

案例 | 社会组织监测与评估创新



资助者圆桌论坛 (CDR)

2020年3月

背景

- ❑ 良好的监测评估是提高社会部门效率的关键。随着新思想、技术、工具、信息平台的涌现，越来越多的伙伴开始重视**服务对象的反馈**，重视**有效/无效背后的机制**，期望监测评估能够**服务于资助对象、基金会、议题或行业**，期望有**更大范围的知识和经验共享**。
- ❑ 新慈善资本（New Philanthropy Capital，简称NPC）是英国的一家慈善智库。2017年NPC采访了全球不同地区的多位评估专家（见附录），探讨了当下社会的发展对监测评估的影响，提出了**8种监测评估的创新方式**。本次我们对该报告进行了编译，供大家挖掘监测评估的更多可能性。
- ❑ 本文编译有删减，原文请参见：www.thinknpc.org/resource-hub/global-innovations-in-measurement-and-evaluation/
- ❑ 拓展阅读：CDR曾编译《案例：重新想象监测、评估、学习》，参见：www.jianguoyun.com/p/DY9dSYsQ6YDCBhig_0g

本文目录

NPC汇总了最有可能长久影响监测评估实践的8种创新方式，这些方式中**有的是由前沿技术驱动的，有的是正被日益广泛应用的想法**，这些想法正逐步超越传统的评估模式，为未来的评估工作带来更多可能。

1. 以用户为中心的评估 (User-centric Evaluation)
2. 共享测量和评估 (Shared Measurement and Evaluation)
3. 理论评估 (Theory-based Evaluation)
4. 绩效管理 (Impact Management)
5. 数据链接 (Data Linkage)
6. 大数据 (Big Data)
7. 遥感技术 (Remote Sensing)
8. 数据可视化 (Data Visualisation)

1. 以用户为中心的评估：将用户置于评估的核心（1/2）

是什么？

- 以用户为中心的评估是让服务对象在评估过程中发挥意义和价值：**在评估的计划、实施和复盘等所有阶段考虑用户的意见，听取用户反馈并根据反馈采取行动。**
- 该方法试图将权力关系从单向关系（仅单方面从用户处提取信息）转变为**用户同样拥有向上问责权力的双向对话。**

为什么重要？

- 能够确保评估的设计、执行和评估结果的收集**更符合需求**，让评估**更有意义**。
- 在评估者向决策者展示结果，**帮助指导决策时**，**以用户为中心的评估结果更有说服力。**
- 能够**让服务对象感受到自己被重视**，感受到评估不仅仅是对资助者负责，也是对用户负责。

优势和机会

- 确保干预措施是符合用户需求的，确保评估关注的是有意义的问题。
- 无论组织的规模大小，都可以以用户为中心开展评估。

劣势和威胁

- 需要组织文化层面的变革，如果只有口头承诺或表面文章则不易实现。
- 如果对用户的反馈置之不理，反而可能会破坏伙伴关系。

1. 以用户为中心的评估：将用户置于评估的核心（2/2）

大家是如何使用的？

- **资助者的应用：**共享洞察基金（Fund for Shared Insight）是一个资助者联盟，其成员认为鼓励和整合反馈是改善慈善事业的关键路径，他们发起的善倾听（Listen for Good）项目帮助非营利组织和资助者更好地收集最终服务对象的声音。
- **人道主义救援工作的应用：**牛津饥荒救济委员会借助智能手机收集难民对难民营的反馈，这些数据每天上传到中央数据库，工作人员依据问题清单进行审查并解决问题。该项目在试点早期时发现社区成员的信任度有所提高。
- **公共部门的应用：**南非政府通过收集公民的监测反馈意见来推动公共服务的改善。

案例：LIFT通过用户反馈了解什么是真正有效的

- LIFT是一家帮助低收入群体结束代际贫困的NGO。该机构此前监测的指标一直不够具体，之后LIFT开始系统地收集服务对象的体验反馈。
- 用户反馈帮助LIFT充分地了解：**用户有需求的时候能接受到服务吗？LIFT是否需要扩大服务范围或执行新的干预措施？用户认为的优先事项与LIFT的假设是否一致？LIFT与服务对象之间是否能够顺畅高效地合作？**
- LIFT借助反馈不断改进，以确保服务达到用户的需求。

