



统计委员会

第四十六届会议

2015年3月3日至6日

临时议程* 3(a)(三)

供讨论和决定的项目：支持2015年后
发展议程的数据：大数据

官方统计使用大数据问题全球工作组的报告

秘书长的说明

秘书长谨根据经济及社会理事会第2014/219号决定，转递官方统计使用大数据问题全球工作组的报告。报告介绍了官方统计使用大数据问题国际会议的主要内容、全球工作组第一次会议的成果和官方统计使用大数据情况的调查结果。报告阐述了全球工作组的职权范围和拟议工作方案，涵盖领域包括培训、技能和能力建设、将大数据与可持续发展目标挂钩、宣传和沟通、访问权和伙伴关系、以及分类和框架等共有问题，并探讨了官方统计使用的特定大数据源，即移动电话及社交媒体数据和卫星成像。报告最后列出了讨论要点。

请统计委员会注意该报告。

* E/CN.3/2015/1。



官方统计使用大数据问题全球工作组的报告

一. 引言

1. 由于技术创新驱动电信及其他设备的广泛和持续使用，数字信息不断生成，例如全球定位系统设备、自动取款机、扫描装置、传感器、移动电话、卫星和社交媒体产生的数据。这些数量大、流速快且种类繁多的数据通常被称为大数据，对于官方统计具有潜在的用处，需要有新的工具和方法进行高效的采集、管理和处理。不过，许多统计机构仍需要通过展示大数据如何及为何有用且相关，对这种创新投资进行业务论证。下文将要详述的官方统计使用大数据问题国际会议在这方面令人鼓舞，因为它展示了许多正在实施的大数据项目的良好实例。

2. 大数据源的潜力在于及时、甚至实时提供通常是以最低成本生成的大量数据。开展与住户和商业调查等传统数据源有关的活动需要时间，而且这些活动往往在数据制作方面较为昂贵。大数据的利用可以补充、减少或替代此种数据收集类型的使用。不过，统计界认识到在将大数据引入官方统计之前，为利用这些创新数据源，包括在可持续发展目标监测和报告方面的应用，需要充分解决与方法、质量、技术、数据访问权、法律、隐私、管理和财务有关的问题，并提供适足的成本效益分析。

3. 在这一背景下，统计委员会在 2014 年 3 月 4 日至 7 日举行的第四十五届会议上确认大数据构成一个不容忽视的信息来源，需要根据自身优缺点进行评估。¹ 为实现这一目标，委员会支持关于设立官方统计使用大数据问题全球性工作组的提议，由工作组清点当前有关大数据使用的活动和实例，解决与方法、人力资源、质量和保密性有关的关切，并拟订不同类型大数据源的分类准则。

4. 统计委员会强调，全球工作组的职权范围和任务应基于战略考虑事项，特别是与 2015 年后发展议程、数据革命倡议和《官方统计基本原则》的联系。委员会还请工作组补充并推进区域委员会和其他国际统计机构开展的工作，并希望工作组的工作方案特别关注发展中国家的具体情况，例如法律框架状况和现有信息技术基础设施的不利之处。

5. 根据统计委员会第 45/110 号决定，官方统计使用大数据问题全球工作组于 2014 年 5 月创建，并在联合国经济和社会事务部统计司和中国国家统计局 2014 年 10 月 28 日至 30 日联合举办官方统计使用大数据问题国际会议之后，紧接着于 2014 年 10 月 31 日在北京召开首次会议。此外，统计司和欧洲经济委员会(欧洲经委会)开展了有关大数据项目及其组织背景的一次联合调查，作为对这些会议的投入和筹备。

¹ 见 E/2014/24，第一章 B 节，第 45/110 号决定。

6. 下一节将概述官方统计使用大数据问题国际会议的情况，然后是全球工作组首次会议的概要(第三节)和大数据项目及其组织背景的调查结果(第四节)。报告最后列出了讨论要点。

二. 官方统计使用大数据问题国际会议

7. 官方统计使用大数据问题国际会议汇聚了 40 多个国家、各类区域和国际组织以及研究界、学术界和私营部门的大约 120 名代表。会议着重讨论了三组大数据源，即：(a) 移动电话、全球定位系统和其他跟踪设备；(b) 卫星成像和其他地理空间信息；以及(c) 推特和其他社交媒体。会议审视了这些来源的类型和具体挑战以及利用这些来源所需要的伙伴关系。会议最后有几场专门讨论了大数据源的常见好处和挑战、引入创新的途径以及对大数据进行业务论证的方式。

