

## I. Project Summary

The goal of the Humanitarian Data Exchange project is to make humanitarian data easily available and useful for decision-making. The project aims to bring together multi-country, multi-sourced, curated data for analytical use through one platform.

Data may not be the first thing that comes to mind when thinking about humanitarian crises like war, famine, earthquakes, and storms but having access to data is critical to answering background questions like: What is the country we're deploying to like in terms of population, geography, infrastructure, economic development and social structures? What resources – financial and human – are available to deploy? Once a relief effort is underway more questions arise: How many people have been displaced this week and where are they now? How many tents have been deployed and how many more are needed? How do we safely get trucks from point A to point B?

This project will create a place to quickly find curated data to answer these questions and others that arise throughout a crisis<sup>2</sup>. Over time, by creating a foundational data infrastructure for the humanitarian community, we will begin to address the quality, comparability and timeliness of humanitarian data. Decision makers will be able to make inferences based on historical data and the current data available, improving the likelihood of a successful response.

For example, following the 2010 Haiti earthquake response, the national mobile carrier Digicel gave anonymized call record data (historical and current) to an organization called Flowminder<sup>3</sup>. The group undertook analysis that showed where people would be moving following the earthquake. If the humanitarian community has regular access to this type of data in real time, responders could quickly deploy resources to where people are moving instead of waiting or guessing about population movements.<sup>4</sup>

This scenario will be normal practice one day. We can imagine that in ten years' time, real-time data will be available at a per-person level (rather than at the cluster or country level) in all formats, on every type of device. Humanitarians and affected communities will know what people need in a precise and accurate way, replacing the guess work that plagues aid delivery today. In a hundred years, the data that we store will still exist and its analysis will continue to inform future generations of humanitarians.

We believe the project presents an excellent opportunity to change the way humanitarians share, access and use data, with positive implications for people who need assistance. It is an ambitious undertaking. It involves engaging a wide range of stakeholders in something that is organizationally, technically and analytically complex. This is a large-scale innovation at a system level.

The following document covers the initial phase of the project, which will run for a twelve-month period and focus on being successful with data from four to five countries. The priority users of this data will be global and field-level information actors in the humanitarian sector, i.e. those charged with using data to

---

<sup>1</sup> This is an abridged version of the original project document that was written in October 2013. This version has been edited for length and clarity. It is being shared through the HDX site to provide background on the project's creation but should not be viewed as an official OCHA document.

<sup>2</sup> For a closer look at the scope of decision maker needs in response see the research by the Digital Humanitarian Network:

<sup>3</sup> <https://app.box.com/s/kneqlcpq99xlkh08w0d6>

<sup>4</sup> <http://www.flowminder.org/>

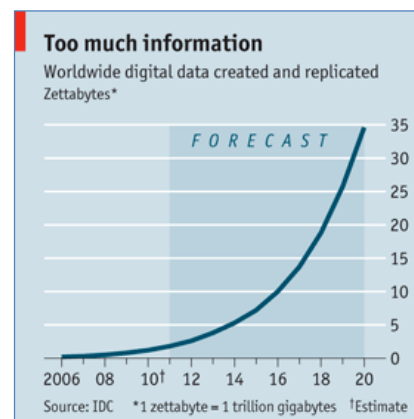
<sup>5</sup> <http://www.pnas.org/content/early/2012/06/11/1203882109.abstract>

make a case for humanitarian activities. This data service will also be a public good with benefits beyond the immediate users.

## II. Project Background

Every few days there is a news article or television programme that talks about the impact of data on our lives. Recent headlines announcing “a data revolution” and “the age of big data” point to the promise of using data to better understand our environment and optimize performance. There are a few drivers contributing to this digital disruption: 1) more data is being generated from mobile phones and sensors than ever before (see chart);<sup>5</sup> 2) new technologies are driving down the time to process data and the cost to store it; and 3) the open-data movement is contributing to an unlocking of knowledge not seen since the Gutenberg Revolution of the fourteenth century.

