

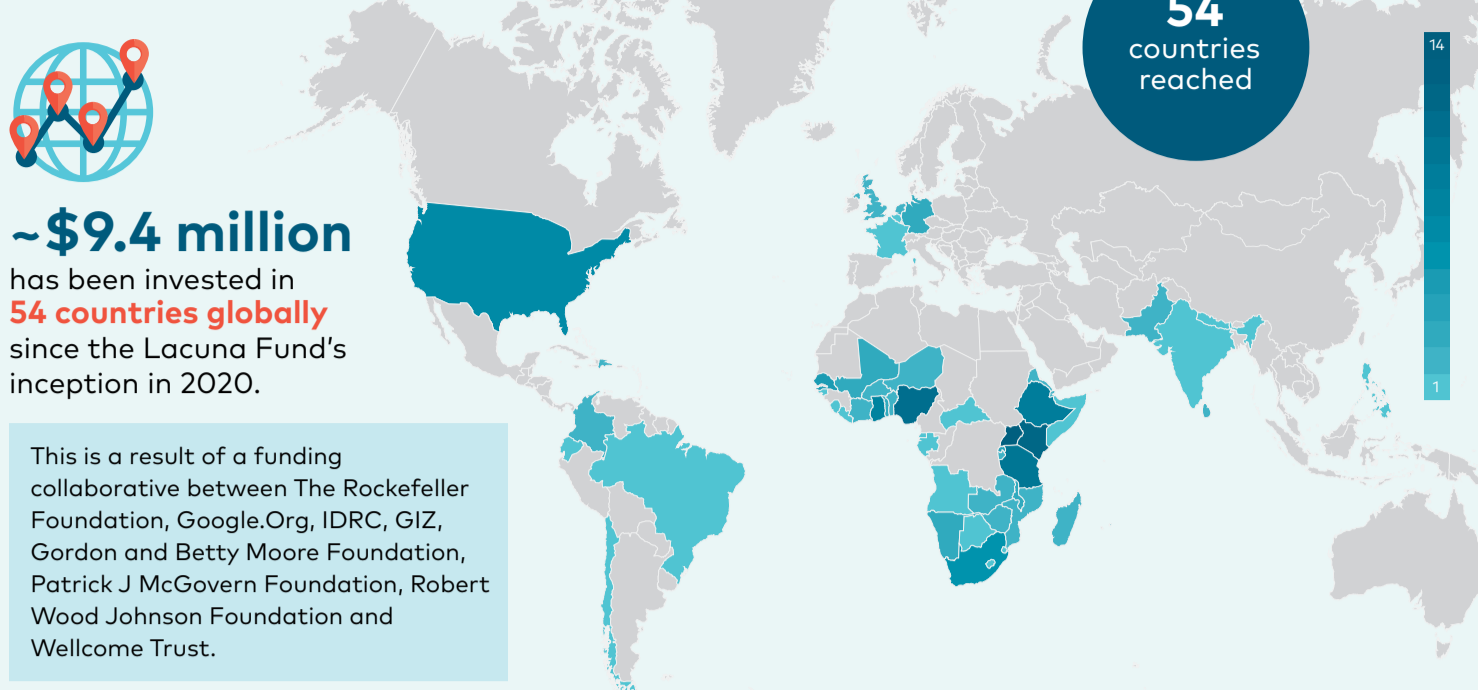
Reflections from Lacuna Fund Evaluation

2020-2022

The infographic highlights a summary of measurable results, lessons, and best practices from implementing Lacuna Fund over 2020-2022, providing insights on achievements and lessons for the ongoing improvement of the fund. These insights are in response to engaging with grantees, thought-leaders, potential dataset users and documentation. The infographic may be used by similar funds, grantees, and funders to support ongoing development in the field of unbiased datasets and responsible AI.

INSIGHTS

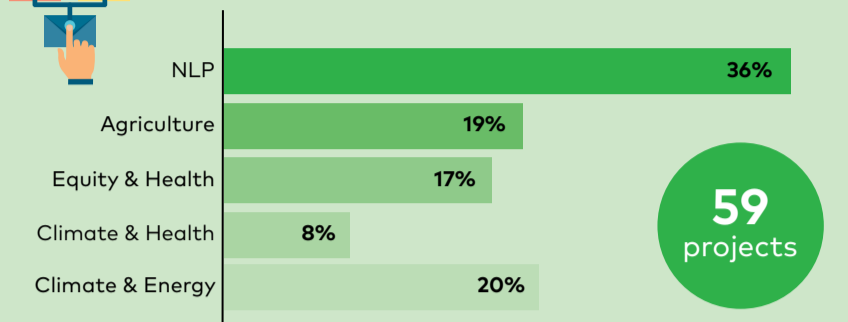
Supporting machine learning (ML) work in low and middle-income countries (LMICs) to impact the ecosystem of datasets



