

REQUEST FOR PROPOSALS

Advancing SDG Outcomes Through the Utilization of Lacuna Fund's Open Datasets

Training and Scaling Machine Learning
Models for the Global South



Application Open:
September 14, 2026

Submission Deadline:
November 13, 2026
23:59:00 HRS EAT

Informational Webinar:
October 15, 2026

Contact details:
s.wanjau@acts-net.org

Implementing Partner	African Centre for Technology Studies (ACTS)
Funder	Wellcome Trust
RFP Opens	14th September 2026
Submission Deadline	13th November 2026
Awards Announced	By end of November 2026
Implementation Period	December 2026 – May 2027 (six months)
Eligible Applicants	Researchers and research teams in Global South LMICs
Available Funding	Up to USD 30,000 (three awards of up to USD 10,000 each)

Introduction

Lacuna Fund supports the creation, expansion, and maintenance of datasets that enable robust and equitable application of machine learning (ML) tools of high social value in low- and middle-income contexts globally. To date, Lacuna Fund has accumulated over 67 open datasets spanning all Sustainable Development Goal (SDG) priority areas across the Global South.

This call for proposals is designed to address a recognized gap: the original grants focused on dataset creation and there is a need for additional funds to support the uptake and use of these resources. The purpose of this RFP is therefore to support researchers and research teams in training new ML models, scaling existing models, or supporting ML and AI applications or tools using datasets already created with support from Lacuna Fund. By doing so, the call seeks to unlock the latent value of existing data sets and translate them into tangible, equitable outcomes across the full range of SDG priority areas.

This call is open to applicants working in any SDG domain represented in the Lacuna Fund dataset repository.

Organizational Eligibility

Lacuna Fund aims to make its funding accessible to as many organizations as possible in the AI space, and to cultivate capacity and emerging organizations in the field.

To be eligible for funding, organizations must:

- Be either a non-profit entity, research institution, for-profit social enterprise, or a team of such organisations. Individuals must apply through an institutional sponsor. Partnerships are strongly encouraged.
- Have a mission that supports societal good, broadly defined.
- Be headquartered in a low- or middle-income country (LMIC) in the Global South. For a list of eligible countries, refer to the Wellcome Trust LMIC guidance (<https://wellcome.org/research-funding/guidance/prepare-to-apply/low-and-middle-income-countries>)
- Institutions based outside eligible regions may participate as partners of the lead institution. Only the lead applicant will receive funds.
- Have all necessary national or other approvals to conduct the proposed work. The approval process may run in parallel with the grant application.
- Demonstrate technical competency, or the ability to build it through a described partnership, in both machine learning model training or scaling or application development and in the specific SDG domain the applicant proposes to work in. Domain competency is a core requirement of this call and will be assessed by the reviewers' Panel.

Selection Process and Evaluation Criteria

ACTS and Lacuna Fund will perform an initial screen for organisational eligibility and technical feasibility. SDG Domain experts, ML practitioners, and stakeholders will then evaluate proposals against the criteria below.

Proposals will be evaluated against the following criteria:

