



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
8 April 2026
Chinese
Original: English

2026 年年度会议

2026 年 6 月 8 日至 11 日，纽约

临时议程项目 9

对开发署的评价

2025 年年度评价报告

摘要

年度评价报告概述了开发署独立评价办公室 2025 年的工作以及促成这些成就的资源。根据执行局第 2020/15 号决定，报告阐述了来自 2025 年完成的一系列评价的关键见解，强调了各项成果和开发署必须应对的反复出现的挑战。报告探讨了整个开发署如何使用办公室的评价，提请注意促成 2025 年所观察到的改善的因素。报告还赞赏分散评价的质量得到提高，同时建议改进其广度和战略使用。报告总结了独立评价办公室与整个联合国系统评价单位和外部伙伴的合作，最后简要介绍了开发署托管实体开展的评价活动的最新情况。

决定要点

执行局不妨：(a) 表示注意到年度评价报告；(b) 鼓励开发署加强分散评价在决策中的运用。



目录

	页次
一. 2025 年对开发署的评价.....	3
二. 来自对开发署的评价的见解.....	5
三. 独立评价办公室评价的效用和使用情况.....	7
四. 评价资源	10
五. 分散评价	11
六. 评价后续行动	14
七. 独立评价办公室对全球合作伙伴关系的贡献.....	16
八. 开发署托管实体中的评价工作.....	18

一. 2025 年对开发署的评价

1. 本年度报告利用一套补充证据来源来评估开发署评价产品的影响范围和使用情况。关于独立评价办公室工作质量与效用的见解，以办公室的分析数据及年度利益攸关方调查为依据，该调查收到了执行局成员、开发署高级管理层、监测和评价官员及合作伙伴的 318 份答复。关于分散评价的证据来自评价资源中心和独立评价办公室质量评估流程。对 2022 年至 2025 年 125 份评价报告的审查，审查了管理层回应和行动计划清晰度和稳健性、促进或限制其实施的因素，以及评价建议在多大程度上为决策提供了信息和推动了方案变革。

战略方向

2. 执行局于 2025 年 9 月批准了新的开发署评价政策(DP/2025/28/Rev.1)，为开发署评价职能设定了到 2030 年的战略方向。该政策根据一项独立审查(DP/2025/6)制定，审查证实了开发署评价架构的稳健性。独立评价办公室主任与执行局之间的直接问责关系，以及政策中界定的资源基准，将继续保障独立性、公正性和完整性，这是可信评价的基础。在政策审查过程中与开发署政策和方案支助局(政策局)之间的协调，加强了与开发署成果管理战略以及更广泛的方案和项目管理架构的联系。

3. 开发署《2025-2030 年评价政策》中的变革理论阐明了该组织对充分推动更强的发展成果的评价职能的愿景。除了致力于提供可信和及时的评价以确保学习和问责之外，变革理论还确定了产生影响的关键先决条件：对评价的机构承诺、能力投资、激励开发署工作人员和合作伙伴应用评价证据，以及为学习和适应创造空间。共同推进上述所有方面，将增强对开发署证据基础的信任，并通过实施造福人类和地球的干预工作加强开发署对可持续人类发展的贡献。

4. 独立评价办公室将这一分析转化为 2026-2029 年多年期工作计划(DP/2026/7)，并制定了三个核心方向，以进一步加强其工作的相关性和效用：

(a) **战略重点。**评价工作将聚焦风险最高且最具战略重要性的领域，引导资源投向那些能为监督与决策提供最大价值的方面。独立评价办公室将优先考虑《战略计划》所有领域的评价，包括在机构转型时期尤为重要的整体评估。独立国别方案评价将重点关注更复杂、波动性更大和所涉价值更高的情境。在较为稳定的环境下，减少评价的频率将有助于更深入地评估长期影响和系统性变革，从而反映出实现发展成果通常所需的时间跨度。

(b) **响应新需求。**独立评价办公室将继续根据执行局和开发署高级管理层提出的新要求调整工作计划，确保评价范围符合不断发展的组织信息需求。人工智能将通过及时生成简明、有针对性的证据产品来支持提升效率。

(c) **受众驱动的互动。**每项评价和知识产品都将以结构化的受众和参与度调查问卷为指导。这一工具从一开始就确定了预期用户及其信息要求，为评价的设计、时间安排和沟通提供信息。该办法将加强评价的相关性和使用，同时充分维护评价的独立性和公正性。

交付情况

5. 2025 年，独立评价办公室落实了核心任务和工作计划(DP/2022/6)，开展了相关的高质量评价，以通过更好的数据和评价洞见帮助决策。成果包括三项专题和整体评价、20 项独立国别方案评价(涵盖 2022 年以来开发署 30 亿美元的交付规模)和四份“反思”综合报告，其中包括过去评价的经验教训，以支持组织学习。评价和综合提供了宝贵信息，帮助开发署、其执行局、各国政府和合作伙伴围绕比较优势领域完善开发署的定位，指出了如何应对交付和成效方面的瓶颈。

表 1
2025 年独立评价办公室报告

整体和专题评价	独立国别方案评价	反思
对《开发署 2022-2025 年战略计划》的评价	非洲(6)：博茨瓦纳、喀麦隆、刚果共和国、厄立特里亚、莫桑比克、津巴布韦	稳定与发展 通过数字生计塑造工作的未来
对开发署供资窗口的形成性评价	亚洲和太平洋(4)：孟加拉国、老挝人民民主共和国、泰国、越南	在最脆弱的环境中适应气候变化
对开发署在选举能力和制度流程方面所提供支持的评价	欧洲和独立国家联合体(2)：阿尔巴尼亚、塔吉克斯坦 拉丁美洲和加勒比地区(8)：加勒比多国办事处(巴巴多斯和东加勒比地区、牙买加、特立尼达和多巴哥)、智利、危地马拉、圭亚那、秘鲁、苏里南	从正义视角审视绿色与公正转型

6. 能力发展举措有助于加强分散评价，并支持在整个开发署内外建立更强有力的评价文化。独立评价办公室更新了其评价方法中心并发布了进行项目组合和影响评价的指南。五次网络研讨会增强了开发署监测和评价协调人在专题评价、评价与人工智能、建议和评价使用等方面的能力。办公室继续通过认可卓越表现和良好做法来促进高质量的分散评价。第五届卓越评价奖重点介绍了五份示范报告，这些报告获选的原因包括它们对性别平等的关注；该奖项还认可了阿拉伯国家区域局下开发署国家办事处在 2021 年至 2024 年间持续的高质量评价。开发署高级领导出席了颁奖典礼，凸显了卓越评价是组织优先事项。

创新和增效

7. 独立评价办公室继续投资于人工智能合乎伦理道德的使用，以加强评价流程并改善开发署在学习和决策中对证据的使用。办公室全面披露工作中人工智能的使用情况，并确保工作人员获得必要的资源与培训，以便以负责任的方式应用人工智能；此举符合联合国评价小组关于在联合国评价工作中利用人工智能的道德原则。

