



MOVEMENT: Minding the data gaps around displacement, funding, and humanitarian needs

JUNE 2017

Executive Summary

INTRODUCTION

In the past year, humanitarian needs have soared, with at least 30 active crises around the world. Over 65 million people have been forced to flee their homes—becoming part of the 152+ million people in need of humanitarian assistance globally,¹ including 20 million people in dire need of urgent food support in the four famine stricken countries.² While humanitarian organisations race to provide lifesaving support, funding levels are faltering, with humanitarian appeals on average barely 30% funded, and aid budgets at risk of taking huge cuts in tough political times.³ In this environment, it's more important than ever to be able to make the most effective and impactful use of resources, and to be able to pinpoint where the greatest needs are to prioritize assistance.

Yet the information architecture of humanitarian aid is not fit for purpose. After a year of digging into the data, meeting with partners and implementing organisations, and surveying existing efforts, we observed that the state of humanitarian data appears grim. It is incomplete, outdated, and siloed. This means that resources can't be tracked to results in real-time, and stakeholders cannot tell if essential services for refugees and displaced populations are being efficiently delivered. Where we have information, it may be dispersed across platforms, in different formats, and doesn't tell a compelling story about how commitments benefit real people. People are likely dying due to the current inadequacies of the system. In the year following the Grand Bargain reached at the World Humanitarian Summit, there is clearly much work to be done.⁴

In order to draw attention to the breadth of needs, and to make decisions on proper resource allocation, there must be one place where data on people's movements, needs and funding levels are brought together to form a complete picture of humanitarian requirements and support. It should be regularly updated in real time—and made an open, public tool for policymakers, implementers and advocates alike to view the trends and gaps in displacement needs and funding. While various organisations have made important progress in improving data availability and standardisation, many of these are multi-year efforts facing an uphill battle, or are geared as internal response management tools. In the absence of a publically available resource, the ONE Campaign embarked on an effort to learn more about the constraints in bringing these disparate data threads together and find out what it will take to get to a more coherent, comprehensive real time system for monitoring displacement data.

ONE's initial efforts have led us to create a unique visual interface, MOVEMENT, which allows users to explore country-level displacement data in a way that hasn't been done before. MOVEMENT integrates existing datasets to show through relatable metrics where displaced people are today, what the greatest humanitarian needs are globally, and where humanitarian funding flows align—or do not—with those places of greatest need. It also illustrates how fragile states bear the brunt of all the above. No existing online effort to-date visually displays

the relationship between refugee/IDP populations, humanitarian needs and funding streams, and drivers of displacement (which we have proxied through the Fragile State Index).

This brief describes the main data challenges encountered while creating MOVEMENT, highlights important efforts to address these issues, and draws attention to critical recommendations to improve data management. These challenges will not come as news to many involved in the humanitarian world. Ample bright minds are currently working on ways to bridge these data divides. We have great respect and awareness of the enormous challenges facing the sector in data management, which are miniscule compared to the life-threatening issues those working in the sector face every day to provide assistance to 152+ million people in need around the world.⁵

Far from attempting to duplicate their endeavours, ONE's work in creating MOVEMENT supports these efforts by advocating for more resources and attention to humanitarian data needs. Making the case for humanitarian and development aid is harder than ever. Backing advocacy with robust data that shows the long-term impact of such assistance on security and intervention outcomes should be a priority for all humanitarian actors and supporters. Our approach is to add support, elevating existing efforts at unifying data across the displacement and humanitarian sphere, while calling attention to areas most in need of better resourcing.

KEY DATA CHALLENGES

The critical sets of displacement data—on population stocks and flows, humanitarian needs, funding flows, and drivers of displacement—are in varying states of disaccord. Displaced population data is accessible at a macro (country) level, but frequently inaccessible at a micro (province or smaller) level. There is no standard repository of data on the stock and flow of displaced populations across all countries, and different organisations collect population information in different ways. When it exists, information on displaced people's needs, such as how much people need access to health, nutrition, shelter, and education, is often siloed between organisations and measured in different ways, making it impossible to compare needs between areas. Such data is usually collected for entire regional or country populations—making assessing the needs of only displaced people near impossible. As the most complete source of needs data, MOVEMENT utilizes metrics from the UN Office for the Coordination of Humanitarian Aid contained in various country-level Humanitarian Needs Overviews, but had to manually scrape 67 reports for current appeals globally. The resulting 356 differently worded metrics were then grouped into similar sets, where possible, to display like information across countries. Meanwhile, various datasets that speak to the drivers of displacement are available, including the Fragile States Index, but the complexity of displacement necessitates drawing on a diverse array of sources, from conflict data to climate change models.

Funding data from donors is not complete or fully transparent, giving only the most superficial picture of available resources to address the vast need. Tracking the "last mile" of funding flows is also nearly impossible. There is no existing way to demonstrate how particular recipient organizations or countries distribute funding at a sub-national level, nor to aggregate how much funding is supporting displaced populations in particular. Although a key commitment in the Grand Bargain is for donors and implementing organisations to improve transparency around humanitarian funding, little progress has been made.

Lastly, almost no one affords displaced people a public outlet to assess and express how their own needs are being met. Data that speaks to how refugees, internally displaced persons, and similar groups of people actually perceive their own needs and how they are being met is spotty at best, and often completely absent.

To overcome the hurdles that limited the scope of MOVEMENT, ONE has developed a series of preliminary policy recommendations that can help catalyse humanitarian data transparency in the near term.

RECOMMENDATIONS

→ For Humanitarian Organisations and National Reporting Organisations:

- Problem: Detailed, timely data on displaced population movements and needs is not currently shared well, nor collected with sufficient granularity.
- Recommendation: Such data should be shared more widely, at a reasonably disaggregated level. Disparate sets of stock and flow data on displaced people created by different organisations should have a baseline of comparable metrics across countries, and be accessible in a centralized, open source database. Current efforts to collect and share data should be better resourced.

