



联合国开发计划署、
联合国人口基金和
联合国项目事务署
执行局

Distr.: General
15 November 2018
Chinese
Original: English

2019 年第一届常会

2019 年 1 月 21 日至 25 日，纽约

临时议程项目 2

评价

独立评价办公室关于其支持评价能力发展的报告

目录

| 章次 | 页次 |
|---|----|
| 一. 导言 | 2 |
| 二. 可持续发展目标、开发署以及独立评价办公室对评价能力发展的贡献 | 2 |
| 三. 提高开发署主导的评价的质量和相关性 | 3 |
| 四. 强化国家评价能力和制度 | 5 |
| 五. 增强全球评价文化 | 11 |
| 六. 未来之路 | 13 |



一. 引言

1. 开发署独立评价办公室的主要职能是开展独立的专题评价和方案评价，除此以外，它根据《2030年可持续发展议程》与国家伙伴和全球评价社群携手为开发署内外评价能力发展作出了贡献。这虽然仅是该办公室整体产出的一小部分，但为其评价工作提供了高度相关的补充。

2. 在开发署内部，该办公室通过提供指导方针、开展培训讲习班、质量评估和其他服务等方式，为加强分散评价能力提供了支持。在开发署以外，该办公室通过举办旗舰国家评价能力会议、支持其他同业交流群，以及开发新工具以便利评价制度诊断和知识共享，推动加强了国家评价能力。该办公室通过与发展援助和评价研究领域的全球思想领袖接触，促进了全球知识共享和伙伴关系建立，并重点关注可持续发展目标、独立监督和相关挑战。

3. 本报告向执行局介绍了截至2018年第四季度独立评价办公室在评价能力发展方面采取的方法和作出的贡献的最新情况。第二章描述了《2030年可持续发展议程》的大背景、开发署以及独立评价办公室变革理论的要素。第三章概述了该办公室在开发署内部为增强评价能力提供的支持。第四章讨论了该办公室为增强国家评价能力提供的支持，特别是两年期国家评价能力会议。2017年是这一系列会议的里程碑：开发署五个区域分别举办了一次会议，首次会议于2009年在摩洛哥举办，最近的一次会议于2017年在土耳其举办。因此，对这一主题给予了特别关注。第五章介绍了该办公室对全球评价社群的贡献，第六章考虑了在加强评价能力方面下一步可能采取的措施。

二. 可持续发展目标、开发署以及独立评价办公室对评价能力发展的贡献

4. 《2030年议程》和可持续发展目标奠定了开发署工作的基础。开发署《2018-2021年战略计划》阐明的愿景是：“通过消除一切形式和方面的贫穷，帮助各国实现可持续发展，加快可持续发展的结构转型，并建设抵御危机和冲击的能力。”¹

5. 在可持续发展目标的执行及开发署对可持续发展目标的支持方面，评价发挥着关键作用。就目标本身而言，《2030年议程》指出，“一个积极、自愿、有效、普遍参与和透明的综合后续落实和评估框架将大大有助于执行工作，帮助各国最大限度地推动和跟踪本议程执行工作的进展，绝不让任何一个人掉队”。² 后续落实和评估工作将“参照各国主导的评价工作结果”³ 并需“加强对发展中

¹ DP/2017/38，第24段。

² A/RES/70/1，第72段。

³ 同上，第74(g)段。

国家的能力建设支持，包括加强各国……的数据系统和评价方案”。⁴ 开发署评价政策还指出，“如果根据国家具体情况和优先事项进行适当调整，评价职能是一项有效的、国家主导的加强公民问责的工具，通过利用土著人民、民间社会、私营部门和包括国家议员在内的其他利益攸关方的贡献，可以加快实现国家可持续发展目标的优先事项”。⁵

6. 为切实支持《2030 年议程》和以统筹兼顾的方式、从经济、社会和环境这三个方面实现可持续发展，《战略计划》需要新的方法以及新的能力、平台和解决方案。开发署比以往任何时候都更需要确保其方案的相关性、有效性和效率，回应复杂和快速变化的世界中各国的需求。在这方面，评价能够作出重要贡献。正如评价政策表述的，“评价有助于作出更好的决策，并促进利益攸关方的学习。评价是帮助开发署、资发基金和志愿人员方案吸取过去经验教训的重要工具，有助于它们更好地了解哪些发展支助在何种情况下会产生良好或不佳效果”。⁶ 除了支持学习外，评价还有助于利益攸关方追究开发署在不同层面为发展成果作出贡献的责任。

