都有使用。

然后，另一个问题是，我们必须决定一个字符是否应该被认为是另一个字符的变体。为此，我们引入了一个评分系统，我们将每个字符与其他字符进行比较，然后给它们之间的关系分配一个数字，例如，数字 1 表示同形，2 表示几乎相同，3 表示可区分，4 表示不同。我们的 7 位核心成员为每个关系都分配了这样的数字。我们决定，对于任何两个字符，只有我们 7 位成员中至少 5 位给它们之间的关系打了 1 分或 2 分，才能判定它们是彼此的变体。在判定某些字符是否应该是变体方面，我们展开了一些相当激烈的讨论，因为有时候这真的很不容易，它取决于字符所使用的字体，我们要看在不同的字体下，这两个字符是否足够相似，能让普通用户无法加以区分，或者，它们是否仍有一些不同的特征。谢谢。

皮提南·古阿莫帕塔纳： 谢谢迈克尔。我感觉我们有点像在环球旅行一样。关于这方面，在继续下一部分之前，整合专家组、生成专家组或其他人最后有什么想说的吗？好吧，那接下来交回给萨尔麦德。谢谢。

萨尔麦德·侯赛因： 谢谢皮提南。也谢谢胜雄、达塔博士、迈克尔和马克带我们了解了根区 LGR 编制过程中的一些事情。为完成这项工作，来自不同语言文字社群的志愿者们投入了成千上万个小时。

现在，随着工作的结束，ICANN 董事会对项目进行了审议。董事会达成了决议，并要求 ccNSO 和 GNSO 在 IDN 的背景下，在他们各自的政策制定流程中考量根区 LGR。接下来，我们将谈谈 GNSO 和 ccNSO 所做的工作，以便更深入地了解社群辛苦工作的成果是如何被用于政策制定流程中的。

我们首先 — 正如我之前所说，很遗憾，唐娜 (Donna) 和贾斯汀 (Justin) 今天都没能来到现场。他们分别是 GNSO IDN 快速政策制定流程工作组的现任主席和副主席。既然如此，我们请 ICANN 工作人员爱丽儿代为分享一些有关 GNSO 如何在其政策制定流程中考量根区 LGR 的细节。有请爱丽儿。

爱丽儿·梁 (ARIEL LIANG)： 非常感谢萨尔麦德。我是爱丽儿·梁，负责为 GNSO 理事会特别是 IDN EPDP 提供支持。其实，我在这个长期项目中只能算是后来者或者新人，但我非常荣幸能够参与进来，与大家一起为 RZ-LGR 获得采用以及国际化域名获得更广泛的使用做贡献。

当我刚开始加入这个项目时，我有这样一个问题。既然最困难的问题之一是定义 gTLD 和变体，而 RZ-LGR 已经编制出来了，他们已经解决了这个问题，那么为什么我们仍然需要启动政策制定流程？原因是，要开始实施 RZ-LGR，将顶级变体变为现实，你必须有共识性政策来采用 RZ-LGR，同时采用与变体管理相关的政策和规则。只有这样，你才能把顶级变体变为现实。这就是我们为什么需要启动 PDP。

下面我想要简单介绍一下 GNSO EPDP 的背景，以及它是如何运作的。GNSO 理事会于去年启动了这项 PDP，工作组从 2021 年 8 月开始讨论。它以一种混合模式运作，也就是说，它不仅有来自社群的代表和参与者，还有来自 ICANN 董事会和 ICANN 组织的联络人。事实上，迈克尔和宏安，他们是这个工作组的重要成员。他们一直积极地参与到这项工作中来。我很荣幸能在 IDN EPDP 上与他们共事。正如萨尔麦德刚才说的，这个 EPDP 工作组是在唐娜·奥斯汀和贾斯汀·丘 (Justine Chew) 的领导之下，唐娜是主席，来自 GNSO 注册管理机构利益相关方团体，贾斯汀是副主席，来自 ALAC。从这里大家可以看出，IDN EPDP 中的多利益相关方代表性。

工作范围上，他们主要负责两个方面。一是定义所有 gTLD 和变体管理机制。二是 IDN 实施指南及其在今后的更新程序。其中，IDN 实施指南与二级 IDN 注册有关，各签约方必须遵守它。