8. 会议展示的大量项目反映了驱动统计界前进的创新精神。将各种大数据源应用到广大统计领域，证明了大数据真正具有改善官方统计的潜力。这些创新工具可以添加到统计系统组合中，从而补充现有应用或向高度相关的政策问题提供更灵活、短期的解决方案。不同的项目还显示如何解决与方法、访问权、隐私和技能有关的一些挑战。下文将介绍会议的一些主要内容。

9. 会议讨论了与移动电话记录有关的技术详情，包括主动和被动的移动电话位置数据之间的区别。在会议上介绍了多种有趣的统计应用和国家经验，其中包括旅游和日间流动统计、估计人口普查数据、绘制贫穷情况图和跟踪疾病爆发情况下的流动模式。由于移动设备在各地包括发展中世界广泛使用，移动电话数据在提供有关发展问题的实时和低成本信息方面潜力巨大。主要挑战仍然是保证保密性、移动电话数据的访问权以及公众对使用数据的信任。

10. 卫星成像有很大潜力在高度细分的级别提供更频繁和更及时的数据，尤其是农业统计。不过，估计方法，例如关于作物产量的估计方法，目前仍在测试中。澳大利亚统计局当前关于卫星成像数据的工作可以补充、甚至部分替代对农作物生产计量的现有调查。由于卫星图像每两周提供一次，使用这些数据可以降低调查的频率(及有关成本)。中国、哥伦比亚和墨西哥等国正在探索官方统计使用卫星成像的类似应用，包括用于生态系统核算领域的检测和试验。

11. 推特、脸书和其他社交媒体来源可以说掌握着有关人类行为的最大的数据集，统计界已经在探索这些来源，以期设计出数据驱动的社会科学应用，例如，在医疗卫生有关问题方面的应用。荷兰提供了一个如何通过脸书和推特数据得出消费者情绪估计数的非常有前途的范例，可提供初步估计数并减少调查需要，因为社交媒体估计数可以更高频率和更低成本生成。在中国和意大利，网络数据采集工具用于估计职位空缺率，可通过提供更好的每月预测数和更精细的区域估计数，支持当前的劳工统计。尽管在方法上存在许多挑战，统计界仍在该领域取得重大进展。

12. 会议还讨论了各种数据源的共同性，既涉及方法和质量关切，也涉及隐私、伙伴关系和信息技术。会议确认每一种不同大数据源都会遭遇现有挑战，迫切需要对这些挑战有一个详细的了解。一旦更好地理解大数据源的挑战，统计界就可以对其共同性加以利用。例如，就可以在不考虑特定数据源的情况下与全球运营公司建立数据访问权伞式协定模板。统计界应当与私营部门协作，而不是竞争，以便促进实现官方统计使用大数据的潜力。同时，应保持公正和独立，并大力宣传利用大量可用数字数据在社会效益方面带来的好处。建立公众信任将是成功的关键。

13. 最后，会议讨论了官方统计界当前和未来的作用。例如，可能会更频繁地要求官方统计人员验证市场提供的信息。会议强调，统计界应保持对自身使命的清醒认识，就相关议题向广大社会、特别是政策制定者提供高品质和公正的信息。创新至关重要，统计界应支持创新，但传统核心业务也仍有需要并将继续保留。例如，大数据应用得出的结果将接受基准测试，而这些基准需要通过传统调查提供。

14. 统计界应确定并开展针对大数据的业务论证。应努力将大数据的使用与 2015 年后发展议程及其可持续发展目标挂钩，因为这些会在未来几年对统计系统产生巨大需求。2015 年后发展议程许多拟议的具体目标不可能全都可以通过基于大数据的指标进行计量，但考虑到大数据的及时性和地理空间详细程度，预计将在这方面发挥重要作用。下一步可探索使用基于大数据的代理指标，提供比调查更为频繁的信息。这样，通过对国内生产总值进行“快闪”估计等方式，大数据就可成为传统系统的补充，而不是替代。