→ For Humanitarian Implementing Organisations:

- Problem: Data on the humanitarian needs of displaced populations is not often separable from the entire population of a country in need—if present—and data on those aggregate needs are normally locked away in static files.
- Recommendation: Organisations should ensure the underlying methodologies and survey techniques used to assess needs are transparent and comparable across contexts. While employing means to safeguard the private data of beneficiaries, additional aggregate information on specific needs should be made public in line with the World Humanitarian Summit Grand Bargain principles of transparency and data sharing. Ongoing reforms towards this end should be bolstered.

→ For Donors Governments and Organisations:

- Problem: Data on planned and actual humanitarian funding is often not available at a subnational level and infrequently updated.
- Recommendation: Donors should provide detailed funding flow data in clean formats to the UN OCHA-run Financial Tracking Service and the OECD in line with their Grand Bargain commitments. They should also provide program implementation and evaluation information by publishing corresponding data in the robust, geotagged IATI Standard and ensuring open access to anonymized underlying data.



Peering into the gaps: The state of displacement data today

Comprehensive humanitarian responses hinge on understanding people's needs, gaps in service provision, and the ultimate impact of funding. We should be able to track the needs expressed by affected individuals themselves as well as the impact of the interventions and policies created to help them. Likewise, we should be able to follow the flow of humanitarian and development funds from donors to recipient countries and implementing organisations.

However, tracking the "last mile" of funding flows is nearly impossible. Humanitarian funding data is generally available at the level of country, appeal, emergency region, or organisation. Unfortunately, it is not possible to demonstrate how particular recipient organizations or countries distribute funding at a sub-national level. Although a key commitment in the Grand Bargain is for donors and implementing organisations to improve transparency around humanitarian funding, little progress has been made.

Additionally, it is extremely difficult to find comprehensive data about displaced people's needs, both at the global level, or at a sub-national level. Like funding data, last mile data on health, nutrition, shelter and other needs is not present across displaced and refugee-like populations. Needs data is collected for entire regional or country populations—not just displaced people—and then broken down by sector. The needs of displaced people may vary from the general populace affected by a crisis, depending on the context, but it is profoundly difficult at present to understand how many displaced people need particular forms of assistance. Additionally, this data emanates from disjointed datasets that a) fail to collect the same information across countries; b) are siloed between organizations; or, c) are kept isolated due to privacy concerns.

Almost no platform affords displaced people a public outlet to assess and express how their own needs are being met. Data that speaks to how refugees, internally displaced persons, and similar groups of people actually perceive their own needs and how they are being met is spotty at best, and often completely absent. Several initiatives are afoot to change this, but they remain geographically limited in scope and scale at present.

Much work is being done to address some of these crucial issues. Within the past year, just to name a few, the International Organization for Migration's Global Migration Data Analysis Centre,⁶ the Joint IDP Profiling Service (JIPS)—working along with Data2x,⁷ the Internal Displacement Monitoring Centre,⁸ and the World Bank have all published on the topic.⁹ Bolstered by such research, efforts at resolving problems around humanitarian data are proliferating. Among others, these include initiatives headed by the United Nations Office for the Coordination of Humanitarian Affairs (UN OCHA), which is undertaking a multi-year process to standardize data management throughout the Humanitarian Program Cycle (.hpc).¹⁰ Meanwhile, the Humanitarian Data Exchange (HDX) is helping in the short and medium term to aggregate more open source humanitarian data streams. That HDX is now managed by a new UN OCHA-run Centre for Humanitarian Data in The Hague, Netherlands, indicates this increased interest in improving "1) data services; 2) data literacy; 3) data policy; and 4) network engagement."¹¹ Other agencies, such as the UN High Commissioner for Refugees (UNHCR), World Food Program (WFP), and the International Organization for Migration (IOM) have all appointed staff to the new Centre.

Rather than replicate ongoing efforts, ONE's work in creating MOVEMENT supports these efforts by advocating for more resources and attention to humanitarian data needs. Making the case for humanitarian and development aid is harder than ever. Backing advocacy with robust data that shows the long-term impact of such assistance on security and intervention outcomes should be a priority for all humanitarian actors and supporters. Through our scoping and research, and our first-hand experience bringing data together through an online platform, we have focused on four related categories of displacement data that have been used in MOVEMENT: 1) displaced population data; 2) humanitarian needs data, particularly for displaced populations; 3) humanitarian funding data; and 4) data about the drivers of displacement. This section explores what data is accessible and usable today, what challenges are inherent in using the data, and what data simply doesn't exist yet. Where applicable we describe the data we've used in the MOVEMENT platform, and what datasets we might be able to add in the future. The most technical aspects of this data assessment can be found in data challenges annex.

DATA ON DISPLACED POPULATIONS

Aggregate refugee and displaced population figures by country—both of origin and of asylum—are relatively easy to obtain. However, there is no standard repository of data on the stock and flow of displaced populations across all countries. The statistics division of UNHCR compiles global data centred on refugees from other UN agencies, humanitarian organisations, local governments, and other sources. It releases population statistics at mid-year intervals.¹² The complementary Internal Displacement Monitoring Centre (IDMC) publishes the gold standard for internal displacement figures annually, which are usually also noted in the UNHCR reports.¹³ Figures from IOM that track various forms of migration and displacement are also incorporated into the IDMC data. IOM's Displacement Tracking Matrix (IOM-DTM), for example, contains high-quality granular data on displaced population needs and movements.¹⁴ A complementary initiative to IOM-DTM is 4Mi, a lightweight approach to collecting and analysing migration flows currently being piloted in the Horn of Africa.¹⁵ However, underlying issues "associated with the collection, compilation and dissemination of statistics" on displaced people are present even with the above aggregate figures and have led to an Expert Group on Refugee Statistics that is drafting new standards to be suggested in 2018.¹⁶ Many of these issues stem from varying legal and political definitions of refugees, internally displaced persons, and populations of concern that are not unified across collectors of data, leading to methodological challenges.

