Session title: **Data and digital solutions for identifying the poor and eradicating poverty**

Overall event theme: 'Accelerating Global Actions for a World Without Poverty - Implementation of the Third United Nations Decade for the Eradication of Poverty (2018-2027)

Organiser: UNDESA in collaboration with the UNECA. The Inter-agency Expert Group meeting

Date and place: 10 - 12 May 2023 in Addis.

Martha Getachew Bekele - Delivery, Quality and Impact Lead at Development Initiatives. Development Initiatives (DI) is a global organisation harnessing the power of data and evidence to end poverty, reduce inequality and increase resilience.

I. Data for identifying the poor and eradicating poverty

How important is data to identify those in poverty and in eradicating poverty?

We need relevant, reliable and disaggregated data, preferably real (or near rea time) data, which is crucial for planning, management, and monitoring the effectiveness and efficiency of public service delivery. More specifically, we need to know who and how many people are in poverty and measure inequalities by understating those pushed back to extreme poverty or experiencing increasing inequality, marginalisation and exclusion. Such data allows us to respond appropriately to any crisis, address needs, and plan better for recovery.

What kind of data are we referring to?

We need to have disaggregated data by age, sex, disability type and also geography beyond the urban- rural binary. The data needs to be collected using simple and standard methods to identify factors keeping people in such states or even factors facilitating movement out of extreme poverty.

How useful is data for local decision makers?

Data on local needs and resources gives local decision-makers access to critical insights. These include for example availability of resources, community assets, locations of boreholes, prevalence of disease, health clinics records and education attainment.

When is data meaningful?

For data to be meaningful, it should go beyond just disaggregation and needs to be inclusive. Vulnerable people in crisis very well know what matter most and the types of support they need to break out from a crisis. This means including affected communities along the data value chain right from the design of data collection instrument to the actual data collection, analysis, and use. Ultimately, whatever level of data one obtains and however inclusive it is, it has to lead to holding decision makers accountable.

Is it true there is no data?
The need to consider diverse data sources: Example on disability

In our studies, we find that there is little evidence to indicate dearth of data. There is data but it is everywhere and inaccessible to many. Data is sitting in departments and agencies of government institutions. It is also with non-state actors that collect for programme interventions and evaluations.

An example on disability data from Kenya: a varied portfolio of data provides insight on the status and outcomes of persons with disabilities in Kenya. Disability is complex, and it is experienced differently by individuals and for this reason, other data sources on persons with disability provide different perspectives and can be used to complement the official statistics to get a more detailed and nuanced

perspective of disability and where policies and practices can effectively tackle discrimination and exclusion.

Over the course of our study on disability data in Kenya, we compiled an inventory of disability data sources that have been published over the last 10 years. The sources were categorised into official surveys, non-official surveys, census data, administrative data, or qualitative data. Using these criteria 44 sources were captured. Of these, 32 were produced by government agencies and 12 were produced by organisations outside of government (the list is not completely comprehensive, and it should be noted that in comparison to government data, non-official data is likely to be harder to locate and access). Kenya National Bureau of Statistics (KNBS) published the most data sources included in the inventory (9) followed by the Kenya National Examinations Council (KNEC) (7).

Table 1: Category of disability data publications (2011-2022)

Data category	Number of disability data publications
Official survey data	15
Non-official survey data	8
Official administrative data	7
Mixed official (survey, census and administrative)	7
Official qualitative data	2
Non-official qualitative data	2
Non-official administrative data	2
Official census data	1

Challenges of data on those living in poverty

Data particularly at the sub-national data ecosystem in is mainly paper based. This limits the use of data in decision making especially at local levels , weak administrative data systems, limited statistical capacity building, rapidly changing, costly and complex ICT infrastructure, over-reliance on external funding for some statistical programmes, and limited exploitation of non-traditional data sources, among others.

Data is not interoperable and not easily accessed outside of the originating institution, limiting its use. In most cases, there are no open data protocols in place. This is the primary reason that data is not easily accessible outside of the institution that collected it, which is something that stifles its wider use. For example, when KNBS carried out a survey on behalf of the National Council for Persons with Disabilities (NCPWD) in the past, KNBS has only been able to share the public report with the Council, and not the granular data, due to data protection measures and limitations on data sharing. This means that the Council has not been able to complete any additional analysis of the raw data and could not generate extra value by creating its own insights.

Non-sensitive data is not provided in an open source. For example, just to be granted access to general reports from the National Education Management Information System, actors have to write a letter to the relevant director requesting permission. Bureaucratic processes, which, on the one hand, are essential to protect sensitive and personal data, but on the other hand, there is little effort to allow access in such a way that does not stifle a broad culture of data use.

Underfunding of data collection. Our engagement with key informants in Kenya on disability data revealed it is significantly underfunded, impacting the quality of data collected. Throughout the interview, stakeholders unanimously stated that disability data is significantly underfunded by government and other actors. In relation to the government, for example, the vast majority of the ministerial-level agencies that we spoke to do not have dedicated budget lines for disability data related activities. This situation is exacerbated by a cyclical challenge where data is needed to justify budget allocation, whilst at the same time financial resources are required to collect data.

