Fall 2021

The Labyrinth of Data Collection for Humanitarian Project Funding and Implementation

Maria Alejandra Pulido
SIT Study Abroad

The Labyrinth of Data Collection for Humanitarian Project Funding and Implementation

By Maria Alejandra Pulido

December 2021 (Fall Semester 2021)

Switzerland: Multilateral Diplomacy and International Studies
Heikki Mattila
Gyula Csurgai

Hamilton College
World Politics Major with a concentration in migration between MENA and Europe
Double minor in French and Hispanic Studies

Acknowledgements

The completion of this research project would not have been possible without the help of all of my SIT advisors and the Academic Director. I would also like to thank all the experts who spoke with me and for all the feedback they gave me. I could not have done this without the help of my family who always supports me in all my endeavours. I would also like to thank all of the NGOs I examined, for the essential role they play in humanitarian assistance. I am a firm believer that they make the humanitarian sector a better place.

Abstract

My research concentrates on four NGOs: IOM, IDMC, JIPS, and OCHA which use different tools to collect data and translate the information into evidence for data-driven decision making (DDDM) for the implementation of humanitarian assistance projects. I focus on the importance, advantages, and various data collection tools which help ameliorate the humanitarian sector since it does not have a current professionalized path to enter the workforce. I incorporated four interviews, attended two conferences and analyzed multiple online sources during my project.

Table of Contents

The Labyrinth of Data Collection for Migrants, Refugees and Displaced People

In the past decade, as a student and adolescent I ventured into the humanitarian sector through working and volunteering opportunities, yet each organization had different approaches to effectively manage various forms of assistance for migrant and refugee populations. The twenty-first century has billions of people in need of food, lodging, and healthcare; however global acquisition to these basic needs varies upon region. Many countries do not have the infrastructure, resources, programs, and services which are necessary to meet every human's basic needs. When countries in the Global South face conflict, food and health insecurity amongst other factors which drive migration, citizens leave the comfort of their home to search for better lives in the Global North, neighboring countries, or in different regions within their own country. This transnational movement is sometimes fueled by emergencies and conflict, but it can also happen within national boundaries. Furthermore, the biggest numbers of displaced people are internally displaced people (IDPs) forced to move within their own countries (Baal & Ronkainen 2017). IDPs make up two-thirds of the total people displaced in the world (Alexander & Parker 2020). The amount of protection that IDPs will be allotted is dependent on the amount of people displaced within the country and past services or data collected to continue providing adequate services.

It is fundamental that international organizations provide assistance to migrants, refugees, and internally displaced people regardless of nationality, gender, sexual orientation, education, class, or religion. In order for organizations to provide adequate services and programs that qualify as humanitarian assistance, it is essential that they have ethical, effective, concise, and rapid ways of collecting short term and long term data on the conditions of migrants, refugees,

and IDPs. Gathering data on IDPs remains a quandary (Baal & Ronkainen 2017). How is data collected by Non Government Organizations (NGOs) for the financing of projects or humanitarian assistance for internally displaced populations? Data is input onto an NGOs spreadsheet which uses a wide variety of tools which differ by organization, and it is financed and coordinated amongst multiparty actors with the goal of a timely response to emergencies. Trustworthy data on migrants, refugees and IDPs is essential to creating confidence in the funds allocated towards humanitarian efforts by multi-party actors. Transparent and reliable information will ensure a successful improvement in the humanitarian assistance provided, consequently increasing standards of living of the receiving population and an improvement in how humanitarian assistance is created, mobilized and delivered to populations in need.

Research Methodology: Goals & Objectives

My research will focus on the different methods, platforms, and tools used to collect data on internally displaced peoples (IDPs). I will only be looking at data collection for humanitarian assistance rather than for development which usually tends to be more long-term. I will not be looking at federal forms of data collection for assistance given to migrants, refugees, and IDPs by country governments. Instead, the specific type of data collection I will be exploring is that provided by the international NGOs: the International Organization for Migration (IOM), the Internal Displacement Monitoring Centre (IDMC), the Joint IDP Profiling Service (JIPS), and the United Nations Office for the Coordination of Humanitarian Affairs (OCHA). My research will incorporate primary data in the form of interviews from a cluster sample of experts chosen according to the relevance of their work's intersectionality with my topic. I looked for experts on data collection within the mentioned NGOs and conducted three formal interviews and one informal interview. Other primary sources include historic books. Secondary sources include

official reports that interpret primary data collected on the methodologies of data collection. My goal is to be able to understand how data is collected, protected and translated into effective projects by NGOs. I want to understand what type of information is fundamental when applying for NGO project funding for humanitarian assistance for migrants, refugees, and displaced people.

I used a variety of qualitative methods to collect data. I relied mostly on quantitative methods derived from interviews such as descriptions of different data collection matrices and platforms. However, I also used certain quantitative methods to include datatables on a small sample of tools and data collection platforms and matrices are available to the IOM, IDMC, JIPS, and OCHA. The ethical considerations I exercised before interviews included letting participants know that everything said in the interview could be used in my research, unless it was explicitly stated as confidential, letting them know that they could be anonymous if they wished to, and that they were free to stop the interview at any point. Participants were also allowed to skip any questions that they did not wish to answer. Making sure there is an ethical approach to the interviews was of extremely high priority since by speaking with experts who may have confidential information I wanted to build trust so they felt comfortable and safe when sharing their information with me.

I expect to find practical information that I can use in my future research, hopefully traveling to deeply explore migration linked topics while helping vulnerable migrant populations. This research will identify tools and mechanisms that researchers can use to conduct their own data collection. Identifying these tools is fundamental to creating transparent, reliable, and honest information on migration which is a topic on the minds of citizens, policy-makers, and

employers. Demystifying the fake news that surrounds migration is essential for the creation of a global general public which has amicable views to different cultures and incoming migrants.

