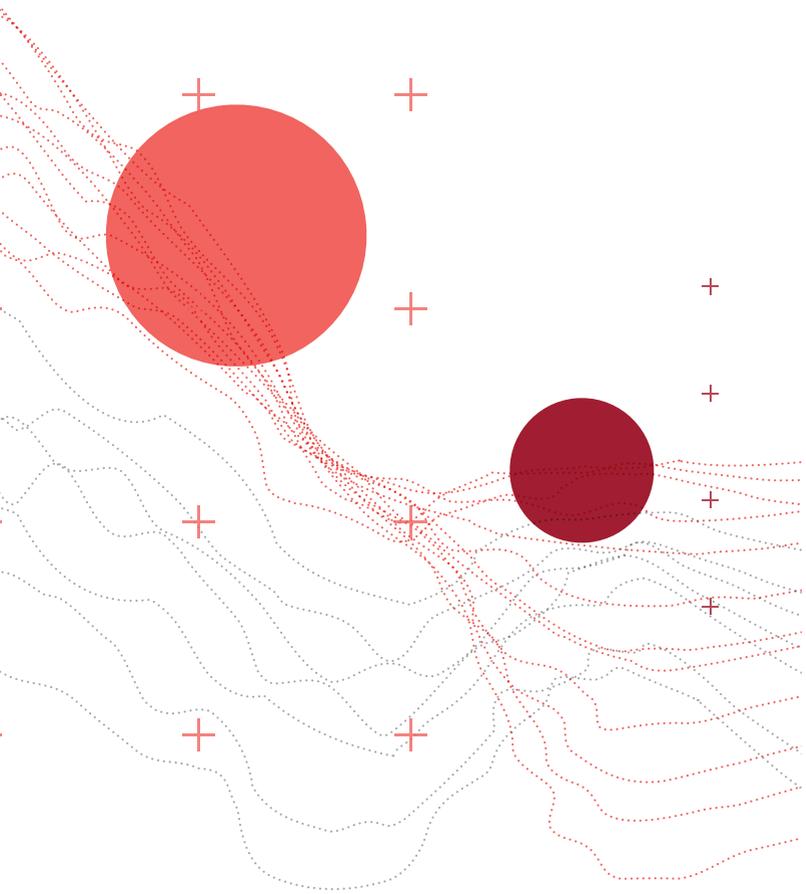


利用创新和知识产权数据进行性别分析指南





为什么要对创新和知识产权数据进行性别分析？

了解女性和男性如何平等地获取和利用知识产权制度,是确保其创造力转化为经济、社会和文化发展的关键所在。

对于国际和国家一级的知识产权局来说,可能很难从知识产权和其他创新数据中找出发明人和创造者的性别。这使得创建为性别政策提供依据所必需的知识产权指标和经济分析变得更加困难。

这份简短的指南总结了编制创新和知识产权性别指标的最佳做法。希望有意开展研究和进行政策设计的政府部门和研究人员能够采纳,以促进实现更均衡的性别平衡。

用于编制创新和知识产权性别指标的路线图

A 收集创新和知识产权性别数据

有许多不同的方法来收集按性别分列的创新和知识产权数据。本指南探讨了这些方法的优点和局限。获取创新和知识产权性别数据的两种主要方法如下：

- **直接收集性别数据**，例如在知识产权申请流程中收集。
- 对现有创新或知识产权数据进行**间接性别归类**。这可以通过使用二手来源、语义规则或性别-人名词典来完成。

B 使用产权组织的性别-人名词典

世界知识产权组织（产权组织）和其他一些国家的知识产权局根据性别-人名词典，如产权组织的《世界性别人名词典》，对知识产权数据集进行性别归类。

C 选择使用哪项性别敏感指标

通过使用按性别分列的知识产权数据，可以制作多项有用指标，尽管这些指标有不同的含义和用途。

A 收集创新和知识产权性别数据

直接收集性别数据

最简单的方法是在收集创新或知识产权数据时要求提供性别信息。例如，就专利数据而言，可以在专利申请表中增加字段供发明人和申请人勾选。

第II栏	申请人	<input type="checkbox"/> 该人也是发明人	性别： <input type="checkbox"/> 女 <input type="checkbox"/> 男 <input type="checkbox"/> 其他
------	-----	----------------------------------	---