8. 2025 年，人工智能促进发展平台进行了重大技术升级，以消除第一代设计的局限性。该平台转向生成式人工智能处理流程，大大提高了人工智能促进发展平台的输出的准确性、情境相关性和实用性。新架构仅需进行极少的重新设计即可完成更新，将运行费用降低了 70%以上。超过四分之三的用户发现该平台使获取和使用评价信息变得更方便、更省时。人工智能促进发展平台的使用率较低，主要归因于文化和制度因素而不是技术因素；这些因素包括认识有限、存在替代的人工智能工具以及根深蒂固的工作习惯，即在筹备会议、设计项目或编制出版物时，尚未经常使用评价证据。为了进一步通过人工智能促进发展平台增加评价在开发署决策制定及方案设计中的应用，独立评价办公室计划围绕有针对性的、以用户为中心的讲习班开展第二阶段的互动参与。这项活动将持续至 2026 年。

9. 独立评价办公室与一家科技公司和罗格斯大学的 SOCRATES(面向技术增强型社会的社会认知机器人技术)国家研究实习项目合作，开发了一个人工智能工具，以提高开发署分散评价质量评估的一致性和速度。采用质量评估人工智能工具带来了显著增效，评估 250 项分散评价的年度成本从约 200 000 美元降至 10 000 美元。审查每项评价所需的平均时间也从 4 至 5 个小时减少到大约一个半小时。该工具始终实现了 80%或更高的准确率，证实了其可靠性。一旦工具得到充分完善，办公室将考虑更广泛地向开发署和其他联合国评价办公室提供该工具，以加强和统一质量评估流程。

10. 2025 年，独立评价办公室还在利用自动化和人工智能优化研究效率和协助开展业务任务方面取得了重大进展。办公室开发了一个 Power BI 模板，可以自动生成项目表并简化财务分析以支持评价工作。办公室还推出了一个定制的应用程序，从开发署透明度门户网站自动下载数据和文件，大大提高了生产力。通过创建工作流程系统，简化了支付和采购流程，使工作人员能够更轻松地完成和监控请求。

二. 来自对开发署的评价的见解

11. 关于 2025 年完成的 40 项独立评价办公室评价和分散评价及元分析的综合分析，揭示了涉及开发署绩效的共同信息：哪些方面行之有效、哪些方面未能奏效，以及背后的原因。

12. 评价确认，开发署加强了在动荡的发展格局中的地位，在危机时刻证明了自己的价值。在《战略计划》灵活且综合的愿景的支持下，开发署始终是各国政府的重要合作伙伴，指导各国应对不断变化的挑战。开发署的全球存在使其能够覆盖广泛国家和地区，了解当地情况，而开发署的广泛任务授权使其能够灵活地应对多样化需求。在危机背景下，开发署即便在极端条件下也能提供基本服务，从而进一步巩固了开发署作为可靠行为体的声誉。人道主义机构将其与开发署在长期危机中的伙伴关系视为加强社会契约和促进包容性恢复的一种方式。

13. 开发署在连续性与变革之间保持了平衡，在完善现有体系的同时，探索新途径。战略创新、发展融资和数字化被融入到方案规划中，帮助开发署应对错综复杂的压力。业务改革——包括推出 Quantum 企业资源规划系统、服务集群和集中成本回收——改善了合规性、标准化流程并使数据能得到更好利用。开发署有效利用值得信赖的伙伴关系和供资机制，包括全球基金、绿色气候基金和全球环境基金，在财政形势严峻的情况下持续开展至关重要的工作。

14. 无论在何种背景或领域，在联合国发展系统内，治理仍是开发署最受认可和最有价值的专门知识领域。开发署加强了公共行政和现代化服务，并帮助维持核心机构在混乱时期的运作——提供技术援助和弥合能力差距。在气候和环境工作中，开发署帮助各国将适应、减缓和生物多样性优先事项纳入国家发展计划。

15. 开发署的价值在于促成跨机构和系统的复杂变革。在开发署开展了以下工作时，所取得的成果最为显著：推动将治理与人类发展及环境优先事项结合起来的综合解决方案，同时协调合作伙伴并保持足够长时间的参与，以便将改革深植于制度之中。开发署作为中立召集方的作用受到重视。

16. 评价指出一些系统性挑战，如果不适当予以应对，就会限制开发署成果的有效性、规模和持久性。这些挑战主要与强调交付的业务模式有关，而该模式与主要依赖预算外专用资金的财务架构紧密相关。

17. 评价经常指出，开发署的项目组合过于分散，建议更加突出重点。开发署在政策、组织或社区层面开展单独的干预工作，这种分散的方法限制了有效性并削弱了影响。相反，当干预工作从多个角度应对挑战并系统性地消除变革的主要障碍时，便取得了更持久的进展。评价明确指出，各方并未期望开发署独自应对每一项挑战。相反，评价建议开发署集中精力实施具有最大潜在影响的干预工作，并与其他合作伙伴进行战略合作。

18. 开发署干预工作的成效常常受到短期和不可预测的财政支助的限制，在信任、问责与正义需要更长时间才能实现的情况下，严重影响了促进持久变革的能力。开发署的交付能力通常依赖于分配给各个项目的资源，因此出现了方案实施所需的基础系统能否维持的问题。评价表明，具有催化作用的融资机制和集合融资机制有助于减少分散的交付，具体做法是围绕共同的成果路径协调合作伙伴，并提供所需的可预测性和灵活性，特别是在危机和转型背景下，此时多年期融资和协调至关重要。

19. 评价指出，虽然开发署推进了成果管理制方法，并使其影响衡量与可持续发展目标保持一致，但监测和评价系统仍然资金不足、分散且过度依赖传闻证据。关于开发署的支持是否改善了机构行为、服务绩效和民众福祉，现有证据并不一致。衡量人类发展具有复杂性，此外，战略规划周期短以及在多伙伴倡议中难以将长期社会变化归因于谁，也阻碍了成果管理制的有效利用和开发署衡量其影响的能力。

20. 在若干情境中，这三个相互关联的挑战还可能导致排斥模式长期延续。如果干预工作过于分散、融资不可预测且学习系统薄弱，那么尽力帮助落在最后面的人的努力往往会被削弱或难以持久。

21. 事实证明，将《战略计划》中“不让任何一个人掉队”的宏伟愿景付诸实践，是一项艰巨的任务。虽然性别平等被越来越多地融入到方案设计之中，但干预工作往往缺乏足够的变革力，不足以消除系统性障碍，例如歧视、根深蒂固的社会规范以及权力失衡。同样，在不同环境下，与残疾人、女同性恋、男同性恋、双性恋、跨性别者和间性者社区以及其他边缘化群体的互动缺乏一致性。来自对数字化和公正转型干预工作评价的证据显示，参与和强有力的保障措施对于保护发展的公平惠益具有重要作用。当改革的推进速度超出了机构与社会的承受能力时，更有可能加剧排斥现象，并削弱公众的信任与支持。