Across the globe, people are connecting through technology at an accelerating pace. This trend was explored in OCHA’s *Humanitarianism in the Network Age*,<sup>6</sup> which argues that access to technology and the availability of data are disrupting how humanitarian work gets done, and taken together they are defining a new age for humanitarian assistance.



Understanding that the ability to organize and use data for business insight can drive competitive advantage, many companies are making the transition from holding data in multiple, fragmented systems to creating enterprise data services and analytic programmes. Private-sector companies, such as Google and Amazon, are leading the way in processing data to better understand customer preferences and increase profit (their market value is US\$290 billion and \$124 billion respectively). Organizations such as the World Bank are also deeply engaged in creating datasets and analytical frameworks that can help explain successful approaches to development, something the World Bank president calls “the science of delivery.” The humanitarian community is on the cusp of making this transition.

## III. Problem Statement

OCHA leadership and staff are aware of the advances taking place in the data and analysis space and have been looking at how the organization and, by extension, the humanitarian system can adapt. The initial work has been around developing a clear and shared understanding of the problem.

As the coordinating body of the humanitarian system, OCHA sits at the centre of multiple information flows from response partners, clusters, donor and recipient Governments, affected people, the media, academia and beyond. OCHA is tasked with going through this information and data, figuring out what is relevant and what it means, and returning it to the community in a more consolidated format, such as through situation reports, humanitarian bulletins and consolidated appeals. In this way, we help decision makers understand what is happening in the overall response environment and what actions need to be taken to help people in need.

<sup>5</sup> Chart taken from: [www.economist.com/node/18226961](http://www.economist.com/node/18226961)

<sup>6</sup> <http://unocha.org/node/11528>

To the extent that information and data can be obtained and standardized, OCHA does this type of synthesis or aggregation well at an individual product level or through a specific field website. However, we do not do it well at a system level. Data and information becomes localized and trapped in documents or web pages. People outside of a field location are unable to access or use the data that is being generated.

**The causes for these bottlenecks include the following:**

- Most of the partner data that OCHA collects and uses is held within individual spreadsheets that do not have standard data fields (e.g. indicator name, source, unit, location, time series, owner). These spreadsheets are stored in varying locations with limited access, such as personal computer files.
- Inconsistent data formats (Excel, CSV file, XML file, pdf report) and terms (shelter kit versus tarp+rope) make data aggregation and consistent analysis both within and across countries slow or even impossible.
- There are no common data licenses for humanitarian data creating a lack of clarity around the terms of use.
- OCHA's information and data architecture is not well understood, making it difficult for partners to connect with us and through us.
- There is no obvious multilateral data-standards body within the humanitarian community. There are some groups within the Inter-Agency Standing Committee that could be used to discuss standards, but no such initiatives currently exist.

**The effects include the following:**

- New ad hoc humanitarian data initiatives regularly appear<sup>7</sup>, but rather than integrating them into an overall system, we develop point solutions to meet a specific need, increasing fragmentation and inefficiency.
- Humanitarians spend unnecessarily large amounts of time (and money) searching for data, and they are often uncertain of the source, quality and terms of use for what they find<sup>8</sup>.
- More sector-specific data is collected than is needed or used, aggravating partners and producing data noise that hides the truly critical information<sup>9</sup>.
- We call upon the volunteer and technical community<sup>10</sup> to process large datasets harvested from social media, but after single-purpose analysis, the data vanishes and its re-use value is lost.
- More and more information products are created with the same or similar information slightly repackaged to meet a specific audience need. OCHA is unable to make raw data available for users to do their own analysis and packaging.
- Decision makers end up with a fragmented and/or thin presentation of specific crises and their relationship to global humanitarian trends.

In addition to the bottlenecks created through poor data sharing, OCHA recently completed a study<sup>11</sup> in partnership with the Economist Intelligence Unit (EIU) that revealed there are currently no agreed

---

<sup>7</sup> Examples of these ad hoc initiatives include cluster indicator monitoring tools such as ActivityInfo (<http://www.activityinfo.org/>); MDG monitoring through Devinfo (<http://www.devinfo.org/libraries/asp/asp/Home.aspx>); geospatial data sharing through local databases (<http://geonode.colombiassh.org/>); NGO aid map (<http://ngoaidmap.org/>); financial data from the Global Humanitarian Assistance (<http://www.globalhumanitarianassistance.org/data-guides>), etc

<sup>8</sup> We heard this in the EIU research (see footnote below) and more recently in the data survey we conducted through ReliefWeb.