Data containing disaggregated demographic information about the composition of displaced populations is difficult to source. As above, there is no central repository of data on different populations across all countries. Privacy concerns limit the publishing of granular data in conflict scenarios, but even state or province-level information is usually inaccessible. For 15 major crisis situations, the UNHCR-administered Operational Portals provide a degree of more granular data.¹⁷ While more timely and nuanced, the Portals are country-specific, and the quality is mixed when compared with the official annual data sets from UNCHR and IDMC. The sites are also segmented by crisis, so comparison across countries is not straightforward.

Data used in MOVEMENT platform: Due to their availability and widespread acceptance, the aggregate official UNHCR figures on various displaced populations by origin and asylum country, including the total population of concern, refugees, and IDPs, are displayed in MOVEMENT.

NEEDS DATA FOR DISPLACED PEOPLE

There is no one comparable data source which shows the needs of displaced people across all countries, in areas such as health, nutrition, education, et cetera. The most comprehensive source of information on humanitarian needs are the Humanitarian Needs Overviews (HNOs) crafted by national UN OCHA offices. HNOs must be sourced from various OCHA websites and are only available for countries or situations that have an outstanding humanitarian aid appeal.¹⁸ Underlying metrics and methodologies of data collection vary widely. This in turn limits accessible macro and micro needs data—countries like Pakistan and sites like the Dadaab refugee camp lack summary needs data. Additionally, the number of people in need is not usually broken down by refugee versus host population, or by any other demographic disaggregation.

Aside from the HNOs, consistent needs or socio-economic data on refugee populations is not easy to find. Legitimate privacy concerns do uphold safeguarding data—but even aggregate, anonymized figures are fiendishly elusive. For example, a 2016 joint policy paper between UNHCR and the Global Education Monitoring Report noted the lack of access to education faced by refugees and IDPs—but also bemoaned the absence of good quality data with which to ground that conclusion.¹⁹ Health data is similarly difficult to access. The UNHCR maintains a standardized Health Information System,²⁰ but the data is only accessible with humanitarian credentials through the UNHCR’s “Twine” platform and public aggregated data is seldom published.²¹ The World Food Program (WFP) also has public datasets on acute malnutrition and similar metrics that impact displaced populations on its VAM Shop platform, but many of their operations products that contain valuable humanitarian information are only released in static form.²²

Data used in MOVEMENT platform: As the most complete source of needs data, MOVEMENT utilizes the HNO metrics, but had to manually scrape 67 HNO and similar reports for current appeals globally. The resulting 356 differently worded metrics were then grouped into similar sets, where possible, to display like information across countries. The need for this manual effort is a major barrier to open data sharing and the reuse of such critical data. Thankfully, UN OCHA is in the process of standardising HNO publishing and data releases for mid-year 2017, but the online platform, a part of the .hpc Initiative, has yet to be published as of May 2017.

FINANCIAL DATA ON HUMANITARIAN AID

One of the largest challenges in displacement research is understanding where humanitarian funding goes and how much of it reaches displaced people. UN OCHA’s Financial Tracking Service (FTS) offers the best up-to-date, accessible snapshot of these humanitarian aid flows globally.²³ It contains constantly updated data on pledged and contributed funds from a range of donors, humanitarian organizations, and national governments. However, there are limits to what FTS can show. The data is only as good as what donors report. Most don’t report detail past the country level, meaning that the “last mile” of humanitarian aid

almost always remains opaque. FTS also allows individual donors to decide what qualifies as humanitarian assistance—a highly discretionary exercise. One country might consider development aid like entrepreneurial training to be such assistance, while another might keep their definition to the more traditional health or shelter types of support.²⁴

Another way to compare more detailed information on aid projects is through the various datasets published in the International Aid Transparency Initiative's (IATI) Registry. 545+ organizations currently have some data in the IATI Standard, while over 80 donors and development organisations have formally signed on to IATI.²⁵ However, while the data standard allows for very detailed project information, the data in IATI 1) lacks the level of financial detail found in FTS; 2) is difficult to access across publishers; and, 3) remains only as good as what implementers and donors report—and unfortunately most publishers do not report particularly detailed information.

Ultimately, it is not possible in either FTS or IATI to precisely identify funds targeted at refugees or displaced populations. FTS allows breakdowns by aid sector, but divining what goes specifically to help displaced individuals is near impossible. Likewise, while IATI's new humanitarian codes allow projects to be tagged, identifying displacement-related projects is deeply problematic through the Registry.²⁶

Data used in MOVEMENT platform: MOVEMENT relies on FTS to display full year 2016 humanitarian aid funding to countries, broken down by donor, and 2017 appeals funding to-date. There are data limitations in FTS that limit MOVEMENT to reflect regional appeals funding.²⁷ As there is no way to identify funding specifically directed towards displaced people and their needs, MOVEMENT relies on country-level humanitarian funding data as a proxy—with the presumption that portions of that funding do, indeed, reach displaced populations. This is a limitation of the field that must be solved.

DRIVERS OF DISPLACEMENT DATA

The majority of displacements occur primarily because of conflict or disasters—and the fragile situations that influence responses to those events.²⁸ The 2017 Fragile States Index developed by the Fund for Peace is an excellent country-level assessment of state fragility that proxies the above drivers through its 12 indicators.²⁹ A comprehensive index that blends quantitative data sets, qualitative research, and content analysis, it is the best overall assessment of state fragility available on a global scale.

