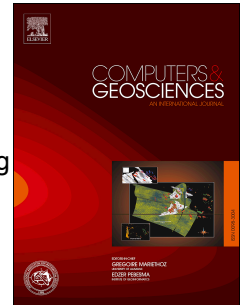


Accepted Manuscript

Bridging the information gap of disaster responders by optimizing data selection using cost and quality

Marc van den Homberg, Robert Monné, Marco Spruit



PII: S0098-3004(17)30571-X

DOI: [10.1016/j.cageo.2018.06.002](https://doi.org/10.1016/j.cageo.2018.06.002)

Reference: CAGEO 4142

To appear in: *Computers and Geosciences*

Received Date: 31 May 2017

Revised Date: 16 May 2018

Accepted Date: 5 June 2018

Please cite this article as: van den Homberg, M., Monné, R., Spruit, M., Bridging the information gap of disaster responders by optimizing data selection using cost and quality, *Computers and Geosciences* (2018), doi: 10.1016/j.cageo.2018.06.002.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Bridging the information gap of disaster responders by optimizing data selection using cost and quality

Marc van den Homberg^{1,2,3}, Robert Monne^{2,3,4,5}, Marco Spruit⁴¹

¹ Netherlands Red Cross 510 data initiative, Anna van Saksenlaan 50, 2593 HT The Hague, The Netherlands

² Cordaid, Lutherse Burgwal 10, 2512 CB The Hague, The Netherlands

³ TNO, Oude Waalsdorperweg 63, 2597 AK The Hague, The Netherlands

⁴ Utrecht University, Utrecht, Princetonplein 5, 3584 CC Utrecht, The Netherlands

⁵ Ortec, Houtsingel 5, 2719 EA Zoetermeer, the Netherlands

Presenting author's email address: mvandenhomberg@redcross.nl

Abstract

Natural disasters are chaotic and disruptive events, with compressed timelines and high levels of uncertainty. Comprehensive data on the impact becomes only available well into the response phase and data is scattered across organizations. Data heterogeneity issues are common. Consequently, responding organizations have difficulties finding data that match their information needs. We investigated the information needs of and the disaster management data available to both national and local decision makers during the 2014 floods in Bangladesh. We conducted 13 semi-structured interviews and three focus group discussions, collecting in this way input from 51 people, transcribed and coded them so that themes of information needs emerged. We mapped the information needs on the available data sets and determined which needs were not, partially or completely covered. We identified seven themes of in total 71 information needs and 15 data sets. The mapping revealed a significant information gap of timely and location-based data. Only 40% of the information needs are covered in time and 75% if no time constraints are considered. Instead of using all data sets, we optimized for coverage -with Integer Linear Programming- combinations of data sets against the costs of extracting data from structured versus unstructured data and against the quality in terms of recency, source and content rating and granularity. Without time constraints, three data sets yield already a coverage of 68%, whereas adding five extra data sets only gives an improvement of 7%. We recommend executing identification and mapping of available data sets on the information needs as part of Data Preparedness. Determination of the optimal combination of data sets can be used to extract data on information needs more efficiently. Currently, we did this manually, but future research will investigate automatic matching of information needs on data sets, by applying intelligent querying and semantic data matching.

Taxonomy: Data Fusion, Big Data Analytics, Natural Disaster, Disaster Management, Humanitarian Operations, Decision Making

Keywords: Disaster Management; Decision Making; Information requirements; Humanitarian Response; Data preparedness, Integer Linear Programming.

Highlights:

- We identify 71 information needs within seven themes and 15 data sets.
- Only 40% of information needs are covered on time and 75% without time constraints.
- We optimize data selection using cost and quality by Integer Linear Programming.
- Without time constraints, three data sets yield already a needs coverage of 68%.

¹ MvdH proposed the overall research question, wrote the manuscript, was involved in the data collection in Bangladesh and the interpretation of the data. RM developed the research design as part of his MSc thesis with support of MvdH and MS. RM performed the acquisition of the data in Bangladesh together with MvdH and labelled and coded the data. MS supervised the research and reviewed the manuscript. All authors were involved in reviewing and approving the final version of the manuscript.

Download English Version:

<https://daneshyari.com/en/article/9952355>

Download Persian Version:

<https://daneshyari.com/article/9952355>

[Daneshyari.com](https://daneshyari.com)