

发展与知识产权委员会（CDIP）

第二十六届会议

2021年7月26日至30日，日内瓦

关于“妇女与知识产权：数据汇编和分享”的报告

秘书处编拟

1. 发展与知识产权委员会（CDIP）在2018年11月19日至23日举行的第二十二届会议上，在“知识产权与发展”议程项目下讨论“妇女与知识产权”这一主题时，作出了以下决定：

“委员会审议了文件 CDIP/22/16 Rev. 2，并通过了本文件附件中所载的墨西哥关于‘妇女与知识产权’的提案。附件将转送2019年产权组织大会。”（[主席总结](#)第9段）

2. 墨西哥提出的上述提案，除其他外，要求秘书处采取一系列行动，并决定“在第二十六届会议上再次审议‘妇女与知识产权’议题的未来方向”。

3. 本文件与文件CDIP/26/8一起，旨在为委员会在本届会议上再次审议“妇女与知识产权”议题提供便利。本文件回应了墨西哥提案中汇编和共享知识产权领域按性别分列数据的要求1和2，而文件CDIP/26/8则报告了秘书处为落实要求3、4和5所采取的行动。报告所涉期间是2019年和2020年。

一、秘书处对墨西哥提案的落实情况

4. 下文介绍了秘书处在落实墨西哥提案的要求1和2方面的行动。

要求1：汇编关于知识产权所有者和创造者性别成分的可比分列国际数据。

5. 国际专利申请不包括关于申请人性别的信息。然而，这一信息对知识产权界很有价值，例如，有助于发现和了解潜在的创新差距。因此，自CDIP/22以来，秘书处继续致力于寻找各种方法，以进

进一步了解性别平等和知识产权的交叉点，并在此基础上汇编按性别分列的国际可比数据。这项工作主要是通过知识产权和创新生态系统部门的经济学和数据分析部（DEDA）来进行的。

6. 统计和数据分析司在其主要旗舰出版物和网络数据中心发布了同期内按性别分列的 PCT 统计数据，具体如下：

- 6.1. PCT 年鉴（2019 年）
- 6.2. PCT 年鉴（2020 年）
- 6.3. 世界知识产权指标（2018 年）
- 6.4. 世界知识产权指标（2019 年）
- 6.5. 世界知识产权指标（2020 年）
- 6.6. 知识产权统计数据中心（所涉期间内定期更新）

7. 创新经济科继续进行关于性别平等和知识产权的方法论和经济学研究。报告期内的研究有两个主轴：（i）进一步完善性别消歧方法；和（ii）改进并扩大知识产权各种数据的性别分类。

8. 在报告期内，创新经济科通过以下方式进一步完善了性别消歧方法：

- 8.1. 在 [Harvard Dataverse](#)（WGND 1.0）和 [GitHub](#)（WGND 1.1）中发布《世界性别名字词典》（WGND），创建更多易访问的版本。关于这一点的更多信息见下文第 10 段。
- 8.2. 纳入更多来自新的国家的来源，提高 WGND 的国际覆盖面（WGND 2.0）。这一新版本发布在 [产权组织网站](#)、[Harvard Dataverse](#)（WGND 2.0）和 [GitHub](#)（WGND 2.1）。
- 8.3. 分析对某些亚洲语言的名字使用性别词典的稳健性。这种分析指出了中文和韩文名字罗马化方面的问题。这种分析的主要意义在于，未来的性别分析将受益于：（i）扩大以亚洲原始字符表示的知识产权数据的范围；（ii）在未来的 WGND 版本中扩大以原始字符表示的亚洲名字的范围。
- 8.4. 分析性别消歧的替代方法，如机器学习插补法。与基础设施和平台部门的先进技术应用中心（ATAC）合作，进行了初步探索。
- 8.5. 构思性别均等预测的系列方法。这些方法可以对何时能实现性别均等的不同情形进行评估。不同的预测情形可以应用于任何新的性别消歧数据。先进技术应用中心根据产权组织性别和多样性专家的要求，定期对 PCT 发明人和产权组织成员国大会代表名单进行了应用。

9. 在同期内，创新经济科通过以下方式对改进、扩大和分析按性别分列的知识产权数据进行了研究：

- 9.1. 利用 WGND 实现 PCT 发明人性别消歧的自动化。构思、设计并全面实施了一种方法，以应用于基于 WGND 1.0 的历史单位记录数据和新的 PCT 单位记录数据。
- 9.2. 基于互补的亚洲原始字符数据，对非罗马化的发明人姓名，实现 PCT 发明人性别消歧处理自动化。构思、设计并全面实施了一种方法来应用 WGND 1.0。该方法需要用到一些额外数据，这些数据包含亚洲原始字符的发明人姓名，由专利和技术部门的 PCT 信息系统司业务和支持科定期提供。

