

Next Generation Disaster Data Infrastructure White Paper

CODATA Task Group

Linked Open Data for Global Disaster Risk Research (LODGD)

Date: January, 2019

Beta Version 0.1

Study Panel

Authors in alphabetical order

Bapon Fakhruddin, Tonkin and Taylor International, New Zealand

Edward Chu, CODATA Task Group on Linked Open Data for Global Disaster Risk Research

Guoqing Li, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, China

Jibo Xie, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, China

Kun-Lin Lu, CODATA Task Group on Linked Open Data for Global Disaster Risk Research

Kunal Sharma, National Disaster Management Authority, India

I-Ching Chen, CODATA Task Group on Linked Open Data for Global Disaster Risk Research

Nuha Eltinay, Arab Urban Development Institute, Kingdom of Saudi Arabia

Saurabh Dalal, National Disaster Management Authority, India

Tengfei Yang, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, China

Acknowledgments

This work is supported by the Committee on Data for Science and Technology (CODATA) and National Key R&D Program of China (2016YFE0122600).

Table of Contents

1. Introduction.....	4
2. Future scenarios and applications	5
3. Big disaster data collection and transmission.....	10
4. Big disaster data processing	12
5. Big disaster data quality control	16
6. Big disaster data visualization	21
7. Conclusion	28

1. Introduction

Due to climate change and global warming, the frequency and severity of extreme weather has been increasing all around the world. According to the Sendai Framework for Disaster Risk Reduction, from 2005 to 2015, over 700 thousand people had lost their lives, over 1.4 million had been injured and approximately 23 million had been made homeless as a result of disasters. The severity of future disasters is expected to surpass the past in the foreseeable future. In addition, huge amount of disaster data collected from different sources, such as local sensors, remote sensing systems, mobile devices, social media, and official responders could easily overwhelm and impair disaster risk reduction related applications, systems, and underneath hardware platforms, especially for large-scale disasters.

In order to build resilience and reduce losses and damages, Sendai Framework prioritize its actions in the following four areas: (1) understanding disaster risk, (2) strengthening disaster risk governance to manage disaster risk, (3) investing in disaster risk reduction for resilience, (4) enhancing disaster preparedness for effective response and to “Build Back Better” in recovery, rehabilitation and reconstruction. In particular, Sendai Framework emphasized that government should strengthen the utilization of media, including social media, traditional media, big data and mobile phone networks, to support nationwide disaster management and damage reduction. The availability of access to multi-hazard early warning systems and disaster risk information and assessments to people should substantially increase by 2030. For this, governments should take into account the needs of different categories of users and data dissemination in order to enhance disaster preparedness for effective response. In addition, space and in situ information, including geographic information systems (GIS), are needed to be fully utilized in order to enhance disaster analysis tools and to support real-time access services of reliable disaster data.

Based on the above discussion and the targets of Sendai Framework, this white paper proposes the next generation disaster data infrastructure, which includes the novel and most essential information systems and services that a country or a region can depend on in reality in order to successfully gather, process and display disaster data, and to reduce the impact of natural disasters. Fundamental requirements of the disaster data infrastructure include (1) effective multi source big disaster data collection (2) efficient big disaster data fusion, exchange, and query, (3) strict big disaster data quality control and standard construction, (4) real-time big data analysis and decision making and (5) user-friendly big data visualization.

The rest of the paper is organized as follows: first, several future scenarios of disaster management are developed based on existing disaster management systems and communication technology. Second, fundamental requirements of next generation disaster data infrastructure inspired by the proposed scenarios are discussed. Following that, research questions and issues are highlighted. Finally, suggestions and conclusion are given at the end of the paper.

2. Future scenarios and applications

2.1 Active emergency response system for living environments

With the rapid development of disaster early warning technology and sophisticated network infrastructure, many countries have already adopted standard warning messages to inform the public that a natural disaster has occurred or will happen. For example, in the U.S., after an earthquake strikes, the official agency adopts the Public Warning System (PWS) to broadcast CAP (Common Alerting Protocol) warning messages to inform people in the affected areas through different media channels, such as radio, television, short message service (SMS), smart phones, internet or electronic signs. In addition, the XML (Extensible Markup Language)-formatted CAP messages are sent to active emergency response systems (AERS) to automatically start the process of disaster reduction, such as stopping elevators at the closest floor, cutting off the gas, opening doors and windows, slowing down high-speed trains and putting factory machines into a protection mode to avoid possible damages.