Data collection is made a costly exercise. Borrowing from the Kenya example, both government entities and non-governmental organisations approach KNBS to carry out surveys and other types of data collection on their behalf. Our interview revealed, in 2014-15, an organisation of persons with disability approached KNBS with a proposal to carry out a programmatic survey on special needs

education, and the Bureau provided an estimation of $800,000 to complete the work. In conjunction with the Ministry of Education, the organisation instead contracted a private entity to complete the work at a cost of $250,000 by using the same frameworks and tools.

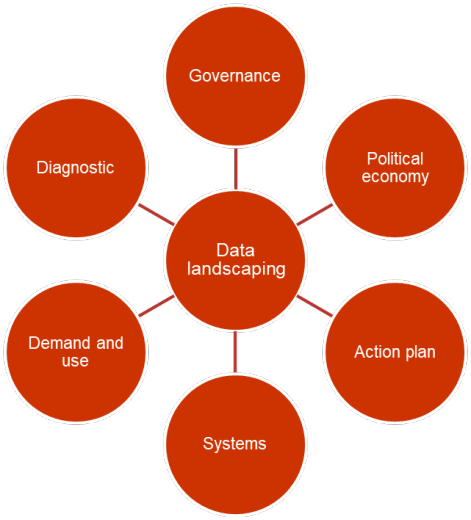
Practical steps to improve data ecosystems to identifying data on those living in poverty

At DI we undertake three activities to support more effective data ecosystems, which we have done in about 15 countries. These are data landscaping, supporting data systems and join-up datasets.

1. Data landscaping

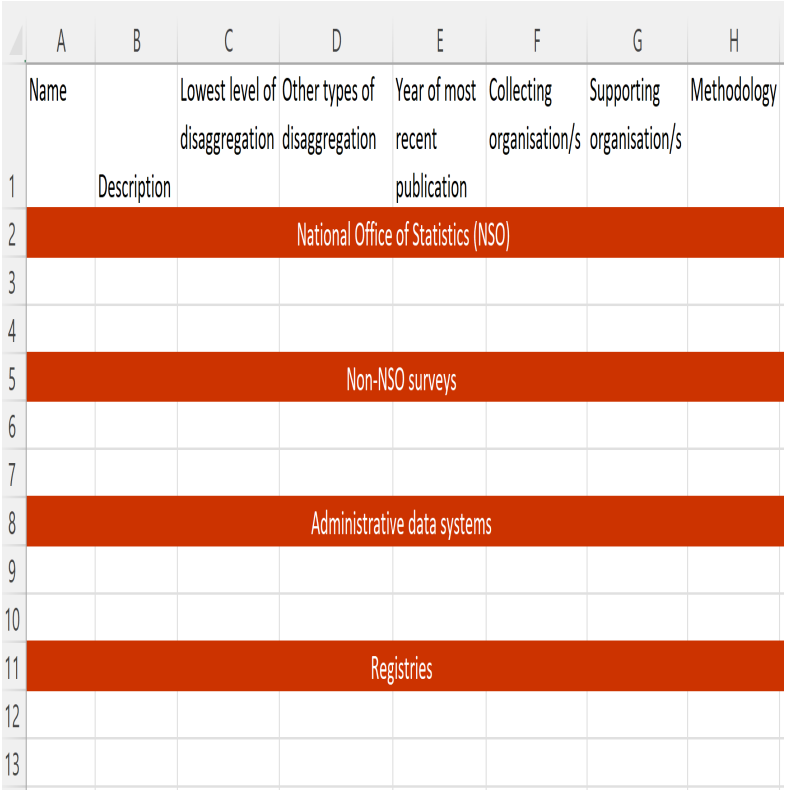
At DI, we have developed a data landscaping model where we try to understand the data ecosystem holistically (Figure 1). These include assessing the political economy and governance of local government data, policies and legislations, data strategies and architecture. The data ecosystem audit also takes a user-centric approach, which allows us to understand data user needs.

Figure 1: Data landscaping of the whole data ecosystem



The data landscaping also allows us to undertake data diagnostic as presented in Figure 2.

Figure 2 Data diagnostic - a template



	A	B	C	D	E	F	G	H
1	Name	Description	Lowest level of disaggregation	Other types of disaggregation	Year of most recent publication	Collecting organisation/s	Supporting organisation/s	Methodology
2	National Office of Statistics (NSO)							
3								
4								
5	Non-NSO surveys							
6								
7								
8	Administrative data systems							
9								
10								
11	Registries							
12								
13								

2. Joined up datasets

There is need to bring diverse datasets together to build a more comprehensive picture. For increased data use at country level, DI encourages joining up data sets. This is demonstrated in our work via the Spotlight on Uganda and Spotlight on Kenya data platforms. These are comprehensive online resources that bring together a range of official publicly available data on Kenya and Uganda under one roof. The platforms provide data and visualisations on key socio-economic indicators and financial flows at the subnational level, highlights key data gaps and supports the increased use of data and information to reduce poverty and promote sustainable development.

