Humanitarian Data Exchange

A project by the United Nations Office for the Coordination of Humanitarian Affairs
Goals

HDX aims to make humanitarian data easy to find and use for analysis.
**Project Elements**

**Repository**
HDX will include a dataset repository, based on open-source software, where partners can share their data spreadsheets and make it easy for others to find and use that data.

**Analytics**
HDX brings together a Common Humanitarian Dataset that can be compared across countries and crises, with tools for analysis and visualization.

**Standards**
HDX promotes community data standards (e.g. the Humanitarian Exchange Language) for sharing operational data across a network of actors.
Component Architecture

Data sources
- System APIs
- Manual entry
- Scraping
- Spreadsheets
- Databases

Data acquisition systems

Dataset repository

Data curation system

Data search engine
- Index data and deliver search results

Data analytics
- Find / analyze curated data
- Upload / download datasets

HDX Blog

Humanitarian Community

Engage with community
System Diagram

Data collection from multiple sources and systems.

Humanitarian Data Exchange

The HDX Process
HDX aggregates, stores, and transforms data for the humanitarian community.

Products

Websites
Phase 1 Priority Users

- Host Governments
- Donor Governments
- NGOs
- UN Agencies
- OCHA
Data Aggregation

300,000+ data points collected so far
Example:
Crude Mortality Rate (past 50 years)
24 OCHA Country Offices

The trend of sum of CMR_1000 for Year. Color shows details about Country Name. The view is filtered on sum of CMR_1000, which keeps non-Null values only.
Example:
Mobile phone subscriptions per 100 people (past 17 years) - 24 OCHA Country Offices
The Common Humanitarian Dataset

• 100+ indicators from multiple data sources
• Includes data from across the programme cycle:
  • Country context (languages, currency, office locations, etc)
  • Preparedness data
  • Operational data
  • Humanitarian financing
  • Geospatial data
Quality assurance framework

Quality dimensions:

1. **Relevance** - the degree to which it meets the current and potential future needs of the clients
2. **Accuracy** - the degree to which the information correctly describes the phenomenon it was designed to measure.
3. **Timeliness** - the delay between when the data is collected and released for use.
4. **Accessibility and interpretability** - the ease with which data can be obtained from the data sources and the availability of the supplementary information (metadata) needed to utilize and understand the data effectively.
5. **Comparability** - the degree to which data can be brought together with other statistical data within an analytical framework.
Real-time, granular data

Towards Humanitarian Big Data Analytics

Phase 1
Current Stage

Phase 2

Phase 3

Phase 4

Real-Time “Big Data”
Project Scope and Timeline

**Phase 1** will develop a functional platform for 3 countries.

**Phase 2** will scale the system to more countries and more users.

The timeframe will depend on resources and data readiness.
HDX HIGH-LEVEL ROADMAP

--- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---

**HDX REPOSITORY + ANALYTICS**
- Continuous development process with closed releases.
- Field visits are conducted.
- Possible prolonged field presence in Nairobi through pilot Data Lab.
- Dataset repository becomes part of HDX platform.

**HDX STANDARDS**
- Working group established.
- First release of draft standard.
- Basic HXL infrastructure.
- The HXL team conducts field visits.

**frog design research**
- Functional in-browser prototype
- First public beta of platform released
- Formal data standards body formed.
- Release version 1.0 of HXL standard(s).
- Working HXL infrastructure.
The humanitarian data exchange
WHERE YOUR DATA COMES TO LIFE

This is an early version of the HDX Repository. Initially, you will be able to find global datasets relevant to humanitarian work as well as local datasets from our three pilot locations - Colombia, Kenya and Yemen. You can also create an account and add your own data to the repository to share privately or publicly. Please have a look around and send us your feedback!

A PROJECT BY: OCHA
RELATED SITES: Humanitarian RESPONSE, reliefweb, FTS
STAY CONNECTED: twitter, facebook
Deliverables for the Field

1. A place to find the data spreadsheets that are shared amongst the humanitarian community
2. An analytic interface for comparing country-based humanitarian data across time and space
3. A data quality review process for a subset of humanitarian data
4. Agreed data standards for a subset of humanitarian data
5. Technical and data analysis support
Data Labs (concept)

A Data Lab is a physical space in a field office that is staffed by a team of data managers, developers, and data analysts dedicated to improving the data environment with OCHA and the wider community.
Supporters
Stay in touch

hdx.rwlabs.org
Email us with questions at hdx@un.org
Follow us @humdata for tweets on #opendata and #HXL