2. 共享测量和评估：与同道中人合作（1/2）

是什么？	为什么重要？
- 共享衡量和评估适用于具有相似使命、行动策略或服务对象的机构，他们可以通过开发和使用相同的监测评估指标，来衡量自身的影响及集合影响力。 - 共享的内容可以包括： - 评估目标和指标 - 评估工具和方法 - 评估数据和结果	- 对于单个组织来说，可以减少自行开发评估工具的时间和成本。 - 通过汇集评估数据、比较评估结果，能够创建一个更大的信息库，让组织对标行业基准，产生更有意义的洞察。 - 信息库还能够帮助组织更好地了解需求和有效的干预策略，并基于此改进服务策略。
优势和机会	劣势和威胁
- 便于比较，便于更快速地积累知识。 - 节省资源，有利于推动部门层面的变革。	- 共享工具的灵活性可能不够，不一定适用于不同的项目。 - 新项目需要先发展一段时间，然后才适合与已有的干预措施进行比较。

2. 共享测量和评估：与同道中人合作（2/2）

大家是如何使用的？

- **共享实证依据：**3ie（International Initiative for Impact Evaluation）开发了证据差距地图，汇总了关于不同议题领域“什么有效”的证据集合。
- **共享评估指标：**全球影响力投资网络开发了IRIS指标，其中包含大约400个社会和环境影响指标，为社会影响力投资者提供了通用的标准化评估工具。

案例：PCOC为整个行业制定评估工具和基准

- PCOC（Palliative Care Outcomes Collaboration）是澳大利亚的一项国家计划，旨在将标准化的评估工具引入国家层面的姑息治疗工作中。
- PCOC提供的服务包括：
 - 与姑息治疗组织**合作开发评估工具，并分享给所有开展姑息治疗的机构。**
 - 合作机构利用共享工具，每6个月向PCOC提交临床数据，PCOC为这些机构提供评估结果以及20个国家共同商定的姑息治疗标准，并**帮助机构根据这些结果和基准改善治疗质量和效果。**
 - PCOC同时举办年度研讨会交流经验，并汇总编写了国家层面姑息治疗领域的月度进展报告，用于行业研究和发展。

3. 理论评估：挖掘有效/无效背后的作用机制（1/2）

是什么？	为什么重要？
- 理论评估不仅关注一个项目是否有效，还进一步挖掘它是如何导致有效或无效结果的。这种方法通常先假设活动将会如何导致变化，然后验证假设。 变革理论是开展理论评估的基础。常见的理论评估，是基于变革理论中的假设，探索不同情境下的不同机制会产生哪些不同的结果。	- 有助于建立 “干预措施是如何发挥作用的详细图景，帮助人们找到触发结果的关键机制和情境，改进干预方案。 - 有助于找到失败的原因，避免重复犯错。 - 评估者可以依据已有的评估结果选择合适的方案，既可以选择复制最有效的变革策略，也可以选择目前证据最少的模式进行验证。
优势和机会	劣势和威胁
- 在复杂的社会系统中，找到合适的变革路径。 - 复制成功的项目。	- 可能得出笼统的无意义的结论，比如“没有一种方法对任何情境下的每个人都有效”。 - 可能出现信息过于密集或复杂的情况。

3. 理论评估：挖掘有效/无效背后的作用机制（2/2）

案例：Well Doncaster 探究项目成功的原因

- 背景：Doncaster通过提供互动交往空间，为人们提供社交机会，建立充满活力和联结感的社区，以期更好地促进健康和福祉。
- 目的：从表象来看，项目评估者们已经能够看出该空间成功地吸引了民众的参与，但他们还希望进一步了解：该空间是否有效地缓解了人际交往之间的隔离感，是否为人们提供了更多社交机会，以及对不同的人群有哪些不同影响。
- 方法：首先评估者对项目提出一系列变革理论假设：什么可能有效、对谁有效、在什么情况下有效。在文献回顾基础上，评估者通过与利益相关方举办研讨会来提出、验证和发展变革理论中的假设。
- 机制探索示例：参加空间活动的人在进门时会被邀请为场内的某位参与者沏一杯茶。最初项目团队认为这只是为了激励人们进门，但评估后发现这一行动被参与者视为一种建立友谊的提议，对于参会体验感和参会者们在空间中的互动有重大影响。
- 成果：最后的评估结果解释了不同的机制（包括方案中的有形行动要素，如参会者提供/接受一杯茶并回应）如何在特定的背景下（如被忽视的社区）运作，最终产生结果（如更紧密联系的社区）。这为如何在其他社区推广该方案提供了宝贵的见解。

