

发展与知识产权委员会（CDIP）

第三十四届会议

2025 年 5 月 5 日至 9 日，日内瓦

“支持政策设计用知识产权和创新数据标准化、
充实化和经济分析系统（2.0 版）”摘要

秘书处编拟

1. 本文件附件载有“支持政策设计用知识产权和创新数据标准化、充实化和经济分析系统（2.0 版）”摘要。
2. 该系统是在发展议程项目“系统化统计数据并制定和实施对采用知识产权制度的影响进行评估的方法”（文件 CDIP/26/4）的背景下开发的。
3. 请委员会注意本文件附件中所载的信息。

[后接附件]

“支持政策设计用知识产权和创新数据标准化、充实化和经济分析系统（2.0版）”摘要

本文件概述了旨在对知识产权和创新数据进行标准化、充实化和分析的系统 2.0 版的主要改进，其目的是提供全面的经济分析，为政策设计提供依据。2.0 版是在发展议程项目“系统化统计数据并制定和实施对采用知识产权制度的影响进行评估的方法”（文件 CDIP/26/4）框架下开发的，并对 1.0 版的基础工作进行了重大改进。主要更新内容来自萨尔瓦多、不丹和印度尼西亚等试点国家的经验，纳入了一个更强大、可扩展和更有洞察力的系统，以满足各国政府和知识产权局的不同需求。

一、2.0 版的主要改进

a. 强化数据质量和严格检查

应用该系统 1.0 版分析参与试点的国家所提供的知识产权数据，凸显了数据质量方面的多种挑战。为此，2.0 版在数据处理的各个阶段加强了数据质量检查。在萨尔瓦多的试点中，发现知识产权数据存在不一致的情况，在吸取了这一经验教训后，现在进行更严格的质量评估，以确保所输入数据在分析开始前的准确性和完整性。

例如，通过交叉参考多个国家来源和国际数据库的数据，该系统能够更好地发现数据中的差异和差距。这些改进保证了系统的产出以可靠的信息为基础，从而提高了从数据中得出的政策建议的可信度。不丹和印度尼西亚的加入进一步加强了这些程序，为不同的数据环境和机构设置提供了宝贵的验证。

b. 协作式数字工作环境

2.0 版的一个重要更新是建立了一个 GitHub 存储库。该存储库可以集中管理所有与系统相关的代码、数据集和方法，为产权组织和参与的各知识产权局提供一个开放的协作式工作空间。通过使用 GitHub，不同地区的利益攸关方可以访问该系统，并为该系统的持续发展做出贡献，促进知识和最佳做法的实时交流。

尽管该存储库目前只供产权组织使用，但在项目结束时将向公众开放。这将使最初试点小组之外的其他国家也能利用该系统对本国的知识产权数据进行分析，从而扩大该系统在全球范围内的影响。这种方法将为知识产权数据管理和分析创造一个透明、开源的环境，为各国知识产权局提供独立管理数据集的工具。

c. 简化数据摄取和清理

在 2.0 版中，由于引入了通用变量词典，数据摄取过程得到了改进。该词典统一了不同国家知识产权数据中使用的标签和分类，确保了数据输入的统一性和可比性。通过在数据摄取时对原始数据进行标准化，该系统大大减少了因标签不一致、分类错误或条目不完整而造成的错误。

新的清理流程会自动应用这些统一的标签，从而简化初始准备阶段。这使该系统能够快速有效地处理大型数据集，从而在更短的周转时间内产生可操作的见解。此外，该系统还能灵活地整合新的国家数据，使未来的试点国家能够更快地采用这些流程。

d. 指标的制定和编纂

2.0 版的突出特点之一是用于指标开发的编纂程序。这一新流程使从原始知识产权和创新数据中创建指标的过程实现自动化，提供了从数据摄取到政策相关见解的直接途径。这些指标提供了系统模块所交付的数据分析处理结果的具体信息，为用户提供了分析工具和资源，以更好地了解国家知识产权制度和创新生态系统的动态。

e. 创新复杂性模块的新功能

[创新复杂性模块](#)在 2.0 版中进行了大幅扩展，纳入了更多可视化内容和指标，使分析师能够提供更多可操作的见解。尤其是以下方面：

- 区分既有能力和濒危能力，帮助各国识别失去宝贵知识的潜在威胁。
- 将关联密度细分为创新的不同维度，以了解一国知识的哪些来源为其带来了知识多样化的机会。
- 区分全球和本地的创新机会，以便根据其所包含的风险增加额外的层次，对多样化路径进行等级化。
- 确定实现一种能力所需的创新成果数量，从而增加创新成果的层次，以实现等级化。
- 表明创新领域对一个单位而言相对复杂。这使得复杂性指标与被分析国家的视角相对应。
- 分析器功能可为任何单位（国家）的组合自动生成一组数字和指标。

f. 项目产出的重新设计和自动化

2.0 版重新考虑了项目产出的总体结构。根据试点国家的反馈意见，该系统纳入了根据产权组织数据为任何国家自动生成指标的功能。这确保了报告、图表和其他产出的高效生成，并将人工干预降至最低。

已考虑重新设计[创新生态系统基础模块](#)，使该模块的内容能够顺利地向用户介绍后续系统模块的技术方面，重点是：性别创新差距、创新热点和创新复杂性。

二、经验教训和下一步工作

2.0 版的开发揭示了采用灵活而又标准化的方法管理不同国家数据集的重要性。在应对 1.0 版中遇到的数据质量挑战的过程中，特别是在萨尔瓦多试点期间，强调了严格验证和统一过程至关重要。不丹和印度尼西亚的经验突出表明，不同的国家数据环境如何需要量身定制的摄取和清理程序。此外，在 GitHub 上建立一个协作性的数字工作环境，强调了知识产权局与产权组织之间透明度和实时协作的价值。事实证明，引入通用变量词典是一个重要的里程碑，因为这样就可以对数据进行标准化分类，而不论提供数据的国家是哪个。这种标准化，再加上对创新复杂性模块的改进，展示了在应用通用工具的同时，也能深入了解各国具体情况的有效性。

展望未来，3.0 版将根据与现有试点国家的持续互动以及后者预计提供的新数据进一步完善。重点将放在指标生成和分析的关键环节的自动化上，使该系统具有可扩展性，可供更多国家使用。此外，开发培训材料的工作正在进行之中，以便在 GitHub 上提供这些材料，预计将在 2025

年下半年为试点国家提供培训。这些步骤对于确保该系统保持相关性、灵活性和适应试点国家及更广泛的全球创新图景不断变化的需求至关重要。

[附件和文件完]