However, drivers of displacement are complex and not entirely understood. Country-level assessments do not fully capture the intra-state dynamics that push people from their homes. These include the capacity of a state to respond to conflict and disasters, economic development, or lack thereof, and many more. Developing data-driven models of displacement that might one day predict why people leave their homes requires pulling nuanced data from a variety of other sources. The 4Mi mixed migration monitoring captures a variety of these drivers, but is geographically limited. In terms of global conflict data, the best and most complete dataset available is from the Uppsala Conflict Data Program (UCDP).³⁰ It covers state- and non-state-based violence from 1975 to 2015 and can be freely downloaded. For current, near-real-time data, the Armed Conflict Location and Event Data Project (ACLED) is also excellent.³¹ However, it currently only covers Africa and Southern/Eastern Asia. On the disaster front, exceptional global information is accessible through the Emergency

Events Database published by the Centre for Research on the Epidemiology of Disasters.³² Other excellent datasets exist on relevant push and pull factors for forced migration, but a comprehensive review falls outside the scope of this overview.³³

Data used in MOVEMENT platform: Drivers of displacement are proxied through the 2017 Fragile States Index developed by the Fund for Peace. Critically, integrating the Index demonstrates the underlying fact that the majority of displaced people both originate from—and are hosted in—the most fragile states in the world.



Anticipating the future: What can be done with better, open data?

IMPROVING ACCESS STANDARDS— WHAT CAN BE DONE TODAY?

Building the MOVEMENT platform underscored the importance of bridging existing humanitarian data gaps. The data silos and barriers to accessing information outlined above, outlined with greater detail in the data challenges annex, prohibited our developing a product that could speak to the needs of and assistance going to displaced people. Developing a view at the country-level proved to be remarkably difficult, let alone displaying the “last mile” of assistance inside countries.

Thus, it is crucial that organizations coordinate with each other regularly about the data they are collecting and the methodology they are using to do so. The danger in not frequently syncing such information manifests itself in data confusion while seeking answers even to our most basic questions. Equally importantly, a lack of high-quality data on displacement and humanitarian aid makes advocating for better resourcing and directing of humanitarian efforts all the more difficult.

Steps towards standardising data collection, reporting, and publishing are in progress—and must be supported. Data architecture must be implementable, built as frameworks that provide standardization across data sets without compromising flexibility or incurring significant operational overhead. A prime example of work in this area is the UN OCHA-led .hpc project, a multi-year process to standardize data management throughout the Humanitarian Program Cycle. The Humanitarian Program Cycle consists of needs assessment and analysis, strategic response planning, resource mobilisation, implementation and monitoring, and operational review and evaluation.³⁵ OCHA aims to boost data transparency and legibility across each stage, beginning with a new platform to access needs assessment data in 2017. Another critical effort to support is the Humanitarian Data Exchange (HDX), an impressive endeavour aggregating open source humanitarian data. The new parent organization of HDX, the OCHA-run Centre for Humanitarian Data in The Hague, Netherlands, deserves attention for its mission to bolster excellence in the field.”³⁶

WHAT BENEFITS CAN BETTER DATA BRING TOMORROW— AND HOW DOES ONE FIT INTO THE PROCESS?

While many broader initiatives like those highlighted above are in progress, most will take years to fully implement. Innovation faces challenges not only from massive resource and capacity constraints, but also from decades of ingrained processes at different agencies that have developed their own methods and standards. It will be difficult to build willingness to do things differently.

In the meantime, there is an urgent need to be able to view funding gaps, hold donors accountable, and advocate for greater support to the areas that need it most. Constructing MOVEMENT is a first step in bringing together available data now, while also allowing in a practical manner to highlight the most glaring data challenges that must be addressed in order to visualise even basic trends. The lessons we have learned are reflected in the recommendations below, and in the data challenges annex.

Building on the gaps outlined in MOVEMENT, we see that ONE might help act as a facilitator in the evolution of humanitarian data by bringing disparate groups to the table, forming high level consensus towards gathering and sharing better displacement data, and empowering those with the best technical skillsets to go about their work. Closer relationships and coalitions will be needed to lobby relevant organizations and donors to publish timely, disaggregated data to organizations like UN OCHA. As mentioned previously, apart from donors, additional partners working on more granular data include IATI, AidData, the World Bank, and implementing organizations like WFP.

In the longer term, better humanitarian and displacement data quality can lead to the ability to collectively analyse data-driven linkages between fragility and security, humanitarian interventions, and mass migration/displacement. Emboldening such high-quality data sharing would facilitate potentially paradigm-shifting deliverables, but fundamentally relies on better displacement data collection and publishing from national governments and state-led agencies.

One such long-term outcome of better data could be the development of early warning systems that predict refugee or IDP flight paths from conflicts or disasters. The machine learning that underpins such work requires high-quality, clean, datasets. Should those be set in motion, there is the exciting potential to convene humanitarian actors with major tech firms like IBM, Google, or ESRI and lobby for applying big-data knowhow to generate further insights into the new, high quality datasets. The UNHCR Digital Engagement team, UNHCR Innovation Lab, World Bank, UN OCHA, and HDX are all ideal partners in this sort of work.

Another long-term outcome could be bolstering data on aid implementation efficacy by putting tools in the hands of refugees and displaced persons to review services (e.g. ground-truthing). Building on data sharing platforms and programs like 4Mi, new technologies such as block-chain ledgers could better enable displacement people to put their own voices first. Virtuous feedback such as this can empower and help those most in need globally. Simply put, after minding the gaps, the next step is finding a way to bridge them.

RECOMMENDATIONS FOR A HUMANITARIAN DATA EVOLUTION

To overcome the hurdles that limited the scope of MOVEMENT—and to move towards the future possibilities outlined here—ONE suggests the following recommendations that can help catalyse action in the near-term. These recommendations aim, in-line with the goals of MOVEMENT, to support and call attention to existing humanitarian data efforts while also suggesting complementary policies.

