SUMMARY OF DATA SOURCES FOR REFUGEE CRISES

TOSCANE CLAREY

11 2018

STRENGTHENING RESILIENCE AND RESPONSE TO CRISES

THE SERVICE

Through the Lot B: Resilience service, DAI offers rapid response, high quality support to UK Government and other donors, across a wide range of development and humanitarian challenges.

We offer support for risk informed design for development interventions across all sectors; risk and contingency financing; understanding changing systems; and strategic integration of humanitarian action and development.

We offer a clear process for users that draws upon a well-established network of relevant expertise provided through over 60 consortium partners. We are able to build strong practical partnerships for rapid and responsive delivery through:

> A dedicated, easy-to-access Secretariat to manage new enquiries and assure delivery

> Consistent end-to-end quality assurance

> A user friendly, customer oriented outlook

> A pro-active approach to knowledge sharing and communication

> A focus on due diligence, efficiency and cost effectiveness.

This document provides a summary of data sources that could potentially be used to develop innovative financing for refugee crises.

The primary source data used for the case studies in this report were the UNHCR Global Displacement Dataset and the UN OCHA Financial Tracking Service.

1 DISPLACEMENT DATA

UNHCR Global Displacement Dataset tracks forced displacement over time in UNHCR target countries. The homepage shows a global visualization of the number of refugees, asylum seekers, IDPs, returnees and stateless persons from 1951 until 2017. The map allows that information to be broken down by country. In the *Time Series* tab one can sort by year (1951-2017, selecting all years if necessary), by host country, by country of origin and by population type (IDP, refugees, stateless persons, etc). *Time Series* gives you a cumulative number of registered refugees per year in any given host country. The *Demographics* tab sorts that information by age and sex per year, host country and location within the host country. *Asylum-Seekers (Refuge Status Determination)* allows one to sort by year, country of origin, host country and phases of an asylum-seeker's refugee status application. *Asylum-Seekers (Monthly Data)* shows the number of asylum seekers in a given country of asylum by year, month and country of origin. It is possible to download UNHCR's mid-year statistics here as well.[1] Critiques of this dataset by Marbach are that 2% of the population's origin country was unknown in 2014 and that one cannot distinguish between a value being zero, not available and not applicable. [2]

The **Marbach Amendment to UNHCR Global Displacement Dataset** is an academic paper that outlines issues with the UNHCR dataset and provides another model, "a Bayesian hierarchical dynamic-count model to interpolate and extrapolate the missing segments in the time-series of displaced-person counts from base UNHCR Data."[3]

The **Global Internal Displacement Database (GIDD)** developed by IDMC allows one to explore, filter and sort data to produce graphs and tables visualizing internal displacement. GIDD has three main tools: *Displacement Data*, the *Risk Model* and the *Displacement Data Exploration Tool*. The *Displacement Data* shows the published figures from IDMC's annual flagship reports. The *Risk Model* explores disaster-related displacement risk metrics and the likelihood of the occurrence of specific displacement events. The *Displacement Data Exploration Tool* helps users produce their own custom charts.[4]

IOM Displacement Tracking Matrix (DTM) tracks internal displacement in 72 countries. Data is collected via surveys, flow monitoring, registration and mobility tracking. The data is shared through GIS products, reports, online visualizations, raw data and customized data exports. The homepage displays a global map visualization that displays internal displacement by cause (natural, conflict, political/economic). The datasets can be sorted by country, operation within that country, round of data collection, component (type of data collection, i.e. survey) and date. The reports are sorted by region, county and keyword. The data visualization gives a monthly evolution of monitoring figures in active operations. Once one selects a year and a month, multiple graphs are

---

[1] http://popstats.unhcr.org/en/overview

[2] http://moritz-marbach.com/download/papers/20160525_unhcrdata.pdf

[3] Ibid.