In the future, the AERS will play a more important role in disaster prevention and become ubiquitous in our living environment. The AERS will support people with decision-making functionality rather than simply trigger actuators to control input and output devices. In addition, customized warning messages will be sent to different recipients based on their identities, spatial locations and the emergency levels of the disaster to assist people to be better prepared for natural disasters. Relevant city services, such as health care and transportation, will also be integrated with AERS to support mass crowd evacuation and emergency medical services.

For people inside a building, unlike existing evacuation systems which provide only static information (i.e., evacuation map, fire equipment location, and emergency contacts), AERS will further provide them with dynamic evacuation instructions, real-time disaster information, and the progress of rescue operations so that people can safely leave danger areas or find a safe place to stay. For on-scene commanders, AERS will provide them with dynamic information of victims, such as their identities, spatial locations and physiological status, as well as the current situation of the disaster. Visualization technology will be used to highlight the severe level of affected areas and the status of both victims and responders. Intelligent decision-making services will also be applied to support rescue operations and health care resource management. In addition, for first responders, AERS will provide them with not only the information of victims, but also indoor navigation and real-time disaster information.

2.2 Crowdsourcing supported disaster information system

With the advance of social network, many research efforts have been made on using social media for disaster management. The main reason is that social networks could provide not only rich information but also near real-time response. Taking the

New Zealand Earthquake in 2016 as an example, the social media provided a large number of photos with geographic information in a short time so that official agencies could conduct a proper rescue action. In the future, social networks will play a more important role in disaster management. Advanced technology such as machine learning big data analysis, and image processing will be further investigated and developed so as to accurately classifying disaster information. Taking typhoon disaster as an example, the fire department requires more detailed information of the injured person than other departments. Therefore, the results of the classification should be filtered based on the user's role and responsibility. In addition, for people nearby the disaster areas could also provide useful information. Data fusion is necessary to have a comprehensive view of threatened areas. The disaster information system should be able to classify the huge information collected from both social media and sensor network in order to provide corresponding personnel with classification results.

2.3 Disaster data quality assurance and control

Active emergency response systems (AEIS) are designed to perform a range of safety related tasks like turning off electricity and natural gas valves, open doors for ease evacuation bring elevators to the ground floor, deliver alerts and instructions to people via their mobile devices etc. Liu and Chu (2015) [1] explains that AEIS are made feasible due to the advancement of four major technology domains: (1) advances in sensor and analysis technologies, (2) emergence of Common Alerting Protocol (CAP) which is an XML-based data format for exchanging public warnings, (3) development of platforms that integrate multiple communication channels enabled to receive CAP messages like Integrated Public Alert and Warning System (IPAWS)-OPEN (FEMA, 2017) [2], and (4) development of building information models and associated digital data exchange standards. Therefore it is evident that an AEIS is primarily a gigantic network of data models designed to perform real time data collection transmission and processing for decision making. Hence the success of an AEIS heavily depends on its quality of data and information.

The quality of information provided by sensor data is a critical concern in an emergency situation. For example, let's consider medical sensing scenario. Medical sensors can be deployed on a patient's body to monitor health related parameters. These data is collected via wireless personal area network for the doctors to monitor the patient's health status in real time. Further the environmental sensors at crisis site such as smoke sensor can detect fire in a building, and it can also work with camera sensors to help determine a rescue route of the patient in a timely manner. The above example illustrates that the utility of sensing in emergency response applications in practice are highly heterogeneous and demand a need for timeliness, prioritization and fault tolerance in sensing. Sachidananda et al (2010) [3] argue that currently the quality of sensor data is usually addressed in isolation by focusing on discrete data processing operations such as raw data collection, in-network processing (compression aggregation), information transport and sink operations for decision making. Further, Qin et al (2013) [4] argue that current research has primarily considered functional aspects of distributed sensor systems

focusing on techniques to sense, capture, communicate, and compute over sensor networks where as in more complex and diverse sensor applications, non-functional application needs (such as timeliness, reliability, accuracy and privacy) become important. Thus it is necessary to perform a quality aware sensor data management in future to make AEIS an effective reality.