→ For Humanitarian Organisations and National Reporting Organisations:

→ **Problem:** Detailed, timely data on displaced population movements and needs is not currently shared well, nor collected with sufficient granularity. Such data should be shared more widely, at a reasonably disaggregated level, to provide the humanitarian community with more current information on ongoing situations.

→ Recommendations

1. Disparate sets of stock and flow data on displaced people created by different organisations should be consolidated and accessible in a centralized, open source database.
2. Organisations gathering data on displaced populations should do so in a consistent format across countries. Recognizing that country contexts vary, a better level of comparability is still possible. Efforts at this are underway and should be better supported.
3. More consideration should be given to tracking and understanding how IDPs become refugees, and how both IDPs and refugees sometimes become returnees.
4. Organisations ought to consolidate access to their various public sources of data, ideally through a single, well documented RESTful API. Such should also be published under the Humanitarian Exchange Language (HXL) to the Humanitarian Data Exchange.³⁷

→ For Humanitarian Implementing Organisations:

→ **Problem:** Data on the humanitarian needs of displaced populations is not often separable from the entire population of a country in need—and data on those aggregate needs are normally locked away in static files.

→ Recommendations

1. Humanitarian Needs Overviews or similar needs surveys should disaggregate the demographics of populations better to include refugees, IDPs, women, children, and other categories.
2. Organisations should ensure the underlying methodologies and survey techniques used to assess needs are transparent and comparable across contexts where possible.

3. While employing means to safeguard the private data of beneficiaries, additional aggregate anonymized information on specific needs should be made public in line with World Humanitarian Summit Grand Bargain principles of transparency and data sharing. This data should be released in the HXL standard by default. Ongoing reforms towards this end should be bolstered.
4. UN agencies and humanitarian organisations that collect situational information on humanitarian crises and release it publically should share their underlying data, even in aggregate form, through RESTful APIs servicing HXL standard files instead of static PDF files.

→ **For Donors Governments and Organisations:**

→ **Problem:** Data on planned and actual humanitarian funding is insufficiently detailed and infrequently updated.

→ **Recommendations**

1. Donors should provide funding flow data in clean formats, ideally in the HXL standard, to FTS and the OECD in line with their Grand Bargain commitments.
2. Donors should provide program implementation and evaluation information by publishing in the more robust IATI Standard and ensuring open access to their underlying data.
3. Donors must push implementing organizations to provide geotagging, even at a state or province level, to better track the “last mile” of funding flows.

ACKNOWLEDGEMENTS

This brief was compiled by Galen Englund under the guidance and direction of Sara Harcourt and David McNair at the ONE Campaign. It incorporates technical contributions based on earlier scoping by Mark Brough (brough.io). Yang Hong of Shoshin Insights also contributed to the technical scoping and the underlying Humanitarian Data Service API that powers MOVEMENT. Amanda Taylor and Jacopo Hirschstein of Tekja Data Visualisation constructed the visualisation framework for MOVEMENT. Kate Vang at ONE has been instrumental in building the data architecture necessary for the platform and also contributed to the suggestions herein.

We are grateful to the many organisations and partners we have consulted with to better understand the data challenges around displaced people and humanitarian funding, and who are working tirelessly to address these constraints with very few resources, while serving populations in need.

Annex: Data Challenges

1. DISPLACED POPULATIONS DATA

AVAILABLE DATA (KEY SOURCES)

- UNHCR Population Stats: The statistics division of UNHCR compiles global data from humanitarian organizations, local governments, and other sources and releases population statistics at mid-year intervals.
- There is an API for year-end figures, but it's not updated for the mid-year statistics.

- UNHCR-administered Operational Portals provide more granular data on 15 crisis situations globally. The data is considerably more timely, as it is collected from the field and pushed to aid humanitarian coordination—but without as much of a formal review process.

- The complementary Internal Displacement Monitoring Centre publishes the gold standard for internal displacement figures annually, which are usually relied on in UNHCR reports.
- Data is accessible through the IDMC API for its Global Internal Displacement Database (GIDD)

- The International Organization for Migration's Displacement Tracking Matrix (DTM) datasets often have detailed data on demographic backgrounds of displaced populations, their access to basic services, as well as overall numbers of displaced individuals in specific settlements.

- A complementary initiative to IOM-DTM is 4Mi, a lightweight approach to collecting and analysing migration flows. 4Mi offers hyper-localized information about migrants' profiles, drivers of displacement, and current needs. It also offers near real-time updates. Developed by the Regional Mixed Migration Secretariat, it is currently hosted by the Danish Refugee Council.

DATA CHALLENGES

- Data is only updated twice a year for most countries.
- Legal and political labels of displacement, such as refugees, internally displaced persons, and populations of concern are not unified across collectors of data, leading to methodological challenges.
- Underlying issues "associated with the collection, compilation and dissemination of [refugee] statistics" are present even in aggregate figures and have led to an Expert Group on Refugee Statistics that is drafting new standards to be suggested in 2018.³⁸

- Data is only on a country-by-country basis, and only for countries affected by chosen crises.
- Data is segmented by crisis, so comparison of statistics across countries is not straightforward.
- Quality is mixed when compared with the official UNHCR annual figures, the IOM-DTM, and 4Mi.
- Data cannot be easily exported currently: there is an API, but data is in the process of being migrated from data.unhcr.org to data2.unhcr.org.
- Improvements are coming: according to conversations with the UNHCR, the new API will eventually feed into a unified API for all of UNHCR's data.

- Data is only updated annually.
- IDMC data focuses on internal displacement.
- The incomplete data collected globally on displacement hampers report completeness.
- There is limited information on IDP demographic profiles and specific locations inside states.
- IDMC has difficulty accounting for transitions from IDPs who become refugees or migrants.⁴⁰

- At present, these are only available for a select few countries, though more are being added.
- Much of the DTM data is not available in a standardized data format across countries.
- As of May 2017, the IOM-DTM team is in the process of updating their website to allow easier access to the DTM data. The organization is also planning to standardize the baseline across countries to allow for more comparable data—though that will be a long-term endeavour.