这种方法的优点是**可以直接从源头获取**与性别有关的信息。自我指称也有一个好处，就是允许用户根据自己想要选择的类别来确定自己的身份。

缺点是**不能追溯性适用**，这可能会影响对过往的分析。它也可能要**花费大量资源**，因为在知识产权申请表中实施一项系统更改并让用户适应，可能需要大量的时间和金钱。此外，由于隐私规定，可能**很难要求申请人披露性别信息**。但是作为非必填字段可能会影响到这项工作的统计范围。

有用技巧

在任何可能的情况下，收集每一件知识产权申请的所有创造者（如发明人、设计人、作者）的性别。

性别差异可能出现在创造团队内部。只有获悉所有创作者的性别时，才能计算出团队指标。

间接归类

使用二手来源

对现有创新和知识产权数据进行性别归类的一种方法是利用含有性别信息的二手来源。包含完整覆盖面和性别标记的常见二手来源包括**国家个人信息登记簿**或**国家就业调查**。

如果每个数据集中的个人都清楚地得到识别，这些来源就可以与知识产权数据关联起来。例如，可以使用**社会保障号码**或**国家公民身份证号码**将个人申请人、发明人和创造者与这些来源进行匹配。

知识产权数据		全国就业调查		
社会保障号码	发明人姓名	社会保障号码	雇员姓名	性别
11111190-2	约翰·史密斯	11111190-2	约翰·史密斯	男
11111198-2	约翰·S·史密斯	11111191-2	约翰·亚当斯	男
11111192-1	玛丽·史密斯	11111192-1	玛丽·S·史密斯	女
11111192-1	玛丽·S·史密斯	11111193-1	玛丽·亚当斯	女

这种方法的主要优缺点与二手来源的质量有关。换句话说，性别归类的质量好坏将与二手来源的可靠性正相关。

优先考虑那些能全面覆盖个人并定期更新信息的来源。这种方法另一个优点是它的实施成本较低。

有用技巧

请注意，一些国家的身份证号码保留了一位数字用于性别信息。在这种情况下，不需要二手来源来确定性别归类。

例如，在大韩民国的身份证号码 (RRN) 中，第七位数字被保留用于指示性别：

641226-4191838 = 女

641226-3191835 = 男

数据

19,136
8,768
21,474

3,220
16,115
10,003
1,908
1,245
11,481
7,644
610

Patent Families

107,12

198,20

21,88,65

37,799

7,8374

7,906

1,731

878

100,72

33,977

987

62,989

14,778

14,253

2,108 +

6,535

18,957

13,716

6,526

489 +

976

3,186

2,583

1,6051

1,4533

1,8794 +

53,794

使用语义规则

另外,可以使用关于个人姓名的**语言习惯和语义**来进行性别归类。

例如,英文中的“Mrs(夫人)”“Ms(女士)”或“Mr(先生)”这样的**姓名尊称**可以明确地指涉某一特定性别。然而,像“Dr.(博士)”或“Prof.(教授)”这样的头衔则需要更多信息来确定性别。

同样,在一些命名习俗中,**人名结构**可以指涉某一特定性别。例如,在西班牙文或葡萄牙文的命名习俗中,以“o”结尾的名字通常指涉男性,而以“a”结尾的名字通常指涉女性。

不幸的是,许多语言没有基于语义规则的明确性别区分,而且数据库可能不收集和存储名衔信息。

这种方法的一个优点是,只要命名习俗不发生重大变化,就可以**追溯性适用**。

有用技巧

请注意,在某些语言中,在复合名中混合女性名和男性名是很常见的,如西班牙语中的“José María”,法语中的“Jean-Marie”,等等。

在这些情况下进行性别归类时,考虑有关性别的语义规则和名的位置往往会有用。

使用性别-人名词典

性别-人名词典是姓名及其最常见关联性别的列表。这些列表可用于根据知识产权数据中所列的发明人和创造者的名来确定其性别。

知识产权数据			性别-人名词典	
发明人的姓	发明人的名		人名	性别
史密斯	苏		约翰	男
史密斯	乔治		乔治	男
史密斯	玛丽		玛丽	女
史密斯	约翰		苏	女