22. 综合来看，2025 年的评价证据描绘出一幅连贯的整体图景。开发署加强了其业务基础，在危机中证明了灵活应变能力，并确立了治理作为其独特的贡献。当开发署采取了以下做法时，所取得的成果最为显著：将社区行动与系统改革联系起来，将数字化工具和融资工具嵌入机构内部，并将包容视为结构性优先事项而非附加措施。为了保持进展，各项评价指出了四项首要的紧迫任务：系统地思考和行动；将公平置于变革的中心；加强开发署的监测、评价和学习系统；更好地使融资与成果保持一致。这样，开发署将更有能力将其比较优势转化为持久、公平的发展成果。

三. 独立评价办公室评价的效用和使用情况

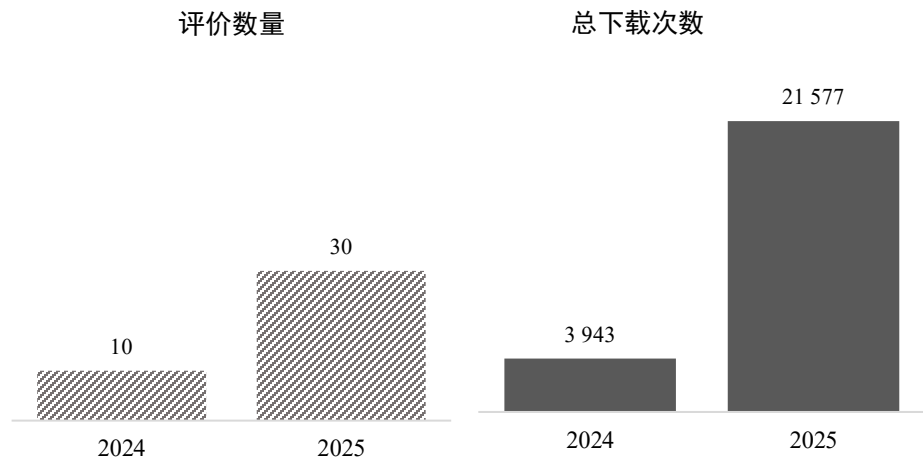
23. 2025 年的分析以及开发署同事和合作伙伴的报告显示，独立评价办公室的工作成果得到更多运用，影响力也在增强。这似乎与一些强化因素有关，包括办公室管理层决定，在评价流程的设计与实施中，更加重视其效用和使用；评价报告重点更突出，质量得到提高；持续努力以更清晰、更有针对性的信息来加强沟通。

下载量和已报告的使用情况

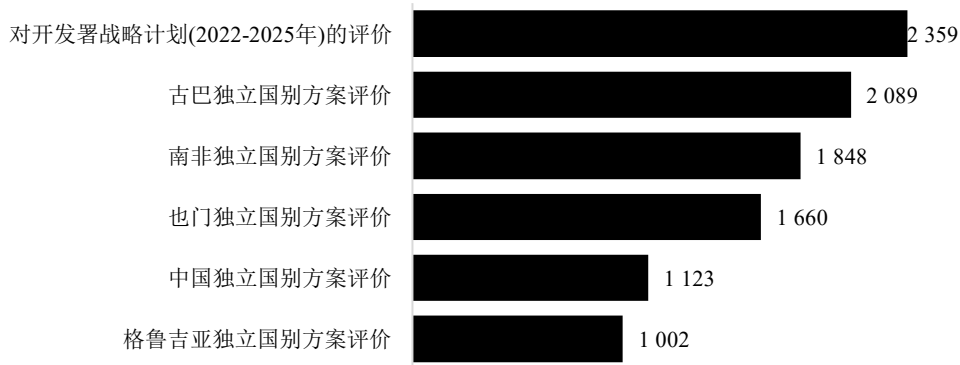
24. 2025 年，独立评价办公室出版物的查阅量显著增加。发布的评价数量增加了两倍，从 2024 年的 10 项增加到 2025 年的 30 项，但下载量增加了四倍多，从 3 943 次增加到 21 577 次。平均而言，每项评价的下载量增加了 82%，其中独立国别方案评价的下载量增幅甚至更高，增加了 154%。

25. 《对开发署 2022-2025 年战略计划》的评价是 2025 年下载次数最多的独立评价办公室报告。完整报告于 2025 年 4 月发布，已下载 2 359 次，近 1 000 名用户在 2025 年访问了办公室网站以获取有关该评价的内容。独立国别方案评价的下载量各不相同，有 5 份报告的下载量超过 1 000 次：古巴、南非、也门、中国和格鲁吉亚。尽管每个独立国别方案评价都专注于一个国家，但其影响范围远远超出了开发署总部和受审查的国家办事处，这表明独立国别方案评价的见解在全球都受到重视。综合报告的下载量达到 600 至 1 200 次。

图 1
评价报告数量增加了两倍，但读者数量增加了四倍多



2025年下载次数超过1 000次的报告



26. 独立评价办公室的评价和综合分析是开发署及其合作伙伴的可靠信息来源。在 2025 年利益攸关方调查中，超过 80%的受访者表示，他们使用了独立评价办公室的评价和综合分析，通过条理清晰的证据掌握信息，获得更清晰的见解。同样占绝大多数的受访者指出，他们利用评价——无论是整体、专题还是国家层面——来确保问责并宣传开发署作为透明组织的形象。就“反思”和综合分析而言，第二常见的用途是促进学习——这符合预期目的。

27. 独立评价办公室的评价有助于制定决策，尽管评价的影响力因类型而异。独立国别方案评价的影响力最大，76%的调查受访者表示，评价结果已被用于更新国家方案文件。专题评价和综合分析与决策制定之间的关联性较不稳定，68%的受访者表示将其用于决策。这在一定程度上可归因于全球评估的范围较广，在国家层面运作中的直接应用较为有限。

推动采纳评价结果的因素

28. 独立评价办公室的所有评价都保持了最高的质量和严谨性标准，确保报告方法稳健、分析可信，并符合联合国规范和国际最佳实践。每份报告都经过了全面的内部和外部同行审议。2025 年，办公室开始将前瞻工具整合到选定的评价中，以应对日益多变、不确定性不断增加的业务环境。这些工具支持“前景扫描”，指导评价建议考虑到开发署面临的新动向、风险和机遇。独立评价办公室对质量的重视反映在 2025 年利益攸关方调查中，约 80% 的受访者认为办公室的评价和综合分析有价值且富有洞察力。

29. 独立评价办公室继续通过使报告更加重点突出和更契合目标受众需求来增强报告的清晰度。对综合分析和独立国别方案评价的模板进行了修订，以实现更清晰的信息传递，增加了关于经验教训和可付诸行动的见解的专门章节。反思文件保留了 2024 年采用的简化格式，以清晰且易于理解的方式呈现证据和国家事例。独立评价办公室在每次评价之初都会开展受众和参与度调查，并在起草报告前进行故事板练习，以更好地了解利益攸关方的信息需求，明确关键信息的优先次序，并构建评价叙事，以实现最大的相关性和影响力。这些做法将在 2026 年全面纳入主流。