- At present 4Mi is only being piloted along the Horn of Africa, but will likely expand further.
- Despite onus on mixed migration flows, it faces limitation in assessing refugee populations due to mandate of UNHCR and institutional barriers.
- Data is not currently exportable through the site.

UNAVAILABLE DATA

- There is no standard repository of data on the stock and flow of displaced populations across all countries.
- Definitional issues prohibit absolute clarity on stock and flow figures from the national statistics office figures used in official IDMC and UNHCR reports.
- Privacy concerns limit the publishing of granular data in sensitive contexts, but even state or province-level information is usually inaccessible.
- Time-series data is difficult to obtain, apart from the UNHCR population stats, as many datasets are new or have undergone changes in their metrics over time.

2. NEEDS DATA

AVAILABLE DATA (KEY SOURCES)

- **Humanitarian Needs Overviews (HNOs)** are the most comprehensive source of information on the humanitarian needs of people globally. Crafted by national UN OCHA offices, they been conducted for every humanitarian appeal and can be sourced from various OCHA websites. The HNOs generally contain information on the number of people in need by cluster, such as health or education, broken down by subnational region.
- **ACAPS**, another humanitarian organization, provides support of multi-sector assessments and publishes an overview of global humanitarian crises that is updated weekly. It provides a simple rating scale to highlight which countries are most in need and provides more detailed situation analyses by thematic area.
- The UNHCR maintains a standardized **Health Information System** that collects data on health indicators.⁴¹
- The World Food Program (WFP) also has datasets on food security, acute malnutrition, and similar metrics that impact displaced populations accessible through its VAM Shop platform.⁴⁴

DATA CHALLENGES

- The number of people in need is not usually disaggregated by demographic background: determining how many of a given population in need are refugees, internally displaced, or women, for example, is not usually possible.
- While the HNOs follow a somewhat standard template, the underlying metrics and methodologies used for data collection vary widely.
- Most HNOs are only published in PDF format and the underlying data is not usually shared, apart from a few hosted on the Humanitarian Data Exchange.
- UN OCHA is in the process of standardizing the publishing of HNOs and underlying data for mid-year 2017, but the online platform, a part of the .hpc Initiative, has yet to be published as of May 2017.
- Underlying data informing the ACAPS rankings is only accessible in topline form, limiting its usefulness for further analysis.
- Health data on displaced populations is difficult to access—UNHCR data is only accessible with humanitarian credentials.⁴²
- Even aggregated, anonymized data is not accessible.
- Targeting data for partner organizations is accessible online through the UNHCR's "Twine" platforms, but public aggregated data is not published.⁴³
- Many WFP operational update products that contain valuable humanitarian information are only released in static PDF form.

UNAVAILABLE DATA

- HNO metrics are only available for countries that have an outstanding humanitarian aid appeal coordinated by UN OCHA. This fundamentally limits the geographical spread of needs data that is easily accessible—though the appeals do cover the majority of humanitarian crises globally.
- Aside from the HNOs, consistent needs or socio-economic data on populations in need is not easy to find. For example, education stats are fiendishly elusive: in 2011 the UNHCR highlighted the lack of reliable data on school enrolment data for refugees. More recently, a 2016 joint policy paper between UNHCR and the Global Education Monitoring Report noted the lack of access to education faced by refugees and IDPs—but also bemoaned the absence of good quality data with which to ground that conclusion.
- Aggregated health data on refugee or displaced populations is not accessible, save for one-off reports or data releases.

3. HUMANITARIAN FUNDING DATA

AVAILABLE DATA (KEY SOURCES)

- The Financial Tracking Service (FTS), run by UN OCHA, offers the best up-to-date and accessible snapshot of current humanitarian aid flows globally. FTS contains data from a range of donors, humanitarian organizations, and national governments. It collects data on pledged and contributed amounts from donors to countries and organizations.
- Disaggregated information, including the targeted sector of funding, is trackable, and the platform updates constantly as donors report specific changes to their pledges over the year.
- The data is available through a full-fledged API.

-
- The IATI Standard offers a way to compare generally detailed information on projects undertaken by the 545+ organizations publishing their data to the Standard. Importantly, IATI is not a data service like FTS, per se. Rather, it is a set of publishing requirements to make data comparable across publishers, coupled with a Registry of who publishes with IATI, and a Datastore to access most of that data in one place.
 - The IATI Registry is also not limited to financial data: sub-national locations, project documents, and results data are included by some publishers.

DATA CHALLENGES

- The data in FTS is only as good as what donors report—and UN OCHA cannot make donors to release better quality tracking data. Data past the country level is not tracked, meaning that the “last mile” of humanitarian aid most always remains opaque.
- FTS also relies on what individual donors to decide on qualifies as humanitarian assistance. One country might consider counter violent-extremism programming to be such assistance, while another might keep their definition to the more traditional health or shelter types of support.
- There are data limitations in FTS for certain appeals (particularly regional appeals), which means the breakdown of data by cluster is not always available.
- Refugee assistance clusters do not adequately reflect total funding of refugees, while much earmarked funding for displaced groups is coded simply as multi-sectoral assistance.

-
- IATI data is not easily manipulated: one must sift through numerous organizations’ datasets to find desired stats via a limited user interface.
 - Further, though many organizations use the IATI Standard, the architecture is 1) somewhat restrictive in what can be published; 2) difficult to start with for many organizations given its comparably heavy reporting standards; 3) requires a high level of upfront processing and manual effort to implement; and 4) introduces complicated overhead to access data as a consumer.
 - In addition, while it retains general development data, IATI does not have the level of comprehensive humanitarian funding data that FTS has, nor the level of programmatic detail that the Humanitarian Data Exchange currently hosts.
 - As such, IATI is not particularly conducive to emergency response situations where time is short and the high burden to conform data to the IATI Standard would detract from the actual relief effort. If adherence to the IATI standard was required for more humanitarian data publishers—not necessarily donors—it could pose an operational overhead challenge to organizations that may already spread thin in the current climate.