15. 总体而言，官方统计使用大数据问题国际会议概述了官方统计使用大数据的前景，有助于更好地理解方法、隐私和访问权方面的挑战。会议推动国家统计局及其他重要利益攸关方和合作伙伴就建立伙伴关系、数据访问权、沟通和宣传等问题以及培训、更新技能和能力建设的必要性分享了经验。

三. 官方统计使用大数据问题全球工作组会议

16. 官方统计使用大数据问题全球工作组在 2014 年 10 月 31 日举行的第一次会议上审查了职权范围(见附件一)，并讨论了 2015 年工作方案。全球工作组现有 28 名成员，分别代表发达国家和发展中国家以及各类国际和区域组织(见附件二)。

17. 在会议间隙，统计司、欧洲经委会、亚洲及太平洋经济社会委员会(亚太经社会)、亚洲及太平洋统计研究所、国际电信联盟(国际电联)和欧盟统计局召开会议，确保在国际组织之间协调大数据工作。其结果是，所有组织承诺分享在该领域的信息和产出。在这方面，欧洲经委会将创建和维护一个活动日历。各方了解，促进合作的一种方式是在各组织有关大数据的网站之间建立紧密联系。

A. 官方统计使用大数据问题全球工作组的职权范围和任务

18. 总体而言，统计界有义务探索使用新数据源，以满足社会在增强产品、改善工作方式和提高工作效率方面的期待。不过，当然也应该在满足 2015 年后发展议程监测和报告要求所提出的新数据需求的背景下，理解官方统计使用大数据问题全球工作组的职权范围和任务。

19. 《官方统计基本原则》完全支持官方统计使用大数据。例如，原则 1、原则 5 和原则 6 分别规定：(a) 应由官方统计机构公正不偏地编纂通过检验证明有实际用途的官方统计并加以公布，以尊重公民的公共信息权；(b) 可为统计目的从各种来源提取数据，不论是统计调查还是行政记录；统计机构在选择来源时应考虑到数据的质量、及时性、成本和给应答者造成的负担；(c) 统计机构为统计汇编收集的个体数据，不论涉及自然人还是法人，都应严格保密，而且只用于统计目的。根据这些原则，统计机构在隐私权和保密性得到严格遵守的情况下为统计目的调查和使用大数据源已不仅仅是一种可能，更几乎是一项义务。大数据有相关且及时的潜质，其收集方式也可能比传统的数据收集方法更具成本效益。

20. 全球工作组的职权范围反映了 2014 年 11 月数据革命促进可持续发展独立专家咨询小组题为：“世界需要数据：发动数据革命促进可持续发展”的报告所载结论。² 报告强调，统计部门需要以更快的速度作出改变并继续调整适应，摒弃昂贵和繁琐的制作流程，纳入新的数据源，包括来自其他政府部门的行政数据，并侧重于提供人机可读、与地理空间信息系统兼容、而且速度之快足以确保数据周期与决定周期相匹配的数据。³ 独立专家咨询小组的报告还明确提到大数据，并在这方面确定必须建立数据科学国家能力，以利用大数据中存在的机会并为官方统计作出补充。发展中国家需要更多国内资源和国际支持，以使数据革命可以促进可持续发展。通过双边和国际协调与合作，必须发展并以透明方式加强官方统计使用大数据的共同和标准化应用，同时证明充分遵循适用法律。⁴

21. 在咨询小组报告中的很多提议中，还有关于创建一个全球“数据创新网络网”的提议，以使该领域的组织和专家能够汇集在一起。这将促进采取改善新的可持续发展目标监测工作的最佳做法，促使确定与数据有关的共同基础设施可以解决能力问题并改善效率的领域，鼓励合作，促使确定重要研究差距并设立对创新的激励措施。

22. 总体而言，统计机构应结合质量、及时性、成本和应答负担等因素选择数据源。大数据源属于在这方面(如上文所述)理想数据源的范围，因为大数据有相关

² 秘书长于 2014 年 8 月 29 日委派这一独立专家咨询小组在 2015 年后发展议程背景下起草一份关于数据革命的报告。咨询小组于 2014 年 11 月 6 日发布了报告。