30. 扩大了面对面和虚拟互动，将评价知识“从报告中提取出来”，直接传递给利益攸关方，以应对时间限制和信息超载的挑战。在整个过程中与开发署同事更密切的合作提升了评价的相关性、增强了信任并促进了评价结果的使用。利益攸关方会议审查了证据及其影响；在国家一级对独立国别方案评价进行审查，在总部进行专题评价审查，包括与开发署执行小组和业务绩效小组一道进行审查。所有“反思”都在研讨会上讨论，将评价学习与开发署实地经验中的良好做法介绍结合起来。此外，在选定的高级别论坛上介绍了评价证据，例如在巴西贝伦举行的联合国气候变化大会；2026 年，这种互动的范围有望进一步扩大。独立评价办公室的区域评价顾问支持在决策中使用评价，但各区域的互动参与程度各不相同。

31. 独立评价办公室扩大了数字参与，在平衡内部和外部沟通的同时，优先开展直接外联，以提升影响力。与开发署同业交流群合作，通过 LinkedIn 和 Sparkblue/Viva Engage 分享了评价报告。对开发署国家办事处开展更系统的传播工作，往往能显著增加评价报告及其他知识产品的读者人数，这也成为带动独立国别方案评价下载量增加的最可靠驱动因素。社交媒体使可见度进一步提升，在开发署国家办事处和评价人员通过各自的网络扩大报告的分发范围时尤其如此。

32. 2025 年 4 月，独立评价办公室重新推出了其网站，采用了新版设计和简化的导航，使用户增加 30%。新网站有三种语文(英文、法文和西班牙文)，使办公室能以更生动直观、更易于理解的方式呈现评价结果。此次重新设计有助于增加对专题评价的参与，提升了在 193 个国家的覆盖范围与影响力。

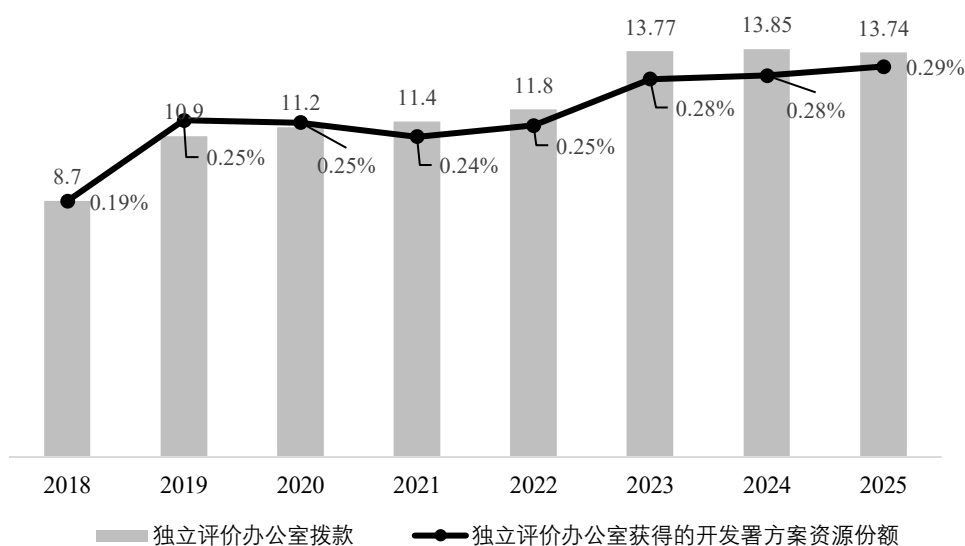
四. 评价资源

独立评价办公室

33. 2025 年，独立评价办公室获得的开发署方案资源比例几乎与前两年相同，略低于《开发署评价政策》规定的 0.3% 的目标。由于开发署预算削减，2025 年财政拨款为 1 374 万美元，比 2024 年减少 0.8%。办公室使用了预算拨款的 94%，剩余 6% 的未使用部分归因于工作人员更替和未填补的职位。办公室根据 1 352 万美元的预期预算制定了 2026 年工作方案。

图二

2025 年独立评价办公室预算因开发署方案资源减少而减少



资料来源：独立评价办公室拨款数据。

34. 独立评价办公室保留了 34 个工作人员职位，其中 28 个为专业职类，6 个为一般事务职类。此外，团队还包括 12 名持国际专业服务合同的长期咨询人。截至 2025 年底，办公室有 8 个空缺职位。

35. 独立评价办公室在其人员中实现了性别均衡(57%为女性，43%为男性)，并确保广泛的地域代表性，以便将多元视角纳入评价。大多数人员仍留在开发署总部，但 30% 位于纽约以外地区：17% 在开发署区域办事处，13% 远程提供研究支持。

开发署

36. 2025 年总体评价资源下降，使开发署偏离了将 1% 的方案资源用于评价的机构承诺。评价资金总额从 2024 年的 4 057 万美元降至 2025 年的 3 782 万美元，专门用于评价的方案资源份额从 0.83% 降至 0.79%。下降的主要原因是分散评价

的供资减少，从 2 672 万美元降至 2 408 万美元。尽管开发署总部和区域局增加了对评价的投资，但开发署国家办事处减少了评价相关活动的支出。

表 2
开发署评价资源(百万美元)

	2018	2019	2020	2021	2022	2023	2024	2025
独立评价办公室拨款	8.7	10.9	11.2	11.4	11.8	13.77	13.85	13.74
分散评价	13.3	14.8	14.5	16.4	25.1	23.04	26.72	24.08
评价资源总额	22	25.7	25.7	27.8	36.9	36.81	40.57	37.82
开发署方案资源中用于评价的份额	0.48	0.58	0.57	0.58	0.77	0.75	0.83	0.79

资料来源：独立评价办公室拨款和开发署报告的数据。

五. 分散评价

执行情况

37. 开发署完成了计划在 2025 年进行的大多数分散评价。在原定于 2025 年开展的 365 项评价中，已完成 314 项(86%)——这一数字与 2024 年的数量十分接近，这可能反映了开发署在一年内开展评价的承载上限。最大的差距出现在非洲(24 项评价被取消，6 项评价仍处于逾期状态)以及欧洲和独立国家联合体(独联体)地区。