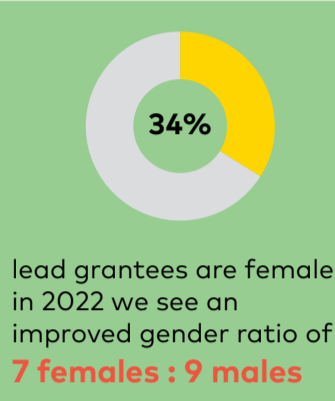
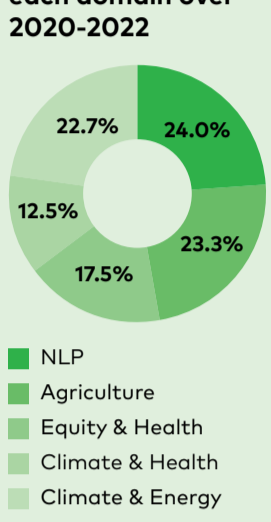
% of Female Project Leads

Natural Language Processing (NLP)	24%
Agriculture	36%
Equity & Health	40%
Climate & Health	33%
Climate & Energy	60%

% of Total Projects



Funding awarded to each domain over 2020-2022



100% of the **59 projects** awarded funds over 2020-2022 benefited communities in low & middle income contexts. Within this, are **2 Equity & Health grants** based in the USA and benefiting low & middle-income contexts.



Prioritizing high-impact domains that represent LMICs through more inclusive and representative datasets

High-impact domains are those that improve the lives of human beings in a significant way. These include:

The agriculture domain which aims to alleviate food security challenges

The NLP domain which seeks to increase the representation of low-resourced languages

The climate domain which aims to restore planetary health

The health domain which strives for quality and universal health services

Domains funded (2020-2022)*

NLP	\$2.3M
Agriculture	\$2.2M
Equity & Health	\$1.6M
Climate & Health	\$1.2M
Climate & Energy	\$2.1M

*With continued donor support new domains are being funded.

The portfolio of African NLP supported by Lacuna Fund laid the foundation to deepen work in NLP, noting that the processes and systems already developed will be a starting point for future grantees. NLP is important because it creates openly accessible text and speech resources that enable operations in diverse languages across LMICs globally.

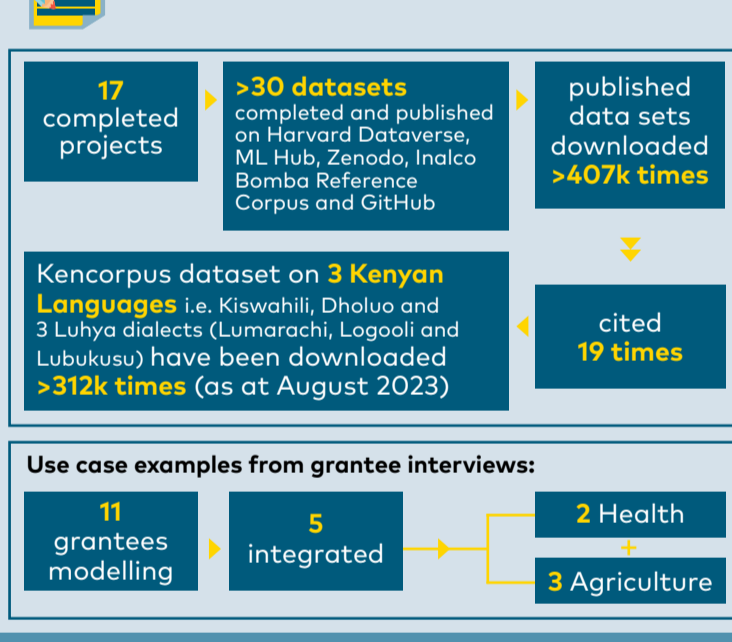
Strengthened principles-based funding approach

Principles of accessibility, equity, ethics, participatory, quality and impact are at the fore of Lacuna's work. These were assessed for what worked and what could be improved. On a scale of one to three where one denotes "average", two denotes "above average", and three denotes "good", the below ranking demonstrates the extent to which the principles are realized.