工作组的章程一共包括 7 个主题、48 个问题，RZ-LGR 是这 7 个主题之一。工作组在考量 RZ-LGR 时，他们注意到一件事，那就是，目前已经存在大量有关 IDN 主题的工作，而且新 gTLD 后续流程 PDP 也提出了很多与 IDN 有关的建议。所以基本上，IDN EPDP 必须建立在现有工作的基础之上，并且考虑这些建议。

关于工作组对 RZ-LGR 的考量以及他们编制的建议草案，我可以稍微给大家介绍一下。新 gTLD 后续流程实际上已经建议在今后生成新 gTLD 和计算它们的变体标签时遵从 RZ-LGR。在实践中，这意味着在未来申请轮次的申请提交系统中，所申请的任何字符串及它们的变体都需要接受基于 RZ-LGR 的算法检查，以确保它们是有效的。EPDP 团队提出了一条与新 gTLD 后续流程建议相符的建议。总的说来，他们建议在考虑现有 gTLD 的变体时遵从 RZ-LGR。这意味着，如果现有 gTLD 注册管理运行机构（例如阿拉伯文或中文 TLD）或注册管理机构希望激活其现有 gTLD 的变体，他们所请求的标签也必须接受基于 RZ-LGR 的检查，以确保它是有效且可分配的。

在这个例子中，IDN EPDP 团队还就这种 RZ-LGR 遵从性提出了一条附加指导。例如，在新 gTLD 申请过程中，应该成立一个域名系统稳定性专家组负责评估所申请的字符串，因为在 EPDP 团队看来，域名系统稳定性专家组是根据 RZ-LGR 审核和确定所申请字符串是否有效的最佳人选。如果申请人对专家组的决定有异议，可使用有限的质疑机制对该决定提出质疑。这一点与新 gTLD 后续流程的建议不符。

我想展示的其他一些例子与 RZ-LGR 运作的技术方面有关。第二个例子是关于不溯及既往原则。RZ-LGR 在它的整个生命周期中预计会不断更新，要么是因为整合了新文字 LGR 而更新，要么是需要对现有文字 LGR 进行更新或修订。这之中可能存在一个非常罕见的情况，那就是，拟议的 RZ-LGR 更新不支持现有 IDN TLD。在这种情况下，IDN EPDP 团队建议对所授权的 gTLD 采取不溯及既往原则以维护根区的稳定性，但他们同时提出了一些其他条件和实施指导。其中一条

实施指导是，生成专家组应该尽最大努力保持与现有 TLD 及获得授权和分配的变体标签之间的向后兼容，以维护根区的稳定性。

我想展示的第三个例子是关于单字符 gTLD。我们知道，现在已经有了中文、日文、韩文 gTLD，所有这些都都在根区中，但我们没有任何单字符 gTLD。对于了解这些语言的人来说，它们的个别性意味着，即使单个字符也有意义，可以独立存在。大家可能对地理位置或姓氏很熟悉。这些就是单字符，但人们都知道它们是什么意思。不过，在根区中，目前还没有任何单字符域名，因为这涉及到很多安全性、稳定性和易混淆性方面的问题。这就是为什么 IDN EPDP 团队也在处理这个问题。他们提出了与新 gTLD 后续流程 PDP 一致的提议，即，建议仅允许在字符为表意文字的有限文字和语言中使用单字符 gTLD。具体来说，这就是指中文、日语和韩语以及汉字，它们可以有单字符 gTLD。

不过，由于可能需要额外的指导或限制，以及需要进一步确定应该允许或禁止哪些字符成为单字符 TLD，我们还有很多工作要做。正因为如此，EPDP 团队希望与中文、日文和韩文生成专家组合作，借助其语言专家的力量来解答这个问题。

以上就是 IDN EPDP 工作的一些例子。可以看出来真的很有意思，未来，我们希望能有更多的社群成员参与进来。我们欢迎大家以观察员的身份参加 EPDP 团队的会议。这些会议都是对公众开放的。我们希望获得更广泛的社群参与，所以欢迎大家，哪怕是作为与会者参加也是可以的。另外，未来，我们还会发布包含初步建议的初步报告，邀请大家进行公共评议。我们欢迎大家就这项工作给出你们的意见和反馈。我就说到这里，接下来交给萨尔麦德。谢谢。