7. 在这方面，独立评价办公室寻求通过一系列互相关联的能力发展活动和干预措施，适度为以下三项直接成果作出贡献：提高开发署主导的评价的质量和相关性；强化国家评价能力和制度；以及更强有力的全球评价文化。这些成果进而应促进产生一项中期成果：评价为旨在实现可持续发展目标的方案和政策提供参考。这将有助于实现预期的长期成果：更具相关性、有效性和效率并且能够产生可持续成果的国家政府和开发署方案——从而达成最终的预期成果，实现可持续发展目标。

8. 该办公室干预的三个领域是互相关联和互相加强的，因为该办公室和开发署的经验为国家评价能力会议和其他全球平台的讨论提供了资料，而从国家评价能力会议和其他国际集会上获得的知识也加强了开发署内部的评价。

三. 提高开发署主导的评价的质量和相关性

9. 独立和分散的高质量评价对确保开发署方案和项目的相关性和有效性，以及对支持实现以人为本且关爱地球的可持续成果而言至关重要。此外，高质量的分散评价是对独立评价的宝贵投入：它们能够为独立的国家方案评价提供实质基础并为独立的专题评价提供相关信息。因此，该办公室的首项能力发展工作侧重于加强管理、生成和使用可信的分散评价的内部能力。其中包括针对分散评价制定评价标准、程序、规范和方法指导；评估分散评价的质量；维护可检索且可公开访问的所有开发署评价资料库（评价资源中心）；以及开发和提供培训讲习班。

⁴ 同上，第 74 (h)段。

⁵ DP/2016/23，第 6 段。

⁶ 同上，第 4 段。

由于以往大部分工作已经向执行局进行了报告，下列各段将概述独立评价办公室的各项举措。

10. 分散评价准则。开发署于2009年发布了《发展成果规划、监测和评价手册》，之后还发布了2011年增编和关于成果层面评价的配套指南作为补充。虽然“黄手册”在概念上仍然可靠且具有相关性，但是自2009年以来，开发署发生了很大变化。因此，独立评价办公室于2017年与政策和方案支助局发展影响小组合作，在瑞士政府的财政支助下，与区域局和国家办事处共同举办了培训和反馈意见讲习班，以获得监测和评价协调中心关于国家办事处需求的意见，以便加强分散评价。⁷ 这被用于修订和更新“黄手册”的评价章节以及编写将于不久后发布的独立评价指导。新准则再次强调开展评价规划和确保开发署国家方案适当的评价覆盖率的重要性，更详细地阐述了评价的预期作用和责任，还纳入了优质评价实例的链接，以期加强今后分散评价的质量和实用性。独立评价办公室以及政策和方案支助局正在制定战略以进一步加强开发署的评价文化和问责制。

11. 培训。在发布新准则后，该办公室将组织新的一系列培训讲习班和网播研讨会，以及为开发署工作人员举办关于评价的最新在线课程。该办公室正在与联合国评价小组合作，该小组正在筹划当前的评价培训举措，以期开发一个能促进评价能力发展和专业化机会的全球信息枢纽。独立评价办公室以往曾为国际发展评价培训方案作出了贡献，目前正在与世界银行独立评价小组探索为今后的课程开发一套用于开发署内外的培训模块的可能性。正如下文更为详细地描述的那样，该办公室积极鼓励开发署工作人员参加最新的两次国家评价能力会议和培训方案，它今后将继续这么做。

对分散评价的质量评估

12. 独立评价办公室于2011年开始对分散评价进行质量评估并持续改进其评估制度。2012年，它开发了一项在线工具以便利开展评估，评价资源中心自2013年中期起提供这种评估。2014年和2015年，质量评估仅限于全球环境基金资助的项目，评价政策有待修订。2016年，质量评估工具进行了改进，同年起，开发署所有分散评价均进行了质量评估（例如见2017年与执行局共享的关于对2016年分散评价的质量评估的特别报告，以及《2017年年度评价报告》——[DP/2018/12](#)）。计划于2019年对质量评估工具进行进一步更新。

13. 质量评估进程为各委托办公室提供了关于如何加强今后评价的反馈意见，近年来，独立评价办公室加强了评价资源中心，进一步便利了从事评价工作的开发署工作人员获取信息。质量评估还被用于指出薄弱领域，这些薄弱领域已被纳入新修订的准则，并且将贯穿培训讲习班。

14. 评价资源中心。独立评价办公室负责维护评价资源中心，该中心是一个公开资料库，可检索开发署的所有评价。在过去一年里，该办公室增强了该中心，

⁷ 见《独立评价办公室工作计划（2018-2021年）》（[DP/2018/4](#)）和《2017年年度评价报告》（[DP/2018/12](#)）。

以便利追踪评价管理对策和重要行动，从而更方便地监测评价计划变更，建立评价者数据库和评价及其质量评估之间的关联，并为负责追踪评价计划及其执行情况的发展署工作人员开发监督看板。

