

---

ICANN73 | 筹备周 — 国际化域名 (IDN) 项目最新资讯

大西洋标准时间 2022 年 2 月 22 日 (星期二) — 09:00 至 10:00

皮提南·古阿莫帕塔纳:

(Pitinan Kooarmornpatana) 谢谢。欢迎大家参加 ICANN73 筹备周的首场会议。大家上午好，下午好，晚上好。我是皮提南·古阿莫帕塔纳，来自国际化域名 (IDN) 和普遍适用性 (UA) 项目团队。

在此，欢迎大家参加今天的 IDN 项目最新动态会。本次会议时长 60 分钟，将涵盖以下几个主题。首先，我们将了解一些关于 IDN 实施指南的更新信息，然后我们将关注 IDN 支持的额外文字以及根据标签生成规则 (LGR) 的最新情况。我先来简单介绍下背景信息。在缅甸文生成专家组的总结中，我们还邀请了一位整合专家组的成员马克·布兰切特 (Marc Blanchet)，来分享一些关于第 5 版根区 LGR 的最新信息。

下面，让我们进入第一部分。IDN 实施指南是关于二级 IDN 注册的政策和实践，旨在最大限度降低安全风险并减少消费者困惑。

gTLD 受合同约束，正如 RA 规范 6 第 1.4 节的规定，注册管理运行机构应遵循 ICANN IDN 实施指南。此外，在 RAA 中，附加条款 3 同样提到了注册服务机构也应遵循 IDN 指南。如快速通道流程所述，ccTLD 经理人同样需要遵循该指南。

尽管该指南目前正在更新中，但之前于 2011 年修订的第 3 版指南提供了关于 IDN 的额外分析、数据、运营专业知识与经验。例如，IDN 术语，即 RFC 7940 LGR 格式下的机读表格。

---

*注意：以下内容为针对音频文件的誊写文本。尽管文本誊写稿基本准确，但也可因音频不清晰和语法纠正而导致文本不完整或不准确。该文本仅为原始音频文件的补充文件，不应视作权威记录。*

最大启动字汇 (MSR) 以及安全与稳定咨询委员会 (SSAC) 开展的用户体验调查。此外，由社群开发的根区 LGR 项目中提出的许多关于如何使用文字的提案同样可用。

然后，社群工作组将指南更新到了第 4 版。第 4 版指南中涵盖了共 7 个主题的 19 条指导原则。其中包括过渡小组、IDN 列表格式、IDN 列表一致性与实践、变体标签、标签相似性和易混淆性、政策和规则的发布以及最后的术语等主题。

这些指导原则是由工作组制定的，经过公共评议后定稿，然后于 2018 年 5 月正式发布。

在那以后，GNSO 于 2019 年向组织提出申请，要求在实施之前对该指南进行研究。董事会通过了该申请。去年也就是 2021 年 5 月，GNSO 启动了 IDN 快速政策制定流程 (EPDP)，以解决 SubPro 报告中未涉及的 IDN 相关的其他问题。其中还包括如何管理变体，以及未来如何更新 IDN 指南。

后来，在 8 月，由于存在一些重合，GNSO 还向董事会申请将采纳指南的时间推迟到 EPDP 完成之后。随后，董事会于 10 月给出答复，要求 GNSO 分析并确定主题中重合的指导原则。

鉴于此，GNSO 于次年 1 月做出回应，同意董事会的建议，并确定了下列重合的指导原则，于是出现了额外的 5 条指导原则。

当前，董事会正在考虑第一点，即如何继续推进工作。这就是实施指南的最新情况。在此，我想先暂停一下，看看大家是否有任何反馈、意见或者问题，然后再继续接下来的议程。

萨尔麦德·侯赛因：

(Sarmad Hussain)

我们没有任何问题。谢谢，皮提南。

皮提南·古阿莫帕塔纳：

谢谢。在会议结束时，我们也会留出一些问答时间，到时候再提问也可以。好了，接下来让我们进入下一个话题，IDN 支持的额外文字。

基于 ICANN 的使命，我们正努力让域名能够支持自然语言。从根本上说，这也是 IDN 产生的原因。而基于统一码 (Unicode) 的国际化域名标准 (IDNA) 也促成了 IDN。