<sup>9</sup> We have done an analysis of spreadsheets of cluster data that is collected by OCHA staff in Pakistan, Yemen and the Philippines.

<sup>10</sup> <http://digitalhumanitarians.com/collaboration-guidance>

<sup>11</sup> The study was conducted in May 2013 and involved 45 hours of phone interviews. Participants included senior-level field staff (OCHA Heads of Office, Humanitarian Coordinators, donors) in several locations (South Sudan, Yemen, Chad, the Philippines, etc.) and HQ-based staff (OCHA section chiefs, donors, humanitarian partners) in several locations (London, Geneva, Rome, Brussels, New York, Washington, D.C). DFID London staff took part in an interview session.

statistical models for how OCHA and its partners determine critical planning figures. These figures include the number of people affected by a crisis and the number of people in need of assistance. In addition, the way that these figures are compiled across sectors leads to either double counting or underestimating individual needs.

#### **IV. Responding to Demand**

This project is responding to an operational necessity: OCHA and its partners need access to trusted data in near real time to understand what is happening in a specific crisis and across crises. Because of its mandate and existing data-sharing agreements with partners, OCHA has access to an enormous amount of (mostly) unused data. We can unlock the latent demand and value of this data only by making it interoperable and universally accessible.

The demand for better access to data has been building for years. There have been several real-time evaluations pointing to the lack of data sharing as being detrimental to a more effective response.<sup>12</sup> In addition to these formal evaluations, there was the groundbreaking study following the 2010 Haiti earthquake – *Disaster Relief 2.0*<sup>13</sup> – which called for making data open and comparable in order to create a common operational picture.

In 2012, the need for data-driven decision-making in humanitarian response was raised in the ECOSOC resolution on strengthening of the coordination of emergency humanitarian assistance of the United Nations.<sup>14</sup> In his recommendations, the Secretary-General requested Member States and the humanitarian community to work with OCHA to develop, agree and adhere to common international standards for the exchange of data to facilitate better preparedness and response to emergencies.

More recently with the EIU study, we conducted in-depth interviews with humanitarian managers to understand how they want to use data. We asked humanitarian practitioners about the type of analysis they do and what questions they most commonly are trying to answer about a humanitarian situation. The most consistent request from this group was for there to be one place to which they can go to find data.

To substantiate this initial study and to create a baseline for this project, we ran a survey for one week in late September 2013 to understand the humanitarian community's data needs and perceptions. The online survey was originally posted on ReliefWeb and OCHA's corporate website, but it was also shared through social media and e-mail networks.

Just under 3,500 people responded to the survey, with the majority of the respondents working for a non-governmental organization and based in Africa. The survey asked respondents to rate statements about whether they found humanitarian data easy to find, timely, trustworthy and easy to compare. We also asked if people regularly analyze data to inform their decision-making.

The only statement that a majority of respondents agreed with was "I regularly analyze humanitarian data to inform my decisions". The majority of respondents either disagreed or somewhat agreed with the statements on the timeliness and trustworthiness of humanitarian data, and a majority disagreed

---

<sup>12</sup> A synthesis report on information management in IASC Real-Time Evaluations was prepared as part of the research for OCHA's *Humanitarianism in the Network Age* publication. The evaluation review include those done for the Pakistan earthquake (2006), the Indian Ocean tsunami (2006), the floods in Mozambique (2007), Cyclone Nargis (2008), the Haiti earthquake (2010), the Pakistan floods (2011) and the Horn of Africa drought (2012), among others.