四. 强化国家评价能力和制度

15. 独立评价办公室支持能力发展的第二个领域关注国家评价能力和制度。开发署评价政策申明，改善国家评价能力将加快实现可持续发展目标的进程，支持国家评价能力本身是一项方案优先事项。开发署方案股和独立评价办公室都寻求强化国家评价能力和制度，作为推进实现可持续发展目标的更广泛努力的一部分。根据该政策，独立评价办公室为“发展同业交流群，并与专业评价网络合作，以提高评价的效用和公信力”作出了贡献。组织两年期国家评价能力会议是此项支持的重要组成部分。

16. 2017 年是国家评价能力系列会议历史上具有里程碑意义的一年：第五次会议于 2017 年 10 月闭幕，这意味着开发署每个区域都主办了一次会议。第一次会议于 2009 年在摩洛哥（阿拉伯国家区域）举行，第二次会议于 2011 年在南非举行，第三次会议于 2013 年在巴西举行，第四次会议于 2015 年在泰国举行，第五次会议于 2017 年在土耳其（欧洲和独联体区域）举行。2018 年，独立评价办公室思考了近十年的经验，并仔细考虑了未来，本报告便是明证。

17. 本报告本章回顾了会议的历史和扩大情况，从 2009 年首次会议仅有 55 名与会者发展到 2017 年最近一次会议有超过 500 名与会者。随后将概述会议评估和审查结果以及对今后这一系列会议的影响。后续章节将介绍该办公室开发的两个新工具，即在线国家评价能力诊断工具和国家评价能力信息中心。

国家评价能力系列会议的历史

18. 开发署组织首次国家评价能力会议的目的是回应开发署执行局关于支持方案国国家评价能力发展的请求。这次会议是 2009 年与摩洛哥国家人类发展观察站合作在卡萨布兰卡举办的，会议的前提是，虽然人们对公共政策业绩评价是一项重要的善治工具已经达成共识，但在许多国家，这一共识与监测和评价制度、流程和工具的实际执行、使用和可持续性之间存在差距。此次会议旨在提供一个论坛，用以讨论开发署方案国伙伴面临的问题，深化其对评价作为有力的公共问责工具的认识，分享其他国家采取的解决办法，并在可能的情况下确认建立具有良好政治和体制基础的相关评价制度的共同战略。

19. 首次会议围绕五个调查领域——愿景、目的、结构和能力、方法以及问责，这些领域之后又被划分为四个国家一级公共政策评价做法重要主题：评价质量、对评价的使用、以及有利于评价的环境。与会者分享和比较良好的公共政策做法，展示初期和较先进的国家监测和评价制度实例，以及强调评价作为问责和决策工具的重要性。

20. 两年之后，基于摩洛哥会议获取的信息，在南非约翰内斯堡举行了第二次会议，重点关注“在关于公共政策和方案的决策中使用评价”。此次会议由开发署评价办公室和南非公共服务委员会联合主办，寻求让国家监测和评价制度处于不同发展级别的国家继续分享经验；查明落实这些制度方面的经验教训和限制因素；并在南南合作和三方合作框架下，确定技术援助的潜在供应和需求以加强国家监测和评价制度的体制能力。

21. 2011年在南非的讨论从展示国家监测和评价制度发展到关注评价管理、行为和使用。期间提供了实例以表明如何利用评价来制定政策和决策。关于各级政府和组织使用评价的讨论和思考阐明了评价质量与使用之间以及评价的“方便用户”和“有效利用”之间的密切联系。

22. 在摩洛哥和南非新出现的涉及评价的独立性、质量和使用的专题推动了第三次会议在巴西圣保罗的召开，这次会议旨在“寻求解决方案以应对与评价的独立性、可信度和使用有关的挑战”。2013年，开发署评价办公室和巴西社会发展和消除饥饿部联合主办了此次会议。

23. 与上一届会议一样，巴西会议也认识到监测和评价是紧密关联的，但同时又更加具体地侧重于评价。在制定国家监测和评价制度的框架内，此次会议强调了与评价有关的三项互相关联的挑战：如何确保其独立性、可信度和使用。此次会议审视了政府在建立被视为具备独立性的监测和评价制度方面面临的挑战，包括在政府内于何处设置以及如何构建评价任务授权和职能。在可信度方面，与会者得出结论认为，可信度取决于评价者的专门知识和独立性、评价流程的透明度，以及评价结果的质量。关于评价的使用，已发现的挑战包括在数量有限的利益攸关方以外扩大评价的使用，这需要进行方便用户的传播、将其翻译为当地语文，以及针对不同公共政策选项和问题表达见解。