4. 绩效管理：持续收集数据并反馈，促进项目迭代（1/2）

是什么？	为什么重要？
- 绩效管理是指将评估纳入战略和绩效管理，定期分析数据并指导项目改进。类似于我们日常所说的监测。 - 这意味着把数据收集作为日常工作，其目的在于学习、优化项目，及应对实际操作中的不确定性，而不仅是在项目结束时报告和问责。 - 这一方法强调反复测试、迭代，近年来常见的说法是敏捷和精益评估。	- 许多社会组织面临的环境是复杂多变的，越来越多的人关注“评估如何指导更好的决策”。绩效管理允许组织持续追踪项目效果，并指导工作改进。 - 绩效管理可以控制在低成本水平，适用于每个组织，无论规模大小。 - 现今世界面临许多不确定性，且时时刻刻都有新的数据生成，绩效管理能够帮助社会组织更好地适应时代发展。
优势和机会	劣势和威胁
- 明确评估是用于改进的。 - 技术的进步正在加快反馈和改进的速度。	- 如果项目不断发生变化，将较难追踪干预措施对结果的影响。 - 敏捷评估方式仅能验证近期数据产生的影响。

4. 绩效管理：持续收集数据并反馈，促进项目迭代（2/2）

案例：精益数据（Lean Data）借助持续的数据反馈来改进服务

- 背景：Acumen是一家社会创投机构，投资于为低收入群体提供服务的初创社会企业。精益数据是Acumen的一项计划，旨在提升社会组织的绩效。Acumen也希望借助精益数据，转变合作伙伴的思维方式，让伙伴重视并发展评估文化。
- 方法：精益数据计划帮助Acumen资助的社会企业运用低成本的技术，直接与最终服务群体沟通，通过收集到的数据指导决策，提升组织绩效。该方法有以下特点：
 - 其重点在于**倾听最终服务者的声音，并提供社会影响力的持续变化数据**，而不在于证明项目的影响力或寻找因果证据。
 - 其数据收集方法强调**高效和快速响应**，并在此基础上尽可能地严谨。
- 技术：多采用短信、电话采访、人机交互式语音技术收集数据。
- 成果：成功的精益数据项目能够**创造出一个用户信息反馈闭环**，使机构能够做出**更好、更明智的决策**，并根据收到的信息迅速采取行动。如某炉灶公司借此了解哪些渠道能够最有效地接触到分销客户；某能源公司借此了解到客户在安装方面存在的问题，进而增加相应的售后支持；等等。

5. 数据链接：连接不同数据源，描绘更完整的图景（1/2）

是什么？	为什么重要？
• 数据链接是指汇总来自不同组织或子部门的某类特定用户群体的相关数据。 • 例如：一个为无家可归者提供住房机会的公益组织，可以将其数据库与地方政府的居住数据相链接，了解目标群体是否成功入住，若未入住，原因是什么。 • 其前提是机构/部门间数据共享。	• 更完整的数据集有助于描绘项目背景、需求、成果的完整画面，更有利于充分地分析干预措施的影响。 • 有助于在项目之初就找到与干预组条件相匹配的对照组，以便更严谨地追踪项目的长期影响。 • 有助于同类议题不同干预措施的相互比较，或评估同类议题多种干预措施的累积效应。
优势和机会	劣势和威胁
• 避免不同机构重复开展某项工作，有利于提高评估效率。 • 更容易追踪长期影响。	• 需要构建合作伙伴关系，访问其他数据源。 • 机构/部门间数据共享可能面临一系列法律和技术问题。

5. 数据链接：连接不同数据源，描绘更完整的图景（2/2）

大家是如何使用的？

- 数据链接最常使用的数据源是公共部门的**公共服务数据库**；
- 其次是公开的数据，即那些任何人都可以自由访问、使用和共享的**开放数据**，如社交媒体数据。

案例：Blackpool Better Start 合作伙伴共同探索促进学龄前儿童健康的有效路径

- 背景：Blackpool Better Start 伙伴网络，希望通过链接当地的多个合作伙伴（包括政府部门、医院、基金会等）数据，分析学龄前儿童健康的最有效干预措施或干预组合。
- 方法：
 - 收集各方数据，包括当地儿童的入院情况、母乳喂养率、产后抑郁症比例等。
 - 要求每个接受干预的儿童及其家庭同意在伙伴系统内分享其数据。
 - 基于多数据源，分析不同干预措施或干预组合对学龄前儿童健康的促进作用。
 - 为前来接受服务的儿童及其家庭提供个人情况的全面分析和个性化干预支持。