<sup>13</sup> [www.unfoundation.org/what-we-do/legacy-of-impact/technology/disaster-report.html](http://www.unfoundation.org/what-we-do/legacy-of-impact/technology/disaster-report.html)

<sup>14</sup> See ECOSOC resolution A/67/89–E/2012/77.

with the statement “It is easy to compare humanitarian data across crises”. The overall findings reinforce our understanding of the humanitarian data landscape: people are doing analysis on data that is not easy to find, not completely trusted and not easy to compare.

We believe that the following project plan will address many of the challenges detailed above. Some of these issues, such as community-agreed data standards, will take longer to address than others. The project’s initial contribution will be to make humanitarian data findable – this is the essential pre-requisite to improving data quality and doing higher level analysis.

**V. Humanitarian Data Exchange Project Plan**

The project plan has two phases: initial development and further scaling. The goal of phase one is to develop a functional data platform using data from four to five countries and with a focus on a priority set of users (see details on users below). These users will be able to search for specific indicators, choose to visualize that data over certain periods of time and across pilot countries, and to download the data into a spreadsheet and take it with them.

Dec 2013 – Dec 2014	Jan – Dec 2015
PHASE 1: A FUNCTIONAL PLATFORM	PHASE 2: SCALING TO ALL OCHA COUNTRY OFFICES
4-5 Countries	20+ Countries

From there, we will scale the platform to bring in data from more countries, more sources and for more users, eventually creating an engaging, easy-to-use, single point of entry for humanitarian data.

To avoid creating yet another humanitarian information silo, the data platform will be accessible from ReliefWeb, e.g. there will be a data tab in ReliefWeb’s main navigation, and the URL for the data site will be [hdx.rwlab.org](http://hdx.rwlab.org). The data platform will also benefit from ReliefWeb’s steady site traffic. In July 2013, ReliefWeb had over 500,000 unique users visit the site. Over 3.5 million unique users visited the site in 2012.

*Project Activities*

Phase one of the project includes seven interrelated workstreams. The following is a summary breakdown of each output area<sup>15</sup>. The workstreams for phase two of the project will be detailed as we make progress on the following outputs.

The phase one workstreams are as follows:

- Output 1: Data collection
- Output 2: Data quality
- Output 3: Data standards
- Output 4: Data systems and exchange architecture
- Output 5: Analytic research and development
- Output 6: User interface research and design
- Output 7: Governance and communications

Although each output has a defined scope of work and deliverables, there are a number of dependencies between them. For instance, we can begin to develop a conceptual architecture but we can’t know how it will really work until we introduce live data into the system. Likewise, we can collect a

<sup>15</sup> A project Gantt chart is available to understand the detailed phase one work areas and their dependencies.

lot of data but it won't be useful until we understand how users want to interact with it through an interface.

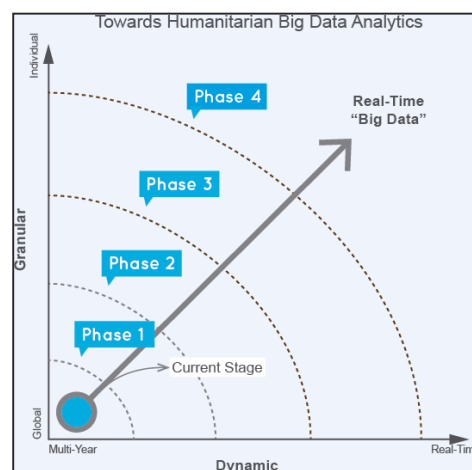
## 1. Data collection

This work involves defining a common humanitarian dataset that brings together multi-sourced data from existing, siloed data domains. This common dataset includes pre-crisis/development data, operational data, financial data and geospatial data. It is important to note that this dataset does not represent new indicators or any negotiation around indicator development, but is simply an aggregation of indicators already in use by the humanitarian community. As the humanitarian community develops new indicators, we can add them to the dataset.

We began the research to understand what data might go into a common humanitarian dataset in early 2013. The research looked at the data that was being used in existing information products such as the consolidated appeals, situation reports, humanitarian bulletins and humanitarian dashboard, among others. The dataset also took in the common operational dataset (CODs) schema that was agreed by the Inter-Agency Standing Committee in 2010.<sup>16</sup> The financial data will come from OCHA's Financial Tracking System.<sup>17</sup>

A spreadsheet listing the indicators for the common humanitarian dataset is available for review. It includes around 600 indicators<sup>18</sup> from over 30 trusted sources.

