Selected Best Practices and Standards to reduce data fragmentation

Data fragmentation between stakeholders and countries remains a key issue regarding data collection on forcibly displaced persons. Improving data governance to address data fragmentation remains a challenge within institutions at the national level, largely due to the absence of an entity responsible for informing standardized methodologies and data policies in order to facilitate an open and transparent space for data collection and data use. Data is often collected independently using different actors and methods which prevents the comprehensive comparability of different datasets. Furthermore, many standard poverty surveys are difficult to adapt within refugee camp settings, which creates some technical barriers to addressing data fragmentation. Many sectors tend to publish data not intended to be used for other purposes.

Process of data fragmentation at each stage

- Inconsistency between terminology, concepts, definitions, and classification, often reflecting country-specific legislation, policies and practices
- Variation in data collection methods
- Difficulty measuring highly mobile populations, especially during ongoing humanitarian crises, which potentially affect accuracy and reliability of data
- Gaps in official statistics, including a lack of socio-economic data
- Use of different available data sources
- Capacity constraints that limit scope and effectiveness of data collection
- Inconsistency in compilation and presentation of data
- Lack of comparability across populations and countries
- Limited ability to analyse across a range of data sets
- Importance of confidentiality and data protection, which limit access and sharing
- Limited capacity for dissemination activities

Throughout a two-day learning event in July 2021, panellists and government officials from six Great Lakes countries discussed best practices to reduce data fragmentation and harmonise rules and practices, not only to promote regional approach, but to address data vagueness and concerns in terms of data security. Three suggestions were proposed including: i) involving a more diverse array of stakeholders, ii) increased standardisation of definitions and concepts related to common indicators and iii) designing a favourable legal framework for responsible data sharing across the region. To illustrate these discussions, the three topics and illustrating case studies from the discussions have been further outlined below.

The experience of Democratic Republic of Congo (DRC) in involving local actors in data collection within sensitive settings shows the benefits of fostering local interventions through the involvement of clusters of partners in data collection, assessment and information sharing. The Expert Group on Refugee, and Internally Displaced Persons Statistics (EGRIS) has provided guidance for standardisation of definition, harmonisation of standards and shared use of terminologies for the countries to develop reliable data sources. The case of Burundi provides us with an opportunity to reflect on the importance of data legislation and the protection of beneficiaries prior, during and after the sharing and dissemination of data.

These three cases present the importance and best ways to reduce data fragmentation between all stakeholders, including the affected communities, to provide tailored services, promote regional integration, and take greater ownership of the proposed development responses.

---

1 This document was developed as part of a year-long series of learning events developed jointly by the World Bank, Samuel Hall, and the Rift Valley Institute on development responses to forced displacement with panellists and government officials from six countries of the Great Lakes region.
Involving More Diverse Stakeholders to Establish Cross-sectoral Data Collection Mechanisms

Further collaboration and coordination on data collection practices between local and regional key development actors, community leaders, UN agencies and institutional stakeholders is needed in order to establish data collection standards. Building on existing and potential data requires a mixed top-down, bottom-up approach involving local actors from various locations, expertise and purposes in the collection process. Establishing cross-sectoral data collection mechanisms with partners is key to avoiding data fragmentation and ensuring transparent dissemination. Therefore, involving more diverse stakeholders would also include sectoral coordination mechanisms, with government counterparts where possible, to lead sector-specific needs assessments, set and monitor intervention. For that purpose, the national statistical systems could bring together data suppliers and data users to harmonise the data sets yet to be disseminated and avoid duplications which equate to wasting resources and time. Within the Great Lakes region, the DRC response to Ebola provides an instructive case study on how to engage stakeholders including local communities in order to help develop tailored interventions aimed at mitigating adverse effects of crises that can be multidimensional.