³ “世界需要数据”，第 9 页。

⁴ “世界需要数据”，第 23 页。

且及时的潜质，其收集方式也可能比传统的数据收集方法更具成本效益，还可将数据周期与决策周期匹配。有关大数据的工作应有助于采用改善 2015 年后发展议程下新的可持续发展目标的监测工作的最佳做法。一些新指标和这些指标的替代物可以基于大数据源，具有更高的及时性和更细的地理空间分类。

23. 在这一背景下，全球工作组制定了职权范围和任务，即：(a) 提供有关官方统计(包括 2015 年后发展议程指标)使用大数据全球方案的战略愿景、指导和协调；(b) 促进切实使用大数据源，包括跨境数据，同时推广现有判例并寻找众多挑战的解决方案；(c) 促进能力建设、培训和经验分享；(d) 促进政策应用使用大数据的沟通和宣传，特别是为 2015 年后发展议程的监测工作；以及(e) 在为官方统计使用私营部门大数据方面建立公众信任。

24. 包括任务在内的全部职权范围载于附件一(如上文所述)。

B. 工作方案和可交付成果

25. 根据在国际会议和全球工作组会议期间的讨论，全球工作组商定，其工作方案将由特定工作队开展的一些工作流构成，并将有一个额外工作队负责整体协调。首先，全球工作组商定，在国际会议所侧重的大数据源的三个类别，即移动电话数据、卫星成像和社交媒体数据方面应继续开展工作。这方面的理解是，促进这三个工作队的工作将包括参与至少一个试点项目。

26. 其次，由于该会议一个反复提及的主题是获取数据和与私营部门和其他社区建立伙伴关系，创建了一个工作队，以更深入地探讨这些相互联系的专题。这个工作队还将调查与全球运作的大数据提供者达成数据访问权的伞式协定的可能性。访问和使用大数据还与公共信任问题和在总体上沟通大数据的裨益和挑战有关系。为此目的，创建了一个宣传和沟通问题工作队，该工作队将审视筹资战略，以使发展中国家可以积极参与试点项目。

27. 第三，鉴于 2015 年后发展议程的背景，全球工作组还商定，让一个工作队专门负责跟踪可持续发展目标监测工作所需的指标和大数据应用之间的关系。最后，拟议并商定再设立两个工作队，一个关于培训、技能和能力建设的工作队(原因是处理大数据的所需技能与统计部门现有的技能不同)；另一个是有关分类和框架等共有问题的工作队。后一个工作队将在欧洲经委会大数据小组提议的分类基础上，在大数据源的分类方面开展工作。这八个工作队工作的总体协调将由工作队领导人组成的协调团队进行。

28. 据指出，这些全球工作组的工作队应汲取欧洲经委会大数据项目的经验，并以欧洲经委会大数据工作队正在制作的材料为基础。欧洲经委会工作队不久后将完成他们的工作，而全球工作组的工作队则正在开始他们的工作。这将使已经开始的一些工作流可以继续下去。每个全球工作组工作队将需要与协调团队密切协商，编制自己的工作方案、时间表和产出。

29. 全球工作组的可交付成果直接与八个工作队的工作有关。因此，预计全球工作组将提出一些提议和报告(具体如下)，并将参加一些试点项目。全球工作组还将继续编制大数据项目清单，这项工作很可能通过其协调团队完成。第一次调查结果将在下节列示。因此，2015年的具体可交付成果包括：

- 为官方统计使用大数据源所需的培训、技能和能力建设的提议和活动；
- 有关大数据和可持续发展目标之间联系的报告；
- 有关官方统计使用大数据的宣传和沟通战略；
- 有关官方统计使用移动电话数据的报告，其中将方法、信息技术和隐私挑战进行分类并提供解决方法，并包括这些数据在2015年后发展议程监测工作方面可能的用途；
- 有关官方统计使用社交媒体数据的报告，其中将方法、信息技术和隐私挑战进行分类并提供解决方法，并包括这些数据在2015年后发展议程监测工作方面可能的用途；
- 有关官方统计使用卫星成像数据的报告，其中将方法、信息技术和隐私挑战进行分类并提供解决方法，并包括这些数据在2015年后发展议程监测工作方面可能的用途；
- 有关加强获取大数据源的途径和与私营部门的伙伴关系的提议；
- 有关共有问题、分类、框架和分类学的报告。