Our biggest lesson from this approach is we need a simple approach to improve the coordination and ownership of existing systems and datasets and making the most of the evidence that is already being collected.

3. Data systems

There is need to improve data governance, including promoting supportive policies, processes, and government structures to enable data use. In our project 'Strengthening sub-national data value chains in Uganda', in partnership with African Population & Health Research Centre, we embarked to fill evidence gaps in national data ecosystems through implementation of catalytic micro-examples for strengthening sub-national data systems with the overall goal of strengthening data systems for planning and decision making at national and sub-national levels. In this undertaking, we are trying to co-create micro-examples of good data systems such as an online system for digitalisation for safer and easier data storage, sharing and internal communication using an open-source platform. So far, in Kayunga district, three departments are digitalising their paper-based archives and migrating to NextCloud where the district planning unit is taking the lead. The next step is supporting an interactive dashboard for presenting infographics based on already digitalised district data. This is helping to promote data use by decision makers where development indicators across departments are

visualised in easy-to-understand charts, graphs and maps. The dashboard also permits comparison of indicators between sub-counties and parishes.

In short, in strengthening data systems particularly at sub-national level, follow an approach that is best suited to different sub-national governance structures. The approach should address gaps across the whole data value-chain and lead to good practices and appropriate entry-points that provide the greatest impact in data uptake and use within sub-national systems. We also need to ensure that the approach is the most cost-effective way of building sustainable data systems to support sub-national level decision-making and ownership.

II. Digital solution for identifying those in poverty and eradicating poverty

Our understanding of digital solutions is digital transformation. That is, the transformation of existing analogue processes, primarily within public administration, to digital processes. Digital data capture at the point of service delivery, such as in health facilities, schools and civil registry offices, is a crucial component of digital transformation. Another big component is digitisation of foundational systems. Digital transformation as a form of digital solution to eradicate poverty also involves building national and sub-national e-government infrastructure. To be sustainable, digital transformation must garner the commitment from government to promote national ownership and management of all infrastructure.

1. Digital Data Collection

At DI, we do not propose radical AI type intervention rather digital data collection. We believe Africa's overriding priority for digital transformation is digital data collection and with digital data collection must come use of that data at point of entry and in local governments.

The data should come from health facilities, schools, registry offices, local planning departments, water utilities, and district agriculture extension officers, among others. Much of this data is incomplete and is still collected at source on paper (being digitised further up the data chain), and some in text, and images, and with no attributes. For most of these data sets, there is no regular frequency for their collection but rather collected on a need-by-need basis. One of the key priorities should be increasing the digital collection of data at source as rapidly and frequently as possible.

2. Digitisation of foundational systems

Apart from collecting data from every public service point, another digital solution that can bring digital transformation is digital civil registration and vital statistics (CRVS). This is critical for the vulnerability and poverty data ecosystem. CRVS incorporates data systems which interlink with the provision of and access to social protection services. The vital statistics produced from civil registration data can be an important source of up-to-date information, which in advanced statistical systems is near real-time. They can help governments to identify people in need, and to understand what their needs are. One challenge we identified in Uganda is vital statistics are not produced using data from civil registration systems rather the national statistics office collects, complies and disseminates vital statistics from decennial population censuses and household surveys such as the Demographic and Health Survey every 5-years. In times of emergencies, such data is of little value for monitoring changes and responding quickly.

Another system under the CRVS umbrella which is linked to social protection services is the provision and use of national identification numbers (NIN). However, there is anecdotal evidence from Uganda which suggests that the national ID system actually creates obstacles that prevent people from accessing services. The difficulty involved with obtaining a NIN, combined with the cost of rectifying any mistakes, means vulnerable people are disproportionately impacted. This has been the experience of many elderly people, when they tried to access the Social Assistance Grants for Empowerment programme and could not receive support they were entitled.

Why deprioritise radical AI?

Coming back to AI, why should we deprioritise AI as a solution to Africa's problems affecting millions of people? Granted AI could have a transformative impact for good but without a foundation of timely, quality, inclusive, and joined-up data, it will fail. For many of the least-developed countries that foundation is missing. There has been an influx of funding aimed at delivering on the promise of AI. However, these projects have relied on large sets of private data. This has brought a lack of clarity on data collection, implicit data biases along gender lines, and potential private sector use of model outputs. Without access to public data and minimal government input, these projects often start with 'what problems *can* be solved', rather than 'what problems *need* to be solved'.

The idea that technological innovation that exploits big data is the solution to low-income countries' data problems is a dangerous concept currently gaining currency in development circles. The data captured in a child's birth registration, a patient's health record and a pupil's progress through school are the kinds of inputs needed for machine learning to develop useful algorithms. In low-income countries this data does not exist in sufficient quantity or quality to be of any use. Yet every month a new data science challenge is launched by donors who appear not to grasp the fact that the data needed for artificial intelligence to work comes from the very systems that artificial intelligence wants to replace.

In summary, we need to identify public data gaps, barriers to data use and models for improving existing data. This is not a quick fix, but it will deliver sustainable solutions for more effective AI-ready data systems in future.