- **Quality** – The team includes qualified experts in machine learning, data management, and software engineering. The proposal provides clear use cases for the trained or scaled model or application and situates the approach within existing work in the domain. The applicant shows that the existing data is sufficient, robust, and complete to support the model or application or provides a plan for augmenting the existing dataset so that it is. Effective and well-reasoned methodology is demonstrated.
- **Domain Competency** – The team demonstrates substantive expertise in the SDG domain they propose to work in. This includes knowledge of the relevant literature, established standards and practices in the domain, and an understanding of the policy or operational context in which the model outputs will be applied. Competency may be evidenced through prior work, qualifications, or meaningful partnerships with domain specialists.
- **Transformational Impact** – The project makes meaningful use of existing Lacuna Fund datasets to generate high-value outputs for underserved populations or problems. Proposals may be considered transformational where they address a particularly important or timely challenge, or have significant reach across underserved communities or geographies. The proposal describes who will benefit from the model or application and how the applicant will measure the impact of project activities.
- **Equity** – The team clearly articulates the equity issue being addressed and demonstrates how the trained or scaled model or application will reduce disparities and increase access to the benefits of ML/AI for vulnerable communities.
- **Participatory Approach** – The team is headquartered in the geographic area where the work will take place to ensure that solutions developed are appropriate in the local context and used by the local community. In-country partners are involved in strategic elements beyond implementation. The proposal describes how affected stakeholders will be engaged and how outputs will be shared with communities.
- **Ethics** – The project has a plan to pass an ethical review probing: privacy concerns, potential for downstream misuse, possible discrimination vectors, and fair and equitable working conditions where paid contributors are involved.
- **Sustainability and Communications** – The project has a plan for the sustainability of the trained model and/or application or tool and its outputs beyond initial funding, including governance, hosting, and potential use cases. Plans should include dissemination activities such as conference presentations, stakeholder workshops, and/or plans to present the model or application to potential users.
- **Feasibility** – The project is feasible in relation to the budget and scope of work proposed. The team might have conducted initial research, proof-of-concept pilots, or feasibility studies that support that the ML model or application will be a success. **Accessibility** – Outputs will be made widely accessible under open-source licensing (Apache 2.0 for code; CC-BY 4.0 or CC BY-SA 4.0 for other intellectual property), or a compelling case is made for more restrictive licensing to protect privacy or prevent harm. The proposed solution is widely accessible to the target audience. For example, the model or application may propose interfaces that make it possible to access the technology in areas with low internet penetration.

Timeline

Milestone	Date
RFP Opens	14th September 2026
Applicant Webinar	15th October 2026
Question Deadline	28th October 2026
Answers Posted	4th November 2026
Full Proposals Due	13th November 2026
Panel Review	14th – 30th November 2026
Awards Announced	By end of November 2026
Implementation Commences	1st December 2026
Final Reports and Model/Output Publication	May 2027

All questions related to this RFP should be **submitted to s.wanjau@acts-net.org** with **"SDG Dataset Utilization RFP 2026 Question"** in the subject line. Questions submitted by 28th October 2026 will be de-identified and answered publicly by 4th November 2026 on the ACTS website.

Purpose of the Call

The purpose of this call is to support efforts to train and deploy new machine learning models, scale existing models, or develop ML/AI applications or tools using the open datasets already created with support from Lacuna Fund. Trained models should be usable via inference and hosted accordingly, either freely through a platform such as Hugging Face or on another instance with a clear sustainability and maintenance plan, and should be accompanied by proper documentation and explanation. The call is open to any SDG domain represented in the Lacuna Fund repository and targets researchers and research teams in Global South LMICs in Africa, Latin America, and South and Southeast Asia.

The datasets currently linked in the Lacuna Fund repository span a broad range of SDG priority areas, including but not limited to health, agriculture, education, climate, water and sanitation, energy, land use, and governance. These datasets were developed with significant investment by local teams and represent a substantial, underutilized resource. This call directly addresses that underutilization by incentivizing their application in model training and scaling and application and tool development.

Eligible Domain Areas

Proposals may be submitted in any SDG domain for which relevant datasets exist in the Lacuna Fund repository. Applicants are strongly encouraged to consult the Lacuna Fund dataset catalogue ([Fair Forward - Open Data & Use Cases](#)) and GIZ Fairforward catalogue (<https://fair-forward.github.io/datasets?view=lacuna>) prior to developing their proposals to identify datasets relevant to their area of work. Across all domains, gender-responsiveness and the inclusion of key vulnerable groups are required. This includes attention to intersecting factors such as gender, disability, age, socio-economic status, access to basic services, and exposure to other risks such as climate risk.