30. 除了不同工作队的切实可交付成果外，全球工作队将参与一些有关实施官方统计使用大数据的试点项目，还将在下一节中讨论的欧洲经委会大数据项目清单以及统计司和欧洲经委会开展的初步调查结果基础上，继续致力于编制大数据项目清单。

四. 统计司和欧洲经委会有关大数据战略和项目清单的调查结果

31. 2014年9月，统计司和欧洲经委会开展了有关官方统计中大数据项目的联合调查。目的是提供有关正在开展的大数据项目的概览，以促进整个统计界内部开展更知情的讨论，并进一步编制全球工作组的方案。向所有参与欧洲经委会大数据项目的统计部门和大数据全球工作组成员发出了调查。

32. 调查有两个目标：收集有关已完成、正在进行和潜在大数据项目的信息，包括关于伙伴关系、数据源和工具以及有关这些项目的组织背景的信息。共分发了两份不同的问卷。第一份问卷旨在收集有关大数据的整体组织、战略和治理结构的信息，而第二份问卷侧重于实际的大数据项目。本次调查采用了一个相当广泛的大数据定义，即“数量大、流速快且种类繁多，需要有新的工具和方法进行高效采集、搜索、管理和处理的数据”。

33. 共向 78 个国家统计局和 28 个国际组织发放了调查。⁵ 32 个国家对有关组织背景的问卷作了答复，而 24 个国家对个别项目问卷作了答复，共提供了有关 54 个项目的信息。此外，国际组织提供了有关 3 个项目的信息。可以认为，这些国家和国际组织参与的项目可能占世界各地已完成和正在进行的大数据项目的很大比例。然而，全球工作组考虑开展全球后续调查，以收集有关大数据项目清单的更多项目的信息。这一调查结果的充分分析报告作为背景文件提供。下面是包含一些主要结论的摘要。

A. 主要调查结果：组织背景

34. 有关组织背景的问卷侧重于处理大数据的一般经验和框架，涉及大数据战略、治理和管理结构、质量评估、隐私和保密问题和技能短缺问题。收到的答复支持了官方统计使用大数据问题国际会议关于国家统计局面临挑战的结论。

35. 调查显示，只有少数几个国家制订了使用大数据的长期愿景，一些国家目前正准备制订大数据战略。此外，一些国家建立了国内实验室、工作队或工作组，并实施了确定大数据是否及如何可作为官方统计来源的试点项目，但大多数国家还没有确定将大数据源和成果纳入工作的业务流程，并且没有固定的大数据项目管理结构。

36. 作为制订大数据战略的一项内容，许多国家正在逐步建立伙伴关系，以探讨大数据带来的机会。为尽量减少探索这些新技术方面的失败风险，各国已决定积极参与国际合作，例如欧洲经委会的沙盒项目、欧盟统计局大数据工作组和根据统计委员会决定成立的官方统计使用大数据问题全球工作组。

37. 只有少数几个国家有大数据的特定隐私框架，而大多数调查的答复者表示，传统统计的隐私框架目前也适用于大数据。考虑到公共形象的重要性，许多组织强调在处理大数据时保护隐私和保密性的重要性，其范围甚至超过了法律作出的严格规定。一些组织在大数据提供者的地点进行所有的数据测试和处理，然后仅向国家统计局转递总数，以此避免隐私和保密性问题。一些组织称，尽管需要严格的数据保护规定，这些规定往往对数据访问权构成障碍。

38. 尽管大多数答复者认识到与信息技术、技能、法律和方法有关的挑战，大多数大数据项目的最大挑战是访问潜在数据集的途径有限或受限制。由于大数据在很大程度上是由私营部门拥有的(例如网上公司、移动电话运营商和银行)，与私营部门密切协作变得极为重要。许多这些行为体是全球公司，因此，全球统计界可使用集体议价力量获得访问这些数据集的途径。