表 3
开发署 2025 年分散评价

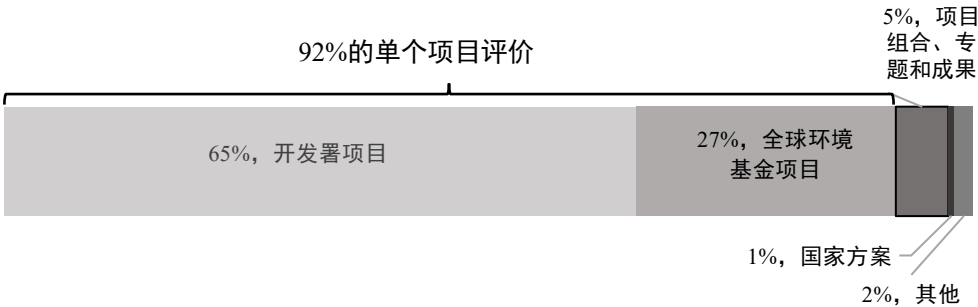
	计划	已完成	取消	逾期	支出(百万美元)
全球*	15	13	2	0	0.56
非洲	104	74	24	6	2.80
阿拉伯国家	47	44	3	0	1.22
亚洲及太平洋	68	67	1	0	2.43
欧洲和独联体	71	58	13	0	1.47
拉丁美洲和加勒比	60	58	2	0	1.49
共计	365	314	45	6	9.97

*不包括资发基金、南合办和联合国志愿人员组织。资料来源：评价资源中心，2026 年 2 月。

38. 开发署的分散评价继续高度侧重于项目层面的举措，从而限制了开展更广泛战略分析的机会。大多数评价(92%)仍在项目层面进行，已完成 289 个项目评价。其中 84 个是对全球环境基金供资的项目的评价(占当年完成的所有评价的 27%)。开发署仅进行了 17 个项目组合/专题/成果评价，以及两项国家方案分散评价，比 2024 年减少了 24 项。这些类型的评价仅占总数的 6%，远低于《2025 年战略计划综合成果和资源框架》中设定的 20% 目标。

图三

开发署 92%的分散评价侧重于单个项目



*不包括资发基金、南合办和联合国志愿人员组织。资料来源：评价资源中心，2026年2月。

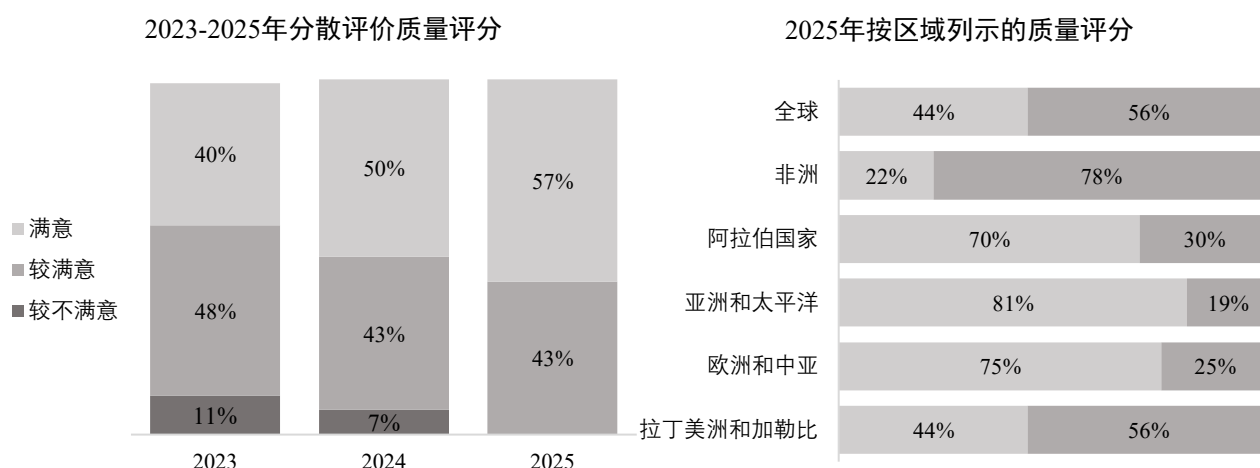
质量

39. 2025 年，分散评价质量继续提高，巩固了近年来的上升趋势。在接受质量评估的 256 项评价中，147 项(57%)被评为满意，109 项(42%)被评为较满意。没有任何评价被评为不满意。满意评价的比例从 2023 年的 39%上升至 2025 年的 57%，而同期较不满意的评价大幅下降。

40. 质量审查显示，整个组织的评价实践基础牢固，但需要投入更多且更持续的关注，以充分整合包容性和跨领域优先事项。72%的评价的工作范围和评价设计被评为满意，体现了坚实的规划与明确的目标。一半以上的评价在报告结构、方法和数据源(58%)以及评价结果、结论和建议(64%)方面达到了满意标准。在跨领域问题，包括性别平等、人权和不让任何一个人掉队方面，仅有 38%的评价被评为满意。加强这一方面至关重要，以确保评价不仅系统地评估有效性，而且系统地评估公平性和包容性。

41. 开发署各区域局的质量模式各不相同。在三个地区(亚洲和太平洋、欧洲和独联体以及阿拉伯国家)，绝大多数评价均达到严格的质量标准，其中欧洲和独联体区域局所覆盖国家的改进尤为显著。相比之下，在非洲、拉丁美洲和加勒比地区以及由总部进行的评价，总体上仍是较满意，非洲区域局所覆盖国家的绩效下降。这些差异突显了需要继续提供有针对性的支持，以在必要情况下加强质量，同时也需要努力在绩效较高的环境中巩固良好实践。

图四
分散评价的质量继续提高



资料来源：独立评价办公室的分析。

促进提高质量的因素

42. 开发署在分散层面的评价职能专业化和强化方面取得了重大进展。截至2025年底，由于开发署政策局和区域评价协调人的持续跟进，几乎所有评价协调人(225人中的220人，即98%)都获得了认证。与此同时，开发署增强了监测和评价协调人的权能，以管理评价流程，并通过绩效审查让他们对分散评价的质量负责；在2025年6月接受调查的监测和评价官员中，超过四分之三的人证实了这一点。

43. 政策局在加强整个开发署分散评价的规划、管理和监督方面发挥了重要作用。根据分散评价强化战略，政策局向业务单位和开发署托管的机构提供持续的技术咨询，支持评价计划的设计和实施。这包括对大约25份新国家方案文件的评价计划实施质量检查，帮助确保这些计划均衡、可行并与方案优先事项保持一致。政策局还组织了五次区域方案和项目管理讲习班，每次讲习班都包括评价会议。为了促进问责制，政策局对照四项机构指标监测分散评价的绩效，并定期向开发署组织绩效组报告。

44. 在独立评价办公室强有力的区域存在的支持下，办公室的指导工作还继续提升评价工作的质量。2025年利益攸关方调查结果证实，各方高度赞赏独立评价办公室的能力建设讲习班，并广泛使用办公室的指南，特别是开发署和全球环境基金的项目评价指南。评价资源中心的“评价方法”页面也获得积极反馈，该页面于2025年进行了改版，并通过一项非常成功的LinkedIn活动进行了推广。尽管分散评价的责任归属于开发署政策局和区域局，但独立评价办公室对特定质量保证请求的响应，有助于建立信任，并巩固了办公室作为建设性、以解决方案为导向的合作伙伴的定位。