IDNA 2008 标准是计算代码点的 [听不清] 机制，同时它也要求根据分析添加更多的限制。然而，该标准并未在结论中执行基于文字的审核。

为此，ICANN 一直在与语言文字社群合作，共同创建安全、稳定的解决方案，这项工作基于根区 LGR 规程。截至目前，在第一轮工作中，我们确定了 28 种语言文字的工作范畴，即最大启动字汇。这里列出的语言文字包括阿拉伯文、亚美尼亚文、孟加拉文，以及泰文和藏文等等。这些都属于为根区确定的 28 种语言文字，这 28 种文字还用于开发二级参考 LGR 和 gTLD。

但实际上，截至目前，在第 14 版统一码中，已经有 159 种文字。这 159 种文字中还包括了一些过时的文字，或者当今社会已经没有使用的文字。统一码 (Unicode) 同样发现，并非所有这些文字都适合用作标识符，而域名也正是标识符的一种。

所以，我们要解决的问题就是，在这 159 种文字当中，哪些是 ICANN 组织应该支持的文字？顶级域名和二级域名都需要解决这个问题。

让我们来详细了解一下统一码。在统一码标准附件 31 中，这 159 种文字被分为了三大类别，即“排除”、“有限使用”和“推荐”类文字。

其中，“排除”类文字有 94 种。从大体上讲，这一类文字是指不常使用、可能已经过时或是在非常有限或小型的社群中使用的文字，并且其中一些文字还存在某些未解决的结构问题。如，呈现问题等。举个例子，婆罗米文、科普特文字和埃及象形文字，这些都属于“排除”类的文字。

另一类是最上面的，应该是“有限使用”类文字，这里打错字了。“有限使用”文字是指在有限社群使用的现代文字。为避免出现安全问题，一些文字可能会被禁用，但在一些情况下仍然可以使用。如巴厘文、加拿大原住民音节文字和切诺基语音节文字等。

最后一个类别是“推荐”类文字，共有 29 种，其中包含一些沿袭下来的常用文字类型，如数字和连字符。这类文字是覆盖范围广，符合当代用语习惯，使用人数众多。相关示例包括阿拉伯文和亚美尼亚文等，正如我们已经在根区 LGR 项目中用到的一组文字。

为回答之前的问题，ICANN 邀请了专家对这些语言文字进行了审核，并在报告中做出了总结。本演示文稿也提供了该报告的链接，总的来说，报告指出，所有“推荐”类文字均可用于根区和二级域名。对于“有限使用”类文字，可能不适用于根区，但视具体情况

可用于二级域名。“排除”类文字既不适用于根区，也无法用于二级域名。

接下来是对文字的评估，为什么“有限使用”类文字应视具体情况而定？因为这类文字仍然可能存在一些潜在的安全性和稳定性问题。因此，语言文字社群需要帮助制定一个安全的解决方案。但正由于其使用有限，我们在制定标签生成规则 (LGR) 时，可能无法找到使用该语言的社群向其咨询。因此，我们可能没有足够的信息来开发参考数据。没有足够的信息，我们就无法制定参考，也就无法对该文字提供有效支持。

这些就是“有限使用”类文字可能带来的一些风险的示例。此处列出的所有文字，切诺基语音节文字、克耶里字母和老傣傣文等，均为“有限使用”类文字。大家可以看到，这里有一些标签示例。其中一些可能看起来像拉丁文的大写字母，但实际上是不同的代码点。看这里的代码点，它们和拉丁文的大写字母 A 有所不同，这可能会让用户陷入混淆或导致网络钓鱼问题。