萨尔麦德·侯赛因：

谢谢爱丽儿非常全面地概述了 GNSO 政策工作组如何基于社群的工作开展自己的工作，将根区 LGR 纳入 gTLD 政策中。

接下来，我们要请的是 ccNSO IDN 政策制定流程工作组的主席胜雄，请他跟我们分享一下工作组是如何考虑将根区 LGR 工作纳入 ccNSO 政策中的。谢谢。

黄胜雄：

好的。谢谢萨尔麦德。我需要转换一下角色，现在的我不是中文标签生成规则的联合主席，而是 ccPDP4 工作组的主席。我很高兴能向大家介绍 ccNSO 是如何考虑根区标签生成规则的。我可以给大家举几个例子。如果大家要考虑根区标签生成规则，它是根区的一项技术要求，而在 ccPDP4 中，我们认为它是针对其他 ccTLD 运营商的一项政策要求，规范了他们运营 IDN ccTLD 的方式。可以说，它更多的是一项政策要求。除了考虑要求以外，总的说来，我们将工作组分成了三个小组。

第一个小组负责变体管理，具体来说就是如何管理你的变体，针对所选 IDN ccTLD，你拥有多少个变体。第二个小组负责取消选择要求，叫做取消选择小组，举个例子，你可以把取消选择看作是一项停用政策。我们确立了一些标准，如果符合这些标准，则 IDN ccTLD 需要激活停用政策，以完成政策制定流程。这是第二个小组。第三个小组是字符串相似性小组。我们还需要关注如何处理字符串的相似性。

除了这三个小组的工作以外，基本上，IDN ccPDP4 还需要研究如何在给定的时间框架内实施这类政策。根据拟议的时间框架，我们应该在 2023 年 1 月提交第一份初步报告。这跟我们的计划差不多。

除此之外，因为我们刚刚意识到，EPDP 也提出了单字符顶级域的需求，所以基于这一需求，我们可能需要重新召集所有专家，讨论我们是否需要对此需求提出任何技术上的考虑，不仅仅是技术上的考虑，我们还要联合语言专家和其他专家，讨论我们可以为单字符顶级域需求提供什么样的政策建议。以上就是 ccTLD ccPDP4 到目前为止所做的事情。预计我们的大多数工作都会在明年初完成。谢谢。

萨尔麦德·侯赛因：

谢谢胜雄。ccNSO 也正在将根区标签生成规则专家组的工作成果纳入政策工作中。ICANN 董事会曾在其决议中提到，希望 GNSO 和 ccNSO 相互合作，确保在实施方面保持一致。既然如此，能不能请 GNSO 和 ccNSO 简单谈谈，你们要如何在这些有关 IDN 的政策制定流程上相互合作？

黄胜雄：

谢谢。正如我所说，在 ccNSO 工作中，我们已经指派了一名联络人参与 EPDP 的工作。阿尼尔 (Anil)，ccPDP4 的联合主席，他就是我们在 EPDP 的联络人。他会及时向我们传达 EPDP 的最新消息，同时将我们的情况带回去给他们。除此之外，总的来说，我们需要与 GNSO 进行协调统一，但这并不意味着我们的每一项政策都应该完全相同。因为 ccNSO 和 GNSO，很显然，它们之间是有一点区别的。所以这并不意味着我们的政策应该相同，但在政策要求方面，在如何使用根区标签生成规则作为技术社群提出政策制定请求时的最低要求方面，我们可以统一一下流程。我也很期待他们的反馈。

爱丽儿·梁:

谢谢胜雄。胜雄刚才提到了 ccNSO 在 GNSO EPDP 中的联络人。同样地，GNSO EPDP 在 ccPDP4 中也有联络人。他是来自注册管理机构利益相关方团体的丹尼斯·谭 (Dennis Tan)。当然，在联络层面，他们进行了很多协调，并且密切关注着对方的工作。另外我想指出的是，ccPDP4 和 IDN EPDP 到目前为止已经举行过好几次会议。最近的一次是 7 月份。在会议上，这两个工作组比较和对比了他们在同一主题下的建议，看看哪些地方相似，哪些地方不同。我觉得，这样的会议对两个工作组了解彼此的进展、发现差异及确定原因非常有帮助。正如胜雄所说，没必要完全相同，但至少我们需要知道为什么会出现差异，然后尽量协调统一。以上就是我要补充的，谢谢。