持续挑战

45. 开发署工作人员使用来自分散评价的证据的能力需要进一步加强。所查明的差距，与其说是与开展评价有关，不如说是与评价结果的运用有关：包括促进知识共享、将评价结果融入项目规划和决策过程，以及在整个评价过程中让利益攸关方参与。2025 年利益攸关方调查的结果显示，各方对开发署工作人员应用评价结果的信心的信心有限，只有不到一半的受访者同意工作人员已具备这种能力。

46. 员工时间和财政资源有限仍然是主要限制因素。在 2025 年 6 月接受调查的监测和评价协调人中，近一半报告称，他们用在评价上的工作时间不到一半，只有少数人(14%)将大部分时间投入到这一职能上。财政限制也被确定为一项重大挑战，影响分散评价的规划和实施，并限制了实地工作的范围。

47. 激励措施薄弱进一步减少了评价证据的系统性使用。许多在项目周期后期进行的分散评价仍然主要被视为问责工具而不是学习工具。尽管大多数调查答复者(57%)认为开发署总体上有强有力的评价文化，但确信工作人员与合作伙伴受到鼓励去实际应用评价证据的受访者的人数要少的多。仅有一小部分人(11%)强烈同意必要的驱动因素已经到位，许多受访者表达了中立的观点。

48. 消除这些限制因素，将需要对评价进行更明确的优先排序，并加强管理层的参与。开发署监测和评价官员强调了领导层参与的重要性，以动员支持，更系统地将评价费用纳入项目和国家方案预算，并使工作量与现有资源更好地相匹配。他们还强调需要酌情调整评价方法，包括在资源不足时减少评价数量，或将有针对性的评价工作作为案例研究纳入更广泛的方案评价中。

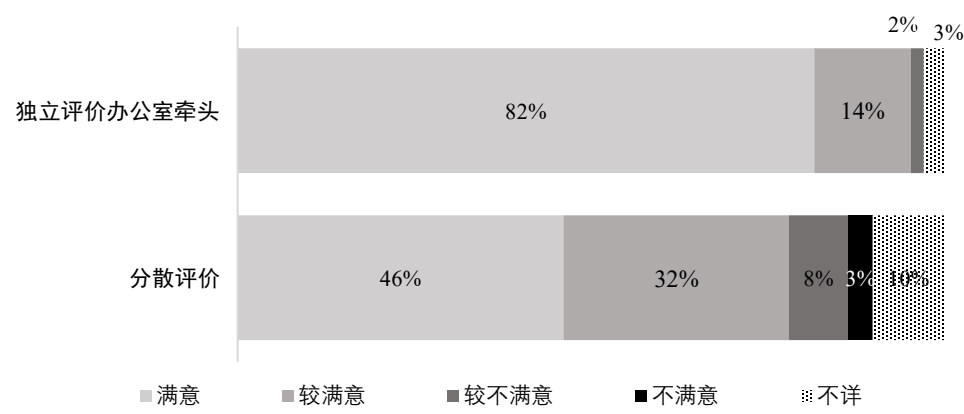
六. 评价后续行动

49. 开发署在评价后续行动方面表现出高度合规，几乎所有近期完成的评价均收到了管理层的回应。2021 年至 2025 年的所有独立评价办公室评价都收到了管理层的回应，95%的分散评价也收到了管理层的回应。各地区的表现始终强劲。相比之下，全球分散评价的后续行动较弱：10%的评价没有收到回应，而且在大多数情况下，回应已逾期六周以上。

管理层回应的质量

50. 开发署的管理层回应及行动计划总体上达到了最低质量标准，与独立评价办公室评价相关的回应及计划，其质量表现则要显著更优。2025 年底对 125 项评价进行的一项研究发现，在过去四年制定的分散评价管理层回应中，仅有不足半数在相关性、清晰度和准确度方面令人满意，其中包括具体、有时限的行动和明确分配的责任。另外 32%被认为是较满意，而 11%没有达到最低质量门槛。

图五
管理层对独立评价办公室评价的回应质量远高于对分散评价的回应质量



资料来源：关于管理层回应落实情况的研究。

51. 在实践中，管理层回应经常重新解释评价建议，往往缩小了范围，而且关于这些调整的依据的透明度有限。即使建议被正式接受，也没有明确迹象表明这些建议在转化为具体行动之前已在最高层进行解读。此外，大约三分之一的管理层回应，合并了多项建议或忽略了具体要素，尤其是在评价报告提出了较长或多层级的建议清单时。这些修改背后的原因很少被记录。

执行情况

52. 开发署报告称，在执行评价建议方面取得了扎实进展，尽管各区域的业绩仍然存在差异，与前几年大体一致。总体而言，73%的行动已经完成，15%已启动，8%未启动或逾期，5%不再适用。与 2024 年相比，所有办事处已完成行动的比例增加。亚洲和太平洋地区表现最好，开发署非洲各办事处和全球分散评价落后较多，逾期或未启动行动的比例较高。

表 4
按区域分列的分散评价行动计划完成情况(2022-2025 年)

	已完成	已启动	未启动	逾期	不再适用
全球	72%	13%	7%	5%	3%
非洲	64%	16%	5%	10%	5%
阿拉伯国家	68%	21%	4%	1%	6%
亚洲和太平洋	82%	12%	1%	0%	5%
欧洲和独联体	74%	15%	6%	0%	5%
拉丁美洲和加勒比	75%	13%	2%	5%	5%
平均	73%	15%	4%	4%	5%

* 不包括资发基金、南合办和联合国志愿人员组织。资料来源：评价资源中心数据，2026 年 2 月。

53. 开发署通常会依据独立评价办公室的建议采取行动，但并非连贯一致。在 2022 年至 2025 年开展的办公室评价所引起的各项行动中，64% 据报已完成，13% 已启动，17% 仍处于逾期状态，6% 尚未启动。逾期的行动涉及 22 项评价，主要是非洲、拉丁美洲和加勒比地区的独立国别方案评价，以及五项专题评价。

54. 尽管大多数行动计划均报告了所取得的成就，但证据基础的质量并非总是坚实可靠。研究发现，81% 的实施的行动达到令人满意的标准，其中 53% 的行动被评为完全满意，28% 的行动被评为较满意。尽管如此，在某些情况下，尚不清楚哪些产出或里程碑实际上已经实现，管理层的回应往往仅止步于意向或未来计划阶段。当行动计划模糊或界定不足时，执行工作往往陷入停滞，范围更加有限，或者被证明难以评估。

55. 当责任由多个行为体(特别是开发署以外的行为体)共同承担时，或者未预见所需资源时，执行差距最为明显。涉及多个利益攸关方、更高组织级别和(或)外部合作伙伴的行动，更有可能被推迟或被视为不再相关。相比之下，分配给国家、地区或总部架构内明确界定的单位的任务更有可能完成。尽管许多后续行动需要额外的能力、技术工作或资金，但管理层的回应往往未能一致地考虑这些资源需求——这阻碍了实施工作的开展，即便已做出明确的推进承诺。