不仅如此，有些代码点看起来像拉丁文的小写字母，例如这里的 `ceo` 和 `new` 等。但从本质上讲，它们只是看起来像小写字母，实际上是不同的代码点。

另一个类似的示例仍然是视觉上易混淆的问题。“有限使用”类文字的一些字符可能看起来像“推荐”类文字，例如，新傣仂文可能看起来像缅甸文或马来亚拉姆文。你可以用新傣仂文单独创建标签，并用缅甸文创建一个看起来几乎一模一样的标签，然而这两者却是截然不同的代码点。此外，孟加拉文是已经用于根区的“推荐”类文字，而使用锡尔赫特文也可以创建出看起来相似的内容。

另一个风险示例就是呈现问题。由于文字本身的使用有限，一些系统实际上可能并没有合适的呈现机制或字体来处理这些文字。举个例子，尼瓦尔文是一种“有限使用”类文字，如果我们从统一码 11420 代码表中复制该文字字符，然后将其粘贴至软件中（这里假设为 Mac 上的 Chrome 页面），很明显，这些字符无法正确呈现，而是变成了一连串的问号。因此，使用没有适当呈现支持的“有限使用”类文字也会产生一些问题。这个表列出了一些我们经测试发现存在类似问题的“有限使用”类文字。

让我们回到先前的结论，该结论概要总结了“排除”类文字可能不适用于顶级域名和二级域名，“有限使用”类文字视具体情况可能适用于二级域名，这些内容均包含在报告中，我们会发布报告以收集公共评议。该程序于 1 月 5 日启动，并于上周的 2 月 15 日才刚刚结束，报告将于 3 月初公布。现在，我们正在将从公共评议中接收到的意见反馈纳入后续步骤的考量当中。

让我们暂停一下。对此大家有任何问题或意见吗？

萨尔麦德·侯赛因：

这里有一些一般性问题。让我来读给大家听。纳比尔·亚辛 (Nabeel Yasin) 问，“你认为 IDN 会导致互联网最终用户的孤立化和互联网碎片化吗？”还有一个来自普林斯·安德鲁·祖塔 (Prince Andrew Zutah) 的问题。“当你们收到不同语言的 IDN 域名问询时，你们是如何克服语言障碍的？”

聊天室中已经给出了一些答复。雷蒙德 玛玛塔 (Raymond Mamattah) 也提出了一个问题：“普遍适用性与 IDN 的主要区别有哪些？”这些问题将在聊天室中得到解答，不过皮提南，如果你想回答，请直接作答。谢谢。

皮提南·古阿莫帕塔纳：

谢谢。让我选择一个问题来回答吧，其他问题会在问答环节继续回答。IDN 和普遍适用性的区别在于，IDN 更像是我们要做好技术准备，以前，我们没有由社群开发的关于如何正确使用特定语言文字的解决方案。现在，在社群的共同努力下，ICANN 极大地促进了社群工作。我们已经有可以使用的解决方案了。

接下来的问题是，如果我们现在从技术上讲已经使用了这一文字，或许互联网生态系统中的其他软件或服务提供商对此并不了解，就可能发生这种情况，比如，使用拥有国际化域名的电子邮件（即 EAI，国际化电子邮件地址），在一些网站上使用该电子-邮件订阅内容，网站可能会拒绝你的订阅，并表示你的电子邮件存在异常。这是继 IDN 之后的后续问题，在技术已经准备就绪的情况下，我们如何能够让该文字为各平台、系统所接受？

萨尔麦德·侯赛因：

谢谢，皮提南。我还分享了一个明天举办的关于普遍适用性的会议链接。丹尼斯·谭 (Dennis Tan) 提出了一个问题。“呈现问题难道不是和普遍适用性存在一定的关系吗？”

皮提南·古阿莫帕塔纳： 是的。问得好，丹尼斯。谢谢。我认为这个问题的答案需要我们一起寻找。是的，我认为让一个系统接受非传统的 ASCII 域名和电子邮件，也正是普遍适用性的一部分。好的，让我们先进入下一项议程，如果最后还有时间的话，我们再回到这些问题的讨论上来。