萨尔麦德·侯赛因:

谢谢爱丽儿。请翻到下一页。我们的时间不多了，但如果大家有什么问题，也许我们可以现场回答一个问题。请到麦克风前面来。然后我们将进入本次会议的第二部分。

莱维·塞瑟克

(LEVY SYANSEKE):

大家下午好或者晚上好。讲得太好了。我叫莱维·塞瑟克，来自赞比亚，是 ICANN75 英才计划的学员。我想我的问题是关于拉丁文 LGR 如何才能恰当地包含这么多种语言。我担心的是，在我所在的国家，对于某些语言，必须省略掉英语字母表中的某些字符，才能将其用于这些语言。那么你们要如何确保在拉丁文中恰当地包含其他语言呢？如果你们去观察，比如说，观察字母表中省略了某些字符的相同部落，他们决定在语言中恢复这些字符后，要如何书写它

们。这难道不会造成某种程度的妥协吗？那么这方面的整合工作是如何进行的呢？这是我的问题。

萨尔麦德·侯赛因：

迈克尔？

迈克尔·巴德兰：

非常感谢你提出这个问题。实际上，我们能够包含哪些字符并不重要，因为我们有所谓的最大启动字汇。它是一个字符列表，里面列出的所有字符在理论上都可以包含在根区中，我们可以从这个字汇中进行选择。

我们工作组还有一位语言专家，他在大学里研究非洲语言的语言系中任职。在与某些字符相关的情况下，他非常擅长帮助我们找到和决定 [听不清]。不过，如果你认为有任何语言未得到代表，可以联系我们，我们后面可以再来讨论这个问题。

萨尔麦德·侯赛因：

谢谢迈克尔。我觉得你也指出了根区 LGR 的一个非常好的特点，那就是，它并非一成不变的。现在我们手上的是第 5 版根区 LGR。这是一个不断累积的过程，如果社群觉得需要加入其他字符，可以重新组建生成专家组，提交更新后的流程。随着时间的推移，语言和文字不断发展，我们会继续支持所需的其他任何字符。

接下来，我们进入本次会议的第二部分，同时请各位专家组成员继续留在这里。有请 ICANN 组织的总裁兼首席执行官马跃然，还有 ICANN 董事会的主席马腾·波特曼上台。请上台，接下来就由你们

来感谢社群在过去十多年里所做的出色工作，另外分享一下你们的观点。谢谢。首先有请马跃然。

马跃然： 其实我希望他能先开始。

马腾·波特曼： 我可以啊。非常感谢你们给我这件 T 恤。它上面写的是“我编制了根区标签生成规则。”我真的很珍惜这件 T 恤，不过这个“我”不是我，而是你们。

我谨代表 ICANN 董事会，感谢你们为编制根区标签生成规则（缩写为“RZ-LGR”）所付出的巨大努力。你们就如何以安全、稳定的方式使用各种语言和文字的域名所撰写的提案，对 ICANN 以及更广泛的互联网社群来说非常重要。在过去的 9 年里，来自 44 个不同国家/地区的许多专家聚在一起，组建了生成专家组，努力为国际化域名开发能够巧妙平衡技术需求和语言需求的解决方案。在这些专家组中，有精通特定文字在不同语言中如何使用的语言专家，有在文字的技术实现方面（包括统一码编码、字体的使用以及如何在不同的操作系统中支持多种文字）拥有丰厚知识的技术专家，还有能够就域名系统和 IDN 提供技术见解的域名行业和计算机网络专业人士。其中不乏来自 gTLD、注册管理机构、注册服务机构和 ccTLD 的代表。另外不要忘了本地社群和终端用户，他们在帮助专家了解终端用户对多语言互联网的需求方面发挥了巨大作用。