UNAVAILABLE DATA

- Untangling humanitarian funding from development funding is fraught with difficulties. The OECD Development Co-operation Directorate (DAC) tracks full-year development funding, but variation in definitions means that the datasets are not necessarily comparable—even when parsed for only humanitarian aid.
- Ultimately, it is not possible in either FTS or IATI to identify precisely funds targeted at refugees or displaced populations. FTS allows for a breakdown by sector, or by targeted cluster for appeals data, but divining what goes specifically to help displaced individuals is near impossible. Refugee assistance clusters do not adequately reflect total funding of refugees, while much earmarked funding for displaced groups is coded simply as multi-sectoral assistance. Likewise, while IATI’s new humanitarian codes allow projects to be tagged by appeal, there is no way of specifically identifying refugee-related projects.

4. DRIVERS OF DISPLACEMENT

AVAILABLE DATA (KEY SOURCES)

→ The Uppsala Conflict Data Program (UCDP) is the best and most complete dataset in terms of conflict data. It covers state- and non-state-based violence from 1975 to 2015 and can be freely downloaded through their API.

→ For more current / near-real-time data, the Armed Conflict Location and Event Data Project (ACLED) is also excellent.

→ The API is basic, but functional.

→ The Emergency Events Database is ideal for the disaster front, affording access to exceptional global information is accessible through the published by the Centre for Research on the Epidemiology of Disasters.

→ The 2017 Fragile States Index developed by the Fund for Peace also works as proxy for many drivers of displacement. A comprehensive index that blends quantitative data sets, qualitative research, and content analysis, the Fragile States Index is the best overall assessment of state fragility available on a global scale. Critically, integrating the Index demonstrates underlying fact that the majority of displaced people both originate from—and are hosted in—the most fragile states in the world.

→ FEWS NET is a famine early warning system that can display areas of greatest concern for food insecurity and famine risk.

DATA CHALLENGES

→ UCDP lacks fine grained data on individual conflict events, but is otherwise an excellent resource.

→ ACLED currently is geographically limited to covers Africa and Southern/Eastern Asia.

→ The Index is only published annually and is not updated in an accessible manner throughout the year.

→ As a composite index, the Fragile State Index is of limited use assessing sub-national trends and events.

→ The index does necessarily take into account disasters—though the societal impact of them may be reflected in some indicators.

→ While excellent in quality, most data must be manually downloaded through shapefiles, rather than from a restful API.

UNAVAILABLE DATA

→ The majority of displacements occur because of conflict or disasters—usually rooted in political problems.

→ However, just under the surface, drivers of displacements are very complex—and not fully understood. These factors include the capacity of a state to respond to conflict and disasters; economic development, or lack thereof; educational attainment; and many more.

→ As such, developing data-driven models of displacement that might one day clearly show why people leave their homes require pulling such nuanced data from many sources—and pairing that with higher quality on displacement.

Endnotes

1. Internal Displacement Monitoring Center, Global Report on Internal Displacement, (2017), <http://www.internal-displacement.org/global-report/grid2017/pdfs/2017-GRID-part-1.pdf>; United Nations Office for the Coordination of Humanitarian Needs (UN OCHA), Humanitarian Needs Overviews, (2017). <https://www.humanitarianresponse.info/>
2. UN OCHA, "Address and Prevent Famine in Four Countries," http://interactive.UNOCHA.org/emergency/2017_famine/index.php, (2017).
3. As of 8 June 2017. See UN OCHA – Financial Tracking Service, "Appeals and response plans 2017," <https://fts.UNOCHA.org/appeals/overview/2017>, (June 2017).
4. For an overview see World Humanitarian Summit, Commitments to Action (2016), https://www.worldhumanitariansummit.org/sites/default/files/media/WHS%20Commitment%20to%20Action_8September2016.pdf; Development Initiatives, Baseline report: Implementing and monitoring the Grand Bargain commitment on transparency (2017), <http://devinit.org/post/baseline-report-implementing-and-monitoring-the-grand-bargain-commitment-on-transparency/>.
5. UN OCHA, Humanitarian Needs Overviews (scraped and summarized data), (2017).
6. International Organization for Migration, "Global Migration Data Analysis Centre," <https://gmdac.iom.int/>, (2017).
7. Data2x, "Data on Displaced Populations," <http://data2x.org/partnerships/data-displaced-populations/>, (2017).
8. See, for example, Ginnetti, Justin, Painting a clearer picture with 'flawed' and 'messy' data (March 2017), Internal Displacement Monitoring Centre, <http://www.internal-displacement.org/library/expert-opinion/2017/painting-a-clearer-picture-with-flawed-and-messy-data>.
9. See, for example, Sarzain, Zara, Stocktaking of Global Forced Displacement Data (February 2017), World Bank, <https://openknowledge.worldbank.org/handle/10986/26183>.
10. UN OCHA, "Humanitarian Program Response," <https://www.humanitarianresponse.info/en/programme-cycle/space>, (2017).
11. UN OCHA, "Centre for Humanitarian Data," <https://centre.humdata.org/about-us/>, (2017).
12. UNHCR figures are published formally on an annual basis, but provisional mid-year updates are available. See, UNHCR, "Population Statistics Database," <http://popstats.unhcr.org/en/overview>, (2016).
13. Internal Displacement Monitoring Center, Global Report on Internal Displacement, (2017), <http://www.internal-displacement.org/global-report/grid2017/pdfs/2017-GRID-part-1.pdf>;
14. International Organization for Migration, "Displacement Tracking Matrix," <http://www.globaldtm.info/v/>, (2017).
15. Regional Mixed Migration Secretariat, "4Mi," <http://4mi.regionalmms.org/4mi.html>, (2017).
16. Expert Group on Refugee and IDP Statistics, International Recommendations on Refugee Statistics (2016), <http://ec.europa.eu/eurostat/documents/7723677/7761703/Draft+outline+of+International+Recommendations+on+Refugee+Statistics>.
17. Not officially part of UNHCR-sanctioned data, the Operational Portals act as a rapid sharing platform for up-to date information across crises. Accessible at <http://data2.unhcr.org>.
18. UN OCHA runs a variety of websites with humanitarian information where HNOs can be found including <http://humanitarianresponse.info> and reliefweb.int.
19. Global Education Monitoring Report, No more excuses: Provide education to all forcibly displaced people, Policy Paper 26 (May 2016), <http://unesdoc.unesco.org/images/0024/002448/244847E.pdf>.
20. UNHCR, "Health Information System (HIS) Toolkit," <http://www.unhcr.org/en-us/protection/health/4a3374408/health-information-system-toolkit.html>, (2010).
21. See, for example, UNHCR, Zaatri Health Information System—Annual Report 2016 (February 2017), <http://data.unhcr.org/syrianrefugees/download.php?id=11541>; UNHCR, "Twine," <http://twine.unhcr.org>, (2017); Gnucoop, Data for refugees: a new data collection system for UNHCR (December 2016), <https://www.gnucoop.com/data-for-refugees-a-new-data-collection-system-for-unhcr/>.