24. 巴西会议的特别之处在于参与式地阐述了 18 项提升国家评价能力的承诺，大体可分为四项战略，包括：通过国内和全球宣传活动促进对评价的使用；确定和加强评价流程和方法；让现有的和新的利益攸关方参与交流与合作；以及探索不同管理评价体制结构的备选方案。

25. 在巴西会议之后，联合国大会在其第一项关于建设国家一级发展活动评价能力的独立决议（69/237）中将 2015 年指定为“国际评价年”。大会于 2015 年 9 月通过了《2030 年可持续发展议程》。一个月后，第四次国家评价能力会议在泰国曼谷举行。

26. 此次会议主题为“将评价原则与发展做法相结合以改变民众生活”。这次会议是在可持续发展目标正式通过之前筹划的，它预计评价能力对于执行新的可持续发展议程而言必不可少。此次会议重点关注政府如何发展必要的国家评价能力以应对可持续发展目标的新挑战。会议上新出现的发展国家评价能力的重要优先事项包括：需要促进国家所有和国家主导的评价并强调使用此类评价以影响政策；制定种种方法以评估实现目标方面的进展，包括评价环境、社会和经济可持续性、社会包容和平等、社会凝聚力和治理；促进更多元的伙伴关系以提高对评

价的认识，增加其使用；以及将对可持续发展目标的评价纳入体制结构——这是一项重大挑战。

27. 后来 2017 年在土耳其伊斯坦布尔举行的会议仍然以可持续发展目标为核心。这第五次会议由独立评价办公室、欧洲和独联体区域局以及土耳其政府联合主办。欧洲评价协会是实质性筹备和执行此次会议的重要合作伙伴。这次全球会议云集了来自 119 个国家的政府代表、评价从业者和发展专家，以审查与“可持续发展目标时代的人、星球与进程”有关的国家评价能力。此次会议寻求延续和深化在泰国启动的关于评价在可持续发展目标后续落实和评估工作中的作用的讨论，例如，更深入地探讨如何确保评价工作“不让任何一个人掉队”，包括环境问题。在此次会议上，与会者围绕自 2009 年首次会议即已提出的主题（例如，体制安排方面的挑战以及关于独立性、可信度和使用的问题）进行了讨论。

系列会议的与会者

28. 这些年以来，国家评价能力系列会议的与会人数显著增加，背景更加多元化，这体现了评价文化在全球的扩张。这些会议吸引了众多国家政府官员以及来自多边和双边发展组织和评价协会的参与者。2015 年以来的近期趋势是，来自全球北方的政府代表以及私营部门行为体、学术界和研究者人数不断增加。

29. 近年来，主要目标群体——政府官员的参与度稳步上升。民间社会的参与高峰出现在 2015 年，当年的会议是与国际发展评价协会共同组织的。鼓励来自开发署区域中心和国家办事处的工作人员参加最近两次会议的原因部分是陪同政府官员以在国内加强会议的后续活动，部分是提高工作人员对评价的认识和评价能力，以补充办事处加强分散评价的努力。这些年来，妇女的参与度显著增加，从 2009 年仅有 23% 增长至 2017 年的近半数（47%）。表 1 概述了这些年来国家评价能力系列会议的与会者情况。

表 1：2009-2017 年国家评价能力会议参与情况

| 国家评价能力会议举办年度 | 2009 年 | 2011 年 | 2013 年 | 2015 年 | 2017 年 |
|--------------------------|----------|----------|----------|-----------|-----------|
| 与会总人数 | 55 | 87 | 156 | 450 | 508 |
| 国家数量 | 30 | 24 | 58 | 100 | 119 |
| 与会者类型 | % | % | % | % | % |
| 政府 | 44% (24) | 76% (66) | 64% (99) | 32% (145) | 33% (167) |
| 民间社会（专业评价者志愿组织、学术界和私营部门） | 18% (10) | 7% (6) | 16% (24) | 48% (220) | 15% (75) |
| 联合国 | 16% (9) | 8% (7) | 8% (13) | 7% (33) | 15% (76) |
| 开发署 | 22% (12) | 9% (8) | 12% (18) | 13% (57) | 35% (178) |

30. 2017 年会议参与程度高的部分原因是，芬兰、瑞典、挪威和荷兰政府提供捐助而使来自符合官方发展援助资格的国家的 71 名与会者能够参会。这些与会者中，近四分之三是政府官员，其他人来自民间社会、评价协会和学术界。