不限于管理层回应

56. 研究证实，管理层回应仅反映了评价如何影响决策的部分情况。在实践中，最大的影响往往产生于评价进程之中或紧随其后，即讨论新结果之时——此类交流往往比正式的后继机制更具催化作用。在某些情况下，管理层回应只是记录了已做出的决定。评价还以未完全体现在正式跟踪系统中的方式影响了变革，最常见的方式是影响战略思维和方向。评价支持了与捐助方的对话，加强了与政府伙伴的合作，增强了开发署的公信力，并为敏感或政治上复杂的决策提供了依据。在一些情况下，评价结果为调整执行模式提供了依据，并推动了更切合实际、与现有资源相匹配的目标设定。

七. 独立评价办公室对全球合作伙伴关系的贡献

57. 2025 年，独立评价办公室明确并巩固了自己在全球合作伙伴关系中的作用。鉴于工作重心始终聚焦于支持共同的发展方向，以及开展学习与知识交流，办公室在有限资源的配置方面更为审慎——此举之所以可行，得益于早年间建立的合作伙伴关系和进行的投资。

国家评价能力发展

58. 独立评价办公室调整了国家评价能力发展的方法，从单独的培训转向加强更广泛的证据生态系统。这一转变与一系列的四年度全面政策审查决议一致，反映在《独立评价办公室 2025-2029 年能力发展战略》中，该战略强调了将机构联系起来、协调合作伙伴和促进国家自主权的重要性。通过与联合国实体、学

术机构和国际伙伴合作，办公室将致力于改善协调、减少重复并更有效地利用现有资源，为可持续评价能力奠定更坚实的基础。

59. 为了实施该战略，独立评价办公室优先考虑加强联合国对国家评价能力发展的集体贡献。作为联合国评价小组国家评价能力发展兴趣小组的牵头方，办公室协助梳理了 35 个国家的相关举措，并确定了多个机构通过培训、政策指导或技术援助提供支持的若干实例。独立评价办公室还与联合国全系统评价办公室、联合国儿童基金会(儿基会)和联合国促进性别平等和增强妇女权能署(妇女署)合作，在经济及社会理事会(经社理事会)业务活动部分会议期间举办了一次“有效问责：面向会员国的联合国系统评价”的正式会外活动。随着重新强调联合国对国家评价能力发展的贡献，独立评价办公室选择辞去由世界银行管理的全球评价倡议联合主席的职务，但仍在开展合作，并帮助制定该倡议的战略。

60. 与此同时，独立评价办公室深化了与德国发展评价研究所和其他联合国实体的伙伴关系，以开发评价系统成熟度模型。设计这一模型是为了帮助各国超越现有的国家评价能力发展努力，这些努力侧重于衡量国家评价能力。该模型将提供切实可行的步骤，推动将评价纳入治理结构，从而随着时间的推移增强决策、问责制和公众信任。

61. 此外，独立评价办公室与开发署合作，将评价作为一个组成部分，纳入治理方案拟订工作。2025 年 4 月，办公室在公共行政专家群体年会的“未来政府”正式会外活动上，介绍了评价整合的案例。

全球可持续发展目标综合联盟

62. 自全球可持续发展目标综合联盟于 2022 年启动以来，独立评价办公室发挥了核心的召集和领导作用，帮助将共同抱负，即获取关于推动可持续发展目标的“行之有效举措”的更有力证据，转化为一个切实运作的多机构平台。该联盟最初仅是联合国各评价办公室之间的一项非正式协作，如今已发展成为广泛的伙伴关系，汇聚了联合国实体、双边及多边组织、基金会和研究机构。

63. 2025 年，在独立评价办公室的支持下，该联盟还能够为关于可持续发展优先事项的综合分析争取到大量技术和资金支持。2025 年，在联合国大会上宣布了用于证据综合基础设施协作倡议的 1.26 亿美元基础投资。这标志着全球综合工作在规模与雄心上均实现了质的飞跃，使联盟得以超越孤立、基于项目的综合，转向协调一致的长期基础设施，从而支持涵盖各项可持续发展目标的政策相关综合工作。

64. 随着联盟进入下一阶段，其治理和运作安排也已成熟。鉴于核心供资得到保障，独立评价办公室将继续担任联盟主席，由资金充足的专门秘书处提供支持。一个代表联盟日益增长的多元伙伴关系的多利益攸关方小组，将提供指导并指引方向。

联合国评价小组

65. 独立评价办公室继续在联合国评价小组中发挥强有力的显著作用，为整个联合国系统的评价规范和方法的进步做出贡献。办公室继续主持联合国评价小组，评价小组的秘书处设在办公室内。独立评价办公室工作人员参加了 13 个联合国评价小组工作组，涵盖性别平等、残疾和人权；气候变化和环境；数据和人工智能；人道主义-发展-和平联系等核心领域。通过这一参与，办公室为集体知识共享和同行学习做出了贡献，加强了联合国评价小组作为整个联合国系统评价质量、独立性和公信力的守护者的作用。

66. 独立评价办公室在选定的联合国评价小组 workflow 中发挥领导作用，这些 workflow 对于加强评价的使用和未来相关性至关重要。办公室协调了评价中的行为科学兴趣小组，深化了关于认知、制度和激励相关因素的实用见解，这些因素影响着评估结果是否被使用以及如何使用。办公室还继续在综合工作组中发挥领导作用，连续第三年担任共同召集机构，带头开展提高整个联合国系统综合标准的工作，包括为此制定指南、举办能力建设讲习班以及与全球证据综合同业交流群进行学习交流。与此同时，办公室还为评价中的前瞻性的方法创新做出了贡献，展示了情景设定如何有助于强化建议。

八. 开发署托管实体中的评价工作

联合国南南合作办公室

67. 2025 年，联合国南南合作办公室(南合办)完成了两项评价。

68. 《对 2022-2025 年南合办战略框架的评价》发现，南合办在促进南南和三方合作方面发挥了重要作用，其政策倡导和能力发展工作在全球和整个联合国系统已取得显著成效。南合办对大会进程、政府间对话和指导的贡献增强了其机构合法性，鼓励了全系统开展南南和三方合作。这些领域的大多数计划成果都进展顺利。

69. 然而，南合办在知识交流、创新和信托基金支持的干预工作方面的表现需要改进：其增值主张不够明确，与国家优先事项的契合度较弱，影响力的证据有限。评价建议南合办明确其催化作用并做出有针对性的调整，以便在快速演变的合作格局中保持适应性。建议包括优先考虑影响力大的干预工作；使雄心与机构能力相匹配；通过更强有力的成果管理制度来加强适应性学习。

70. 《对 2020-2025 年联合国全系统南南和三方合作战略的评价》认为该战略极具相关性。该战略在总部提供了有用的参考，并帮助提升了南南和三方合作在联合国各实体以及国家一级方案拟订中的地位。然而，该战略在联合国系统的集体价值主张和协作优势方面，缺乏明确性。各机构所取得的积极成果，主要反映了各自的任务授权，而不是该战略带来的增值。投资不足、协调人频繁更替以及制度基础薄弱，削弱了连续性和影响力。缺乏基线和目标，限制了问责制和学习，使有效性和可持续性进一步受到限制。