Definitions & Analytical Theoretical Frameworks

The following definitions and theoretical frameworks will assist in further understanding my research. “Humanitarian assistance is defined here as inclusive of humanitarian relief or emergency aid”, however I will not be incorporating development assistance as humanitarian assistance since I will be focusing on emergency assistance (Monico et.al 2014). Internally Displaced People (IDPs) are defined here as persons who are seeking refuge somewhere outside of their home inside the borders of their birth countries. Migrants are defined as internationally displaced people who move into a country outside of which they were born, or exit their country of origin. Refugees are defined as migrants who move to a country outside their birth land due to war, persecution, or natural disasters. The displaced are defined as IDPs, migrants, and refugees. It must be noted that due to the lack of agreed upon definitions by multiparty actors these definitions are my own. Data collection will only incorporate raw data collected directly and on the ground in contact with the displaced. Secondary data such as graphs and maps will be considered metadata which is data that gives information about other data. An aggregator is a program where related items can be linked to each other (Oxford Languages 2021). These definitions are also not universal and should be subject to comparison of other definitions. The state of the humanitarian system (SOHS) is an independent study that compiles the latest statistics on the size, shape, and scope of the humanitarian system and assesses overall performance and progress (ALNAP 2021). This paper will take a teleological approach to humanitarian projects by highlighting the importance and purpose of data for project implementation rather than focusing on any migration drivers.

The Beginnings of Humanitarian Assistance & the Rise of Digital Data Collection

International organizations based in Geneva, Switzerland export a lot of humanitarian assistance, perhaps because of the field formally originating in the country in the late 19th century. Before the 1990s the Geneva Conventions were created in the aftermath of World War II, the Treaty of Versailles introduced a system of international aid, and proper conduct during war was outlined by the Greeks and Romans, and by Chinese General Sun Tzu in *The Art of War* (Rysaback-Smith 2015). Humanitarian assistance grew into a collective societal understanding that people needed guidelines and formal methods of helping each other under various circumstances. This enlightened understanding would never have been reached without the experiences of the creators of international humanitarian organizations. Former SIT student and employee for the International Committee of the Red Cross (ICRC) Elizabeth Rushing explained how international humanitarian law (IHL) originated in 1859 with Swiss businessman Henry Dunant who saw the chaos and blood of the aftermath from the Battle of Solferino (Rushing 2021). Dunant introduced the idea of "Voluntary Relief Societies" which inspired the NGOs which we know today, he asks "in an age when we hear so much of progress and civilization, is it not a matter of urgency, since unhappily we cannot always avoid wars, to press forward in human and truly civilized spirit the attempt to prevent, or at least to alleviate, the horrors of war" (Dunant 1862). The current neoliberal era creates massive gaps between the rich and the poor, which exacerbates migration on a global scale, therefore it is the collective responsibility of countries in the Global North and South to meet basic human needs of all members in society regardless of citizenship status.

Due to the current technological era, data collection is also gathered through satellite data, telecommunications services, and artificial intelligence (AI). The digitalization of data is used alongside more traditional data collection methods such as on the ground surveys, administered registry sources, and interview questionnaires (HDMP 2021). This digitalization of data collection complements the traditional data collection methods in order to create better preparedness and ways of measuring the uncertainties of migration.

The digitalization of data collection became especially important during the COVID-19 pandemic since new data could not be collected on the ground through traditional methods. One of the panelists at the HDMP Conference, Marzia Rango spoke of a new report called the "Sentiment towards migration after COVID-19" which used Twitter data to measure xenophobic sentiments before and after the pandemic (HDMP 2021). Measures like these were important in continuing to understand migration factors during the COVID-19 lockdowns. Even if they were not conducted to enact programs they help analysts and policy makers to get a clear view of the discrimantion migrant communities deal with, which may affect future context.

The Advantages & Importance of Data Collection

Data is necessary in order to fill in the gaps in understanding that are related to migration and displacement. Technology has granted the humanitarian sector great advantages for effective and timely data collection that can be adequately analyzed for policy making and program response. Codes and algorithms will never replace traditional collection methods by humanitarians but "algorithmic humanitarianism" can be beneficial when it works alongside traditional methods, since this tool avoids the "trap of data for data's sake, by making use of digital technologies in a sounder way to reflect our core human values (Meneghetti 2018). Data helps in understanding the needs of the displaced, but it should be collected systematically and

methodologically in order to make a positive impact on the communities. Without much needed data it would be impossible to implement the humanitarian programs.

Data collection is important because it is needed to provide necessary projects and financial proposals from humanitarian organizations. Two-thirds of humanitarian funding goes through UN agencies, and the whole sector received 12 times the amount of financial funds than it did 20 years ago (Alexander & Parker 2020). This means that data collection must be done in a systematic and organized manner so the information created can be reliable, and accessible by participants and those creating and funding the programs. The growth of funding means there are more people being helped and being employed in the humanitarian sector increasing its importance since more lives are involved in what could sometimes be risky operations. "Furthermore, millions of refugees and IDPs will be unable to return home for months or even years. Without investing in better data, we cannot move decisively from short-term responses to humanitarian emergencies to longer-term development support for displaced people and their hosts alike" (World Economic Forum 2020). Emergencies have become more normal in our world but development projects are needed to better the lives of people in the long run since emergency assistance does not usually involve long-term measures but rather immediate help. However, since development plans tend to be longer, there is a larger financial burden, and risk for those actors funding development projects.

Even though there are always risks in the humanitarian field, some of the benefits of today's data collection within the sector are its gearing towards problem-solving values and quick action. This is possible due to data driven decision making (DDDM) and a shift in the population of the humanitarian workforce. DDDM has certain benefits: (1) more confident decisions; (2) becoming more proactive; and (3) realizing cost savings (Stobierski 2019). DDDM benefits

increase the reliability of information. In order to bridge knowledge gaps, and to change the organizational culture by bringing attention to hierarchical gaps perpetuated by traditional power dynamics, interpreter colleagues must work with data science counterparts (Alexander & Parker 2020). These collaborations increase the amount of individuals who will have access to displacement data.

Access to Reliable Displacement Data

More workers in the humanitarian field means more people have access to information on vulnerable populations and are prone to risks. “International humanitarian response to crises employs 210,000 people and accounts for nearly $15 billion in spending globally each year” (Walker et al. 2010). The creation of reliable data is essential for project implementation and closing the feedback loop. With more workers a chain of command is needed for quality assurance. At the IDMC statistics for certain disasters need to be given twice a year, and once all the information is compiled the top 50 major crises’ that occurred that semester are shuffled among experts, meanwhile another colleague examines how the analysis was conducted, and finally if everything has been done well the project coordinators sign off on it (Ponserre, personal communication, 2021). This type of random peer-review system is extremely beneficial to ensuring quality information is being produced. JIPS promotes ethical data standards which means they are responsible for feeding back to the community the data on what is going on so they can participate in formulating solutions (anonymous, personal communication, 2021). This brings the assistance closer to becoming sustainable.