AEIS are heavily reliant on the quality of XML based messages sent to various smart devices for automated processing of disaster management tasks. Efforts have been taken throughout the last decade or so to standardize the emergency related data formats and use them in emergency situations. For example, XML-based EDXL (Emergency Data Exchange Language) (OASIS, 2006) [5] messaging standards, including CAP (Common Alerting Protocol) (OASIS, 2010) [6], enable information exchanges between emergency information systems and public safety organizations, automatic report by sensor systems to analysis centres, and aggregation and correlation of warnings from multiple sources. In early stages users have reported problems when implementing CAP messaging over multiple systems that include commercial satellite and terrestrial network technologies such as C/L/X-Band, GSM, and CDMA in modes of voice and text (Waidyanatha et al, 2007) [7]. However the recent developments has enabled CAP to be used as the standard of emergency alerts and in recent years CAP has been deployed in US, Canada, Australia and parts of Asian Pacific region, including Taiwan and Japan. In the case of AEIS it is necessary that all smart devices used in the system are compatible with the data standards and produce error free quality messages in emergency situations.

Recent developments in building information models, smart/intelligent homes and the concept of smart cities will become important aspects in deploying AEIS. A building information model illustrates the geometry, spatial relationships, geographic information, quantities and properties of building elements such as facilities equipments that can be used for lifecycle management of buildings (Bazjanac, 2006) [8]. On the other hand smart and intelligent homes and environments now offer us devices, applications and services for our comfort, convenience, and social connectivity and also providing services such as monitoring elderly subjects and healthcare (Chan, 2008) [9]. The concept of smart cities uses the availability of the ICT infrastructure and the human capital and use the plethora of generated information in the process of urban development and management (Caragliu et al, 2011) [10]. All the above mentioned concepts based on the smart devices operating in networks to collect, store and process data online and offline. So far not many attempts have been drawn to extend these concepts to support emergency management. Therefore the future challenge is to utilize the data and information from the emerging gigantic network of smart devices known as the internet of things (IoT) for the benefit of emergency response. In this context the next concern is the quality of data and information from the gigantic network of smart devices which is still in the infant stage in the context of emergency management.

2.4 Disaster data standards and format

Disaster data response requires large amounts of data, but the data are produced by different organizations and stored in different formats, which brings difficulties for agencies to perform data sharing and interoperability. Standards and formats of disaster data are the key issues for collecting and using disaster data efficiently. Although the sharing of data resources in networked cooperation has become standard practice in some fields, particularly in the more economically developed countries, in many cases researchers and their institutes experience too much uncertainty and barriers to make the most effective use of the new possibilities[11].

There are many disaster databases around the world. Some of them are playing important roles in disaster reduction, such as the global historical disaster database DesInventar [12] (<http://www.desinventar.org/>), global level emergency disaster database EM-DAT (<http://www.emdat.be/>), space disaster events and loss database in the United States [13] (<http://sheldus.org/>), the reinsurance company database, NatCatService¹ and Sigma (<http://www.swissre.com/sigma/>). The need for systematic data for disaster mitigation and prevention has been an increasing concern of both development and response agencies [14].

Different databases have different standards for disaster data management and storage. The organization that serves the corresponding database usually follows the specific standards to collect and manage disaster data. For example, the EM-DAT disaster database makes entry criteria, the data should meet at least one of the following three conditions: 1) 10 or more people deaths; 2) 100 or more people affected/injured/homeless; 3) Declaration by the country of a state of emergency and/or an appeal for international assistance [15]. Other disaster databases also used their own standards. EMA (Emergency Management Australia) disaster database requests that the disaster must cause more than 3 deaths or more than 20 injured or at least 10 million Australian dollars in total loss. Although these databases have different disaster data standards, these data standards have a clear and concise definition. This is very important for the use of these disaster data. In addition, disaster data standards need to meet demand for disaster reduction activities. And many key components such as regional scale, time scale, accuracy of information, timeliness of information, and the comprehensiveness of disaster should be considered in the standards. Different disaster databases comply with their respective standards, which makes different disaster databases are hard to realize data sharing and interoperability.

To be efficiently shared during the disaster mitigation, the data should be in compliance with a certain standard and be stored in common formats. The categories of disaster data can be divided by three ways as follows. First, from performance point of view, disaster data includes graphical data (such as topographic maps, plan, layout drawings, point notes, structure charts, images, and so on), text data (such as descriptive text, various statistical reports, attribute data related

¹ NatCatService (<http://www.munichre.com>)

to geographical entities, sound, and so on), etc. Second, from data carrier point of view, disaster data includes traditional paper graphics, tables, documents, etc. In addition, they also include a variety of graphics, charts, and documents which are stored in a computer and recordings and videos which are stored in cassettes or CD-ROMs. Third, from data sources point of view, disaster data includes basic geographic data, ground observation data, ground survey data, model simulation data, historical data, census data, disaster report and integration, etc. According to different data categories and data collecting process, there are many regional or global standards due to the diversity and multi-disciplinary of disaster data.