接下来我们将讨论根区 LGR 最新动态。我将为大家简要介绍一下项目概况，然后我会将话筒交给缅甸文生成专家组联合主席欧尹梅 (Yin May Oo)，由她来为我们介绍缅甸文根区 LGR 的一些最新情况，接着由马克分享一些关于第 5 版根区 LGR 的信息。

我来简要介绍下根区 LGR 项目，该项目始于社群对于变体顶级域的需求。当时，从 2009 年到 2010 年，社群共同撰写了一份问题整合报告，该报告指出了为启用变体顶级域需要解决的问题。

报告确定了两个问题。第一个问题是没有对变体的定义。从本质上讲，这取决于不同的文字。第二个问题是没有管理变体的机制。

针对第一个如何定义变体的问题，社群还制定了开发根区标签生成规则的 LGR 规程。从根本上讲，正是这些规则生成了根区标签，并确定了变体的定义。

董事会于 2013 年采纳了该 LGR 规程。自此，各语言文字社群便开始针对自己的文字展开工作，开发根区 LGR。与此同时，ICANN 组织也开始针对变体顶级域的管理提出建议。

到了 2019 年，ICANN 董事会作出决议，GNSO 和 ccNSO 应该将 IDN 变体管理的建议纳入考虑范围内，该建议还整合了根区 LGR。如今，GNSO 和 ccNSO 在政策制定过程中已经考量了这些建议。



2020 年，随着各社群针对众多语言文字的根区 LGR 提出了各种提案，ICANN 董事会还发布决议，要求 GNSO 和 ccNSO 考虑根区 LGR 技术应用的建议。在去年 2021 年初，GNSO 发布了新 gTLD 后续轮次的报告，该报告也整合了根区 LGR 的使用。

社群的工作情况大致就是这样的。在本轮根区 LGR 工作中，我们确定了 28 种语言文字，当前，我们正在完成 26 种语言文字的工作。大家可以看到，一些工作可能会持续好几年，一些工作组已经建立了大约五、六年之久。

但在今年初，所有现有生成专家组，也就是我们称之为生成专家组的社群小组，他们的一切工作都即将结束，那么我们就只剩下塔安娜文（马尔代夫）和藏文的语言文字工作尚未完成。因此，我们希望可以成立一个社群专家组，开展为这两种语言文字制定根区 LGR 的工作。

下面由缅甸文生成专家组联合主席欧尹梅分享一些关于缅甸文根区 LGR 的最新进展。有请！

欧尹梅：

大家好。这些就是我们关注的语言，基于（听不清）语言用户人数从 5.5 亿到 50 万而确定。该语言列表是根据我们缅甸文语言文字社群的语言用户数量来确定的。

缅甸文在缅甸全国使用的历史大约已有 200 年了，在 [文化贸易] 中，我们还借用了大量外来词汇，包括 [听不清] 以及许多不同语言的内容，最终形成了东南亚西海岸的这个大熔炉。

多年来，我们从 [听不清] 藏文中吸收了许多 [听不清]，与 [听不清] 语言用户、泰文和其他许多我们发现的历史语言不断交相融汇。

这就是缅甸文的起源，众所周知，缅甸文属于元音附标文字体系，从印度到东南亚均采用该体系，不仅如此，我们还采用了 [听不清] 符号和许多来自不同种族和文化的其他元素。因此，缅甸文有许多辅音和声调符号，很多都不止一个。一个音节中甚至包含了六、七个声调符号。请播放下一张幻灯片。

一个音节可能是一个辅音，也可以是一个辅音搭配一个中间字符，再添加一个辅音，通过更改重音而形成不同的读音，然后加入不同的语气，使发声更具有节奏感或较高的语调，如长音。所以，缅甸文的语音结构很复杂，我们必须确保合成词遵循统一码规范顺序，避免出现任何呈现规范问题，从而确保标签的安全性和可靠性。这个标签 [听不清] 人类能力，一个辅音添加一个 [听不清] 就会发生变化，然后变出更多的形式。请播放下一张幻灯片。