The reliability of information can be affected by the uncertain, and sometimes unsafe situations that humanitarian workers put themselves in to deploy assistance and collect information. The humanitarian sector is a risky sector “2019 saw a record high for aid worker

casualties, particularly health personnel who accounted for 42% of all aid worker fatalities in 2019" (Alexander & Parker 2020). It must be noted that a risky sector can lead to certain biases especially since high stress situations are more easily prone to mistakes than lower-risk ones. This makes the protection of aid workers and vulnerable populations essential to having more trustworthy data.

Access to data will play an important part in breaking up negative migration stereotypes in the future. However, full access to information on vulnerable populations should not be granted to everyone since the interests behind acquisition of data may not always be amicable to participants on which the data is being collected. Data protection must go hand in hand with accessibility so that there are no ethical issues when dealing with sensitive information. At JIPS most information on sex, age, and disabilities (SAAD) is kept anonymous to protect participants and because household data is usually easier to acquire than individual data (anonymous, personal communication 2021). Data collection without protection is an ethical issue especially when dealing with vulnerable populations such as migrants.

Ensuring the quality of data is even more important than allowing certain people to access the information. Truthful information is necessary to properly deploy resources and enact projects. Reliability is achieved through triangulation (IDMC 2021). At the IDMC analysis and triangulation looks like a master file with all the facts collected attached to specific aggregators along with the analysis techniques that were used (Ponserre, personal communication 2021). Aggregators will facilitate the linking of data that is connected to create a more comprehensible analysis. At JIPS there is a data quality supervisor who reports to the Profiling Coordinator and is responsible for supervising or conducting the upload of all questionnaires administered on the mobile phones/tablets. Peer reviewing is also fundamental to ensure the quality of data and to

present it as reliable when encountering policy makers. Different organizations have different methodologies and mandates, yet the quality of data must be sublime regardless of the methodology.

Organizations and Experts Specializing in Data Collection for Displaced people

Table 1 IOM Data collection platforms:

Name of Data Collection Platform or Matrix	Description of type of data collected
Migration Governance Indicators (MGI)	Qualitative data on migration governance
Global Migration Analysis Centre (GMDAC) Missing Migrants Project	Migrants who have passed away or gone missing on migratory routes worldwide
Migrant Assistance Division (MPA); The Counter-Trafficking Data Collaborative	Victims of human trafficking, anonymizes personal data
Department of Operations and Emergencies (DOE)	Statistical data on refugee resettlement and those granted protection by the IOM
Displacement Tracking Matrix (DTM); the flow monitoring system	Internal and external displacement due to conflict or natural disasters, also provides estimates of irregular migration flows in certain locations
Assisted Voluntary return and reintegration (AVRR) collected by Migrant Protection and Assistance Division (MPA)	case data on migrants in vulnerable situations (unaccompanied migrant children, victims of trafficking and migrants with health-related needs), not publicly available.
Migration Environment and Climate Change (MECC) on the Environmental Migration Portal	qualitative research and data collection on the interlinkages between environmental change and human mobility
Migration Health Division (MHD)	Physical and mental health of migrants, not publicly available

Source of information: IOM website

Different organizations collect data on IDPs, the ones I concentrated on are the IOM, the IDMC, and JIPS, and OCHA since they are in charge of the financing for large amounts of

humanitarian assistance. The IOM has eight data collection platforms dealing with migration, shown by Table 1. Since the IOM has so many platforms and matrixes for various purposes that deal with migration, I focused on the Digital Tracking Matrix (DTM) and was able to interview an employee there. DTM was set up in Iraq in 2004 as part of an emergency response tied with management structures and humanitarian bodies responsible for overseeing all activities in camps (anonymous, personal communication 2021). JIPS was set up in 2009 as an inter-agency body to provide support to governments and humanitarian and development organizations seeking to improve locally owned information and analysis about displacement situations (JIPS 2021). Profiling is essential for strategic decision-making, securing funds and designing effective policies and programmes. The IDMC was established in 1998 by the Norwegian Refugee Council in Geneva focused on monitoring and providing reliable data and analysis on internal displacement (IDMC 2021). OCHA was created in 1991 as part of the UN Secretariat "to ensure better preparation for, as well as rapid coherent response to, natural disasters and other emergencies" (OCHA 2021). All of these organizations serve a specific humanitarian purpose and help each other through collaboration.

I interviewed a DTM coordinator who wishes to be kept anonymous therefore I will not be using their name in this study; they are a point person for academics dealing with partnerships and data. DTM is active in 86 countries and they produce a huge number of analytic outputs such as reports, research, and maps (anonymous, personal communication 2021). Even though DTM produces secondary data the matrix itself is used as a primary source of data on displacement in order to be a primary tool for displacement to program adequate responses. DTM focuses on conceptualizing internal displacement and cross border flows of displacement alongside risks and needs (DTM 2021). Looking at risks and needs alongside each other is essential to grasping

the bigger picture of the humanitarian situation one is analysing or assisting. “DTM is a primary tool on displacement, used for programming responses by collecting data on IDPs or refugees from abroad, operating in complex settings and transit points, such as locations of high traffic like bus stations and border crossing points” (anonymous, personal communication 2021). However, one of the weaknesses of DTM, mentioned by HDMP Conference participant and DTM employee Eduardo, is that DTM only collects data when they need the data but since they deal with vulnerable populations on the ground they will not ask any questions that may cause harm. Other questions that arise with so much data is how to use it and create a balance between a laissez-faire system and intervention.

I had an interview with JIPS regional advisor for the America’s Andres Lizcano Rodriguez. JIPs mandate is “to bring governments and international and local actors together to jointly generate and use evidence that supports displaced and host communities to reach durable solutions” (JIPS 2021). Lizcano explained to me the financialization report on the state of the humanitarian system (Lizcano 2021). I dug more into the financialization report and found that “beyond first level recipients, there is no knowledge about how humanitarian assistance reaches subsequent recipients because it is not adequately reported” (Global Humanitarian Assistance Report 2020). This lack of knowledge means that even with all the data available in our digital world we have not filtered it thoroughly enough to believe in the data’s reliability.

