



统计委员会

第五十四届会议

2023年2月28日至3月3日

临时议程* 项目 3(n)

供讨论和决定的项目：大数据

官方统计使用大数据和数据科学专家委员会

秘书长的说明

依照经济及社会理事会第 2022/324 号决定和以往惯例，秘书长谨转递官方统计使用大数据和数据科学专家委员会的报告。根据统计委员会第 53/124 号决定，本报告概述了委员会各任务组及其区域和部门中心取得的成就、近期的发展和计划开展的活动。此外，还就任务组与各中心之间的系统协作、委员会与地理空间信息界之间的积极协作、隐私增强技术在促进获取私人持有数据方面的作用以及数据科学领导者网络的职权范围提出了建议。请委员会就报告中概述的进展和建议发表意见。

* E/CN.3/2023/1。



官方统计使用大数据和数据科学专家委员会的报告

一. 引言

1. 统计委员会在第 53/124 号决定(见 E/2022/24)中支持官方统计使用大数据和数据科学专家委员会的拟议工作方向,即把大数据和数据科学纳入国家统计局日常工作的主流,建立国家统计局数据科学领导者网络,并加强专家委员会与地理空间信息界的合作。
2. 针对这些议题,专家委员会提出了四项建议,分别涉及数据科学领导者网络的目标和交付成果、加强区域和部门中心、委员会与地理空间信息界的合作、利用隐私增强技术促进获取敏感数据。第二节介绍了委员会各任务组、区域和部门中心以及联合国全球平台取得的进展,并重点介绍了若干方法指南、项目和培训活动的发布情况。第三节描述了 2022 年一些重大活动的成果,特别是第七届官方统计使用大数据和数据科学国际会议的成果。第四节介绍了委员会的四项建议。第五节列出了有待统计委员会采取的行动。

二. 官方统计使用大数据和数据科学专家委员会的成就和计划

3. 专家委员会通过 9 个任务组、5 个中心和联合国全球平台上的各种合作项目开展工作。本节报告了各任务组和中心的一些主要成就和计划。关于各任务组和中心的活动和事件的更详细信息,见专家委员会网站。¹

A. 方法发展

4. 由国际电信联盟牵头的官方统计使用移动电话数据任务组发布了关于使用移动电话数据进行流离失所和灾害统计、动态人口制图、信息社会计量、移民统计和旅游统计的若干方法指南。² 移动电话数据非常适用于测量人员流动性,可以作为调查和行政记录等传统数据源的补充甚至替代,这些传统数据源往往缺乏及时性,频率和精细度不高。任务组还制作了一个视频,宣传其工作。³ 目前正在编制关于使用移动电话数据进行交通运输统计的另一个指南,将于 2023 年发布。
5. 2021 年,地球观测数据任务组发布了一份报告,介绍在农业中使用一些地球观测数据应用所获得的经验教训及相关建议,⁴ 并发布了相应的宣传视频。⁵ 自动识别系统数据使用任务组维护一个关于官方统计如何使用自动识别系统数据

¹ 见 <https://unstats.un.org/bigdata>。

² 见 <https://unstats.un.org/bigdata/task-teams/mobile-phone/>。

³ 同上。

⁴ 见 <https://unstats.un.org/unsd/statcom/53rd-session/documents/BG-3s-3u-AgricultureAndBigData-E.pdf>。

⁵ 见 <https://unstats.un.org/bigdata/task-teams/earth-observation/index.cshtml>。

的实时指南。⁶ 这些数据可用于海洋运输统计，包括具有弹性和低碳运输解决方案的可持续运输统计。扫描仪数据任务组致力于就使用交易数据和网络采集数据等新数据源编制消费价格统计提供指导和支持。

6. 2023 年，隐私保护技术任务组将发布一个新指南，介绍多个隐私增强技术和 18 个案例研究，这些案例所涉范围很广，从概念验证到试点，再到在国家统计局或相关组织实施的生产解决方案。指南还将提供一个详尽的相关标准清单，并说明实施这些技术时必须考虑的主要法律或监管因素。另一份文件将描述隐私增强技术如何通过遵守联合国会员国隐私法规定的法律义务来促进数据共享。指南的目标受众包括律师、政策制定者以及公共和私营部门的数据科学家。

7. 2022 年，统计委员会在第 53/124 号决定(见 E/2022/24)中表示支持创建联合国隐私增强技术实验室，由隐私保护技术任务组具体执行。实验室的工作包括以下三个核心支柱，以加快官方统计界采用隐私增强技术：

(a) 试验：一系列正在开展的试点项目，重点是评价隐私增强技术在官方统计中的实际用例；

(b) 外联和培训：通过培训、公开讲座和教育活动，向更广泛的统计界传播在使用隐私增强技术方面的共同学习成果和见解；

(c) 支助服务：建立一个机制，向希望使用隐私增强技术的统计界人士提供支持和咨询意见。

B. 项目活动

8. 除了先前报告(见 E/CN.3/2022/25 和 E/CN.3/2021/14)中列出的项目外，目前还在联合国全球平台上开展以下项目：

(a) 自动识别系统船舶跟踪数据任务组的几个项目：

(一) 拉丁美洲海事技术合作中心与自动识别系统数据任务组共同开展的项目，生成关于航运活动产生的排放量的高频(每日和每周)统计数据，此外还生成关于巴拿马领水的港口等待和服务时间及运河时间表等增效措施的数据(2022 年 3 月)；

