

《从零开始走向语言数字化》

导读

(Introduction to the Zero to Digital Guidelines)

TRANSLATION
COMMONS



土著语言

国际十年 | 2022 - 2032

目录

语言数字化	2
从零开始走向语言数字化系列	3
从零开始走向语言数字化精编版系列	3
如需查阅本文档的英文原版和其他译本，以及其他 Translation Commons 指南，欢迎访问我们的资源页面。	3
谁适合使用这些指南？	4
Translation Commons	4
联合国教科文组织	4

全球的原住民语言正面临着前所未有的危机。随着代际更替，许多语言正面临消亡，随之流失的还有无价的文化知识与相应的文化认同。当一种语言归于沉寂，世人对该语言社群的传说、生态实践、口述历史、艺术、音乐、药用植物知识、技术、精神信仰、价值观以及饮食习俗，可能仅存模糊的印象，甚至全无。

原住民社群在现代社会中日益边缘化的一个重要原因，是他们无法在数字系统中使用自己的母语。这可能产生严重后果，妨碍社会针对他们的独特需求和观点开展工作。这会限制他们参与一些组织的活动，而这些组织制定的政治经济政策通常会影响他们的生活这阻碍了他们向全人类分享自己的智慧，也阻碍了他们向外界完整、准确地表达观点，甚至在必要时无法有力发声。此外，这也使他们难以获取互联网上丰富的信息资源与工具。而且，成员内部要达成共识也可能面临阻力，尤其是在他们散居各地的情况下。最后，这种孤立还增加了外界曲解其文化和价值观的风险。

语言数字化

将母语数字化是原住民社区应对边缘化的一项重要举措。此举不仅有助于他们更平等地参与世界舞台的事务，也能维护其语言的独特性。数字化使其语言能够获得移动端与桌面软件的支持，让母语使用者可以在网络上交流、分享知识，并使用主流软件与设备。

同样重要的是，每个社群都应主动发起并管理自身语言的数字化进程。这有助于社群保持对自身语言数据的所有权与管理权，并降低文化挪用的风险。

力求实现语言数字化的社群应了解以下最佳实践：

- 收集、标注并存储与本族语言相关的资料；

- 在维护数据主权的同时实现语言数据的共享；
- 加入语言和技术标准组织；
- 推动本族语言在数字系统中的使用，以服务于日常交流、内容出版及母语应用；
- 开发先进的自然语言处理工具，为其语法、正字法，以及机器翻译、人工智能提供支持

从零开始走向语言数字化系列

从零开始走向语言数字化系列是 Translation Commons 为支持联合国“国际原住民语言十年”计划而发起的持续性项目。该系列通过提供一个全面的分步骤框架，协助收集和整理原住民语言数据，让语言走进互联网，以填补信息缺口。

目前已发布的多语种指南包括：

- 《〈从零开始走向语言数字化〉导读》
- 《从零开始走向语言数字化：如何让您使用的语言走进互联网》
- 《语言数据收集指南》
- 《术语指南》
- 《语言数字化的益处》
- 《字体创建指南》

从零开始走向语言数字化精编版系列

Translation Commons 正在制作这些指南的精编版，以便在联合国教科文组织（UNESCO）的资料库中共享。截至当前，已发布的精编版包括：

- 《〈从零开始走向语言数字化〉导读》（即本文档）
- 《语言数据收集指南》精编版

如需查阅本文档的英文原版和其他译本，以及其他 Translation Commons 指南，欢迎访问[我们的资源页面](#)。

谁适合使用这些指南？

本指南面向参与原住民语言复兴工作的个人与组织，包括：

- 原住民语言倡导者及相关社群
- 研究人员与学者
- 政府机构与非政府组织
- 技术开发者

Translation Commons

Translation Commons 是一个非营利志愿者社群，致力于支持语言数字化、指导语言专业人士，并为语言行业提供课程和资源。我们的核心项目之一是**语言数字化倡议**，它为希望实现语言数字化的原住民及少数民族语言群体提供路线图。我们提供教程与工作坊，开发工具与应用，并促进语言倡导者与行业专家携手合作。通过支持公平的数字化访问，我们的使命得以推进——帮助有意愿的语言群体参与全球在线活动，并在母语环境中享受计算机应用带来的便利。

联合国教科文组织

自 2003 年通过《普及网络空间及促进并使用多种语言的建议书》以来，联合国教科文组织始终积极支持语言数字化实践。此举与推动多语内容数字化及数字系统发展的目标一致，旨在确保所有文化，包括原住民文化，都能拥有网络空间并使用自己的语言。

此外，该《建议》敦促各成员国鼓励并支持能力建设，以便在互联网上制作本地及原住民内容。因此，在“国际原住民语言十年（IDIL2022-2032）”框架下，联合国教科文组织致力于推动对原住民语言社群的数字化赋能。