31. 2015 年会议推出了新内容：在为期三天的会议召开之前举办了为期两天的会前培训讲习班。此举反响良好，2017 年会议将预先培训讲习班的内容扩充为 13 场为期一天的研讨会。2017 年这些有针对性的个人能力发展机会使超过 150 人受益。

32. 自 2009 年以来，共有 1 151 人出席了国家评价能力会议，其中 79 人出席了两次或两次以上。

会议评估

33. 独立评价办公室重视与会者的意见。每次会议最后均进行评估调查以及内部讨论、经验教训记录和对今后会议的建议。参加会议的国际评价咨询小组成员也提供了反馈意见和指导。

34. 各次会议的评估格式不同，限制了它们的可比性；但是所有评估均记录了积极的反馈意见。近三次会议的调查提出了类似的关于与会者满意度的封闭型问题，近三次会议有超过 85% 的与会者表示总体满意。2015 年和 2017 年会后调查的被调查者中有超过 88% 的人认为整体互动是有用的，有超过三分之二的人认为最近三次会议的结构和举办有助于促进实现提高国家评价能力的目标。

35. 为补充调查数据，在 2017 年会议期间，独立评价办公室国际评价咨询小组的两名成员与代表政府机构、联合国机构、非政府组织及评价协会的 18 名与会者进行了半结构化访谈。两名顾问结合他们作为与会者的观察，连同办公室汇编的、对 2017 年会议评估调查中关于如何改进会议的开放问题的回复和收到的关于会前培训讲习班的反馈意见，对访谈期间收集的信息进行了分析。下列各段将概述他们分析得出的一些重要信息。

36. 评价咨询小组的评估组得出结论认为，国家评价能力会议探讨了关于评价议程、国际发展和公共政策的多重问题，缓解了相关紧张态势，成为了单一的同业交流群。会议汇集了来自越来越多国家和利益攸关方团体——政府、民间社会和学术评价社群的与会者，以及来自开发署、联合国其他组织和发展伙伴机构的工作人员，以讨论与评价有关的政治、体制和业务挑战。

37. 在专题会议期间进行的访谈和表达的意见表明，国家评价能力会议与其他评价论坛在与与会者类型方面存在差异。来自规划部和其他中央部委的官员、开发署国家办事处的代表、联合国合作伙伴、双边捐助者、多边开发银行、非政府组织和民间社会汇聚在一起，互相学习，分享和讨论各自国家在规划、监测、特别是评价发展政策、方案和项目方面的经验。对于一些受访者来说，这些会议由于受众广泛且具有战略性——包括政府技术人员、评价团体、民间社会和联合国工作人员，因此弥补了目前国际层面上处理评价、发展和公共政策问题方面的缺口。欧洲评价协会委员会一名成员在伊斯坦布尔国家评价能力会议上称：“我们的受

众非常特别。通常的评价会议是学术界、评价者和评价顾问相互之间、基本上是供给侧内部讨论评价问题。在这次会议上，我们真正实现了供给侧和需求侧之间、从业者和决策者之间的对话。这是一次非常不同的对话。”

38. 这一两年期会议被视为讨论建立发展方案和公共政策评价框架方面的政治和体制挑战的重要场所。许多人认为，提供机会以分享和探讨评价框架的政治和体制方面是这些会议的特别优势，使它们有别于评价学术界或专业组织举办的研讨会。2017年调查中有三分之一的被调查者认为，2017年会议最有用的方面在于“分享经验和知识”，21%的被调查者提到，其次有用的是“学习新知识”。

39. 收到的反馈意见明确给出的另一个信息是，目前对技术培训的需求很高。伊斯坦布尔会议之前举办的13场为期一天的技术课程涉及不同种类的定量和定性技术。采用1分（不满意）至5分（很满意）的李克特式量表衡量的整体平均分为4.1分。培训讲习班的高参与率和好评度出乎意料，显示了今后对时间更长、更深入的技术培训课程的明确需求。

40. 访谈和评估结果表明，不同的利益攸关方团体——例如，公职人员（特别是负责公共方案和发展项目的人）——对评价的首要作用的想法可能与来自认知共同体或者双边或多边组织的与会者的看法不同。对于前者来说，评价是学习和改进发展方案和公共政策实效的工具。对后者来说，可能更加强调对公共支出和国际援助的问责和精英化评估。实际上，评价使用和用户的多样性影响了方案内容，要求根据这些受众的差异平衡主旨会议。这些多元见解体现了对会前讲习班和小组会议主题的差异化需求：公职人员和发展方案管理者对从制定和执行方面的挑战角度进行评价与从成果和影响评估角度进行评价同样感兴趣。