Our data-collection work to date has focused on aggregating national-level, pre-crisis data from sources such as the World Bank, UNICEF, WHO and other trusted partners. In collaboration with a company called ScraperWiki,<sup>19</sup> we have collected over 300,000 data points for 240 territories (mostly countries – plus dependent territories such as Martinique, Aruba, Hong Kong, French Guiana and so forth), covering a period of 60 years with 200 different indicators.<sup>20</sup> This data will serve as the foundational baseline data of the common humanitarian dataset from which all other data can be compared. As we make progress, we will begin to include data that is more granular and more dynamic (see chart). Specific to humanitarian operational data, this might include real-time needs-assessment data or cluster-based data at a sub-national level.



We are also exploring data-sharing partnerships with agencies that are supporting Government data-collection efforts. This includes the World Bank's Open Data for Resilience Initiative (OpenDRI), which works with national Governments to open their data, and UNICEF's DevInfo,<sup>21</sup> which is an open-source software that is being used to store local statistical data for indicators related to the Millennium Development Goals. We will also look to learn from the experiences of these initiatives as we create a new data platform.

<sup>16</sup> <http://cod.humanitarianresponse.info/about-codfod>

<sup>17</sup> <http://fts.unocha.org/>

<sup>18</sup> An indicator is a single type of information about humanitarian data. Indicators are the heart of the project's data management process — they allow the system to track information about the kinds of things the humanitarian community cares about, while providing snapshots of different sources and different points in time.

<sup>19</sup> <https://scraperwiki.com/>

<sup>20</sup> See ScraperWiki blog post on this work here: <https://blog.scraperwiki.com/2013/10/us-and-the-un/>

<sup>21</sup> [www.devinfo.org/](http://www.devinfo.org/)



## 2. Data quality

This workstream will look at how data quality is determined. For this, we are taking a two-pronged approach. We plan to curate the data that is part of the common humanitarian dataset. This means that we will have an internal process for reviewing the data quality before we make it available through the website. The project's data managers will take steps to ensure that the data entering the database is correct.

The common dimensions for assessing data quality include the following:

1. **Relevance.** The relevance of data refers to the degree to which it meets the current and potential future needs of the clients.
2. **Accuracy.** The accuracy of the data is the degree to which the information correctly describes the phenomenon it was designed to measure.
3. **Timeliness.** The timeliness of the data refers to the delay between when the data is collected and when it is made available for use.
4. **Accessibility and interpretability.** The accessibility of data refers to the ease with which it can be obtained from the data sources. The interpretability of the data reflects the availability of the supplementary information (metadata) needed to utilize and understand the data effectively.
5. **Comparability.** The comparability of data refers to the degree to which it can be successfully brought together with other statistical data within an analytical framework.

If an incoming dataset fails the quality assurance tests, then a data manager will have to initiate error handling, which may involve getting back in touch with the data source or not importing the data into the system. This quality-control process will be a recurring task as we scale the system. Overtime, we will create models for what correct and incorrect data should look like for a specific data source and we will come to understand the most efficient technical and human processes for handling incoming data.

For data that falls outside of our curation efforts, such as large amounts of processed social media data or the location of roads in a specific county, we are considering using a user-generated quality rating for the data, much like Amazon does for its products. Over time, some datasets and sources will be rated highly and others will not, but it will be left to the community to determine this.

It should be noted that improving the trustworthiness of humanitarian data starts at the point of data collection. Our internal review of data quality will be immediate, but it will take time to create a feedback loop with data sources on data quality. This project will start to address some of these issues, but it is not expected that they will be resolved in the initial phase of the project.

## 3. Data standards

OCHA will receive a grant from the Humanitarian Innovation Fund (HIF) to advance the earlier proof-of-concept work for the creation of a Humanitarian Exchange Language (HXL). The HIF grant will involve community outreach on the HXL Initiative, creating a multi-stakeholder body to oversee the development of the HXL standard, and conducting small-scale pilots of the HXL vocabulary and infrastructure with key stakeholders in up to two field locations<sup>22</sup>. This work is related to the development of the data platform, but the success of HXL is not dependent on the platform, and the success of the platform is not dependent on HXL.