B. 主要调查结果：个别项目

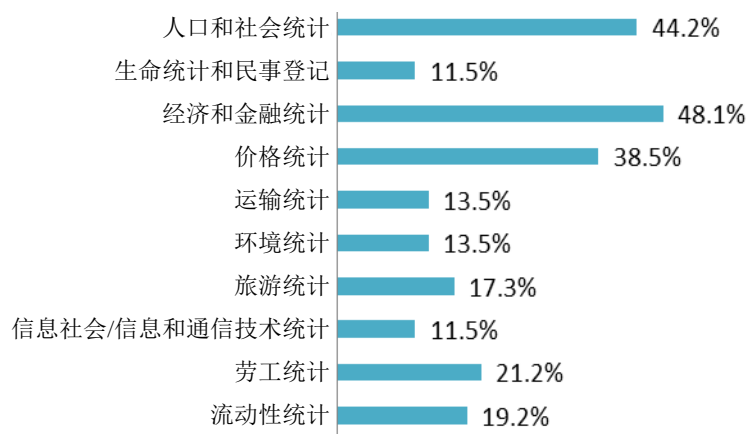
39. 关于个别项目的问卷包括关于潜在使用领域、项目状态、成果和经验教训以及关于伙伴关系安排、数据源和数据分析、使用的工具和技能的详细问题。调查

⁵ 向国际组织发送了问卷形式的电子邮件，主要是告知它们调查正在进行。

收集信息的很多项目是为大数据使用的测试和实验目的正在开展的项目，另外一些项目已经完成。其他项目处于探讨阶段初期或末期。

40. 在潜在使用领域方面，大多数项目涉及经济和金融统计、人口和社会统计以及价格统计(见下图)。值得注意的是，许多项目往往适用于多个统计领域。这些项目将一种类型数据源用于不同的统计领域。例如，墨西哥统计与地理国家研究所收集了在推特上发布的讯息样本，并将数据用于生成不同类型的统计数据，即主观幸福感的指标、旅游和边境流动性的统计数据。

该项目的潜在使用领域



41. 有略多于一半的项目以伙伴关系为基础，最常见的伙伴是商业企业，其次是其他政府机构。最常见的类型是为获取数据访问权与数据提供商结成的伙伴关系，包括与移动电话公司、智能电表数据提供者或社交媒体数据集成者的伙伴关系。例如，为制作日常流动模式，意大利国家统计局与移动电话数据提供商进行了合作。瑞士联邦统计局与该国内四大零售商店结成伙伴关系，每月两次收取扫描数据，制作粮食和类似粮食商品的消费物价指数。

42. 除了与商业数据提供商建立伙伴关系，各国统计局还与许多项目的分析伙伴建立了伙伴关系。这可能是为了获取大数据源的具体专门技能，将原始数据转化为预想的数据结构，或将分析技术用于生成特定的产出。不过，这种伙伴关系也可能是为了获得能够从事劳动密集型任务的人力资源而建立。例如，意大利国家统计局与墨西哥一所大学合作，在大量学生的帮助下对推文内容进行分类，以便开展对主观幸福感的分析。

43. 绝大多数调查答复者选择内部托管解决方案，而不是购买外部托管服务。隐私问题被称为选择内部托管的首要原因。当数据被认为已经存在于公共域，则更经常使用外部托管。另一方面，内部托管可能会成为选择大数据源的一个限制因素，因为有关内部托管的投资将需要涵盖硬件、软件和人的技能，这可能是启动大数据项目的一个显著障碍，特别是鉴于技术在迅速地发生变化。

五. 结论和讨论要点

44. 统计界已开始认真探讨官方统计使用大数据源所带来的机会。在北京召开的官方统计使用大数据问题国际会议显示了几个大数据源在多个不同类型官方统计中的应用，并促进了发达国家和发展中国家之间的经验分享。各国统计局面临许多相似挑战，例如获取数据访问权、隐私关切、方法问题、扩大计算资源需求、以及大数据分析内部技能不足，使得全球协作成为必要。全球工作组已开始构想运作程序和重点领域，并启动了针对当前官方统计使用大数据活动的清点筹备进程。通过对大数据项目的调查，收集了一些可能为其他组织提供重要经验教训的有趣实例。全球工作组要在区域工作组的协作下制定进一步的指导，并开展一些试点项目，以巩固和加速官方统计使用大数据方面的进展。