According to different stages of disaster data acquisition, management, analysis and application, there are several kinds of standards as follows. For disaster classification, there are a variety of disaster events such as flood, earthquake, drought, hurricane, etc. How to define and classify a disaster event is the most important thing for disaster mitigation around the world. The standard such as Peril disaster classification proposed by IRDR (Integrated Research on Disaster Risk) gives a specified classification of disasters. For disaster data collection, in the process of disaster, all kinds of methods are used to collect disaster data. But different people may collect data in different ways and get different data fields. To standardize disaster data collection, the UN Sendai Framework defined a Data Collection Protocol. For disaster data storage, in order to make the data sharing and data exchange of different disaster databases, the standards on data store and management should be defined. The disaster databases such as EM-DAT, DesInventar, NatCatService all defined the database related standards. For disaster data access and interoperability, the international organizations for standardization such ISO, OGC (Open Geospatial Consortium) have defined many standards on geographical data sharing and interoperability, which can be used for disaster data access and data exchange. Some typical disaster data related standards are listed in table 1.

Table 1. Some disaster related standards

Standards category	Name	Organization	Scope	URL
Disasters classification	Peril Classification [16]	IRDR	global	http://www.irdrinternational.org/2014/03/28/irdr-peril-classification-and-hazard-glossary/
Data Collection Protocol	Data Collection Protocol presented in the 5th GP-DRR, 2017	UN Sendai Framework(UNISDR)	global	https://www.unisdr.org/we/coordinate/sendai-framework
Framework	National Disaster Recovery Framework	FEMA	regional	https://www.fema.gov/media-library/assets/documents/117794

OGC data services	WMS,WFS,WCS,sensorweb, Opensearch-geo	OGC	global	http://www.opengeospatial.org/
Flood	FEMA Policy Standards for Flood Risk Analysis and Mapping	FEMA	regional	https://www.fema.gov/media-library/assets/documents/35313
WHO standards on disaster	Classification and minimum standards For foreign medical teams in sudden onset disasters	WHO	global	http://www.who.int/hac/global_health_cluster/fmt_guidelines_september2013.pdf
Disaster loss data sharing	Guidance for Recording and Sharing Disaster Damage and Loss Data	JRC	EU	http://publications.jrc.ec.europa.eu/repository/bitstream/JRC95505/lbna27192enn.pdf
ISO /TC211	Geographic information/geomatics	ISO	global	http://www.iso211.org/
GEO	GEOSS framework	GEO	global	https://www.earthobservations.org

3. Big disaster data collection and transmission

Rapid urbanization and continuous change in climate has resulted in increment of disaster risk. Whenever disaster occurs various incident occurs simultaneously like traffic congestion, road blockage due to landslide or any accident etc. which adheres rescue and response mechanism. The most important part is collection and transmission of situation report or correct to right person or officials so that the time of response can start without wasting single second. The role of data collection and transmission is one of the most important things in mitigating effects of disaster.

The high rate of evolution of disaster risk has led to decrement in proper submission of important resource mapping data. Prevention of disaster risk and hazards assumes an increasingly important role in order to keep these values high enough to satisfy society needs. The possibility of developing and using big disaster data collection and transmission leads to even a single person of society to do its

routine while monitoring disaster risk parameters is a strong advantage in prevention and mitigation mechanism of DRR(Disaster Risk Reduction). This supports the idea that there is a lot of potential in developing big disaster data collection and transmission for remote monitoring. The objective of big disaster data collection and transmission is to develop an effective mechanism for data transmission between a monitoring device, which does the acquisition of DRR parameters, and a remote receiver, which allows storage and access to the acquired data, for instance, by a rescuer geographically distant from the victims .