正如大家所猜想的那样，在该地区，有超过 10 种语言文字正在使用。但是，我们已经进行了缩减，重点关注六种语言。我们排除了 [听不清] 代码点、感叹号和已淘汰的字符，所以为覆盖六种语言，我们的字汇中共有 90 多个代码点。遗憾的是，我们不得不放弃那些用户数量较少且 [没有参考信息] 的语言。请播放下一张幻灯片。

在我们分析了适用于标签的代码点后，我们研究了其他语言，以筛选出可能适用的变体。我们已经发现了拉丁字母、希腊字母、西里尔文和亚美尼亚文中存在这样的变体。由于缅甸文是基于 [听不清] 字符的，这是第一个 [听不清]，因此从该字母中，我们发现世界上许多语言同样使用这种满圆字符。

这是我们的研究结果列表。由于我们有圆形形状，还衍生出了下开口、上开口和侧边开口的形状。其中，我们发现这种形状可能成为其他文字的变体。请播放下一张幻灯片。

我们有基于满圆字符的变体，由于这些变体似乎更接近于西方语言，而我们的语言有大量不同类型的声调符号、元音、中间字符或 [听不清]，我们必须标记出哪些代码点属于哪些组别，并且对于不同的组别，也需要遵循一定的顺序来形成合适的 [听不清]。因此我们必须 [听不清]，并根据形成字母的行为来制定规则。请播放下一张幻灯片。

也就是说，我们必须为每种类型的声调符号编写起始规则，指明哪些声调符号不应跟在其他类型的声调符号之后，否则就可能会造成混淆或呈现问题。这样一来，我们可以消除混淆，让用户不会出现困惑，因为一些用户（大多数用户）当然知道文字内容，但仍然可能有一些用户只能通过有限资源来了解。在我们国家的一些农村地区，可能有一些人并不认识全套语言，只知道一些通用语，而他们也可能会希望使用互联网。所以我们必须确保标签不会影响这些用户的阅读能力。

这里有一些标签示例。对于每条规则，我们必须查看通用语的完整列表，并确保没有形成任何有隐患和不可见的内容。请播放下一张幻灯片。

自 2018 年以来的几年间，我们创建了自己的 GP，并且已经完成了提案。这是我们于 2022 年 1 月开展的公共评议，我们仍然希望能够收到更多的意见。在经历了重重困难和各种各样的情况之后，我们仍然希望能够建立缅甸文体系，让使用这种语言的人们能认识到，

互联网并不只适用于英语语言和 ASCII 字符，因此他们无须担心要去学习其他语言，也可以了解到世界上的其他地区，接触到不同的观点。我们希望让社群知道，互联网已经为降低语言障碍做好了准备。谢谢。

皮提南·古阿莫帕塔纳： 谢谢，尹梅。有没有想问尹梅的问题或意见？

萨尔麦德·侯赛因： 谢谢，皮提南。聊天室中有一个问题，虽然和尹梅没有直接联系，但与之前的演讲有关。这个问题是安德烈提出的。“为什么要支持已经没有在使用的过时字母？如果我没理解错的话。”

皮提南·古阿莫帕塔纳： 谢谢你的问题，安德烈。在刚刚结束的公共评议中，我们实际上也问过类似的问题。感谢你提出这个问题。如果你有什么意见，也请告诉我们。我不确定之前是否贴过公共评议的链接，萨尔麦德，如果可以的话，请贴出该链接，那就再好不过了。谢谢。报告中涵盖了一些关于如何处理不同类别文字的详细信息。谢谢。还有其他问题吗？

萨尔麦德·侯赛因： 还有一些其他的问题已经在聊天室中得到了答复。这里有个问题来自穆卢德·哈利夫 (Mouloud Khelif)：“你们与 UA，特别是普遍适用性指导小组 (UASG) 是什么关系？”还有一个问题：“截至目前，有多少 IDN TLD 在根区中获得了授权？”

