

我们如何知道影响 是否已经发生？

了解和使用影响力证据



简介

目录

简介.....	2
背景设置.....	3
为什么证据很重要?.....	3
什么是影响力证据?.....	3
现有证据.....	4
收集现有证据.....	4
使用现有证据.....	4
新证据.....	5
收集我们自己的证据.....	5
最大化证据的价值.....	7
选择最好的方法.....	7
证据共享.....	7
更多信息.....	8

影响力管理项目由来自不同背景和国家的700多个组织协同运作，旨在商讨影响力的衡量和管理以及我们的目标和绩效方面的基本原则。

上述所说的基本原则用于描述不同基础业务或业务组合对人与地球环境的影响，它有助于价值链上的不同参与者找出管理这些影响所需的信息。在决定改善人类和地球环境之前，首当其冲的便是要收集好影响力的证据，从而弄清楚将会发生哪些实质性的正面和负面影响。

在过去六个月中，内斯塔（Nesta）一直与影响力管理项目保持紧密合作，探索我们应该如何收集和使用影响力证据。本文件就使用现有影响力证据和建立自己的证据基础提供了指导。

本报告的内容是影响力管理项目与业内人士经过一系列探讨讨论的结果。该报告由Nesta和影响力管理项目共同撰写。请将有关本报告的任何反馈意见或进一步询问发送至：

team@impactmanagementproject.com

注：本文原文为英文，故不排除翻译误差。

致谢

影响力管理项目衷心感谢所有项目参与者就影响力证据提供的超凡见解。同时，还要衷心感谢来自Nesta的共同执笔者露西·海蒂（Lucy Heady）。

背景设置

为什么证据很重要？

做决策时要以证据为导向，这似乎无需赘言。为什么不能用我们已知的信息来指导我们的决策？当然，事情不能只看表面。如果撕开表面的外衣，深入事情的本质，本来一目了然的事情似乎立即不那么明晰了。不同领域的人士会优先考虑不同类型的证据。而各领域之间的技术引进有时并没有考虑到其中一些重要的差异。当不同类型的证据不一致或几乎没有证据可以表明事情有效时，放弃证据似乎更容易，更省事。

但没有证据，我们必定会重蹈覆辙。如若凭直觉和冲动行事，例如为了让儿童树立起安全意识远离犯罪就让儿童在监狱中体验一段日子，亦或是为了让青少年避免未婚先孕而扔给他们一个哭闹的婴儿，那么，此番结果必定是有害无益的¹。社会影响力的实现是一段艰难的历程，我们决不能简单认定逻辑似乎优胜于假定。通过建立不同环境不同方法情境下的证据基础，并不断加以学习和调整，我们方可取得积极影响并吸引更多的投资。

使用和收集影响力证据不需要你具备社会科学博士学位或百万美元的预算。以下规范就向大家呈现了以具有成本效益的方式解释现有证据和收集新证据的原则。只要有可能，我们都会提供更详细的资源，以展示这些原则的实际应用。

什么是影响力证据？

我们将“影响力证据”定义为：

“可用于判断所产生影响达到何种程度（或没有产生影响）的可用事实或信息。”²

我们在影响力管理过程中收集了许多不同类型的信息，其中只有部分信息用于了解人类和地球环境，故属于“影响力证据”。影响力证据有助于我们了解已产生的影响，令我们能够降低证据风险水平。³

我们使用证据作出判断时应考虑证据是否站得住脚。有很多特质可以帮助我们确定我们是否可以充分信赖此等影响力证据。类似[证据标准](#)的框架有助于我们根据已收集的证据类型来评估我们对证据的自信程度。

¹ 例如：请参考青少年犯罪预防计划‘现身试法’的[评价](#)

² 根据牛津英语词典的定义，明确适用于判断是否发生了（正面或负面）影响。

³ 在影响力管理过程中用于作出决策的其他类型证据有很多，比如专家意见，但这些并不构成影响力证据

现有证据

收集现有证据

我们常用现有证据来设定影响力目标、选择或设计交付模式（如业务或项目）。为此，我们需要思考在所有[影响力维度](#)中有哪些可用证据。

例如，评估一家现有企业（也许是一名潜在投资方）时，我们可能会考虑数据在所有维度（下图1）中的存在有多广，以及企业实现[一组目标](#)的能力。这有助于我们为受影响的一个或一群人（或地球）找出现有证据基础中存在的缺口或薄弱证据，以及可能需要新证据的地方。