Use of Existing Lacuna Fund Dataset

All proposals must be grounded in the use of one or more datasets from the Lacuna Fund dataset catalogue. Proposals may include, but are not limited to:

- Training and deploying a new ML model using one or more existing Lacuna Fund datasets;
- Scaling or fine-tuning an existing model with Lacuna Fund datasets to improve performance in a local or regional context;
- Developing a baseline model to demonstrate the utility of an existing dataset and facilitate its broader adoption, such as a classification or predictive model;
- Applying existing datasets to develop decision-support tools or predictive applications of direct benefit to underserved communities;
- Conceiving a specific use case or use cases that demonstrate the application of the dataset(s);
- Building an Application Programming Interface (API) on top of AI models for the dataset(s) to increase uptake;
- Designing tools, applications, or platforms (for example, an app or a dashboard) that utilise the dataset(s);
- Developing ways to visualise the dataset(s) to increase accessibility and use;
- Utilising a dataset in the training of future professionals in data science and socially beneficial ML.

We expect a Minimum Viable Product (MVP) at pilot stage as the minimum final deliverable. Proposals seeking to fully develop products that gather a critical mass of users are welcome, provided this is feasible within the granting period.

Proposal Information

Proposal submissions **will only be accepted** through the Submittable application portal at [Link](#) . Applications may be submitted in English, Spanish, French, and Portuguese.

i. Applicant Information

This section will prompt the applicant to provide:

- A 200–250 word proposal abstract;
- Details about the institution(s) and/or team applying;
- The SDG domain and geographic area in which the work will take place;
- CVs for key team members, including evidence of domain competency;
- Information about the affiliated institution's ethical review processes;
- Information about the team's ability to obtain national approvals.

ii. Proposal Narrative

The online template will have the following sections:

1. Qualifications and Domain Competency

Describe the organisation(s) or partnership(s) applying and how they satisfy the eligibility criteria. Set out the team's qualifications to undertake the proposed work, with particular emphasis on demonstrated expertise in the SDG domain of interest. Domain competency is a central requirement of this call and should be evidenced clearly, for example through prior research, published work, practice experience, or named partnerships with domain specialists.

2. Problem Identification and Proposed Approach

Describe the development challenge the proposed mode or application will address and explain how it relates to an identified need in the chosen SDG domain. Identify which dataset(s) from the Lacuna Fund dataset catalogue you intend to use and provide a summary of your proposed model training, scaling, or application development approach. Please include how your project addresses a gap and complements existing work in the field.

3. Specifications and Deliverables

Include the following:

- The dataset(s) to be used, with reference to their source in the Lacuna Fund dataset catalogue ([Fair Forward - Open Data & Use Cases](#)) and GIZ Fairforward catalogue ([The online template will have the following sections:](#))
 - The type of model to be trained or scaled, including the model architecture and intended application;
 - Metrics to be used to assess model performance, fairness, and quality;
 - The format and documentation of model outputs, including reported results on an existing public benchmark or leaderboard for the task where one exists (for example, for language tasks), to provide potential users with a comparable, standardised metric for judging model quality.
- Intended Beneficiaries and Use Cases

Describe prior consultation and proposed collaboration with intended beneficiaries. Outline current and potential future ML use cases for the trained or scaled model in the relevant SDG domain. If applicable, describe what demand the model meets and how signalling should be added to improve its credibility and visibility to serve that demand. This should go beyond documentation alone, for example, through a minimal, replication-friendly kit (code, environment, a pointer to the source dataset, and a one-command inference example) that allows a third party to verify and reuse the model. Such a kit is low-cost given the size of this grant, but is often what drives uptake and findability in practice.

4. Methodology

Provide an overview of the proposed steps and key assumptions for model training or scaling or application or tool development. Include:

- The proposed training pipeline and tools, including interoperability considerations;
- Quality control measures and how the team will address outliers or data quality issues inherent in the source datasets;
- A plan to assess and mitigate bias, including gender bias and other potential discrimination vectors relevant to the domain;
- How existing infrastructure and resources will be leveraged;
- Permissions in place or steps to be taken to obtain required national or other approvals;
- Anticipated challenges or uncertainties and proposed countermeasures.

5. Transformational Impact

Explain how the trained or scaled model will contribute to meaningful impact within the chosen SDG domain. If applicable, describe how the outputs could support further research or commercial application. Note any practical constraints the team may face, such as limited computing capacity or internet connectivity.