(二) 爱尔兰中央统计局的项目，发布基于自动识别系统数据的新前沿指标⁷(2022 年 9 月)；

(三) 将自动识别系统数据平台升级到 NetApp 的最新 Spot 以改善用户管理和提高成本效益的项目(2022 年 11 月)；

(四) 由欧洲经济委员会(欧洲经委会)和英国国家统计局协调的机器学习小组与亚洲开发银行共同开展的识别港口和泊位探测项目；

⁶ 见 <https://unstats.un.org/wiki/display/AIS/AIS+Handbook>。

⁷ 见 <https://www.cso.ie/en/releasesandpublications/fp/fp-pvrts/portvisitsusingreal-timeshippingdata/>。

(ii) 澳大利亚统计局开展的确定港口拥堵情况和船舶实际目的地的项目。

(b) 秘书处经济和社会事务部统计司、经济合作与发展组织和亚洲及太平洋经济社会委员会关于采用.Stat 技术托管柬埔寨国家数据库的项目；该项目处于生产阶段。此外，还为吉尔吉斯斯坦和马尔代夫创建了.Stat 实例。

(c) 统计司和经济合作与发展组织的项目，为涵盖 500 家最大多国企业的全球集团登记册维持一个图形数据库；该项目正处于试验阶段。

(d) 加拿大统计局在加拿大温哥华开展的开源试点项目，旨在计算按性别、年龄和残疾状况分列的城市建成区中供所有人使用的开放公共空间的平均比例(可持续发展目标指标 11.7.1)。该项目开发了一个开源处理管道，并开发了关键投入。

9. 全球便利获取私人持有数据任务组确定了五个全球价值链行业，以便与私营部门建立关系并获取全球价值链相关信息。这五个行业是旅游业(使用案例由马尔代夫牵头)、半导体业(使用案例由荷兰牵头)、电子商务服务业(使用案例由大不列颠及北爱尔兰联合王国牵头)、零售服务业(使用案例由阿拉伯联合酋长国牵头)、园艺业(使用案例由哥伦比亚牵头)。

10. 联合国隐私增强技术实验室启动了一个项目，实施共享、互联的基础设施，以测试和部署安全多方计算和联合学习等网络隐私增强技术。该项目包括一个“学习网络”，由实验室部署的虚拟机组成，每个国家统计局一台，目的是对测试数据进行隐私增强技术试验。此外，项目还包括一个目前正在开发的“测试网络”，其中实验室的每个成员机构在自己的场所部署一台虚拟机，并将其与在统计司的虚拟机相连接。

11. 由南非主持的联合国全球平台财务委员会的宗旨是为平台维持一个可持续的商业模式和营销计划，以管理成本和筹集资金。财务委员会制作了几个宣传视频、一个营销介绍和一本小册子，并与各任务组和区域中心密切合作，以确定哪些项目和培训活动需要资金。财务委员会将继续与发展合作界、慈善基金会和私营部门联系，讨论筹资和合作机会。特别是，财务委员会将呼吁向那些从使用该平台中受益的机构提供捐助。

C. 能力发展

12. 培训、技能和能力发展任务组开发了一个大数据入门培训课程，向使用者介绍官方统计中大数据的主要概念和应用。预计该课程将成为各任务组在各自领域正在开发的基础课程的先修课。此外，还开发了一个大数据培训目录，扩展了地球观测数据任务组早先开发的原型。该目录提供了与发展在官方统计中使用大数据的技能有关的约 300 个培训课程和材料的信息。借助该目录，不仅能便利使用标准搜索功能，而且还能根据用户个人情况推荐课程和材料，确定用户在组织中的职能、目标技能类型和水平以及当前知识水平，从而使用户能够建立定制的大

数据学习路径。这种学习路径方法使用了大数据获取和处理能力框架⁸的概念和维度。

13. 该任务组还管理一个学习管理系统，该系统的建立是为了向公众提供专家委员会不同任务组和其他伙伴开发的所有培训课程。该任务组还为其他任务组开发电子学习课程提供咨询和支持，并积极协助部署这些课程。学习管理系统托管在联合国全球平台上。

D. 大数据和数据科学区域和部门中心

14. 为协助各国统计局建立数据科学技能，并帮助将新技能纳入其统计编制工作流程，2020年和2021年期间在巴西、中国、卢旺达和阿拉伯联合酋长国建立了四个区域大数据中心。这些中心在区域一级将官方统计人员聚集在一起。语言和发展阶段相近的国家可以合作开展联合项目，利用大数据和数据科学编制统计和可持续发展目标指标。此外，专家委员会还建立了一个环境经济核算体系人工智能促进环境和可持续性部门中心，由西班牙巴斯克气候变化中心负责运作，专门处理可持续性和环境问题。