[4] http://www.internal-displacement.org/database

generated representing IDPs vs. returnees over time, number of IDPs/returnees by country, reasons for displacement and trends in displacement.[5]

IOM FLOW Monitoring focuses on migration to Europe and collects data on "the volume and basic characteristics of populations transiting through selected locations - referred to as Flow Monitoring Points (FMPs) – during specific observation hours."[6] Previous transit points, next destination, intended destination, transportation, number, sex, nationality are all recorded. More in-depth Flow Monitoring Surveys (FMS) are also used.

The homepage displays an interactive map of Europe that shows four different visualizations: 1. Number of migrants and arrival trends over time 2. Migrant presence for certain countries 3. Relocation 4. Mediterranean dead/missing. One can search FLOW reports by keyword and country. The datasets can be sorted by country, operation within that country, round of data collection, component (type of data collection, i.e. survey) and date. When one selects the year and monthly time span, the data visualization shows the total number of arrivals then split by sea arrivals and land arrivals, cumulative arrivals and country overviews. The *Latest Figures* tab shows recent trends in arrivals and the main nationalities of arrivals.[7]

IOM's **Migration Data Portal** consolidates key global data on migration from a range of sources into one place. The portal has six main sections: *Data Migration Statistics, Thematic Overviews, Tools and Guidance Materials, SDGs and Global Compact on Migration, Snapshots* and a *Blog.*

1. The *Migration Statistics* can be visualized through an interactive world map that can be sorted by year (1990-2017), by country/sub-region/region and then by a range of indicators and sub-indicators. These include immigration and emigration, migrant flow, vulnerability (migrant deaths, human trafficking), integration and wellbeing (employment and education metrics), forced migration (refugees, asylum seekers, resettlement, IDPs, child refugees), development (remittances), migration policy, and public opinion (emigration plans, migrant acceptance, public opinion on immigration levels, attitudes towards increasing diversity). The source data for the sub-indicators are clearly marked (UNHCR, OECD, UN DESA, etc). On the right there is an overview tab that gives overall trends. Below the map you can see key migration statistics both globally and by country.

2. The *Thematic Overviews* break down migration by theme and recommend data sources and further reading around specific topics in immigration and emigration statistics, types of migration, migration and vulnerability, migration and development and migration policy.

3. The *Tools and Guidance Materials* are a searchable database of best practices and practical approaches to help policy and data collection around migration. It can be sorted by year of publication, topics (ex: age, big data, children, education, etc), data source (census, survey, administrative, innovative), publisher (international organization, government, academia, other), types of information (analysis, guidelines/toolkits/manuals, journal article, report, etc), region of focus.

4. The *SDGs and Global Migration Compact* tab describes the GCM and outlines how migration is linked to various SDG goals.

5. The S*napshots* show key migration statistics by country and link to the Migration Governance Framework, a tool to help various countries assess their migration policies along a range of indicators.

6. The *Blog* tab holds articles and video interviews around migration.

---

[5] https://displacement.iom.int

[6] https://migration.iom.int/europe?type=arrivals&_ga=2.61360428.1481468315.1543256538-631066569.1534964565
[7] https://flow.iom.int/europe?type=arrivals

2   ONSET MODELLING

The **GDELT Database** tracks events and the patterns around those events. Supported by Google Jigsaw, it monitors the world's broadcast, print, web news and social media in over 100 languages every day. The team is working on expanding GDELT's coverage to the year 1800. The information is stored in the *Event Database*, which updates every 15 minutes. There is also a *Global Knowledge Graph (GKG)* that records every person, organization, company, location and emotion mined from the data and constructs a graph that shows context, key actors and feelings around particular events, updated every day. They also have specialized GKG collections that focus on specialized sources or topics. They publish a range of reports such as the Daily Trend Report and the World Leaders Index that ranks heads of state daily depending on the average tone of global news coverage. The data is free and can be explored and downloaded in a range of ways. The *GDELT Analysis* service is a cloud-based service with 14 different visualization tools. *Google BigQuery* allows for real time queries and analysis directly on the database, significant for a dataset of this magnitude. There is also the possibility to download the data in CSV format, though this is for advanced users since the databases are enormous. There is a blog that is a one-stop repository for the latest news, announcements, information and GDELT applications. GDELT posits itself as a solution-oriented tool and outlines its efficacy for gaining increased situational awareness, expanding influencer networks, producing risk assessment and global trends, assessing policy reaction and improve crisis response.[8]