---

<sup>22</sup> The HIF HXL project page can be found here: <http://www.humanitarianinnovation.org/projects/large-grants/UNOCHA>

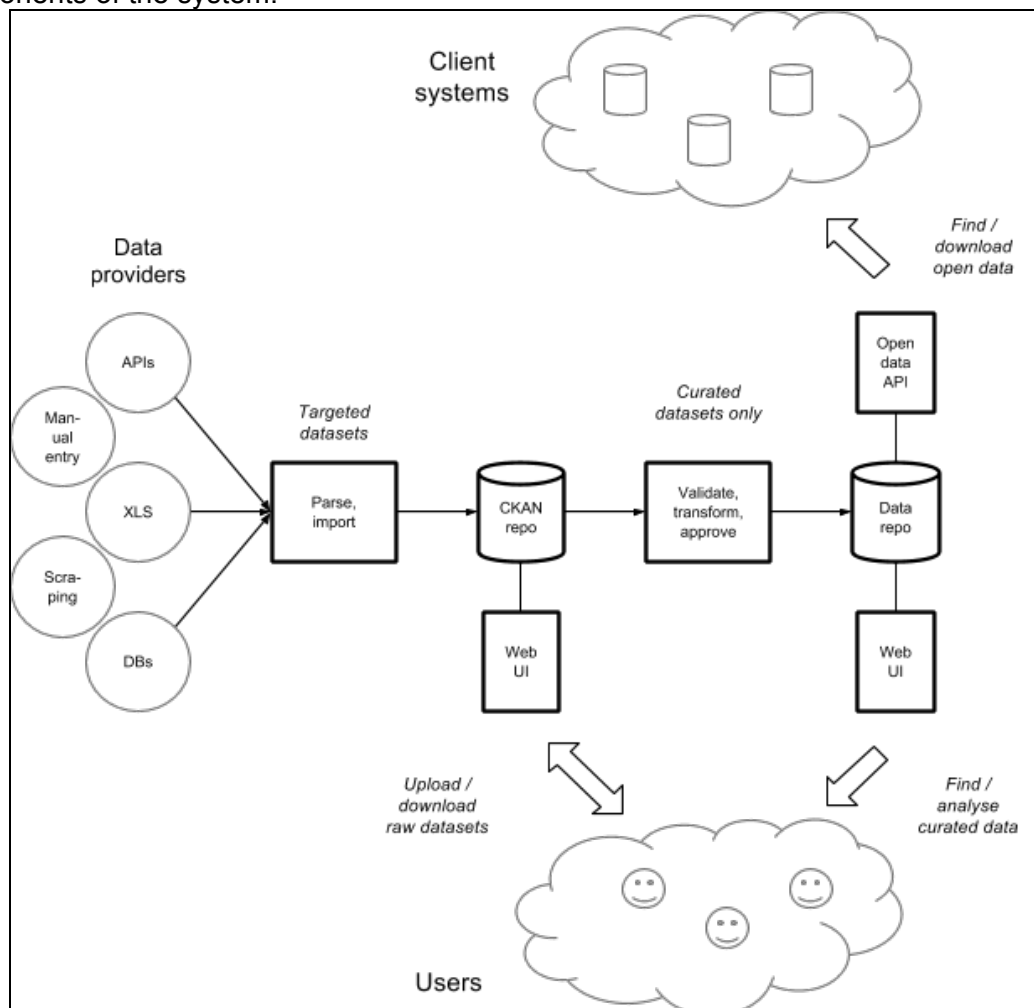
We are excited about the potential of community-agreed standards for data sharing and believe that it will be a crucial part of humanitarian coordination in the years to come. At the same time, we recognize that it will take time for the community to adopt these standards. This will be a slower time frame than the more immediate need to address the user demands listed above.