Now the important parameter understands clarity in state of the art regarding the technology which can be integrated in analyzing big disaster data collection and transmission. The state of the art that can be used are smartphones with wireless communication which is very common nowadays and one in every two person holds this technology with him. The application of android mobile platform and Bluetooth can be easily used in big disaster data collection and transmission. For example, Android mobile platform application is targeted to smartphones and tablets that run the operating system Android. These android devices are becoming increasingly useful and are a symbol of technological progress wherein the compilation of big data and transmission can be done in no time. Android is one of the operating systems used in tablet and smartphones, and now is in possession of the multinational company, Google. There are other operating systems such as iOS, kindle, etc. who's application can also be used in big disaster data collection and transmission. Android is already being used in providing solutions to various users like weather condition, maps, routes, places, early warning systems etc. which is helping to analyze real and correct information. There are some developments with focus on acquisition and visualization of risk & hazard parameters on smartphones .At a commercial level, there are android applications for various purposes and the big disaster data collection and transmission is clearly an area with plenty of incidence. However, the main focus is not telemetry systems and the applications are mainly just informative or a guide for the user in question related to certain steps in disaster risk reduction.

Another application that uses short wavelength radio transmission for exchanging data wirelessly and without internet connectivity over short distances will be very effective in big disaster data collection and transmission is Bluetooth. It is basically low cost , low power , wireless technology for exchange of data information over short distances & it can be another technology which if very much useful in big disaster data collection and transmission from information sharing device to tablets/smartphones. The only thing which is much required in big disaster data collection and transmission is a well-built transmission technology which will make the system much stronger to keep data safe and can be used effectively, efficiently & quickly for taking DRR measures. So, to club devices that have Bluetooth technology, Android and a transmission technology will provide everything to collect big disaster data and transmit it accordingly when it is required.

4. Big disaster data processing

Authorities dealing with disaster risk reduction has to make decisions that range from one-off strategic decisions to monthly and weekly tactical decisions to high-volume, high-speed front-line operational decisions. Usually the effects of primary disaster don't end quickly rather it gets converted into secondary disaster. The common example is result of building collapse due to an earthquake after few days. The requirement of time constraint decision support is the most important parameter in order to mitigate the effects generated due to primary disaster. When the emergency alarm starts the time based mitigating measures can only be effective if the authorities will be able to read the provided data from the database and time based mitigating measures implementation policy works.

4.1 Competence & hypothesis for disaster risk reduction data processing

Geographic Information System (GIS), mobile devices, cloud computing, social media, sensors, and cameras are now ubiquitous and produce massive amounts of changing data. To extract the maximum value from dynamic and perishable disaster risk reduction data, authorities need to process data much faster and take timely action so as to save the lives of thousand people. Whether communicating to first responders and officials, offering proactive support, detecting and preventing risk, or managing the Internet of Things, real-time decision-making is essential. Responding in real time requires systems to make operational decisions automatically. Acute, Systematic, and Robust Decision Management Systems combine proper rules and predictive analytics to render tailored recommendations. Event processing adds correlation & pattern detection on a scale of millions of risk events and big disaster risk reduction data streams in motion at microsecond speeds.

Event-based decision management systems enrich event-based disaster risk reduction data with traditional and big data sources and determine when and why a real-time response might be required. By leveraging decision engines based on DRR rules and analytics, decision management systems can determine what the best and most effective response is. For authorities to respond in real time in case of early prediction or warning, it needs to acquire or develops decision management systems to capture, filter, and analyze data, and make decisions in real time. Such systems need to be able to rapidly determine that a response is required and also intelligently determine what the relevant and appropriate response should be—they need to decide when and how to act. As the world population is increasing rapidly, authorities increasingly need to ensure that response mechanism is delivered in real time so more event centric DRR decision management systems are required.

The combination of DRR real-time decisions and risk event processing in case of disaster, delivers the core capabilities for building a real-time, hazard or risk-based Decision Management System—correlating events, managing decision logic, embedding predictive analytics, optimizing results, and monitoring and improving decision making. Key features of the solution include certain competence and hypothesis as shown in Table 2:

Table 2: Competence & Hypothesis for Disaster Risk Reduction Data Processing

Competence	Hypothesis
Event correlation, Disaster Risk Reduction rules, and predictive analytics in combination	Some real-time response solutions focus on event correlation, on DRR measures, or on predictive analytics. With DRR Event Processing and Real-Time Decisions working together, the solution balances these capabilities, maximizing the flexibility and power of the decision-making systems that can be built with it.
Scalability and flexible distribution	Extreme scalability and deploy ability for disaster risk reduction measures is very much required. The architecture, event detection and correlation are widely distributed and close to the event source. Therefore its monitoring is very much needed for structuring disaster risk reduction strategies. This improves responsiveness and contextual awareness while lowering latency. At any point the Real-Time Decisions can be invoked or used to learn from the patterns detected due to its flexibility of deployment and its support for a mutually shared approach.
Broad support for divergent environments	It is very much required to support externally managed DRR data, content, response & mitigation rules, and predictive analytic models. A wide range of event sources can be handled and event handling is extensible with Java.
Adaptability and robustness in the face of change	The overall solution is very robust in the face of ongoing change. Event Processing allows new patterns and queries to be deployed to a live instance while Real-Time Decisions allows similar changes to DRR rules and analytic models. Real-Time Decisions for disaster risk reduction provides support for analytic models that are based on risk finding & hazard mapping, while automation of the full analytic lifecycle allows professionals to use hundreds of regularly updated DRR based analytic models in mitigating disaster affects.
Extensibility	DRR event processing provides new functions and capabilities that can be made available for pattern-matching analysis for disasters. The support of Real Time Decisions for external DRR rules and analytic models allows the decision makers to achieve extensible prospects in DRR planning.

The first step in developing a DRR real-time response solution is to configure the event processing engine. This involves identifying the message streams in the

environment that comes from disaster spots or locations. These message streams are fed into the optimization engine and are correlated with different disaster situations. Various disasters related practices are defined to process these events, and a DRR event processing network is defined. After the event processing engine is configured, it can be connected to the decision engine. This involves defining the options to be selected between and the hazard & vulnerability measures that will allow the decision engine to choose the best option. Any DRR rules that constrain the choices are specified and analytic models are built to predict behavior and segment risk.

In case of any hazard increment or risk enhancement the DRR decision engine is invoked to determine the eligible choices and identify the best choice for taking mitigation steps, given the performance measures and analytics. These recommendations from hazard or risk prospective patterns can be fed back to the disaster event processing engine or passed out as a response. The DRR decision engine then closes the loop by recording decision performance information. The decision engine automatically divides activity into test and control groups, and the competent authority of the decision decides how the activity will be split between the test and the control group; this logic can be adjusted over time as necessary. Outcomes from DRR decision-making are fed back into adaptive analytic models and used to monitor the overall performance of the system for taking DRR measures.

Social media has become a very fast and efficient mechanism for disaster data collection. For example in case of occurrence of any earthquake the reporting mechanism now days gets activated immediately as soon as any information transmitted from any user which not even the part of disaster risk reduction system. The features of social media used for disaster data collection and transmission are listed as follows.

- The most update to date data: The common public at any place wherever risk or hazard exists source of the information that provides a completely new set of reference to data researchers. With every post, conversation and site or app visit, a user leaves behind pieces of information about themselves. The data transmitted tells simple demographic information to robust details like coping capacity, damage assessment, impact, risk information, early warning system.
- Instant data: Every organizations has some historical data on which basically data transmission is based upon for example census data for fetching male to female ration or how many children exists in any particular area. But in case of disaster to check upon the instant data is very much required to get a clear set of picture of actual situation .So that data produced by local publics using social media plays a vital role about situation analysis and data transmission which is very much correct for starting disaster response and operation mechanism.
- Fine-grained data: Social media should be the best representation of people because it well represents their beliefs, attitudes and actions. No other source of data provides the same kind of granular detail into a person's life in terms of disaster reduction measures. Authorities can use social media to get personal information in an emergency situation to save lives.

4.2 Big data processing for disaster management

Big disaster data is processed and analyzed to develop perception comprehension and projection of the emergency event as explained in section 3. Generally data analytics are classified into four categories depending on its goal and purpose viz. descriptive analytics, diagnostic analytics, predictive analytics, and prescriptive analytics. In emergency response phase, perception of the emergency related incident can be developed through descriptive analytics. To develop comprehension both diagnostic analytics and prescriptive analytics can be used where diagnostic analytics recognize the causes for the incident and prescriptive analytics determine which actions to be taken to manage the situation. Projection can be developed through predictive analytics regarding the future.

Response and recovery phases of disaster situations require effective processing of historically collected information (government data, open data, linked data etc.) as well as big data coming from various channels like sensors, satellites, crowd sourced information (social media feeds, photos, video) and cell phone GPS signals. Therefore, the foremost challenge is to real-time integration of the archived information and big data streams through seamless interaction and collaboration among the different platforms. Further, it is necessary to have data integration capabilities and data fusion (Haghighat et. al 2016) [17] capabilities to integrate multiple distributed and heterogeneous data sources to produce more consistent, accurate, and useful information to support rescue operations. The necessity of handling larger data volumes with different data formats such as structured, unstructured, and semi-structured data with high velocity constraints limitation of human interpretation and overwhelm decision-makers. Hence advanced data querying and analytics such as machine learning techniques and in memory computing will be required.