My third interview was with Sylvain Ponserre, who is in charge of data analysis and risk management at the IDMC. He joined the team in 2017 after working in the UN for 15 years. Ponserre’s work consists of helping countries to have the methodology needed to support monitoring for disasters. The IDMC calculates displacement risk through the Global Risk Model where risk equals hazard multiplied by exposure and vulnerability; they use a humane lens rather

than trying to put economic value to individual risk (Ponserre, personal communication 2021). The IDMC produces metadata which is explained to decision makers in timely manners so DDDM can occur allowing for projects to be implemented. There is no threshold for monitoring, since the IDMC focuses on monitoring anyone displaced regardless of being one individual or a family unit. All of these organizations do essential work which differs in its nature.

Differences in Data Collection and Project Implementation by Organization

Humanitarian organizations vary in tools and methodologies incorporated for the collection of information and implementing evidence into assistance projects. Organizations build on each other, for example DTM is used by the IDMC even if it is a tool of the IOM. The IDMC is not a primary data collector; they do not have any staff on the ground, they have surveys to get qualitative info but they mainly rely on IOM, OCHA, and the government systems which collect data (Ponserre, personal communication 2021). Each organization plays an essential role in bringing together project implementation which could not be done without collaboration.

Table 2: Tools used by NGOs for Data Collection.

NGO	Tools used (not all are listed only the ones analyzed in this study)
IOM	DTM, HXL
IDMC	IDETECT, aggregated social media data, the Global Risk Model, HXL
JIPS	JET, PIM, Secondary Data Review Matrix, Operational Plan on Data Collection
OCHA	FTS, IATI, satellite imaging (UNOSAT), the Information Management Toolbox, HXL

Source: Websites of IOM, IDMC, JIPS, and OCHA

I examined various tools used by NGOs. The tools that I examined by the IOM were DTM which produces metadata reports through a platform for inputting raw data collected on the

ground. DTM has different reports on Mobility Tracking, Flow Monitoring, Surveys linked with migration flows/ displacement solutions/ return intentions/ community perceptions, as well as Registration and region specific data (DTM 2021). The tools that I examined for JIPs were the JIPS Essential Toolkit (JET), Protection Information Management Matrix (PIM), Secondary Review Matrix, and the Operational Plan on Data Collection (JIPS 2021). For the IDMC I explored IDETEC (IDMC 2021). For OCHA I examined Financial Tracking Service (FTS), the International Aid Transparency Initiative (IATI), satellite imaging (UNOSAT), and the Information Management Toolbox. Table 2 shows the different tools used by these organizations, including HLX which I will delve into further in the study.

JIPS uses a variety of tools compiled in JET, PIM, the Secondary Review Matrix and the Operational Plan on Data Collection. JET has different tools which help with: initiating and profiling, establishing the collaborative process, designing the methodology, implementing data collection, processing and analyzing the data, and validating, reporting and disseminating (JET 2021). Designing the methodology and process of the data collection is of utmost importance because it will reflect all of the initiatives that must be taken in order to collect information in ethical and reliable ways which are organized in a manner which facilitates inter-agency dialogue and cooperation. The Protection Information Management Matrix (PIM) provides definitions to relevant methodology so information can be clearly interpreted (JIPs 2021). The Secondary Review Matrix is used "to document and keep track of all relevant data available in one place so that it can be used in the analysis phase for the profiling process (for example to contextualize or compare other findings), to highlight what already exists in the context in time to modify methodology and avoid duplication, to inform the methodology overall by making the topics and indicators included in any primary data collection more relevant for the context (JET 2021). This

Matrix builds on the primary data collected by DTM and other data collection platforms ensuring that agencies need to be able to circulate trust-worthy information amongst each other to step up whatever type of assistance will be given. Finally the Operational Plan on Data Collection explains how the data collection will happen with specific time frames, an organizational structure, roles and responsibilities for those involved in the mission, and a work plan which can later be monitored for logistics (JET 2021). The plan is split into a two-month period because data collection is rather quick without disturbances, but training for at least one week and up to three to four will produce better data (anonymous personal communication 2021). The plan is one of the most essential parts of the whole process because without a good plan there is no good outcome in this sector.

The IDMC uses The Internal Displacement Event Tagging Extraction and Clustering Tool (IDETECT) along with aggregated social media data, satellite imagery, and the Global Risk Model. IDETECT showcases how beneficial technology is to the data collecting effort. The automated program reads world news, UN and NGO reports, filtering for data on internal displacement only, and extracting key bits of information including type of displacement, cause, location, number of people, and reporting units and terms (IDMC 2021). This tool is essential in creating metadata which can be presented to decision makers so actions can be taken to assist in displacement emergencies. Since the IDMC does not have people in the field collecting raw data they rely on other tools to enact humanitarian assistance, one of these tools would be the IOM's DTM. Aggregated social media data became especially useful during the COVID-19 pandemic since it allowed for data to still be collected even during the lockdown period where data collectors could not have access to the participants on which the data would be collected. The IDMC also incorporates a Global Displacement Risk Model into their Global Internal

Displacement Database (GIDD). The Global Displacement Risk Model "is a tool for visualizing disaster-related displacement risk metrics such as how many people are likely to be displaced per country per year or over five- or ten year periods" (IDMC 2021). Anticipating disaster is key to deploying humanitarian resources in a timely manner so DDDM can occur adequately. Calculating risk is extremely important because with so many emergencies increasing humanitarian organizations need to allocate resources effectively so those who are the most vulnerable are insured to get the assistance they need.

Even though there are some similarities, there are key methodological differences between all of these organizations. DTM only collects data that they will be able to transfer onto action since they are dealing with vulnerable populations they "do not ask a person in need questions that won't be transformed into action" (HDMP 2021). This varies from the IDMC which does not collect raw data and compiles metadata, usually for decision makers. Furthermore, the IDMC uses the likelihood of housing becoming severely damaged or destroyed as a proxy for displacement to deploy resources to people between 24 to 48 hours (Ponserre, personal communication 2021). OCHA is probably the most different from all the other organizations in the study, especially because it allocates funds for projects that are then performed by NGOs or OCHA itself.

