INSPIRE

A platform for a network of research bodies to share health data for informing policymaking

This case study was developed as part of *Effective Data Sharing: Beyond Platforms*, a study conducted by Athena Infonomics and commissioned by the Global Partnership for Sustainable Development Data. This project was made possible by a grant from Google.org.

Summary

The Implementation Network for Sharing Population Information from Research Entities (INSPIRE) was established by the African Population and Health Research Center (APHRC) in collaboration with the Analyzing Longitudinal Population-based HIV/AIDS data on Africa (ALPHA) network, South African Population Research Infrastructure Network (SAPRIN), and Committee on Data of the International Science Council (CODATA) to enhance the utility of the data from longitudinal population studies. INSPIRE is a network of data producers who share the data collected from Health and

Demographic Surveillance Sites (HDSS) in five countries in East Africa. Initially, INSPIRE was established to create a large health data repository. It has since evolved into a platform where different datasets are aggregated and transformed to provide data driven-solutions and answer health-related questions in Africa. INSPIRE has established collaborations that have secured multi-country grant applications and opportunities to develop robust data architecture and a platform through which data can be extracted, transformed, and used to better inform health policy, research and decision-making in Africa.

Source: APHRC



Facts and figures

Founded: 2020
Sector: Health
Typology by use: Research
Geography: Africa
Governance Structure: Data Collaborative
Number of Member Organisations: 11 HDSS sites in 5 countries
Ownership: INSPIRE has been hosted by the APHRC, Nairobi, Kenya, for the first five years.

Key Challenge(s) that the initiative was trying to solve

- Establish a network for users and producers of longitudinal population-based health data, and provide them with the infrastructure to securely share the data.
- Create a data hub that creates Findable, Accessible, Interoperable and Reusable (FAIR) data to assist researchers and policymakers address health-related questions.

- Empower the data producers with tools that assist them in data sharing and allow them to reproduce the data.
- Enable data sharing and analysis across Africa to better inform health research and policymaking

How are they solving the problem?

The The INSPIRE network has access to 11 HDSS sites from 5 East African countries. These HDSS sites have data collected from over 20 years of repeated surveys in the surveillance sites. This large data repository has been used to address specific questions that include but is not limited to the following:

- Understanding excess mortality due to COVID-19 in Africa
- Mental Health in Africa
- Establishing the burden of multi-morbidities in African populations
- Establishing the effect of climate change on mortality and morbidity in Africa

Through these projects, the network has successfully developed a working relationship among the data engineers, data scientists, data managers, statisticians, and research scientists to develop pipelines for accessing various analytical datasets required for answering future research questions related to the continent.

"The majority of users of the data from the network are researchers and national stakeholders who intend to use the results to improve the health outcomes and livelihoods of people in Africa."

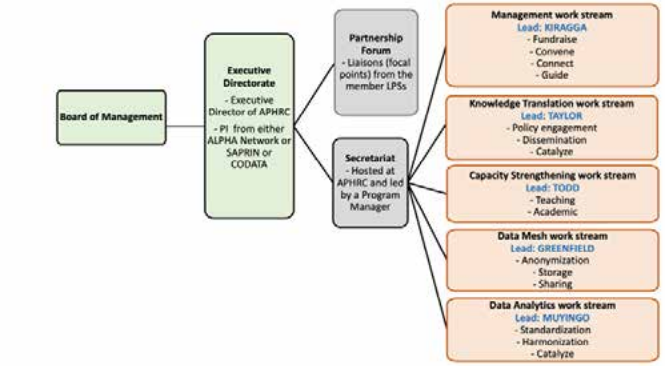
Further, the INSPIRE network realises that accelerating scientific advances and health discovery through data innovation are constrained due to the lack of data science capacity on the continent. Hence, the INSPIRE network has adopted the route of building the technical capacity of the partners as the optimal way of bridging the skill gap. It has engaged Data

Science for Africa to build the capacity by keeping a focus on having an African-led approach to building capacity on the continent. It is crucial to building capacity in the context of Africa rather than importing these services from outside.

What were the key considerations in designing the initiative?

1. Evolve into a service platform from being a repository of data: To achieve its aspiration of a service platform, INSPIRE has developed a detailed governance structure and aims to develop a federated data storage structure. INSPIRE is governed by a multi-layered structure as depicted in the figure below.