使用性别-人名词典有很多好处。只要命名习俗不发生重大变化，这种方法可以便利地追溯性适用，而且实施成本很低。

性别-人名词典可能有一些局限。它的可靠性取决于字典对知识产权数据的人名中体现的语言和民族的调适程度。在某些情况下，将需要接受一定程度的错误。另外，在某些语言中，许多人名既可用于女性，也可用于男性。

在哪里可以找到这些词典？

现存很多人名列表，应该找到最适合您的数据的那个列表。产权组织汇编了若干主要的公开来源，创建了《世界性别人名词典》。可以在[产权组织网站](#)上找到《世界性别人名词典》和所用公开来源的链接。

如何使用性别-人名词典？

在下一节中提供了使用字典对知识产权数据进行性别归类的分步骤指导。

性别-人名词典

B 使用产权组织的《世界性别人名词典》(WGND)

第1步 准备创新或知识产权数据

理想的情况下,知识产权数据将带有一个包含每个发明人或创造者(如果可能的话还有个人申请人)名字的字段。

可以按以下方式准备数据:

将姓和名分开。如果有多个名字,则全部保留,但要分开在单独字段填写。请记住,在许多语言中,顺序可以决定性别。

删除前后空格、双倍行距和不必要的字符(如:'-', ;),并替换重音和其他非拉丁字符(如:é→e, ó→o, ß→ss, à→a)以**清理名的字段**。

保留地址或国家字段提供的**国家或语言数据**。

第2步 使用最合适的《世界性别人名词典》版本

从我们的研究人员性别资源库(R4RGender) **下载**《世界性别人名词典》的最新版本。检查该网址上的其他工具。

根据所在国家(如法国、智利)或语言(如阿拉伯文、中文) **选择**《世界性别人名词典》的**版本**,并将其应用于第1步中的人名列表。

请记住,可以对数据**应用多个版本的**《世界性别人名词典》。

第3步 评估归类结果

检查所得结果! 每个数据集都是不同的。也可以向《世界性别人名词典》中添加新的条目。

分析性别归类的**覆盖范围**。已作了性别归类的记录所占比例是多少?

手动检查归类结果。检查出现频率最高的100个女性、男性和未归类名字,以找出 (i) 错误归类和 (ii) 在所在国家有明显性别特征的未归类名字。

根据(i)和(ii)**创建自己的性别-人名词典**。用它来改进第2步的结果。

将反馈意见发送给我们!

有用技巧

如果找不到您所在国家的性别人名词典,请尝试从使用相同语言的国家中找一个。也可以通过对您知识产权数据中最常见(例如,前200个)的名字手动进行归类,创建自己的词典。



使用
W
G
N
D

C 选择使用哪项性别敏感指标

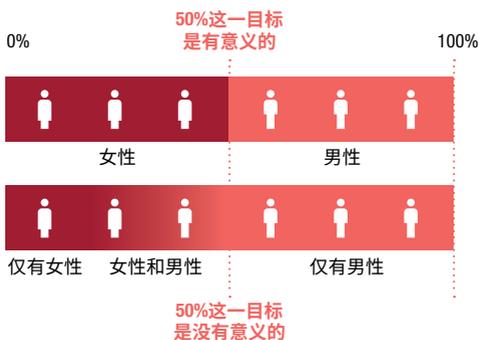
利用按性别分列的创新和知识产权数据，可以制作多项有用的指标。然而，它们有不同的含义和用途。

