

加强数据科学能力和生态系统：实现以数据为中心的公共卫生干预

征求建议书

申请截止时间是美国太平洋时间**2022年1月13日上午11点30分**

背景

COVID-19疫情的规模和范围再次说明获取并分析及时准确的健康数据的必要性，从而对实施控制措施提供支持，并为卫生干预和循证实践提供指导。然而，研究人员赖以获取并分析这些数据的配套基础设施和能力却差异很大。由于多种原因，全球南方的研究人员被排除在数据分析和结果解释之外，他们往往对数据共享的广泛呼吁持怀疑态度。此外，在不同地区部署“一刀切”的分析工具和治理流程可能会产生不良分析结果，如果科学家在分析数据时对收集数据地区的具体情况没有深入了解，甚至可能导致误导性的解释和结论。

因此，数据科学基础设施和能力需要侧重于产生数据的领域，以便制定解决方案，使强有力的循证实践能够指导卫生干预。

挑战

收集、共享、分析和解释与健康相关的数据对于公共卫生实践的规划、实施和评估以及为卫生应急准备提供先进的知识来说至关重要。许多数据资源如果没有充分记录并且没有数据收集地区的具体信息，或者收集方法不当以及没有适当的使用许可，则不能重复使用，特别是在最初生成数据时没有预料到会计划进行二次分析的情况。

本次征求的理想提案是开发“与目标相符的”工具和流程，能够促进低收入和中等收入国家研究人员之间的合作。这项工作的重点是填补数据科学的空白和挑战，其解决方案将针对全球健康问题，同时促进研究人员之间的多学科合作。我们尤其希望为职业生涯早期和中期的研究人员领导的提案提供资助。我们希望每个主要申请人都有高级科学顾问致力于为团队提供科学和技术指导并指明方向，并根据需要促进国际间的联系。

具体而言，挑战的目标是：

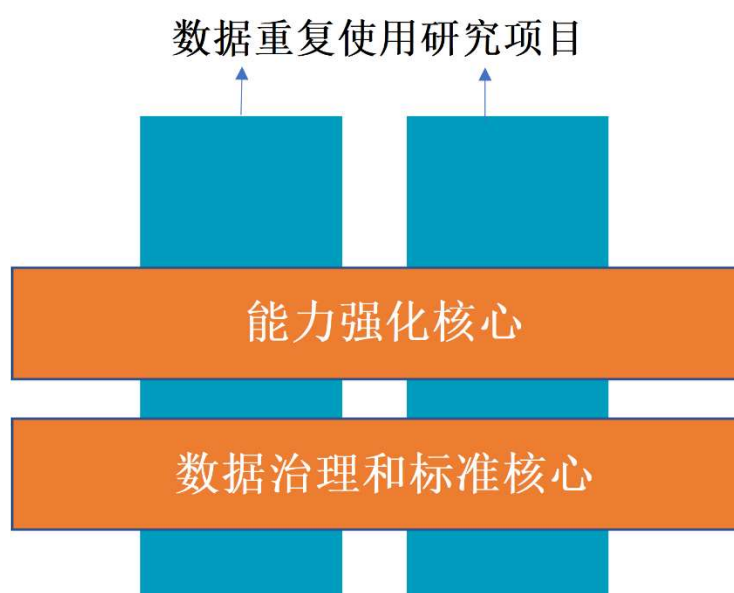
- 开发和改进基础工具、标准和协议，以实现以数据为中心的健康研究、干预和卫生监测
- 为研究和培训目的改进对经过精选和链接的数据集访问
- 实现以数据为中心的南南合作，并继续建立和加强现有的举措和生态系统
- 加强数据科学基础设施、分析能力和地方领导力，以解决公共卫生问题
- 根据各自相应地区的当地利益，通过健康数据促进创新
- 通过适当的数据科学交流工具和渠道的合作与设计，弥合以数据为中心的研究和政策之间的差距
- 通过开发工具和流程来精选、标准化并汇总数据，提高健康数据的可靠性并实现重复使用
- 改进数据收集、共享、治理、监管合规性和分析流程，实现以数据为中心的公共卫生研究

我们期待的提案

我们将制定一个协调计划来支持这些目标。这将包括贯穿各个领域、为该计划提供具体功能的核心，以及侧重于一个更加具体的疾病领域的个别研究项目。

我们邀请来自医疗保健、学术研究、工业界或研究基金会的研究人员提出申请，并将欢迎来自任何以下地区的研究人员提出申请：**非洲、印度、孟加拉国、巴基斯坦、尼泊尔和巴西**。每个提案应由一位主要申请人提交，但是资助金要求至少有来自上述地区的两个研究机构参与，而且不是来自同一国家。