使用现有证据？

用于评估影响力的现有证据是否足够‘有力’的考虑因素有很多：

完整性

- 数据集是否完整和全面？

- 数据是向广泛相关参与者收集的吗？

准确性

- 证据符合质量的既定标准和严格程度吗？⁴
- 证据是由外部审核得出的，还是独立评估得出的？⁵

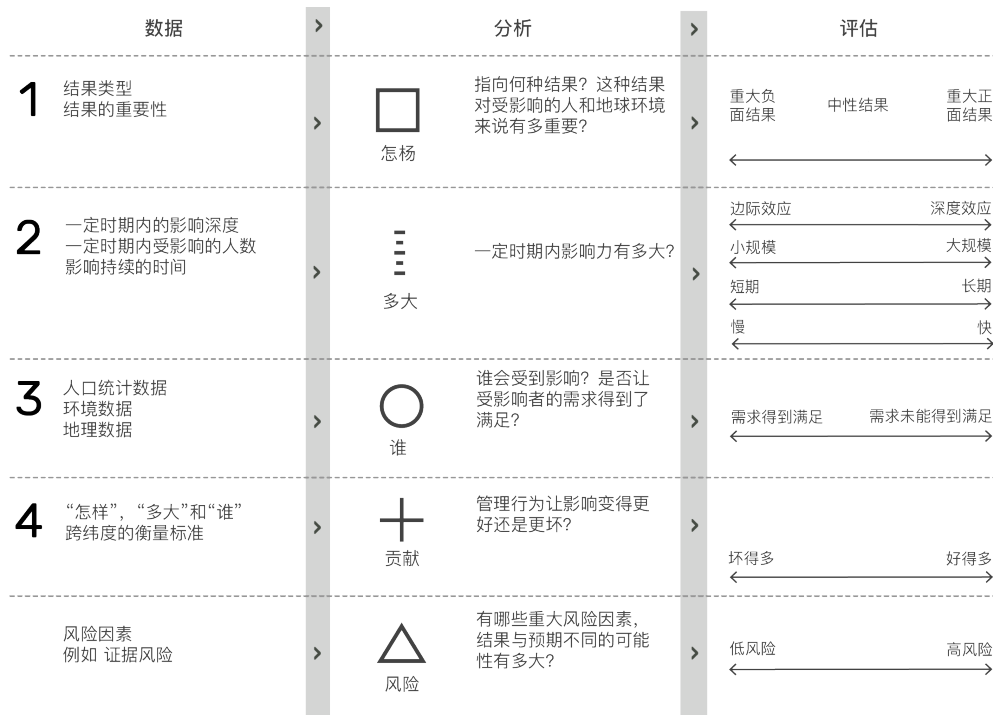
相关性

- 证据直接与所有影响力维度中的相同影响力目标相对应吗？有没有缺口？

在审查一个证据基础时，显示这些特征和结果相同的个体研究（或信息集）越多，证据基础就越有力。⁶

然而，现有证据极少阐明我们需要了解有关既定背景下交付模式可能产生的影响的所有信息。因此，我们需要时不时地收集证据，以重新确认已知信息或填补证据缺口。类似[3ie证据缺口图](#)的资料有助于我们确定这些证据缺口的可能位置。

图1：影响力的维度



⁴ 请参阅：[美国国际开发署数据质量检查表 \(data quality checklist\)](#)

⁵ 请参阅：[Nesta证据标准](#)；[医学证据水平](#)；[什么可行？](#)

⁶ 请参阅：英国国际发展部‘评估证据强度’指南（[‘Assessing Strength of Evidence’ guide](#)）

新证据

收集我们自己的证据

收集我们自己的证据使我们能够证实或反驳现有证据，并进一步了解正在产生的影响以改善人类和地球环境。我们或许还需要了解这种环境在与企业的合作结束后是否会发生变化。我们收集的所有证据都反映了我们制定的战略决策以及设定和重新设定的目标。