性别基准

被列为发明人或创造者的女性是最直观的指标，因为它可以为性别政策提供明确目标，例如50%的女性被列为发明人或创造者。

主要的局限是，知识产权申请可以有多个发明人或创造者，所以总人数并不反映知识产权申请的实际数量。

至少有一名女性发明人或创造者的申请所占百分比与其他常用知识产权指标使用相同单位。



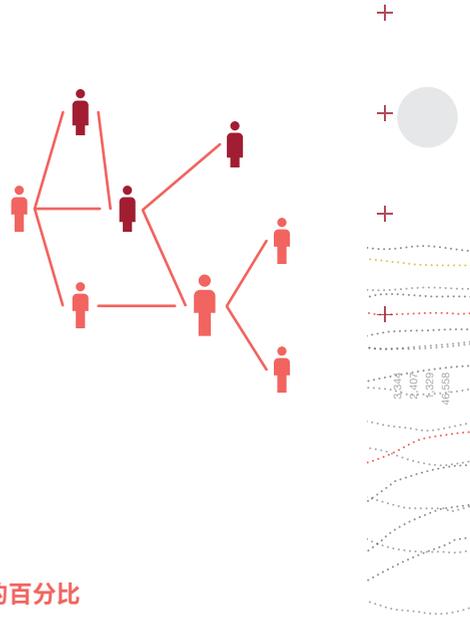
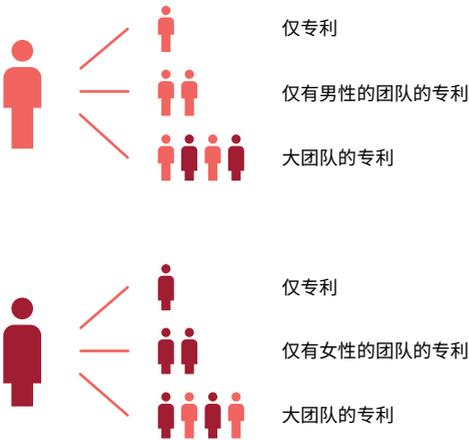
然而，它不提供明确的性别政策目标，并有低估性别不平衡的风险。