国家评价能力系列会议的未来

41. 下文概述了独立评价办公室结合调查、访谈以及国际评价咨询小组的反馈意见提供的信息对上述系列会议的思考。

42. 国家评价能力系列会议一向并且仍然具有宝贵价值，因为它在满足多元化需求的同时维持了一种独特的专业化定位，在评价专家的国际背景日益广泛的情况下，以参与评价的政府官员这一核心受众为中心。

43. 正如最近两次会议体现的那样，可持续发展目标为评价对话和未来会议提供了良好的组织和规范性框架。这是评价社群、政府和联合国之间潜在合作的有形领域，其全部潜力还有待进一步发掘。《2030年议程》的影响表明需要重新审查和完善用于指导评价的原则。

44. 重视实用方面至关重要。开发署国家评价能力会议通过继续建立一个关于可持续发展目标监测、特别是评价的多部门论坛，汇集世界各地在公共政策和方案的战略、工具和做法方面的经验，以及消除专门机构组织的关于各个可持续发展目标的其他主题论坛之间的差距，能够促进《2030年议程》。

45. 2017年会议专门安排了若干会议研究复杂性，复杂性日益被视为发展和实现可持续发展目标的核心要素。目前人们公认，执行难题和未能实现预期成果可

能是因为传统项目设计方法和逻辑框架办法意味着线性进展，而且政策和方案目标应该是非常具体的，以保证干预措施的连贯性及活动、产出和成果之间联系的效率。然而，关注范围狭隘往往导致政策不成体系。部门政策和方案可能是有效的，但由于“协同增效”并不是传统的设计标准，政府或社会工作的整体成功可能会受损。《2030年议程》着重强调了经济增长、社会包容和环境保护这些目标并行不悖，这表明需要在干预规划和评价方面采取一种新视角。这些会议为围绕复杂性及其对规划、执行和评价的影响开展辩论提供了机会。

46. 五次国家评价能力会议的总体评估结论是：这一系列会议创造了一种独特的专业化定位，应当继续保持。从一开始，目标便是创造空间以便与评价者、民间社会和想要加强国家评价职能的国家政府开展对话。随着更多国家政府参与其中、国家评价能力通过吸取错误教训和学习最佳做法而成熟，以及发展环境和优先事项发生变化，这一概念逐步发展。最初的两次会议强调必须建立更好的体制能力，通过评价来管理发展方案。2015年会议针对可持续发展目标的新背景，强调了在国家框架内将监测、特别是评价制度做法加以制度化的重要性，并为国家主导的评价做法和现实铺平了道路。2017年的讨论和文件则强调需要更彻底地将讨论从发展评价转向公共政策评价，切合从事可持续发展目标本地化、评估和后续工作的国家政府的实际。正如该次会议所阐明的那样，许多国家调整或建立了体制安排，将各项目标纳入了国家和地方发展战略和预算编制流程，还基于目标的各项指标评估了数据的可用性。此外，各国目前正在调整监测和评价制度，但迄今为止，对各项目标面临的新挑战的相关评价制度和能力的关注不足。

47. 会议主题不断演变，从为评价创造有利环境和将评价用于决策，到针对可持续发展目标的评价原则；再到评价的制度化；最终到更有针对性地讨论关于各项目标的评估和后续落实对国家政府的意义。然而，促进举办第一次会议的诸多挑战在第五次会议上仍然在讨论。虽然评价会议仅是旨在加强国家评价能力的更广泛的多利益攸关方战略的一个部分，但是它们为各国和各大洲交流在评价制度化和国家评价制度方面的知识和实用经验教训提供了一个独特论坛。

48. 独立评价办公室正在计划将于2019年举办的新会议。这一系列会议已结束一个完整周期，开发署每个区域都举办了一次会议，目前该办公室正在与阿拉伯国家区域局合作以联合组织下一次会议。暂定主题是“通过评价促进可持续和包容性发展”，新的关注点是可持续发展目标和公共政策评价。会议形式很可能效仿前两次会议，先以多种语文提供为期两天的会前培训讲习班，再举办为期三天的会议。

其他国家评价能力发展举措

49. 该办公室根据其举办上述系列会议的经验，扩大了对国家评价能力发展的支持，并创建了用于评价国家发展计划和实现可持续发展目标进展情况的诊断指导。在挪威国际开发署的支助下，该办公室开发了用于评价诊断和战略制定的在线自我评估工具。该工具的主要目的是便利制定或加强实现这些目标的国家评价框架；弥补国家评价思维和做法方面的缺口；以及为国家主导的评价流程和制度提供资料。自我评估工具旨在提供一个灵活实用的框架，使各国政府和其他发展