Integrated Crisis Early Warning System (ICEWS) Dataverse contains 5 datasets for the Defense Advanced Research Projects Agency (DARPA) and Office of Naval Research (ONR) that track events and actors across time. The *Automated Daily Event Data* and the *Coded Event Data* show coded interactions between socio-political actors based on news articles. The *Event Aggregations* produces time series data out of the Coded Event Data, typically at a monthly level. Data parameters specify how events are filtered (e.g., country affiliation, sector affiliation, etc.) and then how they are aggregated (e.g., via count, intensity average) for a given time interval. The *Events of Interest Ground Truth Data Set* tracks the year, month and time that specific events (domestic political crisis, insurgency, international crisis, rebellion and ethnic/religious violence) occur by country. The *Dictionaries* archive actors (named individuals/groups) and agents (generic individuals/groups) and time-dependent affiliations they may have with other actors and agents. Their aliases are also recorded as well as alternate spellings.[9]

The **Armed Conflict Location & Event Data Project (ACLED)** is a mapping project focused on collecting and analyzing data around conflict and crisis. It is event-focused data. There are downloadable datasets for three regions: Africa, Middle East, South and Southeast Asia. The data can be sorted and exported based on time period, event type, actor type, actor, region, country, location and keyword. Events are sorted as: battles, headquarters/base established, non-violent transfer of territory, remote violence, riots/protests, strategic development, and violence against civilians.[10]

Uppsala Conflict Data Program (UCDP)/Peace Research Institute Oslo (PRIO) Conflict Dataset is compiled data for armed conflicts from 1946 until the present day. The *Uppsala Conflict Data Program pages* visualize the newest versions of the dataset through an interactive global map. The "countries in conflict view" shows which countries experienced which types of violence during what time periods. One can sort by time frame (1975-2017) and type of violence (state-based violence/non-state violence/one sided violence) "Fatalities view" allows

---

[8] https://www.gdeltproject.org

[9] https://dataverse.harvard.edu/dataverse/icews

[10] https://www.acleddata.com/data/

for the adjustment of time frame (1989-2017) to see where fatalities were experienced when. Specific actors and conflicts can be sorted for in both views. Clicking on an actor links to more information on that actor.[11]

Correlates of War Project's **Militarized Interstate Disputes** (v4.2) is a dataset about the various facets of war, whether political and trade relations between countries or material accumulation of a country. There are 14 different downloadable data sets: country codes, state system membership (tracks statehood over time), war data (1816-2007, both state and non-state entities), militarized interstate disputes, militarized interstate disputes locations, national material capabilities ("annual values for total population, urban population, iron and steel production, energy consumption, military personnel, and military expenditure of all state members, currently from 1816-2012"), world religion data, formal alliances, territorial change, direct contiguity (registers land and sea borders of all states since Congress of Vienna), colonial/dependency contiguity, intergovernmental organizations, diplomatic exchange, and trade.[12]

Political Instability Task Force Worldwide Atrocities Dataset records the fatalities of non-combatant civilians in wider political contexts. The downloadable data shows date, location, event type/who reported it, latitude and longitude, perpetrator, victims, casualties, modes of violence, description, and data source for each fatality. The current version of the dataset covers three files. They cover January 1995 to December 2012, January 2013 to December 2015, and January 2016 to January 2018. There is currently a complication in the Excel files where some of the ranges for the deaths and injuries are converted to dates.[13]