所收集的证据类型会因其必要程度而异，譬如，企业需要何种程度的证据来了解自身对人类和地球的影响。我们可以从如下两个维度来考虑这种匹配性：

1. 我们收集的数据类型：我们的自信程度通常与我们使用的数据类型有关。

- 直接测量效果的变化程度可以提高确定性。

例如，改善营养状况或减少二氧化碳排放量。

然而，尽管有这些措施的帮助，要想收集所有完整的人口统计和结果类型数据是一件非常困难或极其昂贵的事情，因为产生效应的时间、客户机密性、缺乏适当的测量技术或者费用等因素都可能成为阻力。在此情况下，常见的做法是收集有关最终影响替代指标的数据，比如，住房改善可能是增加收入的替代指标，或者医院候诊名单可能是卫生服务运作良好的替代指标。

- 鉴于活动或输出数据通常比较容易收集，故常用作一项替代指标。

例如，将接种麻疹疫苗的儿童作为该儿童不会染麻疹的替代指标。

在某些情况下，活动与结果之间存在联系的现有证据非常强大，以至于活动或输出数据可以作为结果的替代指标，比如接种麻疹疫苗的儿童人数。当选择我们自己的指标或使用替代指标时，我们面临进行因果关系假设的风险。如果一家企业正采用一个新方法，或者基于输出数据和结果关联的证据是非决定性的、非实质性的或者无法比对的，那么这些替代指标措施将不构成影响力的有力证据，即使它们能提供有用的证据证明部分变化理论是有效的。

- 我们也意识到人类和地球环境可能也会经历其他影响，包括积极的和消极的影响，因此可以借鉴现有的关于可比较方法在可比情境下提供的实质性影响的信息，以帮助我们判断我们可能想要减轻的（消极影响）或增强的（积极影响）其他影响。

例如，通过咨询现有海上风电场效应的研究，我们可以了解到一些海洋动物由于噪音和常规干扰而遭受负面影响。这些信息有助于我们投入资源以尝试收集新信息，从而发现和减轻这些负面影响。

- 通过调查问卷或调查方法发现受影响的人给出的反馈是衡量影响力发生变化的最直接且最可靠的指标。如果现有证据表明人们的反馈与期望的效果密切相关，那么这种方法是最可靠的。这些数据可能是一项替代指标（例如，“你是否按时服药？”），或者表明任何维度正在发生的影响的最佳指标（例如“你的生活质量改善了多少？是幸福感深度的良好指标）。

新证据 续

2. **我们怎样收集证据**：我们使用的方法透露出我们对数据的信心。

- 我们在与企业接洽之前和之后收集数据（直接措施或替代方法）。然而，这种方法仍然存在潜在偏差，因为任何检测到的变化都不大可能是因单次具体合作而引起的，譬如即使没有这次合作，一部分人也会改变他们的状况，或者有其他合作引起了变化。

例如，孕产妇健康门诊追踪到：二零零八年有470名母亲接受治疗，其中89%的人安全地生下健康的孩子，97%的人对服务表示满意。

- 最强大的自信通常是通过外部评估获得的，这个评估消除了选择偏倚、考虑归因及 / 或揭示其他影响。

例如，一项评估表明，在所有可行治疗中，接受补充营养的母亲最有可能生育健康体重的孩子。

- 使用对照组的评估（如随机对照试验或倾向指数匹配）有助于我们建立因果关系。虽然创新方法的引用降低了这种技术的成本费用，但它们通常仍然需要较高的财政投入和更多的技术知识。

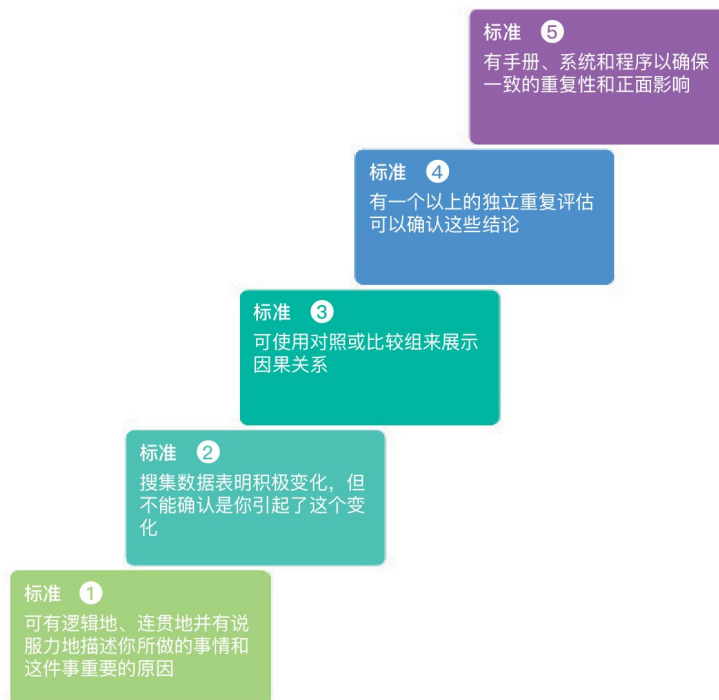
例如，对邻近地区一组类似情况的母亲进行的调查报告表明，同一时期安全和健康分娩率为65%，而孕产妇死亡率比区域平均值低43%。