15. 巴西区域中心于2021年11月在拉丁美洲和加勒比经济委员会美洲统计会议上启动。该中心在2022年举办了一个关于使用网络采集进行价格统计的在线讲习班和一个面对面讲习班，并与专家委员会地球观测数据任务组合作举办了为期三天的关于卫星图像用于土地利用和地表覆盖的会议。2023年，它将组织更多关于将卫星图像用于可持续发展目标指标的讲习班，以及一系列关于使用移动电话数据进行官方统计的网络研讨会。它还将与区域内各国一道就大数据和数据科学的使用情况开展后续调查，并开展隐私增强技术和电子商务研究项目。

16. 中国区域中心于2020年12月7日在杭州举行的仪式上正式启动。该中心由中国国家统计局管理，致力于利用大数据和数据科学开发新的统计方法、技术和管理方式，并为来自国家、区域和全球组织的统计人员和数据科学家提供能力建设、培训和分享最佳做法的平台。该中心已举办了三次国际大数据研讨会(分别在2020年12月、2021年9月和2022年11月)；开展了以扫描仪数据、遥感数据、移动通信数据、行政记录和电子商务平台交易数据为中心的创新数据解决方案研究，并在一些领域取得了进展和成果；确认了门户网站和数据融合计算平台等应用系统的开发框架；与数据科学企业和研究机构开展了交流，尝试建立合作机制。

17. 中国政府致力于托管一个全球大数据中心，专家委员会对此表示支持，统计司正与中国国家统计局探讨如何将该区域中心发展为全球中心，使其在使用遥感进行农业统计和利用交易数据进行电子商务分析的研究方面成为全球领导者。国家统计局表示希望在疫情缓解后，统计界的许多专家能访问该中心，开展联合项目和研究，并在这方面邀请有兴趣的各方参加将于2023年4月在杭州举行的第四届联合国世界数据论坛。

⁸ 见 https://unstats.un.org/bigdata/task-teams/training/UNGWG_Compency_Framework.pdf。

18. 非洲区域中心于 2020 年 3 月通过经济和社会事务部、卢旺达政府和非洲经济委员会之间的谅解备忘录建立，由卢旺达国家统计局与非洲经委会共同管理。其主要目的是推动非洲大陆官方统计和可持续发展目标指标的创新。该中心提供的服务是研究如何利用大数据和数据科学在非洲开发新的统计方法和技术；为统计人员和数据科学家开展能力建设方案和培训；举办关于数据科学和官方统计现代化的国际研讨会和讲习班。该中心在 2021 年举办了两次大数据和数据科学在线讲习班，2022 年举办了一次面对面讲习班，各国统计局聚集一堂，讨论它们在现代化方面面临的挑战以及在非洲开展可扩展项目的合作领域。

19. 阿拉伯联合酋长国区域中心也于 2020 年 3 月通过经济和社会事务部与阿拉伯联合酋长国政府之间的谅解备忘录建立，由阿拉伯联合酋长国联邦竞争力和统计中心管理。该中心的主要目标是促进大数据和数据科学在项目中的使用，推动在新开发的方法、算法和工具方面进行知识分享，并为中东和北非区域的官方统计师群体提供培训。该中心有四个工作流，即学术界和人才，让大学、学生和青年参与到与大数据和官方统计有关的项目中；私营部门的参与，例如社交媒体公司和金融服务公司的参与；社区参与，包括学生和数据科学家实习计划；与政府当局的互动协作。

20. 2022 年 4 月，美国地质调查局、统计司、巴斯克气候变化中心和多诺斯蒂亚国际物理中心签署合作书，在环境经济核算体系人工智能促进环境与可持续性部门中心的运作方面开展协作并提供支持，以支持联合国全球平台。根据设想，该部门中心将成为全球公认的协作型和以行动为导向的知识、技术和创新中心，把人工智能和大数据使用方面的创新技术和数据科学方法结合起来，并提供一个急需的平台，进一步提高环境经济核算和可持续性领域数据和模型的互操作性。部门中心设在位于西班牙 Leioa 的巴斯克气候变化中心，由该中心管理。该中心和统计司正在讨论是否有可能将该部门中心发展成为联合国环境经济核算和可持续性中心，以支持气候行动和可持续发展目标监测。

21. 环境经济核算体系人工智能促进环境与可持续性部门中心在 2022 年的主要活动和成就是：开发了一个关于该工具如何支持海洋核算的测试应用；为世界银行即将出版的题为《不断变化的国家财富》的出版物汇编以实物和货币单位计算的全球碳储存基线账户；在卢旺达为 6 个非洲国家的从业人员⁹ 举办了关于利用环境经济核算体系人工智能促进环境与可持续性工具编制生态系统账户的培训讲习班；为博茨瓦纳、加纳、肯尼亚、菲律宾、卢旺达、塞内加尔、南非和乌干达提供生态系统账户编制方面的技术支持。就 2023 年，正计划开发社会经济用水基准模型，以便对与水有关的生态系统服务进行货币估值，并与欧洲航天局和美利坚合众国国家航空航天局等利益攸关方合作，整合新出现的可核算数据集并使之具有互操作性。