这些生成专家组完成的志愿者工作量是巨大的，为此投入的精力也是巨大的。自 2014 年第一个生成专家组成立以来，你们共组建了

17 个生成专家组，吸引了 270 多名语言文字社群志愿者的加入，志愿者工作时间超过 10,000 小时，启动了 30 多项公共评议程序。你们总共完成了 26 种文字的提案，分析了超过 386 种语言，而且覆盖的语言比这更多。如果没有你们的奉献和专业知识，这项工作根本不可能完成，它是多利益相关方模型发挥作用的最好证明。

因此，我要感谢和赞赏每个文字生成专家组长期以来为编制 RZ-LRG 做出的巨大贡献。请在座的或 Zoom 会议室里的每个生成专家组的成员在听到我叫专家组名字时举一下手。阿拉伯文生成专家组、亚美尼亚文生成专家组。你不打算在叫到每一种文字时都举一下手吗，萨尔麦德？中文生成专家组、西里尔文生成专家组、梵文生成专家组、埃塞俄比亚文生成专家组、格鲁吉亚文生成专家组、希腊文生成专家组、希伯来文生成专家组、日文生成专家组、高棉文生成专家组、老挝文生成专家组、拉丁文生成专家组、缅甸文生成专家组。我知道新婆罗米文生成专家组在现场。僧伽罗文生成专家组，还有泰文生成专家组。谢谢。请大家和我一起，用热烈的掌声感谢那些为这项工作做出贡献的人。

我还想感谢我在董事会 IDN-UA 工作组的前同事和现任同事，包括宏安，他从 2011 年以来一直在监督这项工作的开展。虽然现在，所有活跃生成专家组在根区 LGR 方面的工作已经完成，但我希望这不代表你们结束了对这项社群工作的参与，我希望你们能继续参与到 ICANN 的 IDN 和普遍适用性相关工作中来。ICANN 董事会和董事会 IDN-UA 工作组将继续推进国际化域名以及所有域名和电子邮件地址的普遍适用性，力求让世界各地的用户能够更轻松地了解互联网。这是我们的首要任务。非常感谢。

马跃然： 在马腾 (Maarten) 后面发言的一个好处就是，他讲的大部分内容都是我要讲的，所以我可以不用 —

马腾·波特曼： 上一次会议你先发言的时候，我也是这么说的。

马跃然： 噢，好吧。谢谢。也许我们应该稍微对一下发言稿。我会尽量简短，不会在说了这句话之后再讲个 15 分钟。

从本质上说，你们所做的事情就像造房子一样，你需要基石，你需要从某个地方开始。你们所做的工作，这项持续的工作就是对这座房子来说最重要的基石之一。没有这样的基石，它就不能称为房子。它就不能成为我们可以使用的任何东西。

下次能不能做件大点的 T 恤来满足我这样正常身材的人啊？这已经是特大号了。可能是我的问题吧，不是 T 恤尺码的问题。

但我要说的是，你们就像是无名英雄。很多人都能够从普遍适用性中受益，他们可以使用自己的文字上网等等。但是想想看，这个事情实际上是从这里开始的，如果不把这些东西放在互联网本身的核心中，其他什么也不会发生。

以前，谢林 (Cherine) 曾经说过，“没有钱的预算只是梦想。”当我在念发言稿时，我现在就在从我的大脑中读取发言稿，我在想，我们可以尽情谈论互联网的包容性和多样性，但如果我们不把这些东西加入进去，它就只是梦想。所以可以说，你们所做的事情让我的梦想成真，为此，我非常感激大家。我会在减肥的时候穿这个。非常感谢。

萨尔麦德·侯赛因：

谢谢跃然 (Göran) 和马腾。我们想给所有生成专家组成员拍张合照。有请参与其中及为之做出了贡献的所有人上台。另外请通过 Zoom 远程参会的各位打开你们的摄像头。麻烦停止展示演示文稿了，分享一下远程参会者的视频，然后大家都到台上来，我们所有人摆好姿势准备拍照。谢谢。

拍完照后，大家可以相互交流。很抱歉，由于疫情限制，我们无法为各位的交流互动提供鸡尾酒。但我们仍然希望大家花点时间，与你们各自语言文字社群的成员多多交流，他们为编制根区 LGR 付出了巨大的努力。那么，大家一起上台拍照吧。谢谢。可以停止录音了。看镜头。

[会议记录结束]