Table 3: OCHA's Information Management (IM) Toolbox:

Section	Information and more tools available by section
IM Overview	Principles of Humanitarian IM, Guidance Documents, IM Process in 6 steps: Plan, Collect, Process, Analyze, Communicate, and Feedback.
OCHA IM unit	Management of functions, associated processes, strategy (office performance dialogues), work plan, and staffing and capacity development.
Common Operational	CODs are authoritative reference datasets needed to support operations and decision-making for all actors in humanitarian response, there are two

Datasets (CODs)	types: (1) Core CODS which are COD-AB (administrative boundaries), COD-PS (population statistics), COD-HP (Humanitarian Profile); and country specific codes (2) based on natural disaster CODs or Complex Emergency CODs. COD governance and responsibilities such as CODs in the first 48 Hours, Info management in Working Group Role, Guidance on Common Operational Datasets, COD Roles and Responsibilities, COD Governance Model, and the Inter-Cluster Coordination Group. P codes are unique geographic identification codes, there can only be one per administrative unit and the taxonomy (names and P-codes) should be available as a gazetteer which is a geographic dictionary.
Technology: Tools, Software, Platform	Data Collection Tools are: GPS, KoBo Toolbox (free open-source tool for mobile data collection) which guides in the designing of a form, collecting the data, and analysing it by exporting the data to Data Analyser or Excel. Data Processing is done through Excel where Power Query can be used to connect data across a wide variety of sources, and this is all done through HXL. Mapping resources are: Humanitarian OpenStreetMap, the Crisis Mappers Network, with ArcGIS map templates, Data Driven Pages template, Quantum Geographic information System (QGIS), Powerpoint, Satellite Imagery, and Unmanned Aerial Vehicles (UAVs). Visualization resources are Adobe Creative Cloud (Illustrator, InDesign, Photoshop), Business Intelligence, social media, and messaging apps amongst others.
IM Products	The most common products are: Coordination Products (Contact Lists, Meeting Calendar, Who does What Where (3W)- Product), Humanitarian Reports (Situation Reports- Sit/Rep, Humanitarian Bulletin -HB), Maps/Infographics (Humanitarian Snapshot, Humanitarian Dashboard, Maps, Funding Graphics), and HPC Products (Humanitarian Needs Overview, Humanitarian Response Plan, and Periodic Monitoring Report).
Coordination	OCHA's Information Management Working Group (IMWG) is the forum for strategic and technical discussions and collaboration on information issues related to humanitarian response and preparedness. Working groups will then utilize cluster reporting cycle and information flows, coordinated data scramble, humanitarian access, and cash and information management alongside a Global IM functional Team (GIFT).
Humanitarian Programme Cycle (HPC)	Emergency Response Preparedness, Risk Analysis, INFORM Subnational Risk Assessment, Assessment Design and Analysis, Prioritization Ranking, Assessment Registry or Survey of Surveys, Humanitarian Needs Overview (HNO), Needs Comparison Tool (NCT), Rapid Impact analysis, Severity analysis, Analysis of Workflow and Spectrum in a Humanitarian Setting, Strategic Planning, Humanitarian Response Plan (HRP), Indicator Registry, and Resource Mobilisation. Monitoring with Gap Analysis, and Who does What Where (3Ws). HPC Tools are: Response Planning and Monitoring Module (RPM),

	Humanitarian InSight (brings global data on needs, response and funding to a single interactive viewer, HPC Projects Module, and FTS)
Types of Humanitarian Responses	All humanitarian responses, regardless of the specifics, should have the basic IM products in place: Contact lists, meeting schedules, operational information sharing platforms, Core CODs, and 3Ws. In addition IMOs should be familiar with: Center for Humanitarian Data, Global Disaster Alert and Coordination System (GDACS), UNOSAT, UN Seamless Product Information, Data Exchange, and Repository (SPIDER), Current Emergencies (emergencies which require system-wide scale up response and corporate emergency response), and IASC Scale-up Responses. Actors should also understand complex emergencies, natural disasters, and infectious disease events

Source of Information: OCHA's Humanitarian Atlas 2021

OCHA uses FTS, IATI, satellite imaging (UNOSAT), and the Information Management Toolbox. "Each government donor has a page on FTS, which is constructed to display funding from the donor's perspective based on: the donor's source year, rather than the year the funding is used by a recipient organization" (OCHA 2021). It is important to know where the funds are coming from in order to ensure that there is transparency in how humanitarian efforts are funded. UNOSAT provides "UN funds, programmes, and specialized agencies with satellite analysis, training and capacity development, at their request" and Member States with satellite imagery analysis over their respective territories, and provide training and capacity development in the use of geospatial information technology" (UNITAR 2021). Satellite technologies come in handy especially in disaster and conflict zones where there are no means for conducting field work. Panelists at the HDMP spoke about using satellite imagery to count the number of cars crossing a border to get an estimate of the amount of people externally displaced from a country and or region. OCHA's Information Management Toolbox is "an online space for OCHA's information Management staff to access current and curated information tools, services, and systems to support humanitarian response and preparedness coordination" (Humanitarian Atlas 2021). This Toolbox is essential for outlining the response and preparedness work in different phases of the

project creation and implementation process. Table 3 looks at all the different features of the OCHA Information Management Tool. Place codes (P codes) are used in humanitarian preparedness and response by giving a unique ID to different places with the same name, P-code and names are used in data collection tools (ex:3W, RPM), they are used to amalgamate data from different sources, they are used as a dimension for analysis (geographic severity, priority, etc.), and they can be applied to other datasets so that their geographic location is known if it can be a spatial file (Humanitarian Atlas 2021). This hashtag coding makes part of a wider Humanitaria Language Exchange system that facilitates communication amongst humanitarian actors to facilitate the deployment of resources.