6. 大数据：在更广泛的层面上产生新见解（1/2）

是什么？	为什么重要？
• 大数据通常指的是在数字交易和互动的过程中产生的无结构数据。 • 大数据通常是实时产生、不断变化的，包括人们在网上说什么或做什么（尤其是在社交媒体上）；金融交易、网络搜索或其他网络活动留下的痕迹等。	• 数字时代的到来让数据不断增多，而技术的发展允许人们处理更复杂的数据。 • 大数据有助于在地方、国家甚至全球层面对项目的需求或有效性产生新见解。 • 大数据分析有助于帮助我们识别趋势、模式和行为等。对于社会部门来说，大数据可以非常迅速地提供传统上难以获取的信息，比如公众对人口流动的看法或数据。
优势和机会	劣势和威胁
• 为定量分析提供更多机会。 • 为群体层面的洞察提供更多机会。	• 需要较高级的软件和技能来管理数据。 • 可能面临道德挑战。

6. 大数据：在更广泛的层面上产生新见解（2/2）

大家是如何使用的？

- 大数据通常对于理解用户需求最有用。但是当干预措施影响到大量人群时，大数据同样有用，比如可以通过分析社交媒体中人们对某主题的讨论内容，评估干预措施对公众认知的影响。

案例：联合国通过分析社交媒体数据，理解公众认知，指导全球性运动的开展

- 背景：2013年，联合国千年运动和供水与卫生合作协会曾开展全球卫生运动。为了使运动尽可能有效，他们希望了解网络上谁在影响关于卫生设施讨论的方向、公众关注的卫生设施相关话题有哪些，并监控项目干预的影响。
- 方法：
 - 寻找合作方开发一款工具，监测社交媒体上的讨论和网络新闻内容。这一工具从社交媒体中提取对话内容，识别其中与卫生设施主题相关的文字。
 - 根据分析结果，确定最具影响力的社交媒体账户、公众最关注的话题、评估不同卫生议题的讨论热度持续时长。
 - 借助这些发现，进一步指导卫生运动的宣传和资金筹集重点方向，提高公众参与度，最终改善卫生条件。

7. 遥感技术：追踪远程数据，评估项目影响（1/2）

是什么？	为什么重要？
- 遥感技术是利用一系列技术（如移动电话、传感器、卫星等）收集传统方法无法收集到的远程信息。 - 该方法在私营部门和公共部门的应用包括绘制客户行为地图、交通流量等，它能够实时提供高度详细、准确的数据。	- 有助于从孤立或不同的地点收集数据。 - 几年前无法收集，或需要花费大量成本才能获取的数据，如今随着技术的发展，可以用非常经济的方式收集到。 - 遥感技术是被动地收集信息，而非通过人类输入信息，因此能够降低研究者对结果的影响，减少偏差。
优势和机会	劣势和威胁
- 自动采集信息的形式降低了人为的影响。 - 可以访问远程数据。	- 对许多慈善组织来说，适用性有限。 - 可能面临道德挑战。

7. 遥感技术：追踪远程数据，评估项目影响（2/2）

大家是如何使用的？

- 遥感技术经常在国际发展部门中得以应用，既能测量复杂的数据结果，又能收集到人工难以进入的地区信息（见下方案例）。
- 在健康领域，随着可穿戴设备的发展，遥感技术能够让健康干预类慈善机构的评估者随时了解受益人全天健康情况变化和运动数据，评估干预的影响。
- 遥感还可以直接与居民家中的设备连接起来。例如清洁炉灶项目向居民提供能够使用低污染燃料的炉灶。项目组通过在用户家中安装烟雾传感器，测量用户是否真的使用了炉灶，以及使用者家中的烟雾水平。

案例：Global Environmental Facility (GEF) 利用遥感技术监测多地森林退化程度

- 背景：GEF是由18个国际组织组成的网络，其预防森林退化的项目涉及全球1700余个项目点，很多分布在偏远的地区，这使得传统的评估方法（如人工前往项目地记录当地森林退化情况）变得几乎不可实现。
- 方法：GEF利用卫星获取项目干预地区的详细的森林退化图像，追踪森林退化速度数据，并将其与未接受干预的对照地区森林退化情况进行比较。通过这种方式，GEF能够详细评估其干预工作产生了怎样的影响。