计划结构



数据治理和标准核心

我们期待为一个数据治理和标准核心提供资金。该核心将：

- 支持通过该倡议资助的研究项目，包括能力加强核心
- 将得到资助的研究项目作为用例，来开发和完善数据共享流程和治理
- 通过制定流程来实现南南合作分析，寻求加强全球社区和中低收入国家 (LMIC) 研究人员的多学科合作
- 制定适用的标准或标准操作程序，以解决包括匿名化和扰动协议在内的数据共享问题
- 开发可重复使用的工具、协议和标准，以便收集、清理、协调和共同构建与一个或多个疾病领域相关的个人数据和元数据，从而提高健康数据的可靠性并实现重复使用。新工具和流程开发应利用并建立在现有工作的基础上
- 开发数据信息处理流程，以克服收集完整和准确信息的障碍；以及与数据和信息治理以及数据重复使用相关的挑战

能力强化核心

我们期待为一个能力加强核心提供资金：

- 支持通过该计划资助的研究项目和数据基础设施及管理核心
- 通过培训和最佳实践开发，弥合数据科学家、建模者、监管机构和政策制定者之间的差距
- 开发可重复使用的培训材料和手册
- 通过不同形式（研讨会、黑客马拉松、大型聚会等）领导和协调协作分析工作
- 针对中低收入国家数据科学培训方面的现有差距并以现有举措为基础，领导能力建设（在可能的情况下重复可用资源以避免重复工作）
- 开发数据信息处理流程，以克服收集完整和准确信息的障碍，以及与数据和信息治理以及数据重复使用相关的挑战

数据重复使用研究项目

我们期待为两个数据重复使用项目提供资金，其研究工作将侧重于更具体的疾病领域之一：

- 构建数据驱动的公共卫生平台、工具和仪表盘，能够为政策制定和公共卫生干预提供指导
- 确定、使用、精选和链接明确定义的数据集（管理、临床、环境等），以解决政策差距并提供有关全球卫生研究重点的见解和概念和/或技术飞跃，包括：
 - o 母婴和儿童健康
 - o 传染性疾病
 - o COVID-19 和疫情准备
- 能侧重于特定疾病领域的提案，但需要证明承诺提供对所有研究产品的访问，包括各种已开发的工具和流程

资金类别

成功提案将获得最高250,000美元的资助金，资助期限为18至24个月。虽然提案需要来自位于任何上述地区不同国家不同机构的至少两名研究人员的合作参与，但提案应由一个主要申请人/机构提交。虽然每个机构的每项申请仅限一个主要申请人，但研究人员可以作为合作者参与多个合作申请。

获选入围的提案将成为协调计划的一部分，该计划旨在加强数据科学能力和生态系统，以实现以数据为中心的公共卫生干预。我们期待建立一个研究人员网络，与现有的[数据科学大挑战网络](#)及其合作伙伴共同合作，以吸取不同地区的经验教训，并进一步开发流程、工具和治理，以最大限度地提高以数据为中心的研究的影响力，以及确定并解决与全球数据相关的挑战。

予以考虑的成功提案应该：

- 在对开放科学、数据共享以及构建协作和分析基础设施的共同承诺的推动下，使这些能够造福世界各地的人民

- 寻求加强中低收入国家研究人员的全球社区和多学科合作，并使患者和更广泛的公众能够参与
- 使位于不同国家的至少两个研究机构之间能够实现实质性合作，包括非洲、印度、孟加拉国、巴基斯坦、尼泊尔和巴西。所建议的合作应该是推进项目目标和产生深入见解的关键，如果没有共同产生的方法就难以获得这些深入见解。然而，申请可以包括与位于其他地理区域的机构的合作
- 有可能在提出的预算范围内并且在18-24个月的时间范围内对解决公共卫生问题产生影响
- 数据重复使用提案应使用有明确定义的数据集，直接解决核心研究问题，而且有明确的治理计划，并能够展示该数据如何在以后的阶段使更广泛的社区受益。
- 展示对影响政策和公众的承诺

以下提案恕不考虑：

- 旨在通过这笔资金收集和生成新数据的研究
- 由任何下列地区以外的机构牵头的提案：非洲、印度、孟加拉国、巴基斯坦、尼泊尔和巴西
- 只涉及单个研究机构或仅涉及来自同一国家多个机构的合作者的提案
- 未展示出对开放科学的明确承诺并使他们的发现、流程和/或开发的工具可访问和可重复使用的提案
- 未展示出如何利用现有工作和网络的提案
- 未展示出如何从该协调计划资助的其他项目中受益并支持其他项目的提案
- 在资金期限内无法完成的提案