71. 评价建议下一个战略界定明确的愿景和集体价值主张；加强成果问责制和领导监督；将机构间机制重新定位为伙伴关系平台；促进联合拟订方案和资源调动；投资于更强有力的体制安排。

72. 这两项评价旨在为《南合办 2026-2029 年战略框架》和《联合国全系统南南和三方合作战略》提供信息。时间安排实现了同步，因此证据可以直接影响战略选择、成果框架和机构定位。这两项评价之间的协调一致，使数据收集更为高效，评价人员与规划人员之间的协作更为紧密，并为决策构建了更为连贯的证据基础。各小组之间的密切接触确保了评价结果和建议在整个规划过程中得到整合。南合办还举行了有针对性的简报会，与利益攸关方分享成果，并采取了参与式的管理层回应办法，增强了责任和问责。

联合国志愿人员组织

73. 2025 年，联合国志愿人员组织完成了对其《2022-2025 年战略框架》的最终评价。此次评价重申了联合国志愿人员组织对问责制、学习和创新的承诺，为 2026-2029 年规划周期提供了宝贵的证据基础。

74. 此次评价发现，《联合国志愿人员组织战略框架》与不断演进的联合国发展议程和可持续发展目标保持一致。

75. 志愿人员人数从 2022 年的 12 408 人增加到 2024 年的 14 631 人，对本国联合国志愿人员以及拉丁美洲和欧洲志愿人员的需求不断增长。这一增长主要是由于这些地区持续存在的危机，联合国志愿人员组织能够通过吸引合格人才来应对危机。由于维和任务减少，西非和中非的联合国志愿人员人数略有下降。联合国志愿人员组织扩大了与联合国实体，特别是开发署、儿基会 and 世界卫生组织的互动协作，并因其灵活性、多样性和当地参与而受到认可。联合国志愿人员组织推进了性别平等和残疾包容，女性志愿人员数量增加，残疾志愿人员的数量也越来越多。

76. 数字化转型提高了联合国志愿人员组织的业务效率，尽管在任务匹配、外地支助职能的人员配置以及有效部署获全额资助的志愿人员方面仍存在挑战。95%的合作伙伴对征聘和部署工作感到满意。

77. 联合国志愿人员组织的财务可持续性得到成本回收和捐助方伙伴关系的支持。由于特别志愿人员基金的捐款略有下降，因此需要进一步实现捐助方多元化。

78. 评价建议：

(a) 加强志愿服务和联合国志愿人员组织的参与同联合国系统优先事项和规划框架之间的战略一致性。

(b) 增强联合国志愿人员的权能，积极支持他们做为可持续发展目标倡导者发挥作用，从而提高志愿服务的影响力。

(c) 通过改善数字集成、用户体验和系统响应能力来提高业务效率。

(d) 通过一项支持一致性、能力和响应能力的全组织综合战略，使志愿人员的动员更加有效。

(e) 提高业务灵活性和响应能力，以更好地满足联合国实体和社区不断变化的需求。

(f) 通过加强联合国志愿人员组织的外联、参与和沟通战略，增强志愿服务在发展框架内的制度性整合并提升可见度。

(g) 扩大全额供资方案和特别志愿人员基金的捐助方群体，反映会员国对联合国志愿人员组织任务和志愿服务的支持。

(h) 更有力地将跨领域优先事项，特别是性别平等和残疾包容纳入志愿人员组织的服务范围。

79. 联合国志愿人员组织继续推动新评价人才的融入，以支持联合国系统内评价能力的发展。联合国志愿人员组织与联合国评价小组合作，于 2025 年向 8 个国家(肯尼亚、塞内加尔、匈牙利、日本、巴拿马、卡塔尔、泰国、乌拉圭)的 6 个联合国实体(联合国教育、科学及文化组织、开发署、联合国人口基金、儿基会、联合国大学、世界卫生组织)派出 10 名新晋青年评价员。

80. 为了避免与 2024 年中期审查和《2022-2025 年战略框架》最终评价(两者都涉及与性别平等相关的方面)的范围重叠，计划中的性别平等评价被推迟到 2026 年。修订后的时间表将确保关于上述两项评价(包括与性别平等相关的结果)的管理层回应的落实，并有助于对工作成效进行更有力的分析。此次专门的性别平等评价将重点关注性别平等主流化的制度层面，涵盖联合国志愿人员组织的内部人员队伍以及该组织动员志愿人员的实践做法。

联合国资本发展基金

81. 联合国资本发展基金(资发基金)重申对强有力评价的承诺，在 2025 年大幅扩大了评价工作，使评价数量增加了两倍，并投资了 99 万美元用于独立评价。在已开始的 9 项评价中，3 项已完成，6 项将迟于 2026 年年中最后完成。这些评价覆盖 37 个国家和大约 5 亿美元的投资组合。

82. 对青年创业方案的最终评价，评估了该举措对得不到充分服务的青年获取金融服务的影响。评价强调了资发基金在利用数字金融服务改善经济包容性方面取得的成功。尽管该方案的相关性和效率很高，但在开展干预工作后，金融服务提供商向可持续商业模式的转型仍是一项挑战。各项建议强调，需要更长的实施时间框架和更系统的知识管理方法。

83. 对“最后一英里”融资信托基金的形成性评价证实，资发基金在高风险市场承担早期风险的方法，可以有效地为较大规模的金融机构铺平道路。这项形成性评价的建议，重点关注资发基金如何借鉴其他基金的经验教训和最佳做法，为最不发达国家设计一个新的投资融资平台。所涉要素将包括：包含变革理论和明确的发展成果框架的战略价值主张；基金和工具层面的关键绩效指标；有

资金回流的多年供资战略；清晰的财务与成果报告；健全的治理与监督机制；明确的知识职能。

84. 基比拉和平避难所项目的最终评价分析了项目在高冲突、高风险环境中的相关性，重点关注环境保护与建设和平的交汇点。此次评价赞扬使用混合融资，通过“可融资的基比拉”模式使社区主导的项目具备可融资性。该模式通过为当地社区提供有利于森林的可持续生计，成功地减少了社会冲突。需要进行更强有力的监测，以更好地了解这些金融投资对当地和平与稳定成果的直接贡献。

85. 随着资发基金过渡到新的《2026-2029 年战略框架》，其评价职能正在演变，以满足更复杂的发展融资格局的需求。2026 年的主要重点将是加强对金融工具和投资计划的评价，确保资发基金作为催化型金融实体的独特任务得到严格证据的支持。这一演变包括纳入新设立的影响衡量、核查和学习部门。这一转变旨在加强组织的整体数据收集和影响衡量能力，从而提升可靠证据的可获得性，为未来的评价提供依据，并推动组织学习。