Similarities in Data Collection: HXL and International Data Collection Cooperation

Similarities between organizations range from the collection of metadata to interagency cooperation and the use of HXL. Humanitarian Exchange Language(HXL) is used within the humanitarian sector to facilitate interagency communication and faster data analysis for proper deployment of resources in emergency situations. It is a great idea, but in strategy not all NGOs use it, since it requires users to jump over a certain technical threshold and within the NGOs most employees never have the need for it (anonymous, personal communication 2021). The IDMC asks colleagues from the IOM and the UNHCR to review sensitive figures before publishing them to make sure their figures are as accurate as possible (Ponserre, personal communication 2021). Peer reviewing across agencies is needed to ensure the quality of the work across the humanitarian sector, and to grasp necessary information which could benefit a different government agency or NGO. Similarly, organizations like the IDMC train government officials on how to use DTM so this methodology can be applied on their own, for their own data collection (Ponserre, personal communication 2021). Having governments trained in DTM is an

advantage to staying on top of very politicized issues such as migration, and creating the appropriate ways and means for policy change that can benefit the displaced.

Interagency cooperation is sometimes facilitated by a global NGO network for principled and effective humanitarian action (ICVA), which serves to bridge the gap between humanitarian actors (ICVA 2021). Lead ICVA members on IDPs, the Norwegian Refugee Council, and UN agencies participate in the Senior Network on Internal Displacement, which helps with networking and collaboration within the humanitarian sector (ICVA 2021). Establishing dialogue between NGOs and humanitarian actors is extremely important so the quality of work, and clarity on fundamental differences in how the UN and NGOs view priorities can be reached. A good way to establish dialogue is by having an operational language like HLX taught by training.

HLX allows for faster and more effective interagency cooperation through the use of hashtags as markers within data collection sets that can be used by NGOs. HLX started in 2013 by OCHA and other partner organizations including the IOM and UNHCR with "aims to improve coordination across agencies responding in humanitarian crises, through a more efficient and effective system of collecting and sharing data" (Warner & Obrecht 2016). HLX started with the idea of linked open data where computers interpret hyperlinked data, but there were issues with implementing this process, therefore hashtags were created. HXL distinguishes individuals from households by adding "hh" to the existing hashtag, and Humanitarian Data Exchange (HDX) facilitates the sharing of data collected by NGOs through HXL and IATI (HXL 2021). This measure will professionalize the humanitarian sector since it is not fully standardized like the medical sector. HDX tools include Quick Charts, HXL Tag Assist, and Data Checks (humdata 2021). Another tool is how IATA aids OCHA in publishing humanitarian funding data which is supported by the Grand Bargain amongst donors and aid agencies.

The Grand Bargain

The Grand Bargain allows for donors and humanitarian organizations to improve the efficiency and effectiveness of the humanitarian sector. The “Grand Bargain 2.0" reframes the overall objective to achieving “Better humanitarian outcomes for affected populations through enhanced efficiency, effectiveness, and greater accountability, in the spirit of Quid pro Quo as relevant to all” (IASC 2021). Having a quid pro quo mentality incentivizes donors to provide the necessary funds. “Self appointed “champions” would take up specific actions from the Grand Bargain 2.0 framework and proactively and independently recruit other key stakeholders to work together in closed format” to allow for discussions which would “then be presented for further debate or adoption with the other Signatories” (IASC 2021). This system of directly reaching out to stakeholders allows for the creation of connections where participants are allowed to express their needs and concerns, consequently increasing the likelihood of success through clear communication. The pillars of the grand bargain are: (1) Flexibility, predictability, transparency, and tracking; (2) Equitable principled partnerships; (3) Accountability and inclusion; (4) Prioritisation and coordination (IASC 2021). These pillars showcase the necessity of having values especially when dealing with money for a cause since many rich actors do not always acquire their funds in ethical ways. Delineating the relationship between stakeholders and those humanitarian workers putting the projects of assistance together is essential for transparency and accountability.

Table 4: The 9 thematic work streams of the Grand Bargain

9 thematic work streams of the Grand Bargain
1.Greater Transparency
2. More support and funding tools to local and national responders

3. Increase the use and coordination of cash-based programming
4. Reduce Duplication and Management costs with periodic functional reviews
5. Improve Joint and Impartial Needs Assessments
6. A Participation Revolution: include people receiving aid in making the decisions which affect their lives
7.Increase collaborative humanitarian multi-year planning and funding
8. Reduce the earmarking of donor contributions
9. Harmonize and simplify reporting requirements

Source: Agenda for Humanity and 5 Core Responsibilities 24 Transformations

The Grand Bargain delineates specific ways and means for financing humanitarian projects. The Grand Bargain was first proposed by the former UN Secretary General's High-Level Panel on Humanitarian Financing in its report "Too important to Fail: addressing the humanitarian financing gap" (Agenda for Humanity 2020). Having a financing mechanism for humanitarian projects facilitates the work for those creating the projects after analysis of the evidence-turned data. The project builders can concentrate on the analysis and methodology for building a project as seen with JIPS, tools such as the Operational Plan on data collection mentioned previously. The Grand Bargain sets out 51 commitments distilled in 9 thematic work streams which are shown in Table 4. These thematic work streams delineate the most important factors that need to be considered for financing humanitarian projects. However, the Grand Bargain does not have any efforts to dismantle the negative media attention that gets politicized around displacement and migration. Even though financing the projects is the priority there should be a sub-body within the Grand Bargain which finances an effort to depoliticize migration and better its media image which is easily manipulated. This sub-body should not be a priority but it should at least exist.

Limitations associated with collecting data on vulnerable populations

With so much data needing to be analyzed, formatted, and shared the creation of professionalized pathways for the humanitarian sector have not been fully developed. Humanity creates more than 2.5 quintillion bytes of data each day, it's never been easier for businesses of all sizes to collect, analyze, and interpret data into real, actionable insights (Stobierski 2019). The large amount of data in our contemporary society takes more time to process, hindering how fast emergency humanitarian assistance measures can be enacted. Technology brings much data to our fingertips but AI is not perfect and so much technology could be a liability by deterring people from collecting data since it is meticulous even with technological assistance. Digital tools are used to collect information, and big data is already such a big part of our everyday lives, yet humanitarian AI could be prone to epistemic injustice. "The general threat of AI, in humanitarianism and elsewhere, is not the substitution of humans by machines but the computational extension of existing social automatism" (Open Democracy 2018). In other words social automatism, or the holes in the system where some benefit and others do not, would be perpetuated by calculations of data science that would count more than testimony from the subjects of the data. Taking the word of a computer over that of a human is bound to raise ethical concerns within societies, there must therefore be a balance between progress and ethics.