行为体能够制定一种系统性方法来确定评价国家发展战略和可持续发展目标的关键领域、途径和参数。它将评价做法置于其他公共部门反馈机制的框架内，并认识到评价做法并不是一系列产出，而是会随着时间推移不断发展。

50. 在线工具由一系列问题组成，这些问题使各国能够确定建设国家评价能力的要求。它帮助用户理解在不同国家有利于国家评价制度和体制能力的环境。该工具旨在供政府机构使用，而不论其评价制度、流程和能力如何。诊断流程包括四个阶段：诊断流程的规划和建立所有权；为诊断流程编写必要信息；进行诊断、分析、报告和建立所有权；以及最终开展经确定的后续行动。

51. 2017年，在伊斯坦布尔举行的国家评价能力会议推出了自我评价工具，目前已可使用。2018年6月，开发署非洲区域局和独立评价办公室在卢旺达基加利联合组织了一场为期两天的讲习班，这为来自14个国家的代表讨论国家评价能力发展方面的区域经验教训、为共享在线工具，以及为确定试点和推行国家及合作伙伴提供了机会。塞内加尔和乌干达这两个国家表示有兴趣开展在线工具的试点，在起草本报告之际，塞内加尔正在使用该工具进行一项评估。同时，在亚洲及太平洋区域，尼泊尔启动了试点诊断。

52. 基于会议方面的经验、与会者的请求以及可持续发展目标背景下知识共享对于发展评价能力的重要性，独立评价办公室正在建立一个方便用户的在线信息中心，用以提供按国家分列的关于国家评价能力的有用且有序的文件。该中心将成为国家评价利益攸关方（特别是政府）及开发署和更广泛的国际评价社群的综合信息来源。开发署将有能力组织和获取信息以更好地查明国家能力需求、了解不同国家的背景和目前对评价有利的环境，以及设计今后的会议和其他干预措施。该中心将有能力促进宣传工作，同时为评价能力较弱的国家提供激励。在线中心将于2018年底投入运营。

53. 独立评价办公室在其他国家评价能力发展举措中与开发署和其他合作伙伴开展了互动交流。2015年在泰国举行的国家评价能力会议结束后，该办公室促成举办了一次关于发展国家评价能力的亚洲及太平洋区域培训讲习班以支持执行《2030年议程》，开发署和联合国儿童基金会随后推出了一项联合倡议以开展国家案例研究，目的是确定该区域有关促进国家发展战略以及《2030年议程》审查、后续落实和评价流程的关键推进手段的新做法。该举措仍在执行过程中，它是2017年会议的一项会议主题。该办公室将与非洲区域局参加一个由法语国家评价网牵头组织的、重点关注法语国家评价制度化的联合会。目前正在四个国家开展诊断研究以评估制定评价制度化项目的准备情况。该办公室为参与评价和相关纪律工作的国家政府官员参加诸如非洲评价协会会议、“亚洲评价周”和欧洲评价协会会议等国际会议提供了便利。例如，该办公室支持博茨瓦纳的两名官员出席了2018年“亚洲评价周”并共享了审计长办公室和国家战略办公室（评价牵头机构）关于依据《2030年议程》进行审计和评价的经验。

五. 增强全球评价文化

54. 独立评价办公室支持能力发展的第三个领域处于全球层面。过去几十年间，发展方法发生了急剧改变，特别是出台了《2030年可持续发展议程》，这对作为一项进行批判性反思的独立专业做法的评价提出了新挑战。在国际发展领域，评价通常是指“发展评价”，或是对开发署等机构实施的发展方案的评价。但是，正如上文所述，2017年国家评价能力会议表明，必须将讨论从发展评价转向公共政策评价，以切合从事可持续发展目标本地化、评估和后续落实这些复杂工作的国家政府的实际。在这一新背景下，越来越多的行为体参与了对构成《2030年议程》的雄心勃勃且相互关联的多级问题的审查。因此，评价者必须对参与跟踪实现可持续发展目标的进展的越来越广泛的行为体作出回应。

55. 尽管《2030年议程》确认了评价，但对提交给可持续发展高级别政治论坛的自愿国别评估报告的分析显示，国家一级后续落实和评估工作主要关注各个指标框架以及监测实现定量目标的进展。评价的作用包括生成关于在何种条件下何种举措对何人有效的知识，这些作用仍需加强。必须进一步激励全球评价社群，与其互动交流，向其学习以及对其产生影响。在这方面，该办公室将发展援助和评价研究领域的全球思想领袖汇集到多个论坛，从而为全球辩论贡献经验，继续建设自身能力，确保为开发署提供的分散评价指导符合国际良好做法，并巩固其前沿的国家评价能力会议。