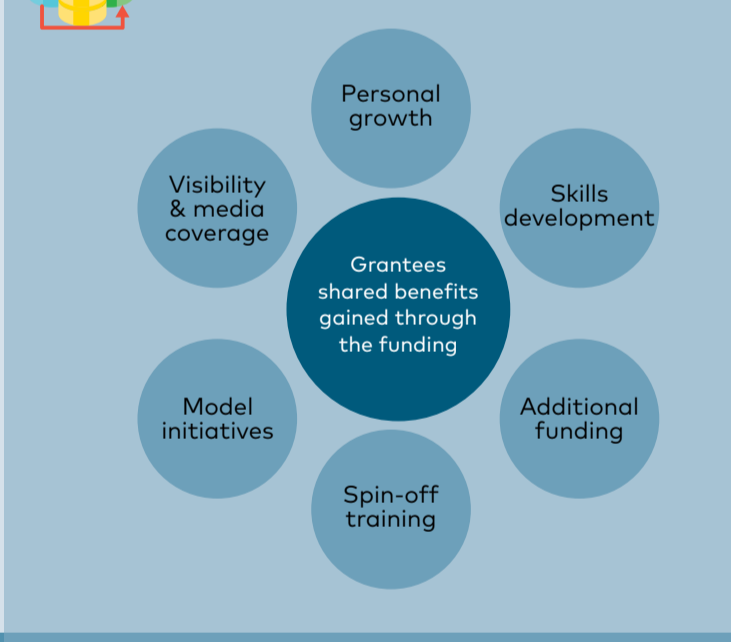


ACCESSIBILITY	EQUITY	ETHICS	PARTICIPATORY	QUALITY	IMPACT
<p>Dispersing calls for proposals in multiple languages and requiring grantees to make datasets openly available.</p> <p>● ● ●</p>	<p>Grantees are filling data gaps and making ML and AI more equitable. Increased funding representation of LMIC regions over time and improved female ratio of grantees.</p> <p>● ● ○</p>	<p>RFPs require applicants to describe steps to ensure ethical approaches to data collection. Lacuna Fund Conflict of Interest policy made available to grantees.</p> <p>● ● ○</p>	<p>RFPs ask applicants to consult or collaborate with potential data providers and users. Governance structure & processes also participatory and inclusive.</p> <p>● ● ●</p>	<p>TAPs review applications for quality. A Data Quality Advisor was contracted in 2023 and is reviewing applications, grant reports and datasets, and providing feedback on quality.</p> <p>● ● ○</p>	<p>It will take time to realize the impact gained by communities. Monitoring of use cases for impact will continue.</p> <p>● ○ ○</p>

Increasing unbiased datasets availability and use case development



Influencing intended benefits to grantees and their research communities



Use-case example: Lacuna Fund funded a project for High-Accuracy Maize Plot Location and Yield Dataset in East Africa, where their model looks at a field and determines the crop yield for better planning. From 10,000 fields, 2,000 were hand labelled and a machine learning model was created to relabel the remaining 8,000, which can be applied to any field now.

[[AFRICA/MICCA]] are sponsoring our AI summer school, which is essentially training people on how to create models on that Lacuna data set. They are supporting us to have students come to Vancouver in Toronto this year and present their models. The Lacuna Fund has opened up a floodgate.

- Grantee from the Climate & Health domain

LESSONS

Process-related reflections

Grantees require a **collaborative platform** that facilitates **knowledge sharing** and remaining in contact with each other; they also want a more **refined funding process** reflecting alignment between key milestones and funding allocation.

Dataset development has proved to **take longer than originally anticipated** so grant periods are to take this into consideration for future funds.

Improvement in sustainability plans is required, this includes considering licensing, unclear use cases for community data, absence of an African data distribution platform and respective infrastructure.

Multiple use cases have been made possible from **datasets**, particularly in the agriculture and NLP domains.