- 9.3. 将性别消歧扩大到 PCT 汇编数据中的个人申请人和法定代理人。构思、设计并全面实施了一种方法，将其应用于基于 WGND 1.0 的历史单位记录数据和新的 PCT 单位记录数据。目前正在对该方法的结果进行进一步分析。
- 9.4. 将性别消歧扩大到国家知识产权汇编数据中的专利和实用新型发明人。构思、设计并试行一种方法，以便利用 WGND 1.0 应用于 DOCDB/PATSTAT 单位记录数据中的历史数据。这项研究的初步结果已在《世界知识产权指标》系列中发布。
- 9.5. 分析海牙工业品外观设计单位记录数据的性别消歧的可能性。应用 WGND 1.0 对海牙体系的数据进行了收集和分析。作为这一分析和后续分析的初步结果，得出的结论是要扩大 WGND 来源的国际覆盖面（见第 8.2 段）。使用 WGND 2.0 对相同的数据进行了再次分析。对海牙体系单位记录数据的进一步分析，表明有两个主要制约因素：（i）新的海牙成员在不同时间段的加入，导致样本产生很大波动；和（ii）一些海牙成员国的外观设计法不要求说明设计人。这两个限制因素限制了对全球性别平等趋势（包括几个国家的趋势）所进行的解释。在分析过程中，我们认为在公布结果之前还需进一步开展工作。
- 9.6. 分析马德里商标单位记录数据的性别消歧的可能性。和海牙一样，应用 WGND 1.0 和 2.0 对马德里体系的数据进行了收集和分析。对马德里体系单位记录数据的分析，表明有一个主要制约因素：商标数据只采集申请人的信息，而申请人可以是自然人和非自然人。只有自然人（即商标申请人个体）的性别可以用性别词典来消歧。对申请人个体的性别分析为马德里体系性别均等性提供了不完整的特征描述。然而，这可能是对使用该体系的小型实体（如企业家或小微企业）的潜在特征描述。这后一种解释需要在公布结果之前进一步分析。

要求 2： 分享按性别分列的数据收集方法和过程、对指标的使用、监测和评估方法，以及知识产权相关的性别差距分析。

10. 在同期内，秘书处采取了若干举措，以分享对妇女和知识产权相关主题的分析和方法。特别是，创新经济科探索了几种方式，公开分享性别消歧方法第一版（基于 WGND 1.0），例如：
 - 10.1. 与学术网络、联合国实体（如联合国妇女署和教科文组织）以及私营公司的相关公众直接接触。
 - 10.2. 使用哈佛 Dataverse 存储库，这是一个免费的数据存储库，向哈佛社区内外任何学科的所有研究人员开放（[Dataverse 存储库](#)）。
 - 10.3. 使用 GitHub 存储库与更广泛的受众分享 WGND 数据，并在此基础上创建工具（[GitHub 存储库](#)）。
11. 在上述分享经验的过程中，秘书处注意到，这些渠道对技术受众有很好的效果，但在接触更普遍的受众方面有一些局限性。具体来说，直接分享渠道对一些知识产权局（如美国专利商标局和加拿大知识产权局等）和学者产生了影响，但这些主管局和学者大多在发达国家。数字共享的经验表明其传播范围要广得多，在产权组织网站上有 1,400 多次下载，在哈佛大学 Dataverse 有 1,353 次下载。¹

¹ 到目前为止，要评估 GitHub 的使用情况还为时过早。

12. 为了支持更广泛、更快速地推广性别方法和分析，创新经济科编写了一份新的“指导原则”文件草案，重点介绍如何判断知识产权中的性别。这些指导原则介绍了知识产权数据性别消歧和相关性别指标方面的基本知识。

13. 作为对这些指导原则的补充，创新经济科创建了一系列关于“[创新、创造和性别](#)”的专门网络内容，向更多的普通受众解释秘书处的经济和方法论研究。²初步反馈表明，这不仅可以用来宣传产权组织在这一主题上的工作，还可以用来突出其他国家——如美国、英国、加拿大或智利——的知识产权局以及其他合作伙伴，如联合国妇女署或教科文组织所做的工作。目标受众包括成员国的政府官员和国际政府机构、世界各地的学者和学生以及对该主题感兴趣的任何人。

二、下一步工作

14. 从秘书处在报告期内的经验来看，有一系列的具体步骤需要考虑。

15. 首先，秘书处需要进一步简化与性别有关的方法论内容，让技术性不强的受众能够获得。这将允许向更多的成员国传播这些方法，并建设本地能力，使之形成对性别作出反映并按性别分列的分析。这些传播工作将包括发布关于“创新、创造和性别”的研究论文和网文，以及组织有针对性的能力建设讲习班和信息分享会。后者是在 CDIP 关于“妇女参与知识产权”的决定所推动的分享会的框架内。

16. 第二，秘书处将继续对扩展和分析关于知识产权所有者和创造者性别的可比分类国际数据的方法进行研究，特别是那些旨在将分析扩展到其他形式的知识产权（如实用新型、工业品外观设计或商标）的方法。

17. 请委员会注意本文件中所载的信息。

[文件完]

² 在起草本文件时，这些内容正在往产权组织网站上添加。