45. 请委员会就以下问题发表意见：

- (a) 官方统计使用大数据问题全球工作组的职权范围和任务；
- (b) 全球工作组的拟议工作队，以及在下列方面的对应可交付成果：培训、技能和能力建设；将大数据与可持续发展目标挂钩；宣传和沟通；移动电话数据；社交媒体数据；卫星成像；访问权和伙伴关系；以及分类和框架等共有问题；
- (c) 统计司/欧洲经委会的调查结果以及将此种调查扩展成全球调查的计划；
- (d) 全球工作组为巩固和加速在官方统计使用大数据方面取得进展，打算与区域工作组协作开展的试点项目。

附件一

官方统计使用大数据问题全球工作组的职权范围和任务

统计界有义务探索使用新数据源，以满足社会在增强产品、改善工作方式和提高工作效率方面的期待。官方统计使用大数据问题全球工作组的职权范围是基于这个原则，但肯定也应当在满足 2015 年后发展议程监测和报告要求所提出的新数据需求的背景下理解。

千年发展目标监测指标的主要不足是缺少及时性(或可用性)。为改善这一状况以实现 2015 年后议程的各项目标，2015 年后发展议程问题高级别知名人士小组在 2013 年 5 月 30 日报告(“一个新的全球伙伴关系：通过可持续发展实现消除贫穷和经济转型”)中呼吁发动数据革命，利用现有数据源和新数据源将统计充分纳入决策、促进数据的开放访问和使用、并确保加强对统计系统的支持。

本职权范围还参考了数据革命促进可持续发展问题独立专家咨询小组的报告，其中强调统计部门需要作出改变并继续调整适应，摒弃昂贵和繁琐的制作流程，纳入新的数据源，包括来自其他政府部门的行政数据，并侧重于提供人机可读、与地理空间信息系统兼容、而且速度之快足以确保数据周期与决定周期相匹配的数据。^a

在这一背景下，大数据源被认为是支持 2015 年后发展目标监测工作所需数据革命的一个重要部分。大数据有助于改善统计质量的一些方面，例如及时性和相关性，同时不影响公正性和方法的稳健性。

《官方统计基本原则》^b 也鼓励使用大数据等新的数据源，其中规定：

- 应编纂通过检验证明有实际用途的官方统计，由官方统计机构公正不偏地加以公布，以尊重公民的公共信息权。(原则 1)
- 可为统计目的从各种来源提取数据，不论是统计调查还是行政记录。统计机构在选择来源时应考虑到数据的质量、及时性、成本和给应答者造成的负担。(原则 5)
- 统计机构为统计汇编收集的个体数据，不论涉及自然人还是法人，都应严格保密，而且只用于统计目的。(原则 6)

根据《基本原则》中的这些战略考虑因素，以 2015 年后发展议程和数据革命为基础，并参考统计委员会第 45/110 号决定，工作组制订了以下任务：

^a 独立专家顾问小组，“世界需要数据：组织数据革命促进可持续发展”，第 9 页

^b 见 A/RES/68/261。

(a) 为官方统计使用大数据全球方案、包括 2015 年后发展议程各项指标提供战略愿景、指导和协调；

(b) 促进大数据源包括跨境数据的切实使用，同时利用现有范例，设法应对当前众多挑战，包括：

- 方法问题，包括质量关切和数据分析；
- 数据源访问权的法律及其他问题；
- 隐私问题、特别是与使用和再使用数据、数据链接和再次确认有关的问题；
- 数据的安全、信息技术问题和管理，包括数据公布、云计算和存储评估以及成本效益分析的先进方式；

(c) 同时促进能力建设、培训和经验分享；

(d) 促进政策应用使用大数据的沟通和宣传，特别是用于监测 2015 年后发展议程；

(e) 建立公众对官方统计使用大数据的信任。

附件二

官方统计使用大数据问题全球工作组成员

国家

澳大利亚

孟加拉国

喀麦隆

中国

哥伦比亚

丹麦

埃及

印度尼西亚

意大利

墨西哥

摩洛哥

荷兰

阿曼

巴基斯坦

菲律宾

阿拉伯联合酋长国

坦桑尼亚联合共和国

美利坚合众国

机构

亚洲及太平洋经济社会委员会/亚洲及太平洋统计研究所

欧洲经济委员会

欧盟统计局

国际电信联盟

经济合作与发展组织

海湾阿拉伯国家合作委员会统计中心

联合国经济和社会事务部统计司

联合国全球脉动

万国邮政联盟

世界银行