#### 4. Data systems and exchange architecture

This work involves creating the technical back-end for the system. It includes creating a data model for the common humanitarian dataset and a system of exchange for data to move from spreadsheets and databases into a data hub and from the hub into the user interface. Our goal in phase one is to create a small system that works end-to-end rather than over-specify what might be needed in the future.

The architecture will have two components: a traditional relational database for the common humanitarian dataset; and a “bazaar”-type model that will enable the sharing of multiple ad hoc datasets from the community. For the latter, we plan to use CKAN, an open-source software for data storage and sharing. CKAN is being used by dozens of national Governments<sup>23</sup> (UK, US, Australia, Germany, etc) and municipalities (Ottawa, Santa Cruz, Buenos Aires, etc) to share data with the public.

We have created a high-level, component architecture for the system. The diagram below shows the main components of the system.



<sup>23</sup> <http://ckan.org/instances/#>



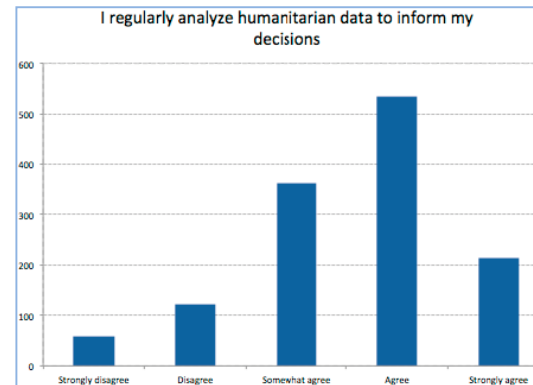
Given the public nature of ReliefWeb and the data platform, we will ensure that robust data-security protocols are in place for both data storage and sharing where appropriate. During the first phase, we will not include any sensitive or personally identifiable information in the data repository. All data in it will be suitable for public release. We will design the architecture and implementation to support sensitive information in future phases, including the ability to clearly mark and distinguish secure data from other data in our repository.

## 5. Analytic research and development

The analytic research began with the extensive interviews that we conducted with the EIU. The questions that humanitarian managers have about a response and the decisions that they need to make are not much different than they have always been. Managers need to know who requires assistance, where they are, what they need, who is responding and what are the gaps. Over a longer time horizon, they need to know if a situation is getting better or worse, and what cyclical risks or human vulnerabilities can be mitigated through humanitarian response efforts.

The data that we will provide through the platform will make it easier for people to answer these questions by making the data findable and comparable. The common humanitarian dataset contains the essential humanitarian data-building blocks. This includes the baseline data mentioned above, but also the critical figures on the number of people affected by a crisis, the number of people in need of assistance, and the number of people targeted for assistance by the humanitarian community. Normalizing this data across countries and crises will be challenging, but it is a pre-requisite to showing the value of combining multi-sourced data from several countries.

Once the data becomes comparable, users will be able to develop their own analytical models to explore the questions they have about humanitarian crises. We know from the recent ReliefWeb survey that people are already regularly analyzing humanitarian data (see chart). The platform will make this easier by offering data-visualization functionality so that users can easily see the trends and correlations of these indicators on charts and geo-referenced on a map.



The project will also seek to work with other parts of OCHA (specifically our Programme Support Branch in Geneva) and the humanitarian community to define the underlying methodologies for key humanitarian figures, such as the number of people affected and the number of people who need assistance. This effort will take place in slower time and will not be addressed in phase one of the project.

## 6. User interface research and design

We will be using a user-centered design approach to the interface design. We will invest in research and consultation with priority users of the data to understand how they want to access it and how they intend to use it. It is only by delivering a compelling user experience that this platform can drive a new set of behaviours around data sharing and achieve critical-mass adoption.

The user research will focus on phase-one targeted users of the data. These users include staff from OCHA, UN agencies, NGOs, and donor and recipient Governments across the following dimensions: field/HQ; technical/non-technical; and manager/working level. We will carry out the user consultations through face-to-face meetings in New York and in select field locations.

The following table shows two typical scenarios of data use for phase-one of the platform.