📊 指标：被列为发明人或创造者的女性百分比

合作者和团队规模

同一件知识产权申请中的共同发明人和共同创造者的数量可以提供关于合作性质和性别分布情况的信息。

合作情况可以通过网络数据可视化来说明。



 **指标:** 仅有女性或女性占多数的知识产权申请的百分比

 **指标:** 按性别划分的平均团队规模

 **指标:** 按性别划分的共同发明人网络

指标

行业和部门

许多知识产权分类（例如，国际专利分类、尼斯分类和洛迦诺分类）能够分析女性在**行业或技术领域**的代表性。

有关各部门性别差距的信息可以来源于对**公司和高校**的专利申请中被列为发明人的女性代表性的分析。

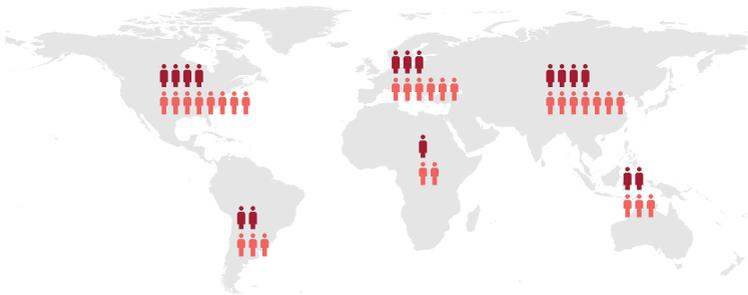
 **指标：按类划分的被列为发明人的女性百分比**

 **指标：按申请人类型划分的被列为发明人的女性百分比**

地理分布

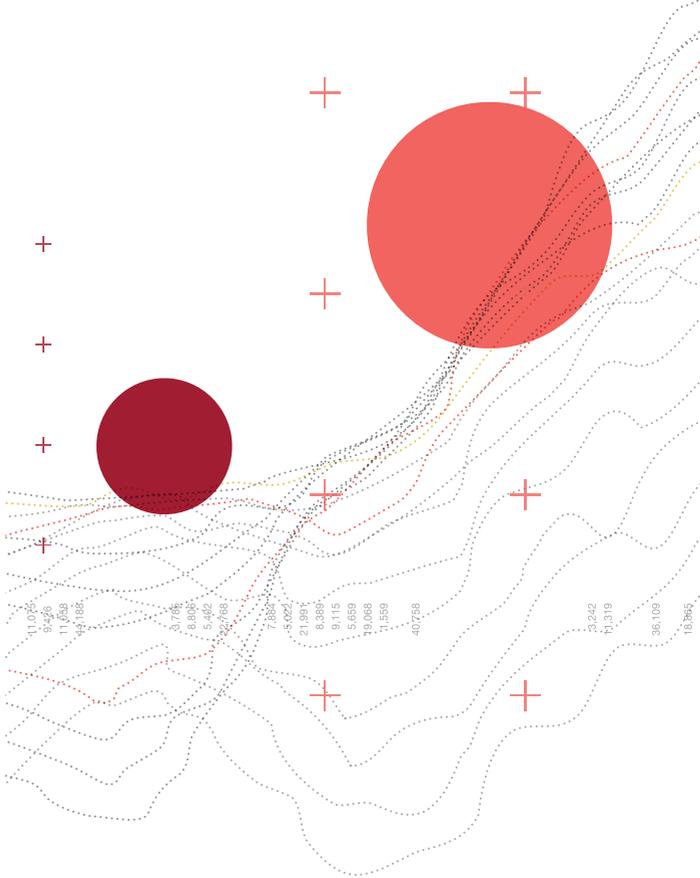
用发明人或创造者的地址能够按国家、地区或城市对性别差异进行更深入的分析。

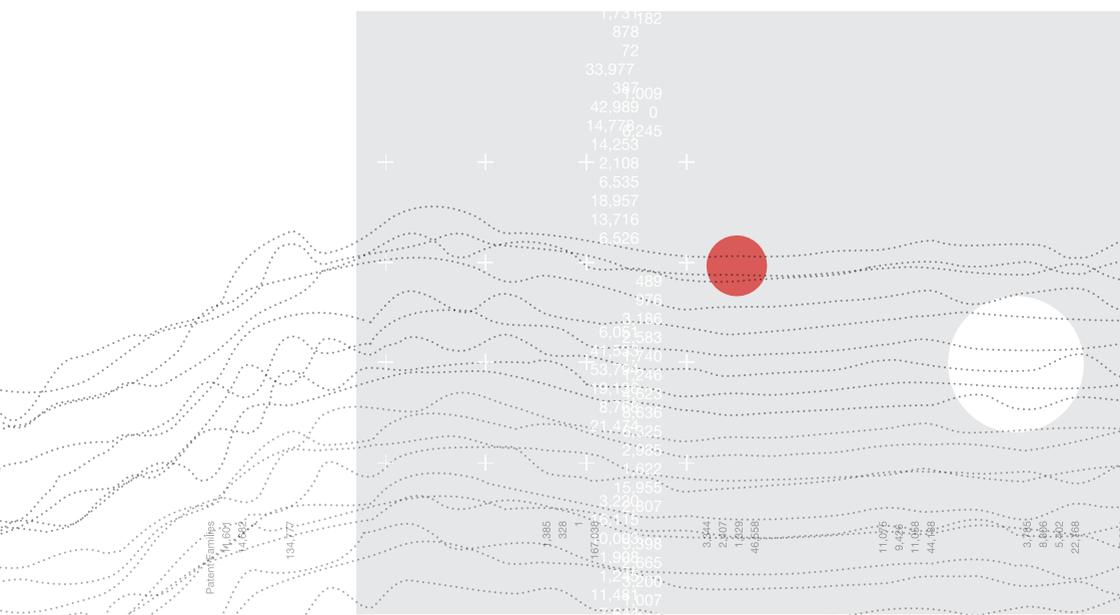
地理分布可以用**数据生成的地图**来显示。



 **指标：按州或省划分的被列为发明人或创造者的女性百分比**

 **指标：按城市划分的被列为发明人或创造者的女性百分比**





世界知识产权组织
World Intellectual Property Organization
34, chemin des Colombettes
P.O. Box 18
CH-1211 Geneva 20
Switzerland 瑞士

电话: +41 22 338 91 11
传真: +41 22 733 54 28

产权组织驻外办事处联系方式请见:
www.wipo.int/about-wipo/zh/offices

© WIPO, 2022年



署名 4.0 国际
(CC BY 4.0)

CC许可不适用于本出版物中非产权组织的内容。

产权组织第1073C号出版物
DOI 10.34667/tind.45432
ISBN 978-92-805-3414-6