⁹ 见 <https://teebweb.org/workshops/development-of-ecosystem-accounts-using-aries-for-seeca-training-for-country-practitioners/>。

三. 官方统计使用大数据和数据科学专家委员会的外联活动

22. 与 2020 年和 2021 年一样，2022 年举办了多个有专家委员会参加的在线网络研讨会，涵盖各任务组及区域和部门中心正在开展工作的所有专题。不过，也开始逐渐举办一些面对面活动。其中一些活动已在区域和部门中心的活动中介绍中提到。本节介绍其他一些活动。

A. 2020 年阿拉伯联合酋长国迪拜世博会，2022 年 1 月

23. 2022 年 1 月底，官方统计使用大数据和数据科学专家委员会与阿拉伯联合酋长国联邦竞争力和统计中心在 2020 年世博会上组织了一次为期三天的活动，主题是“调动大数据和数据科学为可持续发展目标服务”。活动包括举行阿拉伯联合酋长国中东和北非区域中心启动仪式；启动联合国隐私增强技术实验室；专家委员会任务组和区域中心举行进一步全体会议、对话和炉边谈话及实践会议，包括关于机器学习的会议，在“青年圈”举行主题为“COVID-19 对人们生活的影响”的特别会议。

B. 第七届官方统计使用大数据和数据科学国际会议，2022 年 11 月，印度尼西亚日惹

24. 第七届官方统计使用大数据和数据科学国际会议于 2022 年 11 月 7 日至 11 日在印度尼西亚日惹举行，主题是“全球挑战与相关和及时数据的重要性”。会议包括关于可持续性和全球经济复苏、粮食安全和获取相关数据的高级别小组讨论。作为会议的一部分，举办了几个小型讲习班，内容涉及利用地球观测数据进行农业统计、利用移动电话数据进行旅游估算和人口统计、利用自动识别系统数据进行海运统计、利用机器学习技术进行官方统计和利用隐私增强技术获取敏感数据。会议的其他亮点是联合国大数据黑客马拉松和联合国隐私增强技术实验室黑客马拉松，全球有近 2 000 名年轻数据科学家和统计人员参加。会议促进了数据创新，推出了多项新倡议，并展示了黑客马拉松团队为帮助应对全球挑战而提出的创造性解决方案。专家委员会还在会议第二天举行了年度全体会议，不仅审查和讨论了各任务组、区域和部门中心的工作和各种项目活动，还商定了《日惹宣言》，详见附件一。关于会议和黑客马拉松活动的更多详细信息，可见背景文件。

C. 2022 年期间组织的其他活动

25. 2022 年 6 月举行的 Mobile Tartu 2022 国际会议¹⁰ 聚集了主要学者、研究人员和专业人士，讨论在人口流动和移动大数据方面的最新研究和应用。会议包括主旨发言、演示介绍、小组会议和热烈讨论，内容涉及借助移动大数据测量流动性的概念、方法和经验要素，以及如何理解相关的社会空间和人类环境的相互作用。移动电话数据使用任务组举办了一次主题为“使用移动电话数据进行官方统计的方法指南”的会议，展示了在灾害和流离失所指标、旅游和移民统计、信息社会指标和动态人口制图方面的应用。

¹⁰ 见 <https://mobiletartu.ut.ee/avaleht>。

26. 作为其五周年庆祝活动的一部分，英国国家统计局数据科学园在 2022 年 7 月举办了为期三天的活动，以促进国际合作并庆祝科学园的工作产生的国际影响。来自 26 个国家和国际组织的专家聚集在英国纽波特，其中包括来自巴西、加拿大、约旦、黎巴嫩、卢旺达和泰国的合作伙伴。活动包括一系列冲刺会议和讲习班，重点讨论英国国家统计局、其他国家统计局和联合国参与的项目。这些活动涵盖能力、技术、新方法和数据源等几个核心主题，由欧洲经委会的蓝天思维网络、区域中心及欧洲经委会和英国国家统计局协调的机器学习小组组织。活动还包括数据科学领导者网络和欧洲统计系统协作网络的网络智能网络的筹备会议。

27. 2022 年 11 月 2 日至 4 日在首尔举行了 2022 年官方统计国际研讨会，主题为“数据管理和安全共享数据”。¹¹ 研讨会由韩国统计局与统计司共同举办，讨论数据管理专题。数据管理办法促进公共部门组织之间及公共部门和私营部门组织之间共享数据，允许定期使用相关数据来生成官方统计，同时保护隐私和保密性。该办法旨在通过更好的方法和获取途径改善国家数据生态系统。研讨会方案涵盖数据治理、公平和包容、共享和协作。研讨会还讨论了与保护数据隐私同时确保充分利用数据有关的问题。举办了一次特别会议，介绍大韩民国公共和私营部门数据治理和数据安全技术的前沿案例。在潜在敏感信息的收集、处理、分析和分发这一整个生命周期中，可使用隐私增强技术降低隐私风险，并提供可证明的隐私保证。