皮提南·古阿莫帕塔纳： 谢谢。普遍适用性指导小组是由社群领导的指导小组，我们与 ICANN 国际化域名和普遍适用性项目以及 UASG 保持着密切合作。当 UASG [听不清]，UASG 下面还设立了多个工作组。在明天和主会议周上，我们将专门举行关于 UA 的会议。大家有兴趣的话，也可以参加。但总的来讲，国际化域名和普遍适用性项目也会与社群领导的 UASG 合作。

下一个问题是，有多少国际化域名 (IDN) 顶级域 (TLD)... 大约有 150 个 CC 国际化域名顶级域和通用顶级域。一会儿，我将给出具体数字。我会发布一条链接。好的，谢谢，马克，是 154 个。

好了，现在让我们继续吧。下面把时间交给整合专家组成员马克·布兰切特，他将为我们分享下一版根区 LGR 的更新信息。

马克·布兰切特： 谢谢，皮提南。我要讲的内容是关于自上次 ICANN 会议后，整合专家组的工作内容。我们的目的是完成专家组的本职工作，即整合。因此，我们会承担生成专家组 LGR 的工作，并将其整合起来，形成社群用于顶级域和根区的根区 LGR。请播放下一张幻灯片。

我们当前正在开展第五版根区 LGR 的相关工作。第四版根区 LGR 已于 2020 年秋季发布。第四版根区 LGR 共包含 18 种语言文字，如阿拉伯文、孟加拉文、汉字、梵文、埃塞俄比亚语、格鲁吉亚文、古吉拉特文、果鲁穆奇文、希伯来文、埃纳德文、高棉语、老挝语、马来亚拉姆文、奥里雅语、僧伽罗文、泰米尔文、泰卢固文和泰文。

我们还打算加入六种新的语言文字，即希腊文、日语平假名、日语片假名（该语言文字还包含了汉字，但这些汉字已存在于第四版 LGR 中）、韩文（包括特有的语言文字朝鲜文，以及一些已经出现在第四版 LGR 中的汉字）、拉丁文和缅甸文。

有趣的是，许多年前我们收到过提交上来的两种语言文字，亚美尼亚文和西里尔文，但我们决定推迟加入这两种语言文字，也就是说，当时我们并没有将它们整合到根区 LGR 中，因为它们与拉丁文和希腊文有着密切的联系。我们在争取将这四种语言文字整合到一起，然后将其纳入最终版，因为它们有相同的变体。在整合过程中，我们必须切实将不同的跨文字变体，从一种语言整合进另一种。。也就是说，我们需要耐心等待。不仅如此，我们还需要一些时间，以确保所需的不同的语言文字之间存在连贯性。

因此，第五版 LGR 现拥有两套完全整合的相关语言文字，即 CJK（中文、日文、韩文）和 LGCA（拉丁文、希腊文、西里尔文和亚美尼亚文）。从某些方面来讲，第五版 LGR 将是最为复杂的整合版本。因此，我们花费了大量的时间来验证，确保整合内容的正确性。请播放下一张幻灯片。

我们当前正在制定的第五版 LGR 将包含 25 个要素 LGR。要素 LGR 是指针对某一特定语言文字的 LGR。我们有 25 个要素 LGR，是因为我们有 26 种语言文字，而日语的两种文字语言对应一个 LGR，即 26 减 1。与此同时，我们还将生成一个通用 LGR，这是由各要素 LGR 机械生成的。

我们计划于 3 月 24 日发布第五版根区 LGR，以供公共评议。这意味着我们希望尽快收到尹梅刚刚提到的缅甸文版本，以便将其整合进来。

我们衷心 [希望] 大家能抽出时间查阅第五版 LGR，因为再次强调一下，这将是最为复杂的一个版本，而且也最有可能在今后的轮次中以及其他类似情况下使用。这一版本基本上整合了我们收到的所有现有生成专家组 LGR。请播放下一张幻灯片。

跨文字变体将是此次发布版本中非常重要的一部分内容。只有通用 LGR 会纳入全部的跨文字变体。正如我之前所说的，通用 LGR 是从要素 LGR 中机械生成的，我们从不同 LGR 中提取出所有的跨文字变体，并将其整合到一起，同时运用传递性，创建起完整的通用 LGR。通用 LGR 主要用于有效标签的冲突测试。