Fearon and Laitin are zipped files that contain replication data for James D. Fearon and David D. Laitin's "Ethnicity, Insurgency, and Civil War," American Political Science Review 97, 1 (March 2003): 75-90. repdata.zip can replicate Table 1 of their article, which describes different estimation methods for civil war through per capita income, population, % mountains, non-contiguous state, oil, new state, political instability, democracy, ethnic fractionalization and religious fractionalization. mkreptable1.co reproduces the five regressions from Table 1. egroupsrepdata.zip contains the list of ethnic groups and the variables for ethnic and cultural fractionalization by country discussed in James D. Fearon, "Ethnic and Cultural Diversity by Country," Journal of Economic Growth 8, 2 (June 2003): 195-222.[14]

3   HUMANITARIAN ASSISTANCE AND FINANCE

UN OCHA's **Financial Tracking Service** is a centralized source cataloguing humanitarian funding flows. It covers affected countries, donors, and appeals/plans.

The homepage shows an interactive map showing where funds are from, where funds are going and progress on appeals from 1980 until 2018. The *Global Overview* spans 1999 to 2018, and there is a summary and a data tab. The summary has graphs showing total reported funding by affected country, funding from top 5 donors, reported funding by sector, and funding trends inside/outside a response plan/appeal for any given year. The data tab sorts based on country, donor, donor type, recipient, recipient type, and sector. In addition to the global data, one can search HA flows by country. The summary tab for any given country gives a range of graphs such as trends in reported funding over time, funding by source, funding by sector, largest UN recipients, etc.

---

[11] https://www.prio.org/Data/Armed-Conflict/UCDP-PRIO/

[12] http://www.correlatesofwar.org/data-sets/MIDs

[13] http://eventdata.parusanalytics.com/data.dir/atrocities.html

[14] http://web.stanford.edu/group/ethnic/publicdata/publicdata.html

The data tab for each country is presented by year. One can sort the data by individual flows, donor, donor type, recipient, recipient type and sector. The table generated shows flow ID, source org, destination org, description, sector, amount (USD), and funding status. It allows further personalization of the table through other metrics such as destination/source emergency, flow date, original amount, reporting organization, among others. The *Appeals/Plans Overview* gives visualization of funding requirements, funding progress by appeal, largest sources of funding, CERF allocations, etc, by year. The dataset can then be sorted by appeals/plans, sector, donor, donor type, recipients, appealing organization, recipient types and appealing organization type. The table shows appeal/plan, plan type, original requirements, current requirements, funding (USD), percent covered, and pledges (USD). FTS allows the user to delve further into HA flows by donor by exploring summary information (snapshot for a given year, trends in reported funding, funding for affected countries, funding by sector, funding to recipient organizations) and data sorted by individual flows, emergency, country, plan/appeal, recipient, recipient type, sector.[15]

ODA Humanitarian Assistance Database tracks and measures humanitarian aid assistance flows, focusing primarily on ODA. The information can be seen at a glance by recipient or by donor from 2007-2015. Various bar graphs are then generated showing trends, aid by sector, donor, recipient and instrument (ODA grants, ODA loans, Other official flows). The *Query Wizard for International Development Statistics (QWIDS)* has more complete databases showing humanitarian assistance by DAC donor (2005-2016) and by sub-sector (2005-2016), which is sortable. There is also an interactive map that allows spatial comparisons by donor and recipient view, as well as a sector.[16]

Development Initiatives' **Development Data Hub** is a source for financial resource flow data that takes poverty, social and vulnerability indicators into account. On the homepage there is a global overview through eight interactive maps showing poverty, vulnerability, government finance, international finance, international official finance, humanitarian finance, data revolution and forward-looking ODA (UK budgets) across time and space. The information is drawn from 50 smaller downloadable datasets. There are individual country profiles to explore, as well as five multilateral profiles. There are visualizations that show funding priorities for ODA and other official flows (OOFs). The Development Data Hub has a spotlight on Kenya and Uganda showing poverty and vulnerability, population, education, health, water and sanitation, and public resources.[17]

---

[15] https://fts.unocha.org/

[16] https://www.oecd.org/dac/stats/humanitarian-assistance.htm

[17] http://data.devinit.org