6. Model and Output Management and Licensing

Please describe:

- Plans for licensing model outputs to maximise responsible downstream use, in line with Lacuna Fund's Intellectual Property (IP) Policy;
- Any anticipated issues related to the use of the source datasets and how these will be managed;

- Plans for documentation of the trained model, including a model card or equivalent;
- The hosting platform to be used for the trained model and associated documentation. In addition to hosting, outputs (the model and its model card) should be registered in a discoverable index, so that the signal reaches demand rather than producing an artefact that cannot be found. This closes the loop with the repository already identified to applicants as a source of datasets.

7. **Risks, Including Ethics and Privacy**

Identify potential risks and describe steps to mitigate them. Specifically:

- Provide a reflective statement on the communities and identities represented within the team, and how these may influence the work;
- Describe how informed consent was or will be addressed in relation to the underlying datasets and their application;
- Describe how equity in project labour will be ensured, including fair compensation where applicable;
- Describe how gender diversity and other demographic considerations are incorporated in the team composition and model development process;
- Present a plan for anonymization of any personally identifiable information (PII) and compliance with applicable privacy laws;
- Discuss potential adverse impacts in the production and use of model outputs and steps to mitigate them, including human rights risks and energy consumption implications.

8. **Sustainability and Communication Plan**

Describe how the trained model and its outputs will be maintained beyond the initial funding period, including governance arrangements and potential use cases that could sustain the work. Where sustainability involves a revenue or service model, applicants must demonstrate how this will preserve the open-source nature of the dataset and model, rather than enclose it (see BMZ Digital.Global, "Code for Community: How Open Source AI Sparks Business and Impact"). At this grant size, a credible, lightweight signal of sustainability is sufficient, such as identifying who will host the model afterwards, confirming that it will remain findable, and naming a maintainer; a full business plan is not required. Explain how the model will follow FAIR principles (Findable, Accessible, Interoperable, Reusable). Outline communications activities to spread awareness, including networking with potential users, conference presentations, and stakeholder workshops.

III. **Project Timeline and Deliverables**

This section will prompt the applicant to submit a table with a timeline for major activities and deliverables. The timeline may include, but is not limited to, team onboarding, model training, application or tool development, quality assurance, validation, and model/application/tool publication.

Submission of Institutional Review Board (IRB) approval must be included as a deliverable for proposal seeking to launch a medical or diagnostics device. The team will have to confirm that the dataset has the necessary informed consent in order to conduct the proposed uptake and utilization work.

All timelines must include a date by which the trained model and accompanying documentation will be publicly available.

Note: All proposed projects must be completed, model outputs published, and final reports submitted by **May 2027**. For planning purposes, agreements will be finalized and work may begin by December 2026.

Budget

Provide a budget for the proposed work, submitted through the submittable portal using the budget template available in the applicant portal.

Funding Summary

- Total pool available: up to USD 30,000
- Number of awards: 3 grants of up to USD 10,000 each
- Target regions: Africa, Latin America, South and Southeast Asia (one award per region is anticipated)
- Indirect Costs are strictly limited to 13% of direct research costs for LMIC-based institutions.

The Expert Panel will assess the feasibility and suitability of the budget as well as the linkage between the budget and the grant narrative. Budgets may include, but are not limited to, costs for:

- Capacity building related to model training, quality assurance, and validation;
- Computing power and data storage;
- Model documentation and publication;
- Licensing;
- Open-access publication of results;
- Workshops and stakeholder engagement activities;
- Communications activities, including conference attendance for up to two events to present outputs.

Funds may not be used for the direct payment of any customs, import, or other duties or taxes levied in respect of the importation of goods or equipment.

END



Contact Us

African Centre for Technology Studies (ACTS),
2nd floor, Konza Complex
Nairobi-Mombasa Road,
P.O. Box 45917 - 00100 Nairobi, Kenya.
+254710607210 || +254737916566
info@acts-net.org