One of the limitations of data collection in the sector as a whole is the lack of specific SAAD information. This information may not be essential but it is still important in regards to gaining information about gender inequalities and how many unaccompanied minors are being displaced. SAAD is only in like two percent of the data input by the IDMC (Ponserre, personal communication, personal communication 2021). Acquiring this information can be essential to having more specific humanitarian measures that are customized for specific needs. However,

the high needs showcased in the humanitarian sector make it hard for workers and data collectors to focus on this when there are so many other essential needs that need to be met.

Having a humanitarian sector that is not professionalized is a liability, but thankfully tools such as HXL have facilitated the analysis of data across multiple agencies, when used. "With no professional body yet in existence, individual organizations have built their own competency frameworks and are providing in-house training, often in collaboration with academic institutions" (Walker et. al 2010). Since different organizations within the humanitarian sector perform a variety of services it would be difficult to have a global blueprint or fully standardized method for delivering humanitarian assistance. Having uncertainty over how to be part of the humanitarian process could deter many people from taking this career route. The closest tool in the professionalization of Humanitarian Assistance are master's-level programs which are universally recognized and the only current routes to international recognition (Humanitarian Atlas 2021). Creating set pathways of entering the humanitarian sector could benefit the sector as a whole since it could attract potential employees. But, since humanitarian assistance does not have a professionalized body or career path individuals must be certified through other ways. "Professional certification of individuals occurs through three general routes:a portfolio, or the collection of someone's experiences and work; competency, proven through either examination or documented experience; and curriculum, or passing courses" (Walker et al. 2010). In other words, experience and building a strong resume is the best foundation for working in the humanitarian sector.

Conclusion

The way that data is collected, analyzed, and financialized is a labyrinth that never seems to end. There are so many tools which vary by organization and each organization specializes in

different parts of the process. However, together different NGOs help make humanitarian assistance missions possible for migrants, refugees, and IDPs. More attention within the assistance and data collection sector needs to be paid to IDPs since they make up the largest percentage of displaced persons. DTM and JIPS have specific tools which assist with data collection, meanwhile the IDMC creates metadata and OCHA does all of the above while focusing on financing humanitarian projects. Together these organizations alongside many others make humanitarian projects a possibility. One of the tools I found to be most intriguing is HXL since it facilitates data collaboration and has made me want to learn more about how to use it myself. The use of technology such as satellite imaging and AI displays the benefits that technology can have in data collection. OCHA's IM toolbox is especially useful since it is very specific and provides so much information on how to properly organize and implement humanitarian projects.

My research was successful in showing a glance into the labyrinth of data collection for the implementation of humanitarian projects. However, since I only looked at 4 organizations, and didn't look at all the services provided by each organization there are limits to my research. I wish to venture more into the sector to be able to learn more about how to change the negative culture of news and politicization which surrounds migration. Bringing attention to the large numbers of IDPs could be beneficial in changing the negative xenophobic perspective that has come along with populism. If it is easier for individuals to identify with being displaced then maybe their perspectives towards foreigners will ameliorate and we will create a more tolerant society. Similarly, if there are set ways for people to enter into the humanitarian sector it will attract a larger workforce which can collaborate with all the challenges and meticulous details which need to be taken into account for a project to be successful. We have more than enough

tools to successfully create, finance, and enable humanitarian projects, success is a matter of navigating the labyrinth of tools and organizations which are at our fingertips. Just like data collection is at my fingertips everytime I embark on a humanitarian mission.

Abbreviations

AI-Artificial Intelligence

DDDM- Data-Driven Decision Making

CODs: Common Operational Datasets

FTS-Financial Tracking Service

GDACS- Global Disaster Alert and Coordination System

GIFT- Global IM functional Team

HDMP- Harnessing data innovation for migration policy (also HDIMP) in Europe and Africa Conference

HDX- Humanitarian Data Exchange

HNO- Humanitarian Needs Overview

HPC- Humanitarian Programme Cycle

HXL: Humanitarian Exchange Language

IASC- Inter-Agency Standing Committee

IATI- International Aid Transparency Initiative

ICRC- International Committee of the Red Cross

ICVA- International Council of Voluntary Agencies

IDMC- Internal Displacement Monitoring Center

IDETECT- The Internal Displacement Event Tagging Extraction and Clustering Tool

IDPs- Internally displaced persons

IHL- International Humanitarian Law

IM- Information Management

IMWG-Information Management Working Group

JIPS- Joint IDP Profiling Service

JET- JIPS Essential Toolkit

QGIS- Quantum Geographic Information System

NCT: Needs Comparison Tool

NGOs- nongovernmental organizations

OCHA- United Nations Office for the Coordination of Humanitarian Affairs

PIM- Protection Information Management Matrix

RPM- Response Planning and Monitoring Module

SAAD- sex, age, and disabilities

SOHS- State of the Humanitarian System

SPIDER- UN Seamless Product Information, Data Exchange, and Repository

UNITAR- United Nations Institute for Training and Research

UAVs- Unmanned Aerial Vehicles

3Ws- Who does What Where

Bibliography

Agenda for Humanity. (2020, February). *The Grand Bargain*. https://agendaforhumanity.org/initiatives/3861

Alexander, J., & Parker, B. (2020, September 9). *Change in the humanitarian sector, in numbers*. The New Humanitarian. Retrieved November 23, 2021, from https://www.thenewhumanitarian.org/maps-and-graphics/2020/09/09/25-years-of-humanitarian-data

ALNAP. (2018). *The State of the Humanitarian System | ALNAP*. ALNAP. https://www.alnap.org/our-topics/the-state-of-the-humanitarian-system

Colombia, Government of 2021 | Financial Tracking Service. (2021). Fts.Unocha. https://fts.unocha.org/donors/4839/summary/2021

Dunant, H. (1862). *A Memory of Solferino*. ICRC.