8. 数据可视化：借助直观的视觉效果呈现见解（1/2）

是什么？	为什么重要？
• 数据可视化是以图形化的、生动的和引人入胜的形式呈现数据。 • 数据可视化已经不算是新鲜事物，但是随着技术和工具的发展，数据可视化变得越来越容易，Tableau等工具为数据可视化提供了很多便利。	• 有助于帮助人们看到传统数据呈现方式下看不到的数据趋势或模式。 • 这种方式鼓励数据阅读者参与到对结果的交流讨论中，并使交流更高效。 • 有助于评估者向服务对象或领导层清晰地展现评估结果，激励下一步行动。
优势和机会	劣势和威胁
• 以更易于访问的方式呈现数据。 • 让数据中的趋势或模式更显化。	• 可视化可能一定程度上掩饰不良的评估结果。

8. 数据可视化：借助直观的视觉效果呈现见解（2/2）

大家是如何使用的？

- 目前，数据可视化最常见的用途是使信息对观众更有吸引力和更容易获取。比如通过交互式工具以热图、气泡图、散点图、时间序列等形式呈现数据变化。
- 数据可视化还被用作分析，例如网络分析可帮助理解不同的人或组织之间的联系。

案例：Trussell Trust（TT）借助可视化绘制数据地图，确定优先干预区域

- 背景：TT支持一个由当地400多家食物银行慈善组织组成的网络。TT希望了解各地是如何使用这些食物银行的，为什么使用，能否预测需求。
- 方法：
 - TT首先链接了各地人口普查数据和TT收集的一系列服务数据。但由于数据量大，过于复杂，无法在电子表格中查看，于是通过绘制服务地图，呈现出了食物银行在全国的分布情况、使用情况。
 - 通过地图识别有相似人口信息的不同区域的食物银行使用情况，**从中找到需求没有得到充分满足的地区，借此预测使用需求和优先干预区域。**
 - 未来还能在机构制度或国家政策发生重大变化的时候，衡量政策变化带来的影响。
- TT正将此模式向更多食物银行慈善组织推广。

开始行动

- 本报告中提出的创新方法为我们的监测评估工作带来了一些新的视角。如果你认为对于小型组织来说，先进的技术或复杂的方法太具挑战性，那不妨从简单的地方做起。
- 比如，如果你认为开发一套共享评估指标太难了，那么可以**先开始分享你的评估结果、经验和教训**；如果你认为让服务对象参与到评估的各个环节太耗费资源了，那么可以**从收集用户反馈，并基于反馈改进服务方案开始**。
- 除了本文中提及的理论和案例，原文中还提供了8个趋势相关的阅读文章和网站，以及提升基本评估能力的资源，感兴趣的伙伴可查阅原文。

附：受访名单（1/2）

编号	姓名	机构
1	David Ameyaw	International Centre for Evaluation and Development
2	David Bonbright	Keystone Accountability
3	Megan Campbell	Feedback Labs
4	Julia Coffman	Center for Evaluation Innovation (CEI)
5	Jenny Dibden	Department for Business
6	Alnoor Ebrahim	Harvard Business School
7	John Gargani	Gargani + Company
8	Ian Goldman	South African Presidency's Department of Planning, Monitoring and Evaluation
9	Karim Harji	Purpose Capital
10	Andrew Means	beyond.uptake

附：受访名单（2/2）

编号	姓名	机构
11	Jo Puri	3ie
12	Linda Raftree	Independent Consultant
13	Nissa Ramsay	Comic Relief
14	Lucy Bernholz	Stanford Center on Philanthropy and Civil Society
15	Adrian Brown	Centre for Public Impact
16	James Goodman	Forum for the Future
17	Penny Hawkins	Independent Consultant
18	Sarah Mistry	Bond
19	Jeremy Nicholls	Social Value UK / International
20	David Pritchard	Independent Consultant
21	Patricia Rogers	RMIT University
22	Shubh Sharma	Dasra



资助者圆桌论坛（CDR）

CDR是面向中国资助者的技术支持平台，其使命是“服务中国资助者，探索有效公益，推动社会问题解决”。

cdr4impact.org.cn

版权声明：本报告版权归原作者所有，CDR将本报告编译，为公益同行学习分享所用。