四. 前进方向

28. 在日惹举行会议期间，专家委员会还举行了年度全体会议。除了审查各任务组、区域和部门中心以及联合国全球平台的活动，委员会成员还讨论了四项建议，涉及：(a) 各任务组与区域和部门中心之间开展更系统的合作，以提高方法发展与项目和培训活动之间的协同增效作用；(b) 官方统计使用大数据和数据科学专家委员会与统计和地理空间信息整合专家组积极合作；(c) 数据科学领导者网络的职权范围；(d) 具体和实际使用隐私增强技术来获取敏感数据。委员会商定了以下提案。全体成员名单见附件二。

A. 任务组与区域和部门中心之间的系统性合作

29. 在过去八年中，专家委员会各任务组举办和参加了许多活动，如独立的网络研讨会或面对面或在线研讨会、讲习班或会议。所有这些活动都或多或少涉及能力发展。虽然这些活动对每个任务组的产出(以及委员会的总体产出)起到了促进作用，但如果这些新的工作领域采取更系统、更协调的能力建设办法，并将这些活动与全球各国现有的统计方案相结合，则会产生更大效益。委员会的区域中心有其独特优势，即这些中心的活动可以与区域统计委员会的工作方案挂钩。例如，巴西大数据和数据科学区域中心的活动可与美洲统计会议的区域工作方案挂

¹¹ 见 <https://unstats.un.org/bigdata/events/2022/data-stewardship/>。

钩。就非洲大数据和数据科学区域中心而言，非洲经济委员会是该中心的正式合作伙伴，因此可以确保与非洲统计委员会直接协调。

30. 根据其任务规定，培训、技能和能力发展任务组将发挥各任务组与区域和部门中心之间系统合作的组织者作用。该任务组帮助其他任务组编写培训材料，包括编写电子学习课程。它还在英国国家统计局数据科学园的支持下，就每个中心计划开展的活动与各区域中心进行协商。此外，近年来任务组还开发了若干工具，帮助各国评估其在统计编制中采用创新办法的准备情况以及如何发展必要能力和技能。可通过区域中心利用这些工具来评估区域内国家的能力发展需求。因此，建议区域和部门中心在该任务组的支持下，与相应的任务组合作，在使用地球观测数据、移动电话数据、自动识别系统数据和农村获取服务、扫描仪数据和网页采集以及使用隐私增强技术等方面，规划和开展能力建设活动。能力发展的方式可不限于常规的讲习班和研讨会，可包括基于项目的活动和统计局结对。

B. 官方统计使用大数据和数据科学专家委员会与统计和地理空间信息整合专家组之间的合作

31. 如第 53/124 号决定所述，委员会支持关于加强专家委员会与地理空间信息界合作的提议。在这方面，建议委员会主席和统计和地理空间信息整合专家组共同主席签署换文，表示双方承诺就共同关心的问题开展合作，比如在以下领域开展合作：地理参照数据点的质量框架；将全球统计地理空间框架¹² 落实到地球观测数据、移动电话数据、自动识别系统数据和农村获取服务等任务组的工作中，并开展相应的联合培训活动。区域和部门中心可与专家组和任务组的成员一起规划和开展这些培训活动。委员会在 2022 年 11 月 8 日在日惹举行的全体会议上同意了加强合作的提议，专家组在 2022 年 12 月 1 日和 2 日在圣地亚哥举行的第七次会议上支持这一提议。

C. 组建数据科学领导者网络

32. 数据科学领导者网络是专家委员会的一项倡议，在 2022 年得到统计委员会第 53/124 号决定的支持。该网络召集各国统计系统内的数据科学领导人，包括区域和部门大数据中心的领导人，就大数据和数据科学问题进行战略讨论，分享经验和知识，加强各国统计局的决策和领导力，并加快研究合作和技术伙伴关系。网络将为委员会制定官方统计使用数据科学领域的总体愿景、开展协调、提供指导、确定优先次序和指引方向等任务提供投入。预期产出是具体战略和路线图，以期在这一领域实现协调一致的综合工作方案。

33. 数据科学领导者网络的职权范围详细说明了网络的宗旨、主要职能和活动、范围和责任、治理和报告关系、成员和工作方式。职权范围的一些要点是，该网络将推动就国家统计局在为国家统计局系统各机构和其他政府机构提供数据科学服务方面的当前和未来作用的关键问题进行战略讨论。网络还将讨论实验数据和指标的现状、道德和法律考虑因素、数据共享和获取、数据科学研究的治理和政

¹² 见 https://unstats.un.org/unsd/statcom/51st-session/documents/The_GSGF-E.pdf。

府各部门的能力建设、数据科学家行为准则以及大数据和数据科学的质量框架。它将发起新的倡议和项目，同时将相关活动的执行交给现有小组。网络还将就在哪些领域有必要通过积极主动的共同投资来弥补能力差距提供指导。网络向专家委员会报告工作，由一个主席团领导。英国将担任网络的第一任主席。网络成员面向国家统计局、区域和部门中心及国际组织的数据科学领导人或同等角色的人。主要议题将在专门的冲刺会议上讨论，邀请网络所有成员参加。附件三载有职权范围全文。