下面显示的内容实际上就是我们在变体处理方面决定要做的工作。因此，如果要素 LGR 中对某一变体给出了专门的定义，则我们便会将其 [列出] 并整合。出现在 [完整] 文字字汇中的任何文字内部变体，始终会在相关和合适的 LGR 中列出。ASCII 变体始终会列出。拉丁文、希腊文、西里尔文和亚美尼亚文组合的要素 LGR 将列出该组合中的所有变体，但只有拉丁文 LGR 会列出组合之外的变体。

对于之前延期发布的语言文字，西里尔文和亚美尼亚文，其整合产生的变体将添加进第五版 LGR 中，一些其他的要素 LGR（如马来亚拉姆文）因加入新的语言文字而产生的特定变体也将纳入第五版 LGR 中。我们在此呼吁相关生成专家组，如西里尔文、亚美尼亚文、马来亚拉姆文以及其他语言文字的专家组，

对我们所做的整合工作进行验证，这件事非常有必要，其实际目的是确保我们在整合过程中，只向你们的要素 LGR 中添加了新的变体，而不存在任何漏洞或问题。这实际上是对所有 GP 的呼吁，但整个社群都应该仔细检查我们的工作，这样我们才能在根区 LGR 的后续阶段中始终保持良好状态。请播放下一张幻灯片。

这个问题上次已经提到过了，我就不再过多赘述了。我只简单地提一下，通常而言，如果你想要处理或验证一个带有根区 LGR 的标签，应首先使用要素 LGR 验证该标签是否有效，然后使用通用 LGR 验证是否存在冲突，如果仍然没有问题，则应该重新使用该语言文字的要素 LGR 来计算可分配变体，这样一来，你即可得到各式各样的标签，也就是原始标签的变体。

这就是使用根区 LGR 的方法。当然，一些工具，如 ICANN 提供的 LGR 工具，可帮助处理整个流程。请播放下一张幻灯片。

对于整合专家组的后续工作，再次强调一下，我们计划于 3 月 24 日发布第五版根区 LGR，也就是大约一个月以后。我们衷心希望社群，尤其是生成专家组能够检阅并验证我们即将开展的工作。如果你们没有发现任何问题，也请向我们的评论网站发送一条消息，告诉我们你们验证的语言文字不存在任何问题，这将对我们大有帮助。这对我们来说非常重要。如果你们抽出时间进行了验证，请务必告诉我们。

我们暂定于 2022 日历年第二季度的公共评议后定稿。如果社群成立了藏文和塔安娜文（马尔代夫）生成专家组并向我们提交了相应的 LGR，以供最终的整合，可能还会有下一版本的根区 LGR。但是，正如皮提南的演讲所述，这些专家组暂未形成。因此，我们还没有



---

任何关于新版本（即第六版 LGR）的日程安排。我们还没有任何关于这项工作的日程安排。这应该是最后一张幻灯片了，我的演讲到此结束。谢谢。

皮提南·古阿莫帕塔纳： 谢谢，马克。萨尔麦德，聊天室中有什么问题吗？

萨尔麦德·侯赛因： 是的，确实有几个问题。比尔·尤里斯 (Bill Jouris) 提出了一个问题。为何不列出西里尔文、希腊文和亚美尼亚文中出现的跨文字变体，即使拉丁文 GP 并不认为它们属于变体？或许我们可以先回答这个问题，再继续下一个问题。

马克·布兰切特： 好的，我们制定的政策基本上都适用于通用 LGR，通用 LGR 包含所有的跨文字变体，但我们没必要向所有 LGR 中添加大量与该语言文字本身并不相关的变体。实际上，我们决定限制某些要素 LGR 的实施，以确保语言文字的可读性。不过，通用 LGR 将始终包含跨文字变体的完整列表。