Case Study: DRC - Local intervention: Cluster of local partners within fragile contexts

Key features

- Parallel tracking programming where local actors and governmental representatives inform each others on need gap analysis
- Context analysis through in-depth needs assessment in fragile contexts in collaboration with cluster of local partners
- A community based approach to information sharing and a close coordination between INGOs and local actors on the ground for diligent action
- Multi-Stakeholder governance aiming at gathering all involved sectors to plan an adequate response

The consequences of internal armed conflicts coupled with an increasing number of related natural disasters have accentuated the population movement in the DRC. As a result of these two major push factors, some areas were rendered challenging to access. However, the DRC proved institutional assets as per extensive experience in data collection within fragile settings to respond to those identified challenges. The DRC involves a range of specialised structures in the process at the institutional level, such as the Ministry of Planning through the National Institute of Statistics, the Ministry of Interior through the national commission for refugees, which collect indicative data on those areas. These institutions also work in close collaboration with the General Directorate of Migration and UN agencies across the country.

To foster local intervention, the government also developed a cluster approach—clusters of partners or thematic groups met to discuss the most appropriate way to intervene, considering risk factors and the extent of potential benefits. This approach considers the local realities when planning for local intervention, as well as assesses objectives that integrate the sustainable development goals based on the principles of peacebuilding and state-building. This facilitates greater information sharing for both stakeholders and authorities about the gaps, adjustments and opportunities in implementing justice and security foundations, as well as access to livelihood and basic necessities. In relying on an array of local partners and grassroots expertise, stakeholders can ensure the relevance of intervention and promote a development response based on empirical data.

During the Ebola outbreak, the DRC proved that a community-based approach to information sharing constitutes a credible response within fragile contexts. When combined with close coordination between
INGOs, local actors on the ground and key workers within the communities, this approach served to mitigate the pandemic’s adverse effects on social life and public health, positively informing the health response. Additionally, the Engagement Commission led by the Ministry of Health encouraged communities to develop their own strategies to provide adequate health and social responses based on their own context and realities and provide adjustments when needed. This engaging strategy proved itself critical in gathering feedback and testimonies from communities – that were initially hard to reach – through entrusted community leaders. Data is analysed by partners and development actors and expedited to governmental counterparts for their consideration before (re)designing a comprehensive intervention at the local level.

**Main Impact**

- Comprehensive response strategies developed that were evidenced based on local realities documented by grassroots experts, community leaders and local partners.
- Encouraged communities developing their own strategies to provide adequate health and social responses.
- Improved reliability of data sources and data access for stakeholders within fragile contexts.
- Facilitated information sharing between stakeholders and the authorities about the gaps, adjustments and opportunities in implementing justice and security foundations as well as access to livelihood and basic necessities.

**Standardisation of definitions, concepts, and methodology**

The standardisation of concepts and definitions to ensure appropriate standards at a regional level was central to the discussions on reducing data fragmentation and allowing standard operating protocols to provide similar understandings of key concepts, while developing a harmonised methodology for all stakeholders and key development actors. EGRIS has led international recommendations on refugee and IDP statistics as discussed below. For example, the figure below presents the scope of refugee and refugee related populations, divided into three classifications. A standard statistical concept of these populations allows for better comparability and opportunity to understand potential differences that exist within statistics. This will help actors involved in data collection improve consistency and facilitate the systematic collection of well-identified categories depending on the intended area of action/intervention.

As a region, the first step toward harmonisation of data procedures would be to understand the core definition of each concept and discuss persons of concerns’ specificities in terms of needs, rights, and vulnerabilities. Once the standards are set and the data collected, further and closer collaboration is needed to ensure a common use for homogenised deployment of human and material resources in designing and implementing routine administrative data collection and profiling.
Case Study: EGRIS – Embracing international standards

Many resources and international support remain at the disposal of countries affected by forced displacement in order to help identify categories of people, define vulnerabilities and support their data collection effort. EGRIS, led by The United Nations Statistical Commission (UNSC) has recently released reports addressing the needs for better standardisation of protocols and methodologies including on definition and categorisation of persons of concern.