D. 具体和实际使用隐私增强技术来获取敏感数据

34. 保护数据免受未经授权的访问、处理或分发是隐私增强技术的唯一目标。国家统计局需要在收集、处理和传播敏感数据的过程中保护个人和企业隐私。此外，国家统计局可能需要保护私营部门输入数据的商业相关内容。输入数据的精细度越高，数据对于政策目的的价值就越大，但数据也就越敏感。在数据收集、各方之间传输、存储、处理到最终与决策者和公众共享等许多环节，数据的隐私和安全性都可能会受到损害。为了减轻数据生命周期中每个阶段的潜在风险，国家统计局可使用不同工具。

35. 加密、身份认证、授权和验证可用于确保数据不会在无意中暴露给不适当的一方。尽管在传输过程中和静止状态下的加密技术已经成熟，但在许多方面数据仍可能安全不保，也无法保证数据的使用方式。隐私增强技术是用于安全处理和分享敏感数据的技术。存在两大类隐私增强技术，即增强输入隐私的技术和增强输出隐私的技术。输入隐私关注的是一方或多方如何以保证数据不会在严格限定范围之外使用的方式处理数据。输出隐私侧重于修改计算结果，使输出数据不能被用来对原始输入进行反向工程。通过明智地使用这些技术，可以构建安全的数据生命周期，实现数据协作和信任，并为数据主体提供信心。

36. 隐私保护技术任务组在三个核心支柱，即试验、外联和培训及支助服务方面加大了努力，以加快官方统计界采用这些技术。试验是通过一系列积极的概念验证和试点项目来推进的，重点是评价隐私增强技术在官方统计界的实际使用案例。外联和培训则是通过培训、公开讲座和教育材料，向更广泛的统计界传播在使用隐私增强技术中获得的共同经验和见解。最后，建立一个机制，使那些使用或打算使用隐私增强技术的人能够与任务组的专家接触，以获得支持和咨询意见，以此提供支助服务。

37. 隐私增强技术实验室是将隐私增强技术的使用引入官方统计工作的一种实际合作方法，实验室推动开展试点项目试验，在实践中学习，并为那些希望成为隐私增强技术早期采用者的人提供支持服务。虽然在高层次上描述假定隐私技术所能创造的价值很容易，但重要的是深入研究潜在项目的细微差别，以了解和展示这些技术在当前关注的现实世界问题方面的全部价值。了解这些好处有助于统计界更好地评估实施成熟项目所涉及的风险回报计算。此外，在数据治理的任何领域都有许多利益攸关方，每个利益攸关方都有不同的观点和权衡因素，从技术可行性到数据安全到法律因素。通过在安全的环境中运行实验和项目，这些利益

攸关方可以表达他们的关切和观点，并减少生产使用前的障碍。专家委员会建议具体和切实地鼓励使用隐私增强技术来获取和分享官方统计中的敏感数据。

五. 有待统计委员会采取的行动

38. 请委员会：

(a) 承认并支持专家委员会各任务组、区域和部门中心及联合国全球平台所做的工作；

(b) 支持和鼓励区域和部门中心在使用地球观测数据、移动电话数据、自动识别系统数据、扫描仪数据和网页采集、农村获取服务以及使用隐私增强技术进行官方统计方面开展能力建设活动；

(c) 支持官方统计使用大数据和数据科学专家委员会与统计和地理空间信息整合专家组加强合作，制定地理参照数据点的质量框架并将全球统计地理空间框架落实到专家委员会各任务组的工作中，同时规划相应的联合培训活动；

(d) 核可数据科学领导者网络的职权范围；

(e) 鼓励使用隐私增强技术来获取和分享官方统计中的敏感数据；

(f) 支持《日惹宣言》。

附件一

日惹宣言

官方统计使用大数据和数据科学专家委员会

回顾官方统计使用大数据和数据科学专家委员会的任务¹ 是为官方统计使用大数据和数据科学全球方案，包括为编制可持续发展目标的指标提供战略愿景、指导和协调，并促进大数据来源的切实使用，同时借鉴现有先例，为许多现有挑战找到解决办法，

又回顾统计委员会 2017 年 3 月第四十八届会议通过了《开普敦可持续发展数据全球行动计划》，² 以支持落实《2030 年可持续发展议程》，³ 这需要地方、国家、区域和全球各级以及多个利益攸关方收集、处理、分析和传播数量空前的数据和统计，也需要全球统计界在实现国家统计局系统现代化和加强国家统计局系统这一战略领域采取行动，重点是实现治理和体制框架的现代化，利用统计标准和新的数据架构进行数据共享、交换和整合，促进在统计生产过程中使用新技术和新数据来源，

重申 2017 年 11 月《波哥大宣言》，⁴ 其中专家委员会建议联合国全球平台降低所有统计局合作开展大数据项目的准入门槛；汇集可信数据、方法、服务和应用供统计界共享；与私营和公共部门组织订立透明的伙伴关系协定，

重申 2019 年 5 月《基加利宣言》，⁵ 其中专家委员会建议统计界在世界各地建立多个大数据区域中心以开展联合活动，促进官方统计使用大数据和数据科学方面的能力建设，并在一个明确界定的治理框架下开展工作，

