

Title

The Humanitarian Data Exchange - UN-OCHA's data first responders

Contributors

Francis Irving

CEO, ScraperWiki Ltd.

A technology leader, Francis created the original TortoiseCVS, which has improved version control for tens of millions of people. He was a founder of TheyWorkForYou and WhatDoTheyKnow, which show the world how to use scraping and good user experience to improve democracy. He then became CEO of ScraperWiki, a company created to share data from the internet. Irving has a BA in mathematics from Oxford.

Type of the presentation

In-use contribution

Abstract

Data-driven responses are not just for commercial interests, when disaster strikes the following data are critical:

- Baseline data - population, levels of poverty, road networks and airports etc, gives the context for the disaster;
- Impact data - how many people are affected, where and so forth is critical for gauging and managing the required response;

The Humanitarian Data Exchange(HDX) is a initiative by the UN which brings these data together and provides a platform for sharing. It has already delivered value in the West African Ebola and Nepal earthquake crises.

Extended Abstract

UN-OCHA (United Nations Office for the Coordination of Humanitarian Affairs) coordinates responses to natural disasters and emergencies. This coordination covers both international funding and also on-the-ground responses. This can include anything from the Syrian refugee crisis to the Ebola virus and the Nepal earthquake. Data is central to this work — using information about refugee movements, the weather, and NGO capabilities along with baseline data for the country OCHA decides who can best use money to save the most lives.

The Humanitarian Data Exchange (HDX) is a new project with a goal of increasing the reuse of data in the Humanitarian world. UN-OCHA is building a data hub. It's based on the Open Knowledge Foundation's CKAN product. ScraperWiki is supporting the data collection process and providing technical project management.

It all starts with baseline data, basic information for all countries derived from various sources like the UN itself and the World Bank. This data includes things like population, levels of poverty, mobile phone usage, it says something about a country before disaster struck. This data can be collected online at leisure, it is transformed to a standard format before it is loaded to the HDX.

Once disaster strikes the key data is on the impact of the crisis: numbers of people affected, the severity of damage, the disposition of responders. Speed is of the essence here, HDX is the repository for data first-responders. Provenance and control are also important, in a crisis unreliable information can surface and push out that which is reliable. So HDX supports the inclusion of provenance data, along with a political decision that it is to be a definitive source.

Impact data comes from OCHA's many country offices, and also partners, such as the Red Cross. They tend to store it in spreadsheets on their own systems, or share it with the world in PDF reports. In the necessary rush of emergency work, it's hard to find more time to put into data sharing.

The presentation will answer these questions:

- What motivates people to share data?
- What barriers must be overcome in data sharing during a crisis?
- What features in a data hub do on-ground workers need?
- How can a non-profit make best use of crowd sourcing from its volunteers?
- What metrics best measure how useful a data hub is?

We will provide some case studies including the Ebola virus and Nepal earthquakes.

Lessons from the front line of humanitarian aid are invaluable in industry too. Every business is saying today: "It needs to be easier for our staff to share data with each other."