An important use of crowd-sourced data is event detection where significant incidents are recognised by analysing crowd-sourced (Saran et. al 2017) [18]. In emergency response systems real-time event detection should be performed very fast without compromising for precision. The advancements in location inference techniques using geo-tagged social media data would provide useful tools for precise identification of the locations (Laylavi, F., 2017) [19]. Further, the information regarding events and their locations should be used to develop the crisis maps to conduct rescue operations and plan recovery strategies. This geo-referenced information plotted on maps should get updated continuously as new information received and events unfold (Beatson et. al 2014) [20].

The effective management of the events also requires the collaboration and coordination of a range of government decision makers, emergency response stakeholders, and community-based non-government organizations. Therefore, the availability of real-time location-aware information, as well as the capabilities to effectively integrate and utilize available information with different autonomous agencies is key to effective decision making and resource deployment to respond to crises (Fosso Wamba et al., 2012) [21]. In most situations response teams have to take decisions based on incomplete and inaccurate information. This may be due to limited

availability of data network resources. In the case of Tohoku Earthquake, it has been reported that the number of outages of communication facilities such as access lines and cellular base stations increased during the first 24 to 48 hours after the strike (ref). Therefore the mechanisms should be deployed to achieve high availability of the networks and systems. Re-configurability of the network, network virtualization, cloud based systems that are not affected by the damaged infrastructure etc. are essential aspects to consider in developing emergency management systems.

Predictive data analytics can play a wider role in disaster readiness and reduction phases. Natural disasters are extreme and unexpected phenomena resulting from natural processes of the earth and atmosphere. Therefore, predictions of natural disasters are a major aspect in improving the readiness. Though a massive amount of data available in almost all disciplines related to disasters such as geosciences, weather medical insurance etc. the knowledge extraction from such massive data cannot always be performed by using standard statistical techniques. Therefore, it is necessary to use new approaches such as recognizing relevant patterns of natural disasters through automated machine learning techniques etc. to make predictions. Such work is not trivial

5. Big disaster data quality control

Big data is a new concept, quality concepts related to big data are still in its infancy and hence the quality criteria (dimensions), the data quality management principles and methodologies are not yet matured (Cai and Zhu, 2015) [22]. Since big data has the specific three characteristics volume, velocity and variety, specific challenges are faced when assuring the quality of data. Hence traditional data quality dimensions and frameworks become obsolete and new challenges emerge.

5.1 Challenges of disaster data quality control

The large volume and high velocity of crowd-sourced data requires an instant quality assurance process within a shorter period of time. Data from social media such as Facebook, Twitter Instagram and various other communication channels such as Viber, Whatsap multimedia messages and text messages are so huge and therefore, Harvard Humanitarian Initiative (2011) [23] reveals that a raging river of information is created. In order to use them for decision making, the information should pass the quality checks pertaining to dimensions such as accuracy, currency, completeness etc. Therefore, the disaster responding systems should have the capability to validate the data and information before use for decision making. Sources such as GIS data and extra information of the device and the timestamp of the message can be used to perform validations. Further, the verification services can be built to send the collected data back to general public and requesting feedback from them.

Due to the rapid changes happening in the environment during disaster situations, some big data remains valid only for a very short time (for example readings of thermal sensors, flood level etc.). If such data is not collected and used in real time, then they may obtain outdated and invalid information. Processing and analysis based

on these data will produce useless or misleading conclusions. Hence special capabilities should be built to handle the quality of time sensitive data. Further, due to multiple sources and the rapid changes, the necessity for entity resolution and record linkage will grow, and matching of entities relies on good reference information for similarity scoring and linkage. Therefore, it is necessary to develop common reference domains and provide an environment for capturing and sharing disaster related terms, data element definitions and semantics.

The diversity of data sources brings variety of data types and complex data structures and hence increases the difficulty of data integration. For example, unstructured data, such as documents, video, audio, etc. and semi-structured data, including software packages/modules, spreadsheets, and financial reports and finally structured data from various databases. Semi-structured and instructed data play a major role in disaster situations and hence need innovative quality criteria to assess them. Another aspect in using crowd-sourced data is translations and interpretation of meaning in multilingual environments. The quality of data depends on how well the translations and interpretations match with the real world disaster context that help saving lives and relief operations effectively and efficiently.