重申 2020 年 9 月《首尔宣言》，⁶ 鼓励包括私营部门在内的所有利益攸关方群体在联合国全球平台上分享数据、方法、技术和学习材料以造福社会；推动利用该平台编制快速指标，为关于新出现的问题和关键优先议程的政策提供信息；鼓励所有统计机构利用平台的能力；促请统计界增加实物和财政捐助，以确保平台的长期可持续性，

1. 祝贺印度尼西亚统计局于 2022 年 11 月 7 日至 11 日在日惹出色地主办了第七届官方统计使用大数据和数据科学国际会议；

¹ E/CN.3/2015/4，附件一。

² 见《经济及社会理事会正式记录，2022 年，补编第 4 号》，第 48/102 号决定。另见 <https://unstats.un.org/sdgs/hlg/cape-town-global-action-plan/>。

³ 大会第 70/1 号决议。

⁴ E/CN.3/2018/8，附件二。

⁵ E/CN.3/2020/24，附件二。

⁶ E/CN.3/2021/14，附件。

2. 强调及时和相关的数据在为决策者处理经济、社会和环境领域新出现问题提供信息方面的重要性，鼓励与私营部门和其他利益攸关方合作，获取和探索使用新的数据源；
3. 敦促所有利益攸关方群体支持大数据和数据科学区域和部门中心开展项目和培训活动；
4. 促请地理空间信息界与统计界合作，将全球统计地理空间框架落实到官方统计使用大数据和数据科学的方法发展中；
5. 鼓励所有统计机构加入数据科学领导者网络，集体讨论有关数据科学中心在国家统计系统中的地位和责任等战略问题；
6. 敦促所有利益攸关方群体支持联合国隐私增强技术实验室实际使用隐私增强技术，宣传隐私增强技术对官方统计的好处，并使统计机构能够采用隐私增强技术来获取和分享敏感数据。

附件二

官方统计使用大数据和数据科学专家委员会成员

国家

澳大利亚

孟加拉国

巴西

喀麦隆

加拿大

中国

哥伦比亚

丹麦

埃及

格鲁吉亚

德国

印度尼西亚

爱尔兰

意大利

墨西哥

摩洛哥

荷兰

阿曼

巴基斯坦

菲律宾

波兰

大韩民国

卢旺达

沙特阿拉伯

南非

瑞士

阿拉伯联合酋长国

大不列颠及北爱尔兰联合王国

坦桑尼亚联合共和国

美利坚合众国

组织

非洲开发银行

加勒比共同体

亚洲及太平洋经济社会委员会

非洲经济委员会

欧洲经济委员会

欧盟统计局

联合国粮食及农业组织

全球脉动

国际货币基金组织

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

亚洲及太平洋统计研究所

秘书处经济和社会事务部统计司

万国邮政联盟

世界银行

附件三

数据科学领导者网络的职权范围

1. 引言

提供相关、及时和可用的统计数据 and 指标对于各国政府确定优先事项、作出知情决定和实施更好的政策至关重要。数据科学革命正在通过新的方法、技术和数据来源(如大数据和机器学习)，帮助改善此类统计数据和指标的编制。

随着许多国家统计局迅速扩大其大数据和数据科学的能力和活动，有机会加强国家统计局之间的国际合作，使全球和国家统计系统现代化。通过在领导层分享知识和确定关键优先事项，各国统计局数据科学领导者可以提供战略眼光和领导，指导这一领域的许多多边合作倡议。一个新的国际性、包容性的数据科学领导者网络可以提供深刻见解、战略建议并争取各方支持，以推动创新，最大限度地提高效率，并通过数据科学加快整合统计生产。

数据科学领导者网络是官方统计使用大数据和数据科学专家委员会根据统计委员会 2022 年通过的第 53/124 号决定提出的一项倡议。

2. 宗旨

数据科学领导者网络召集各国统计局的数据科学领导人，包括区域和部门大数据中心的领导人，就大数据和数据科学问题进行战略讨论，分享经验和知识，加强各国统计局的决策和领导力，并加快研究合作和技术伙伴关系。

该网络将为委员会制定官方统计使用数据科学领域的总体愿景、开展协调、提供指导、确定优先次序和指引方向等任务提供投入。预期产出是具体战略和路线图，以期在这一领域实现协调一致的综合工作方案。

3. 主要职能和活动

数据科学领导者网络将：

(a) 推动就国家统计局在为国家统计系统各机构和其他政府机构提供数据科学服务方面的当前和未来作用的关键问题进行战略讨论，特别是确定数据科学如何能够支持决策过程，从而补充官方统计的提供；

(b) 推进就数据科学在官方统计中的作用的关键问题进行战略讨论，如实验数据和指标的现状、道德和法律考虑因素、数据共享和获取、数据科学研究的治理和政府各部门的能力建设、数据科学家行为准则以及大数据和数据科学的质量框架；

(c) 通过全球和区域用户协商，使统计界参与关于新出现的挑战和数据科学研究议程优先事项的讨论；与国家统计局及其数据中心建立网络联系，扩大分析能力和数据科学能力的范围和提供；