Targeted phase-one user	Scenario of data use
OCHA Head of Office	A new Head of Office arrives in a country. She will be able to easily find all of the pre-crisis data about the country and compare this with other countries in the region and across the globe. She will also be able to find up-to-date response figures, such as the number of people targeted for assistance, how many are IDPs or refugees, and where they are located. From this data, she will be able to better assess the severity of the humanitarian context that she is working in and what geographic locations require her attention.
Donor Government official	A humanitarian adviser will be able to develop his own analytical model based on a set of indicators in the common humanitarian dataset (for example: under-5 mortality plus number IDPs, refugees, etc). These 10-20 indicators could be compared and visualized across a number of humanitarian settings to understand the severity of a crisis and to answer programmatic questions.

## 7. Governance and communications

### *Project Governance*

A project this complex requires strong governance. A governance plan has been developed and will be applied over the course of 2014. It includes governance at the executive, technical and analytical levels.

We also recognize the value of having independent, objective advice on the project outside of direct line management. An Independent Advisory Board is being created and will be responsible for the project oversight, with a specific focus on domain areas that are especially complex. These areas include technology, standards, innovation and advanced analytics.

Board members include the following experts:

- Peter Walker, Professor and Director, Feinstein International Center at Tufts University
- Nigel Snode, Google.org, Head of the Crisis Response Program
- Linus Bengtsson, co-founder of Flowminder
- Tariq Khokar, Data Scientist, World Bank
- John Adams, Technical Lead for DFID's open data work, member of IATI Technical Advisory Group
- Anthony Craig, Chief of the Emergency Preparedness and Response Coordination Support Branch, WFP

### *Communications*

We will develop an outreach strategy to ensure that staff and partners understand what we are doing, why we are doing it and how they can get involved. Outreach material will include slide shows, case studies, short films and a glossary of data terms. The main forum for the outreach will take place through ReliefWeb Labs,<sup>24</sup> which is a virtual space for experimentation and communication on the ReliefWeb website. We hope to create an active early user community that will help shape the decisions we make and essentially co-create the data platform with the project team.



As part of the recent data survey, we asked people to leave their e-mail address if they wanted to stay engaged in the development of the platform; over 2,700 people (75 per cent of all respondents) shared their address. Our outreach will also include more focused, face-to-face meetings with humanitarian partners.

<sup>24</sup> <http://labs.reliefweb.int/>

## **VI. Project Milestones**

### **The project has three initial milestones:**

1. Secure sufficient funding and have a few critical staff in place to lay the technical foundation for the platform and scale the system.
2. Agree with stakeholders and users on the phase one priority countries, data, reference lists and analytic outputs.
3. Agree on the phase one data model, system architecture and supporting technologies.

### **The next set of milestones will follow from the planning and initial modeling work:**

4. Collect majority of the data in the common humanitarian dataset for three to four countries. (This will be done through both manual and automated data collection processes).
5. Establish local data-sharing systems and platform workflow for pilot countries.
6. Complete research on the user design requirements for the global web interface.
7. Release periodic iterations of the phase one prototype of the global platform with data from pilot countries.
8. Develop plan for field-level data-sharing systems for more field locations in phase two of the project.

For all of this work, we plan to take an iterative approach to development. This is in contrast to waiting until the end of phase one to release a finished product. We realize that we cannot have a “build it and they will come” mindset. We will adopt a service attitude, providing high-touch interaction with lead users to speed adoption. We will certainly experience setbacks, have to re-think initial plans and milestones, and ultimately be adaptive in order to reach our goals.

## **VII. Project Benefits**

### **The benefits of this project include the following:**

- There will be one entry point for data from multiple databases and spreadsheets. This will be accessible from the web and eventually from a mobile device.
- The source, quality, terms of use and technical standard for the data will be clear.
- It will be possible to conduct trend and comparative analysis of data from different countries and humanitarian crises.
- There will be open-data standards across a network of humanitarian actors. This will enable an open-data ecosystem that bridges the disaster and development cycles.
- A wider group of actors will be able to use and add value to humanitarian data and analysis, tapping into the potential of volunteer communities and the public to support humanitarian response.