Gillsater, B. (2021, November 15). *Why we need more data on the state of the world's refugees*. World Bank -UNHCR Joint Data Center on Forced Displacement; World Economic Forum. Retrieved November 23, 2021, from https://www.weforum.org/agenda/2021/11/refugees-internally-displaced-people-data/

Humanitarian Atlas. (2021). *Data Collection: KoBoToolbox - IM Toolbox - Confluence*.

https://humanitarian.atlassian.net/wiki/spaces/imtoolbox/pages/42046660/Data+Collection+KoBoToolbox

Inter Agency Standing Committee. (2021, June). *The Grand Bargain 2.0*. The Grand Bargain 2.0 Endorsed Framework and Annexes. Retrieved December 1, 2021, from https://interagencystandingcommittee.org/system/files/2021-07/%28EN%29%20Grand%20Bargain%202.0%20Framework.pdf

International Council of Voluntary Agencies (ICVA). (2004, August). *International Council of Voluntary Agencies Annual Report 2003*. ICVA. https://www.icvanetwork.org/resources/icva-annual-report-2003

Internal Displacement Monitoring Centre. (2021). *How we monitor*. Harnessing Data Innovation for Migration Policy (HDIMP) in Europe and Africa. Retrieved December 5, 2021, from https://www.internal-displacement.org/monitoring-tools

International Organization for Migration & Internal Displacement Monitoring Centre. (2021, November 29). *Harnessing data innovation for migration policy (HDIMP) in Europe and Africa* [Virtual session]. Harnessing data innovation for migration policy (HDIMP) in Europe and Africa, Geneva, Switzerland.

International Organization for Migration, Internal Displacement Monitoring Centre, & Platform on Disaster Displacement. (2020, November 24). *Assessing the state of data and research on human mobility in the context of disasters, climate change and environmental degradation: Where are we and what comes next?* Disaster Displacement. Retrieved November 23, 2021, from https://disasterdisplacement.org/wp-content/uploads/2020/11/111120-DKWG-concept-note-marketplace-FINAL_compressed.pdf

International Organization for Migration (IOM). (2021). *Welcome to DTM | DTM*. Displacement Tracking Matrix. Retrieved November 23, 2021, from https://dtm.iom.int/

JIPS. (2021). *Transforming Data into Action for Solutions to Internal Displacement JIPS Strategy 2021–2023*. Retrieved November 23, 2021, from https://www.jips.org/uploads/2021/03/JIPS-Strategy-2021_23-vf.pdf

Lindskov Jacobsen, K., Bergtora Sandvik, K., & Martin McDonald, S. (2018, October 25). *Humanitarian experimentation*. Humanitarian Law & Policy Blog. Retrieved November 23, 2021, from https://blogs.icrc.org/law-and-policy/2017/11/28/humanitarian-experimentation/?utm_source=ICRC+Law+%26+Policy+Forum+Newsletter&utm_campaign=f38cc696d5-EMAIL_CAMPAIGN_2017_11_28&utm_medium=email&utm_term=0_8eeeebc66b-f38cc696d5-79046821&mc_cid=f38cc696d5&mc_eid=d97815e8e7

Manifesto on algorithmic humanitarianism. (2018, Apr 04). *OpenDemocracy,* Retrieved from https://www.proquest.com/magazines/manifesto-on-algorithmic-humanitarianism/docview/2021485463/se-2?accountid=11264

Meneghetti, L. (2020, February 7). *3 ways big data is changing the humanitarian sector*. World Economic Forum. Retrieved November 23, 2021, from https://www.weforum.org/agenda/2018/01/3-ways-big-data-is-changing-the-humanitarian-sector/

Monico, C., Smith Rotabi, K., & Abu Sarhan, T. (2014, April 9). *International Aid, Relief, and Humanitarian Assistance*. Encyclopedia of Social Work: National Association of Social Workers Press and Oxford University Press. Retrieved November 23, 2021, from

https://oxfordre.com/socialwork/view/10.1093/acrefore/9780199975839.001.0001/acrefore-9780199975839-e-957

Norwegian Refugee Council & Internal Displacement Monitoring Centre (IDMC). (2019, May). *Global Report on Internal Displacement*. IDMC. https://www.internal-displacement.org/sites/default/files/2019-IDMC-GRID-Summary.pdf#:~:text=IDMC%E2%80%99s%202019%20Global%20Report%20on%20Internal%20Displacement%20has,Foreign%20Affairs%20and%20Liechtenstein%E2%80%99s%20Ministry%20of%20Foreign%20Affairs

OCHA. (2021a). *IM Toolbox - Confluence*. IM Toolbox OCHA. https://humanitarian.atlassian.net/wiki/spaces/imtoolbox/overview

OCHA. (2021). *United Nations Satellite Centre UNOSAT | UNITAR*. UNITAR UNOSAT. https://unitar.org/sustainable-development-goals/united-nations-satellite-centre-UNOSAT

Oxford Languages. (2021). *aggregator definition - Search*. https://www.bing.com/search?q=aggregator+definition&qs=LT&pq=aggregator+d&sc=8-12&cvid=CDD2B2EEDDD948BE810BCFB1A34A9BCB&FORM=QBRE&sp=1&ghc=1.https://www.bing.com/search?q=aggregator+definition&qs=LT&pq=aggregator+d&sc=8-12&cvid=CDD2B2EEDDD948BE810BCFB1A34A9BCB&FORM=QBRE&sp=1&ghc=1

Rushing, E. (2021, September). *Why and How IHL Matters/ The ICRC 2021* [Guest SIT Lecture]. SIT Guest Lectures International Security, Peace and Stability Class, Geneva, Switzerland.

Rysaback-Smith, H. (2015). History and Principles of Humanitarian Action. *Turkish Journal of Emergency Medicine*, *15*(1), 5–7. https://doi.org/10.5505/1304.7361.2015.52207

Stobierski, T. (2019, August 26). *The Advantages of Data-Driven Decision-Making | HBS Online*. Business Insights - Harvard Business School. Retrieved November 23, 2021, from https://online.hbs.edu/blog/post/data-driven-decision-making

The Centre for Humanitarian Data. (2021). *Data Services – The Centre for Humanitarian Data*. The Centre for Humanitarian Data Data Services. https://centre.humdata.org/data-services/

United Nations High Commissioner on Refugees (UNHCR). (2021). *The Global Compact on Refugees*. The Global Compact on Refugees | Digital Platform. Retrieved November 23, 2021, from https://globalcompactrefugees.org/article/global-compact-refugees

UN Refugee Agency, Baal, N., & Ronkainen, L. (2017, January). *Obtaining representative data on IDPs: challenges and recommendations*. Https://Www.Unhcr.Org/598088104.Pdf. Retrieved November 30, 2021, from https://www.unhcr.org/598088104.pdf