- 明确背景如何影响这些效应的交付，从而实现可复制性。

例如，促进孕产妇健康的积极成果的因素有很多，包括营养和卫生条件，而不同背景下的人他们在这方面的需求满足程度不同。这将影响产前和产后护理的有效性及 / 或适用性。

证据标准是一个有用框架，帮助我们根据所使用的方法评估我们对证据的信心。

图2: 证据标准示例⁸



⁸ Puttick & Ludlow (2013) ‘[Nesta: 证据标准](#)’。这些标准是基于马里兰科学方法测算表和Project Oracle证据标准 (Project Oracle Standards of Evidence) 制定的。

最大化证据的价值

选择最佳方法

资源最密集最复杂的证据收集方法并非放之四海皆准。所以，我们需要找到一个有用的出发点，即，考虑我们对现有证据的自信以及各方愿意承受的影响力风险水平。

使用成本较低的方法来检验战略执行过程中发生的影响是否合理，然后在确保战略合适后再加大投资。

收集、分析和传播证据都是有代价的，因此最重要的是我们收集的证据既有用又具成本效益。

以下几点有助于我们收集有用的证据：

1. 证据是否有明确的目的？

这可能是：

- 改善人类或地球环境
- 增加有关针对特定人群及 / 或结果起作用（或无效）的公开可用证据
- 降低风险和不确定性，以增加影响按预期实现的可能性

2. 证据价值是否与其成本相称？最有价值的证据可能是：

- 有利于我们管理人类和地球受到的最实质性影响的证据
- 注重建立最具不确定性的证据基础

3. 该方法是否适合所寻找的证据类型？⁹借鉴现有证据基础来了解：

- 现有证据缺口以及它们存在的原因
- 可能涉及的风险（如道德、文化或现实）
- 在这种背景下经常使用哪些数据收集方法

4. 谁会从证据中获益最多？

这应公开：

- 采用的证据标准
- 谁承担费用¹⁰

证据共享

证据共享通过让更多企业学习和改进交付模式来最大化其价值。如何使用证据直接关系到它如何共享。

为了使证据价值最大化，我们在报告证据时应遵循以下原则：

- 透明度：报告应该包括测试内容、时间和结果、方法、数据质量、背景和参与者。
- 及时性：应尽快分享证据惠及他人，并尽可能考虑数据可能影响的决策时间表。
- 综合性：报告应清楚说明证据如何证明在企业缺失的情况下会发生什么，并报告达成预期影响所涉及的任何负面结果和任何假设。
- 背景化：新证据应始终在其他现有证据和初始目标的背景下呈现。这使我们能够看到结果不同之处和学到新经验。

⁹ 请参阅：[IDInsight基于决策的评估 \(IDInsight decision-focused evaluation\)](#)