萨尔麦德·侯赛因： 谢谢，马克。瓦迪姆·米哈伊洛夫 (Vadim Mikhaylov) 还有一个问题。“第五版 LGR 中的拉丁文并不仅仅是基本的拉丁文，还包括延伸的拉丁文，我的理解正确吗？如果是这样的话，延伸包含了哪些内容，统一码标准中的 A、B、C 也包含在其中吗？”

马克·布兰切特:

其实，这个问题与拉丁文生成专家组更为相关，比尔今天应该也来到了现场，或许可以由他来给出答复。总的来说，每个生成专家组的工作就是通过公认的特定标准，即 [听不清]，来识别所有语言的使用级别。他们会根据文字 LGR 支持的所有语言来确定工作范围，从而创建供审核的代码点集。

与此同时，他们还需要查看最大启动字汇 (MSR)，这个集合通常更大，移除了所有不再使用的过时字形。由此便建立了供审核的初始代码点集。总的来说，这就是支持的所有拉丁文，大约有 200 种拉丁文。

萨尔麦德·侯赛因:

谢谢。萨蒂什·巴布 (Satish Babu) 还提出了一个问题。第五版是完全向后兼容第四版的吗？

马克·布兰切特:

是的。重申一下，最终版本还未定稿，但它应该是完全兼容的，目前我们没有理由认为它不是向后兼容的。如果确实出现不兼容的情况，那么公共评议中将重点说明，以便供大家审核。

萨尔麦德·侯赛因:

谢谢。纳比尔·亚辛还提出了一个问题，已经在聊天室中得到了答复，不过我还是把它读出来吧。“是否所有的 DNS 系统，如 BIND，都能够将 IDN 域转化为 IP？”聊天室中有留言称，在 DNS 系统中，所有的标签都是 A-标签格式，所以我猜这些内容最终会转换成 ASCII 字符，以供内部决议处理。马克，请你来补充一下。谢谢。

马克·布兰切特:

总的来说，这个问题与国际化域名应用更为相关。我们有域名和国际化域名两种形式。其中一种形式是在文字内，用户可以看到真实值，这种字形串通常是在诸如浏览器等场景中呈现给用户的。然而，该字形串并非实际上被编码到 DNS 系统中的内容。DNS 系统仍然采用 ASCII，这样设计的目的是无需更改世界上所有的 DNS 系统。

因此，用户用自己的文字编写的字形串实际上会通过一种叫国际化域名编码 (Punycode) 的算法被转换为迈克尔所写的 A-标签，该标签以 xn-- 开头，后接一个 ASCII 字符列表。在通常或一般情况下，用户不会看到这些过程。这就是 DNS 用于解析 IP 地址的方法示例。

萨尔麦德·侯赛因:

谢谢。雷蒙德 玛玛塔 (Raymond Mamattah) 也提出了一个问题。是否有关于埃维语字符的工作？这是西非国家，尤其是多哥和加纳的主要语言之一。

马克·布兰切特:

我对这方面不太清楚。我不清楚皮提南和萨尔麦德是否了解，不过我没有任何印象。好，迈克尔说拉丁文中包含了这种语言。好的。谢谢。

皮提南·古阿莫帕塔纳:

谢谢。还有其他问题或意见吗？还剩下一分钟时间。好吧，让我们来快速总结一下，在指南更新方面，仍然有一些后续工作要做。

---

对于刚刚结束的公共评议，也就是关于支持用于 IDN 的其他文字，如果大家还想提交一些意见，请随时提交，不过我们即将制作最终的意见汇总文件。还有关于缅甸文的公共评议，该评议将持续到 3 月 3 日。即使跨文字变体看似和大家没有多大关联，该变体中也包含了 O 形和一些常见的字符形状。大家有兴趣的话也可以看一看。

最后，关于第五版根区 LGR，其公共评议有望于 3 月底发布，我们希望到时候大家能帮忙审查并积极发表意见。这就是今天会议的全部内容了。如有任何问题或意见，请随时联系我们：  
IDNprogram@icann.org。

在此，衷心感谢大家参加本次会议。预祝大家度过愉快的一周，ICANN 会议圆满成功。谢谢。

[会议记录结束]