(d) 发起新的倡议和项目供现有小组执行。在履行向各执行实体提供指导的作用时，网络可查明国家统计局存在哪些能力差距，可通过共同投资办法加以弥

补。网络将提供一个空间，对各国统计局发展这些能力的“商业模式”进行分析，并在必要时确定和推广共同投资办法；

(e) 就关键问题开展战略讨论，支持和倡导国家统计局建立数据科学团队，并强调这种投资的短期和长期效益；

(f) 通过全球和区域用户协商提供指导，说明如何以创新方式让青年统计人员和数据科学家参与塑造官方统计未来的全球和区域协商；

(g) 对区域中心的主要优先事项和交付成果提供指导并施加影响。

4. 范围和责任

数据科学领导者网络将主要关注大数据和数据科学在全球和国家统计系统现代化中的作用。这一关注重点将涵盖统计局可用来提高其生产流程的质量和成本效益的所有新数据来源、方法、基础设施和技术，从而产生更及时和准确的信息，改善决策，造福全球。

在这一职权范围内，议程的范围应尽可能广泛，以确保该网络涵盖新的和正在出现的领域以及更成熟的主题。在必要时，可逐步建立分任务组，集中就具体专题进行战略讨论和知识交流。这可能包括下述讨论议题和交付成果。

拟议讨论议题

拟议议题包括：

(a) 国家统计局数据科学传播战略(如何讲故事，如何展示数据科学的相关性和影响力以获得支持，消除各自为政现象和障碍)；

(b) 围绕数据采集的社会许可和道德规范；

(c) 法律和道德框架；

(d) 将大数据和数据科学纳入国家统计局并使之主流化；

(e) 国家统计局发展到更高的数据科学成熟水平；

(f) 向国家统计局的年轻统计人员和数据科学家提供辅导；

(g) 区域中心的主要优先事项和交付成果。

拟议交付成果

拟议交付成果包括：

(a) 载有网络第一年主要交付成果的工作计划，明确所有权和预期影响；

(b) 在专家委员会网站上提供年度报告，其中载有会议讨论要点和主要产出、成果和影响；

(c) 就如何确定优先事项、机会、风险和挑战以及如何应对挑战以产生影响，向国家统计局提供指导方针；

(d) 商业模式或商业计划,就在哪些领域有必要进行积极主动的共同投资以弥补国家统计局的能力差距提供指导;

- (e) 附带活动
 - (一) 头脑风暴会议
 - (二) 研究合作项目
 - (三) 专题小组活动
 - (四) 技术援助
 - (五) 领导人辅导方案

5. 治理和报告

数据科学领导者网络通过其主席团、咨询委员会和年度全体会议机制向专家委员会报告工作。

网络由一个专门的主席团领导,主席团向专家委员会报告。网络的运行是开放式的。专家委员会将就以后是否延续网络向统计委员会提供咨询意见。

主席团成员和主席将在适当时候通过与表示有兴趣成为成员和(或)担任主席的国家统计局协商确定。

预计成员和主席会随时间而变化,使尚未担任成员或主席的国家有机会参加和(或)领导网络。

可能担任主席团成员的有巴西、加拿大、丹麦、印度尼西亚、荷兰、大韩民国、卢旺达、瑞士、阿拉伯联合酋长国、大不列颠及北爱尔兰联合王国、国际货币基金组织、非洲经济委员会、欧洲经济委员会和秘书处经济和社会事务部统计司。

6. 成员

数据科学领导者网络的成员面向国家(国家统计局)、区域(区域中心)和全球(部门中心和国际组织)的数据科学领导人。这些领导人可能包括数据科学和大数据中心的主任及研究和创新主任。

对于具体事件和活动,将邀请政府或国际机构、学术界、私营组织、技术提供商和民间社会等各界的专家参加。

7. 工作方式

会议的频率和形式将由主席团与网络成员协商确定。

预计会议将由主席团组织和管理,以确保有效和最佳利用时间。工作方式将包括:

(a) 在专门的冲刺会议(即密集的工作会议)上讨论主要议题,邀请网络的所有成员参加;

- (b) 将知识交流会议与讨论会议分开；
 - (c) 考虑在不同时区举行会议或举行区域会议；
 - (d) 记录所有知识分享会议并通过专家委员会网站分享；
 - (e) 确保网络的目标和工作计划与其他论坛和网络的目标和工作计划保持一致，填补空白，避免重复劳动；
 - (f) 认识到各国统计局的数据科学能力和对数据科学的采用可能处于不同阶段，并根据需要调整讨论；
 - (g) 考虑如何将不同成熟度的国家统计局安排到会议中以进行最佳讨论(例如，较成熟的国家统计局可以支持较不成熟的国家统计局)；
 - (h) 主席团成员出席并领导定期会议，缺席时任命其代表出席，并获得资源，以确保网络的顺利运行；
 - (i) 主席团成员确保网络专注于战略讨论，将任何潜在活动的执行工作转交给现有的委员会任务组或由成员组建的新小组。
-