Key Features

The international EGRIS is composed of participants from national authorities, international statistical organisations, and other technical experts. The group aims to respond to the need for an international statistical framework for forcibly displaced populations. The international recommendations on refugee statistics (IRRS) and international recommendations on IDP statistics (IRIS) aim to:

- Support countries and regions establishing statistical frameworks standardising terminology on specific contexts to ensure that key concepts meet international standards
- Assess the reliability of data sources
- Improve statistical coordination at the national, regional, and international levels

The International Recommendations on Refugee and IDPs Statistics provide material assistance to guide the measurement of collected data on these populations. The recommendation established a statistical framework standardising relevant terminology and classification that can be used as a foundation in countries’ prospects for regional responses to forced displacement. EGRIS provides recommendations on terminologies and definitions and supports statistical institutions with sufficient methodological tools to capture and analyse data at a different level, whether national, regional or international. Furthermore, EGRIS supports implementing these recommendations by assessing the quality level of data produced by countries based on international recommendations and ensuring it provides enough substance for comparability assessment with regional partners to help reduce data fragmentation at a regional scale. The ultimate goal of these recommendations is for the countries to take full ownership of these expected standards for more conventional use, sharing, disaggregation and dissemination of data.

Many countries are taking steps toward implementing the recommendations. For example, in 2019, integrating EGRIS recommendations in their data collection strategy, Kenya's Bureau of Statistics, in close collaboration with UNHCR, included refugee and stateless people in their national survey. In addition, following the conduct of the fieldwork and an analysis of socio-economic data collected from a stateless community identified as the Shona community, the Kenyan government acknowledged the need to recognise this population by providing citizenship documents, including birth certificates. This practice will facilitate their prospects for livelihood and foster more effective integration. Indeed, the formal legitimacy accorded by the granted documentation shall come with a better sentiment of belonging, improving the social well-being of previously marginalised communities.
Data sharing, regulation and protection standards

Data regulation and protection standards must be incorporated in data collection protocols on forcibly displaced populations. Designing a responsible framework for data protection and confidentiality is essential to meet protection standards, especially at the regional level. Frameworks play a central role in the voluntary involvement of persons of concern in the data collection process. However, scaling up data sharing from the national to regional level presents several challenges, including standardised security measures, such as drafting rules and regulations for data sharing and drafting ethics procedures and practices. One of the key questions posed during discussions on data collection was - How can countries in the Great Lakes region strengthen the statistical information sharing system on displacement without compromising data protection? In the effort for cross-border data harmonisation and reduction of data fragmentation, looking at the legal framework and existing laws across countries on the matter of data to determine what can be shared and what cannot be shared for security reasons is inevitable. Following the emergence of data digitization, technological advancements and innovations must always be considered when developing strategies that aim to address security concerns and data protection standards.

Case study: Burundi – Protection of beneficiaries and data legislation

Key Features

In Burundi, the legal and institutional frameworks are dictated by the 'Law of 2007' on the organisation of the statistical system in Burundi. Article 1 defines the fundamental principles and the institutional framework regulating the activities, services, and organisations in charge of the production and dissemination of public statistical data throughout the country. Furthermore, it implies supervision of the general functioning of the national statistical system as well as its coordination. Burundi’s legislation considers data collectors’ ethical duties, including suppliers, and protects the past or potential beneficiaries of future interventions, programme monitoring exercises, and evaluations, as covered by Articles 14–18 on the obligations of data users.

On a broader scale, Article 17 of the Social Protection Code establishes a single social register to identify and manage beneficiaries of social protection programs. The ongoing government-led project consists of setting up a responsible method of data collection and statistics for the beneficiaries of social protection programs. These
beneficiaries, including internally displaced persons and returnees, are guaranteed protection and confidentiality by regulation measures.