22. World Food Program, "The VAM Shop," <http://vam.wfp.org/>, (2017); See, for example, World Food Program, "Situation Reports," <http://executiveboard.wfp.org/situation-reports>, (2017).
23. UN OCHA, "Financial Tracking Service," <https://fts.UN OCHA.org/>, (2017).
24. See the data challenges annex for more detail on this.
25. IATI, "IATI Registry," <http://iatiregistry.org>, (2017); Simbodyal, Rohini, International Organization for Migration joins IATI (2016), IATI, <http://www.aidtransparency.net/news/international-organization-for-migration-joins-iat>.
26. The CRS codes used in the 2.02 IATI Standard contain three codes under "emergency response": Material relief assistance and services, Emergency food aid, and Relief co-ordination; protection and support services. Again, these do not specifically identify funds to support refugees.
27. See data challenges annex; Global Humanitarian Assistance, Humanitarian assistance to education for the Syrian emergency (2016), <http://www.globalhumanitarianassistance.org/wp-content/uploads/2016/02/Humanitarian-assistance-to-education-for-the-Syria-emergency.pdf>.
28. Internal Displacement Monitoring Center, Global Report on Internal Displacement: 9 (2017), <http://www.internal-displacement.org/global-report/grid2017/pdfs/2017-GRID-part-1.pdf>.
29. The Fund For Peace (FFP), "Fragile States Index," <http://fundforpeace.org/fsi/>, (2017).
30. Uppsala University, "Uppsala Conflict Data Program (UCDP)," <http://ucdp.uu.se/?id=1>, (2017).
31. Armed Conflict Location and Event Data Project, "About ACLED," <http://www.acledata.com/about-acled/> (2016).
32. Centre for Research on the Epidemiology of Disasters, "EM-DAT," <http://www.emdat.be/>, (2017).
33. See, for example, IDMC, Understanding the root causes of displacement: towards a comprehensive approach to prevention and solutions (2015), <http://www.internal-displacement.org/assets/publications/2015/20151208-root-causes-displacement.pdf>; and, World Bank, Forced Displacement and Development:11-17 (2016), <http://siteresources.worldbank.org/DEVCOMMINT/Documentation/23713856/DC2016-0002-FDD.pdf>.
34. The Fund For Peace (FFP), "Fragile States Index," <http://fundforpeace.org/fsi/>, (2017).
35. UN OCHA, "Humanitarian Program Response," <https://www.humanitarianresponse.info/en/programme-cycle/space>, (2017).
36. UN OCHA, "Centre for Humanitarian Data," <https://centre.humdata.org/about-us/>, (2017).
37. HXL is a lightweight data standard created by the Humanitarian Data Exchange designed to be a flexible columnar metadata format to walk the line between high overhead of standards adherence vs complete lack of standards. More information at <http://hxlstandard.org/>.
38. Expert Group on Refugee and IDP Statistics, International Recommendations on Refugee Statistics (2016), <http://ec.europa.eu/eurostat/documents/7723677/7761703/Draft+outline+of+International+Recommendations+on+Refugee+Statistics>.
39. Not officially part of UNHCR-sanctioned data, the Operational Portals act as rapid sharing platform for up-to date information across crises. Accessible at <http://data2.unhcr.org>.
40. Internal Displacement Monitoring Center, Global Report on Internal Displacement: Overcoming data shortfalls, (2017), <http://internal-displacement.org/global-report/grid2017/pdfs/2017-GRID-part-3.pdf>;
41. UNHCR, "Health Information System (HIS) Toolkit," <http://www.unhcr.org/en-us/protection/health/4a3374408/health-information-system-toolkit.html>.
42. See, for example, UNHCR, Zaatri Health Information System—Annual Report 2016 (February 2017), <http://data.unhcr.org/syrianrefugees/download.php?id=11541>.
43. UNHCR, "Twine," <http://twine.unhcr.org>, (2017); Gnucoop, Data for refugees: a new data collection system for UNHCR (December 2016), <https://www.gnucoop.com/data-for-refugees-a-new-data-collection-system-for-unhcr/>.
44. World Food Program, "The VAM Shop," <http://vam.wfp.org/>, (2017); See, for example, World Food Program, "Situation Reports," <http://executiveboard.wfp.org/situation-reports>, (2017).