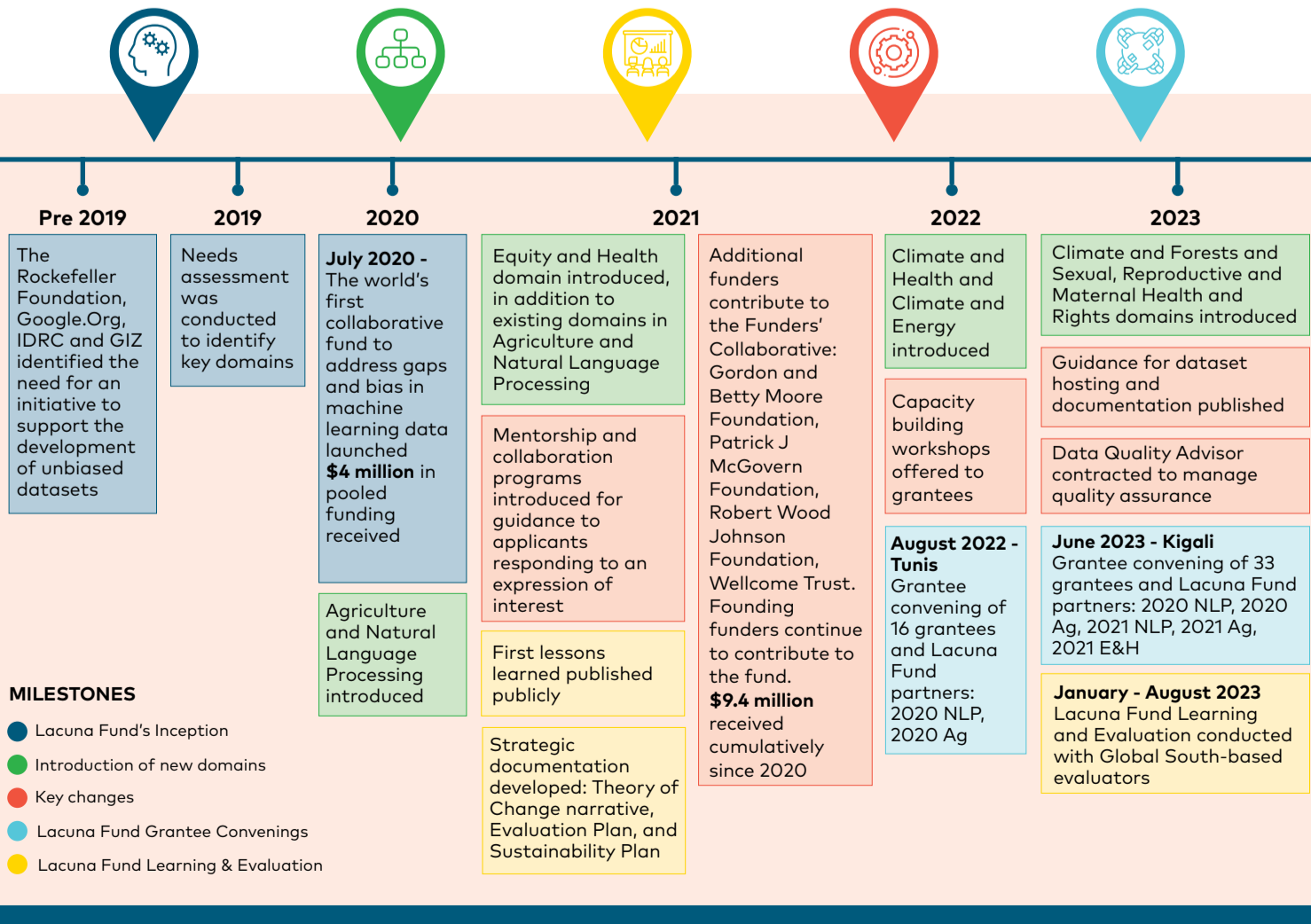
Looking to the future

Creating datasets for machine learning in high-impact domains will produce **positive solutions to global challenges** as experienced by LMICs.

Agriculture remains a **key domain for dataset development**, in addition to other emerging domains such as Climate & Energy, Disaster Warning and Engineering.

5 emerging trends shared by thought leaders: the demand for ethical data for future scenario modelling will increase; alternative business models for data access and sharing; generative AI; integrated knowledge for early warning; and graphical process unit hoarding.

It is recommended to **support grantees** that are capacitated to establish hubs for strengthening collaborations, for a decentralized approach to supporting the development of datasets.



ABOUT LACUNA FUND
Lacuna Fund is the world's first collaborative effort to fund the development of training and evaluation datasets for social impact in LMICs. Supported by an experienced Steering Committee, Technical Advisory Panels, and Funders, Meridian Institute is the backbone support to the Fund; led by the Secretariat that manages the processes and distribution of funds.

The infographic is based on evaluation data collected from 59 grantees, 13 bellwethers, 48 users, and more than 130 documents. The full evaluation report can be found on Lacuna Fund's website.

Contact info: secretariat@lacunafund.org
Website: <https://lacunafund.org/>