¹⁰ W如果采用资源密集型方法来回答有助于建立该领域的问题，那么企业就不太适合独自承担费用。

更多信息

推荐资料

发现有力证据

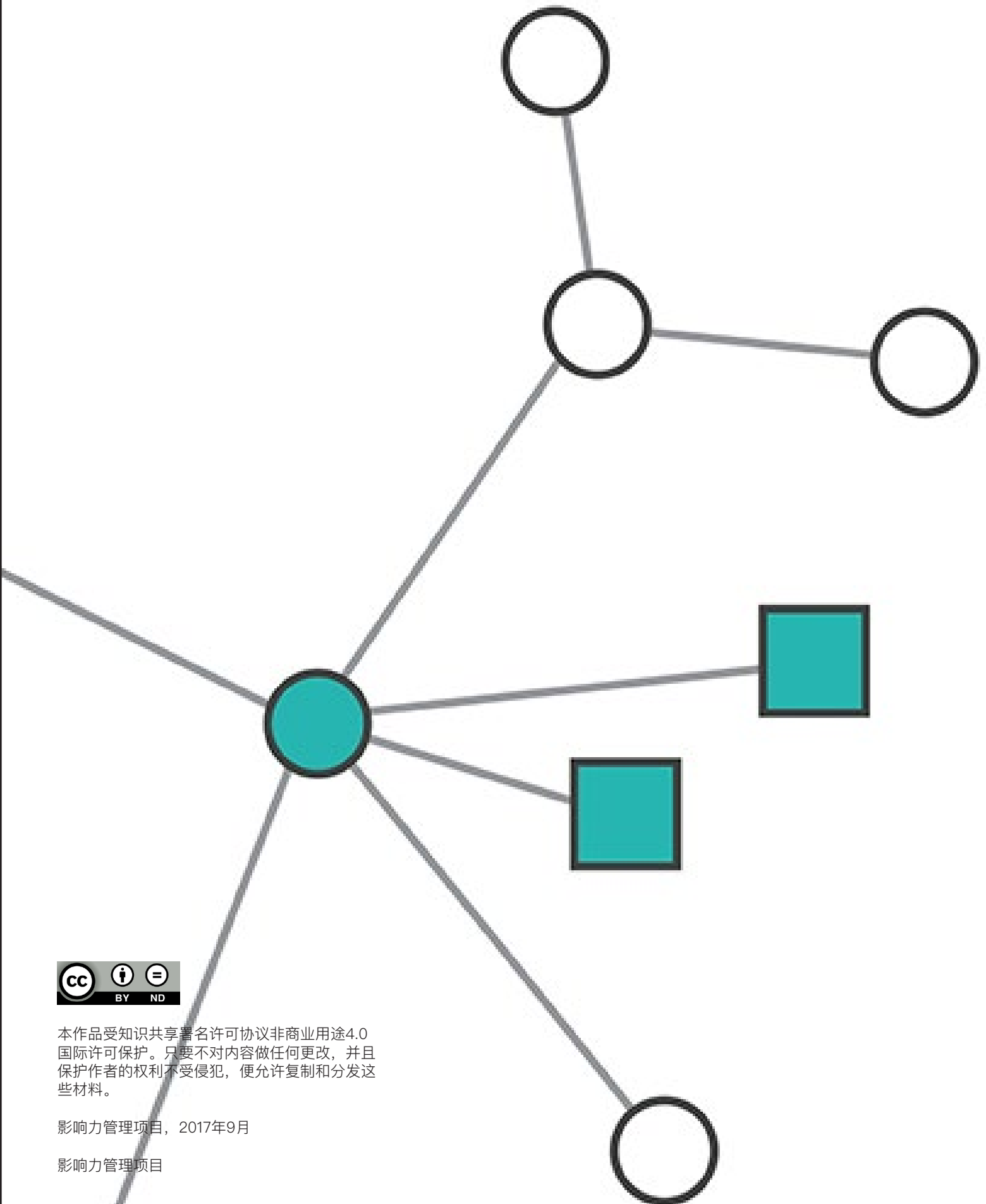
- Bond [评估证据质量的原则](#)
- 英国国际发展部[评估证据强度指南](#)
- Nesta [证据标准](#)
- 美国国际开发署[数据质量检查表](#)
- 循证医学[资源](#)
- 有效证据联盟 [使用研究证据](#)
- ImpactMatters [影响力审核标准](#)

搜查现有证据

- Dartington [社会研究单位](#)
- Beam Exchange [证据图](#)
- 3ie [证据缺口图](#)
- What Works Centre [论文集](#)
- 坎贝尔协作网 [图书馆](#)

选择方法

- IDInsight [基于决策的评估](#)
- CDI [审查评估方法](#)
- IPA [研究资源](#)
- 斯坦福社会创新评论 [影响力投资者如何衡量影响力](#)
- OHRC [收集伦理数据](#)



本作品受知识共享署名许可协议非商业用途4.0国际许可保护。只要不对内容做任何更改，并且保护作者的权利不受侵犯，便允许复制和分发这些材料。

影响力管理项目，2017年9月

影响力管理项目