56. 该办公室是联合国评价小组的现任成员；办公室主任负责领导可持续发展目标工作组，其他工作人员参与专业化工作组和分散评价利益组，这些小组与该办公室旨在加强开发署内外评价能力的战略之间存在明确的联系。

57. 2018年，该办公室与关注国际安全、繁荣和正义的联合王国外交和英联邦事务部的执行机构威尔顿公园合作举办了一次多学科高级别活动。活动目标是分析《2030年议程》和不断变化的发展背景在何种程度上引发了一系列新的发展挑战；为关于问责和学习方面评价的目的和使用的全球辩论作出贡献；以及探索在不损害严密性和可信性的情况下更好地实现评价的目的和促成完善方案制定和政策决定的途径。关于对开发署未来评价能力发展工作的意义的讨论中出现了一些新的要点，包括：进一步强调国家对评价的所有权以及更多地让公民参与辩论和确定可持续发展计划和项目，这意味着评价者在投资创建有能力解读和批判性地评估事实和证据的受众方面担负着更大的义务。鉴于可持续发展目标的复杂性和跨部门性质，评价需要花费更多时间调查各项系统性努力以及社会变革如何起作用。由于目标范围广大，发展工作变得日益多元化，评价“对象”天差地别。在这方面，选择对什么进行评价及为什么对其进行评价成为了一个越来越重要的问题。另一个结论是，对评价者的培训需要着重强调评价工作就是作出判断以激发各种对话，从而帮助引起社会变革。这些要素以及其他许多要素都将纳入未来开发署主导的评价培训讲习班和国家评价能力会议。

58. 该办公室于2015年在范围较小的地方一级启动了“系列午餐会议”，让发言者分享评价和其他发展相关专题方面的知识和经验。这些会议推动了辩论，并

加强了内部能力以及开发署方案同事和联合国其他评价单位同事等受邀伙伴的能力。2018年的发言者包括联合检查组一名检查专员、前任美国国务院外国援助资源主任办公室评价和援助实效主管、乐施会监测、评价和学习小组以及欧洲联盟总司的一名国际发展专家。

六. 未来之路

59. 2019年,该办公室将继续加强能力发展战略,寻求更大程度地实现其多项能力发展举措之间以及这些举措与评价工作之间的协同增效。

60. 关于加强开发署分散评价能力,该办公室将继续与政策和方案支助局发展影响小组合作,以进一步增强开发署内的评价文化,并强化分散评价的能力和问责,包括推出和传播分散评价准则、其他培训活动和讲习班,以及即将开设的在线课程。

61. 2019年,该办公室将与阿拉伯国家区域局合作举办第六次国家评价能力会议,吸取前几次会议的经验教训,加强技术培训内容,为国家经验提供更大的重要平台空间,以及不仅继续倡导独立性、可信度和使用原则,而且继续主张在制定和评价发展政策和方案时奉行可持续发展价值观的意义。该办公室将在会议设计、举办和后续工作中寻求与区域局进一步互动交流,包括继续鼓励开发署工作人员与其国家对口部门一道出席会议。目的是建立工作人员对评价的理解以及便利启动或加强国家办事处为强化国家评价制度提供的支持。即将推出的国家评价能力信息中心将设置关于开发署国家级各项评价的链接,并有机会突出强调致力于加强国家评价制度的开发署项目。

62. 该办公室将与各区域局合作,尽可能为旨在加强国家评价制度的区域举措提供支持,包括扩大和复制开发署-儿基会亚洲及太平洋区域举措,以确定和分享在制定响应《2030年议程》的国家评价制度方面的良好做法,参与法语国家评价网的制度化项目,以及在执行工作中传播在线诊断工具和经验。

63. 该办公室将寻求与其他发展伙伴一道加强举办的会议和培训课程的协同作用和互补性,包括与世界银行合作编写未来各版的《国际发展评价培训方案》,以及与联合国评价小组及其成员合作在区域和国家层面针对联合国工作人员和合作伙伴开展共同培训。

64. 独立评价办公室的主要职责仍然是开展独立评价以帮助开发署应对不断变化的发展状况及其合作伙伴不断演变的需求。促进加强开发署内部概念化、委托和使用高质量分散评价的能力,将对该办公室开展的评价以及更广泛的组织产生积极影响。通过为加强国家评价能力作出贡献(尽管贡献不大),特别是举办国家评价能力会议,该办公室提供了一个平台,使开发署及其他各方能够与国家合作伙伴携手共同加强国家评价制度,将其作为善治工具和实现可持续发展目标的要素。通过促进关于评价议程和职业的全球辩论,独立评价办公室提高了开发署作为牵头发展组织的知名度,并为今后的学习和发展带回了经验教训。