Data on forced displacement and refugees are collected by relevant institutions including the National Statistical Office, INGOs and UN agencies (UNHCR, IOM) under the supervision of the Ministry of Interior and are published on an annual basis in the Ministry's statistical yearbook. The dissemination process involves various stakeholders, such as UN agencies and government counterparts. The legal steps taken by the government of Burundi to protect beneficiaries of programmes and persons of concern should be considered a best practice to be built on and further developed at a regional level as it constitutes a strong baseline for statistics legislation as recommended by international standards.

Strengthening the legal framework will have a positive impact on the quality of data collected, improve consistency, increase confidence between all stakeholders through transparent practices and platforms in line with set regulations. It should also serve as a baseline for further revision of the existing regulatory framework aimed at addressing identified breaches while improving implementation of set regulations, as outlined by IRIS recommendation 319.

**Main Impact**

- Centralised dissemination, verified and published on an annual basis in the Ministry's statistical yearbook
- Increased confidence of beneficiaries that took part in the survey
- Provided a legal baseline for responsible data collection and statistics on the beneficiaries of social protection programs
- Facilitated evaluation and monitoring of past interventions without compromising confidentiality rules

**Additional Issues and Questions to Think About**

*Developing ICT structures and knowledge*

The idea of establishing an integrated shared web-based platform to provide the same amount and level of data for all key stakeholders across the region was shared during the learning event. E-practices could increase capacities by saving material resources and facilitate data processing, from the collection stage via telephone screening, use of mobile phones and tablets for immediate data storage to diligent data analysis through internal sharing. Such practices can track beneficiary information when needed and aid with geographic location of samples via the GPS functionality at the same time. For this, a sufficient level of technical expertise, IT structures and logistics would be required from countries willing to implement innovative systems at a regional level. From the digitisation of data collection processes to the computerisation of the electronic filing system, it is necessary to reinforce ICT structures to conduct innovative data management protocols based on technological trends that would save money, resources, and time.

Moreover, following the emergence of new technological capabilities – from collection to dissemination – more efforts should be deployed on cyber-security procedures in relation to data security to ensure those data are stored, safe and under the custody of relevant authorities.

*Involve applied research to ensure evidence-based methodologies*

Beyond regular data collection, it is essential to invest in basic applied research to standardise definitions and methods and increase awareness of policy trends, by doing so, developing techniques that are only based on empirical studies and up to date content on the matter of refugee and internally displaced statistics, quantitative methodology and migration studies that inform on best practices around the world. Academia and research can play a central role in the harmonisation effort to reduce data fragmentation in setting up regular indicators for potential re-conceptualisation/re-definition when needed and according to current trends, regional or global contexts and seminal works.
Applied research assists policymakers in gathering best practices throughout the region and following fast-evolving trends on the matter to improve public policy-making skills. It can also help map migration dynamics beyond the traditional “push and pull” approach to embrace transnational realities within regions like the Great Lakes. For example, in Burundi, Rwanda, and Uganda, refugees and asylum seekers move across borders for short periods of time. Further research on how to categorise these flows could be incorporated within future development responses, in order to identify new opportunities, present within forcibly displaced populations.

Regional technical working groups to ensure regular update and harmonisation of practices

As previously discussed, the regional harmonisation of texts on demographics and socio-economic aspects of forcibly displaced persons requires a common approach to the information collected in the databases between the different countries and stakeholders at the regional level. Technical working groups across borders and on the regional level between involved countries must be installed to ensure smooth communication. Maintaining communication channels between governments and key stakeholders would aid the installation of a monitoring structure, on efforts deployed to reduce fragmentation of data (including standardisation of text, legal framework, data sharing protocols etc.), where a regional team composed of decision-makers, researchers and development actors could evaluate, monitor and fill identified gaps promptly. The Great Lakes region would benefit in further developing a regional space to share policy solutions, best practices and challenges in SDG implementation. In this case through establishing data-sharing agreements or information-sharing networks to share micro and/or aggregated data, as and when appropriate.