Source: INSPIRE



It is governed through a secretariat at the APHRC in Nairobi, Kenya. And there are five work streams, with a hub leader for each stream, handling activities such as capacity strengthening and standardisation of data in the INSPIRE network. The INSPIRE network is working to build a governance structure that will help its evolution from a repository of data to an integrated data platform.

Further, INSPIRE currently stores data centrally in the cloud. However, it plans to move towards a federated storing structure where the data producers store their data and retain its custody. INSPIRE's platform shall mine the data remotely for particular use cases.

"We don't want to get the data and put it in some kind of box and be the custodian of the data. We want to explore a federated structure where we are not hoarding the data"

2. Adding value for data sharers: INSPIRE finalises the use cases for shared data through two mechanisms. The first is through interest from funding bodies. Once a funding body publishes requirements for a specific data-driven solution, INSPIRE reaches out to its partners in the network to work together on the agreed use cases. The second is through the recommendation from INSPIRE's executive or Secretariat members, who guide INSPIRE in working towards addressing specific use cases. Either way, INSPIRE works toward the value desired by the member. It is a conscious step to avoid making the data sharing process a one-way extractive process and make its working more collaborative and aligned with the members' objectives.

3. Provide tools to build the technical capacity of the members: INSPIRE aims to build capacity for the partners and give credit to the partners and data producers for their work. To achieve the goal, it provides tools and services to the data producers that enhance data sharing, reproducibility, and the ancillary data production processes. These tools will also assist the data producers in discovering, accessing, and appropriately using the shared data and metadata. INSPIRE aims to build an in-house technical capacity for data producers which will enable them to participate in the extraction, transformation and loading of data into common data models.

Financial Sustainability of the initiative

INSPIRE relies on various research grants to sustain its activities. Currently, it does not charge any fee for the shared data as it focuses on putting in place the right infrastructure, technical team and platform to provide a competitive service. Subsequently, it plans to charge fees for services such as data harmonisation.

Lessons Learned

- **Building data capacity within African countries is a crucial step to effective data sharing on the continent.** INSPIRE is building the in-house technical capacity of the partners to establish trust. It realised that conventional data-sharing programs can be extractive. The practice of not involving the data producers at the stage of analysis can resemble data colonisation and create an atmosphere of distrust among the local data producers for future data-sharing programs. Further, such analysis is susceptible to mistakes since the data is being analysed by people who are removed from the context of the data. Hence, INSPIRE aims to build in-house technical capacity to ensure trust among the stakeholders remains intact.

"I want this to be a platform that combines multidisciplinary data-sets from Africa and collaboratively addresses problems – behind the dataset is a team of well-equipped, well-resourced and skilled data scientists/engineers who are based in Africa, to get outputs that inform livelihoods. I want to build capacity to analyse this data in Africa"

- **It is possible to enlarge the scope of a research and data partnership to solve bigger problems.** This approach of building a collaboration of people and HDSS sites spanning multiple countries has a huge potential to develop multi-country research proposals and bring on-board new partners and data collectors that can expand beyond the initial scope of a data sharing collaboration. INSPIRE started with HIV datasets but now also includes climate /environmental data. INSPIRE is slowly looking to enlarge its scope today, based on a network of trusted partners, a good data platform and data analysis capacity.

INSPIRE was piloted to address health questions around HIV, but given the utility and the need for the service, it soon extended its scope to various health subjects. However, it has not restricted itself to just data sharing among the networks. As an equitable measure, it has thoughtfully put its resources into building technical expertise and data capacity among the members of the network.

Sources

Stakeholder Interviews

Wellcome's Longitudinal Population Studies Working Group, Longitudinal Population Studies Strategy, 2017

What is INSPIRE about?

Michelle Mbuthia, Strengthening Africa's COVID-19 Data Ecosystem using Artificial Intelligence, 2021

Jim Todd Presentation_INSPIRE Network, 2020

Implementation Network for Sharing Population Information from Research Entities in East Africa (INSPIRE-EA), 2021

Joseph Beyene et al., A Roadmap for Building Data Science Capacity for Health Discovery and Innovation in Africa, 2021

Victoria O. Kasprowicz et al., African-led health research and capacity building-is it working?, 2020