Tapia et al. (2013) [24] argues that in disaster context crowd-sourced data may never meet the standards of quality required in situations such as search and rescue operations, while in others, such as resource and supply management they may be useful as long as appropriately verified and classified. Goodchild and Glennon (2010) [25] reveals that many aspects of quality problems in data when responding to disasters need further research since not much attention has been paid on this aspect.

Identifying the purpose and who is going to use big disaster data is the most important component of big disaster data quality control and assurance. The methods of examining should be very much precised and should be directed by purpose of monitoring program and the quality of disaster data that somebody needs. The level of disaster data quality assurance will depend on its end uses at the time of crises, and the measures taken to check reliability of disaster data for its proposed purpose. For some monitoring groups, the main objective is identifying earthquake risk for the local community or school, where the focus is on earthquake mitigation awareness to an issue rather than producing high quality disaster data on some other types of disaster. However, groups that collect data to inform decision makers or as part of an integrated monitoring program with the authorities, research organizations, regional bodies and state agencies must take measures to ensure the data are credible and reliable. When appropriate quality assurance and quality control measures are implemented on big disaster data, we can be confident that competent authority's decisions are based on sound and reliable data.

Certain parameters of consideration in dig disaster data quality control and assurance is disaster data from contamination, illegal data filters, outlier detection, checking test assumptions with normal probability plots, diagnostic measures and GRUBBs test. All these parameters should be taken into consideration for quality control and assurance for big disaster data. Data used for disaster risk reduction projects for visualization, analysis, compilation, and sharing should meet a defined

standard for quality. Data quality requirements vary from project to project. Requirements for how accurate or complete a dataset needs to be are based on how the data will be used. Requirements are also influenced by technical, product, and location requirements. It is very much required to perform quality control and quality assurance test, properly work through the identified issue and quality control related report generation. All these issues can be implemented as a tool for assessing big disaster data quality control and assurance.

Apart of certain tools GIS can become a very helpful platform in big disaster data management. GIS provides a platform for the management of geographic data and disparate documents (plans, photographs, etc.) necessary to meet the emergency management mission. GIS provides a capability to access information based on the geographic location to which it pertains, allowing users to get various types of information from the map display. GIS will be an efficient tool in data quality control and assurance and it will help to get a real time picture for at kind of disaster risk mitigation exercise. The graphs and pictograms also help to check big disaster data quality. Dynamic data (camera feeds, weather, traffic, hospital status, automated vehicle location (AVL), incidents, sensors, etc.) to provide situational awareness for decision support are some points used in GIS which provides better quality disaster data and hence assurance of data can be gained up to a certain levels.

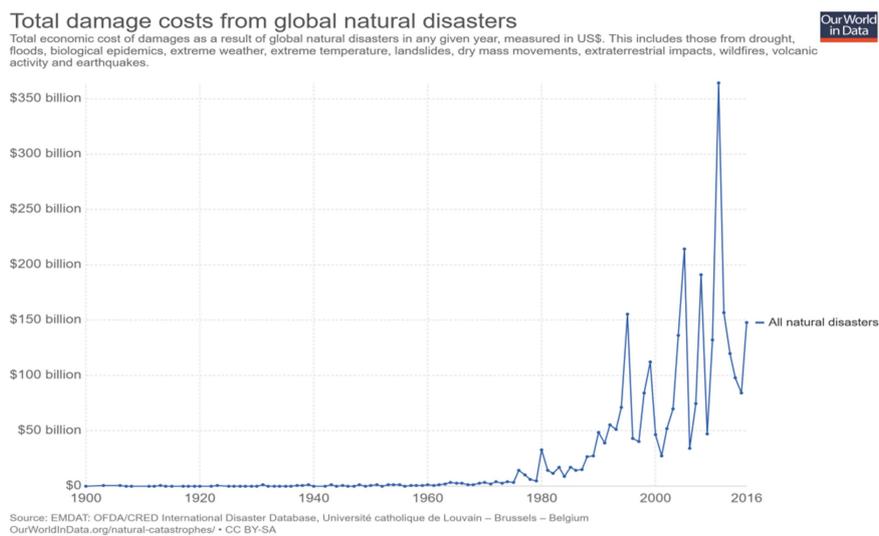


Figure 1a: Total damage cost from global disasters

Source: EMDAT: OFDA/CRED International Disaster Database, Université Catholique De Louvain-Brussels- Belgium

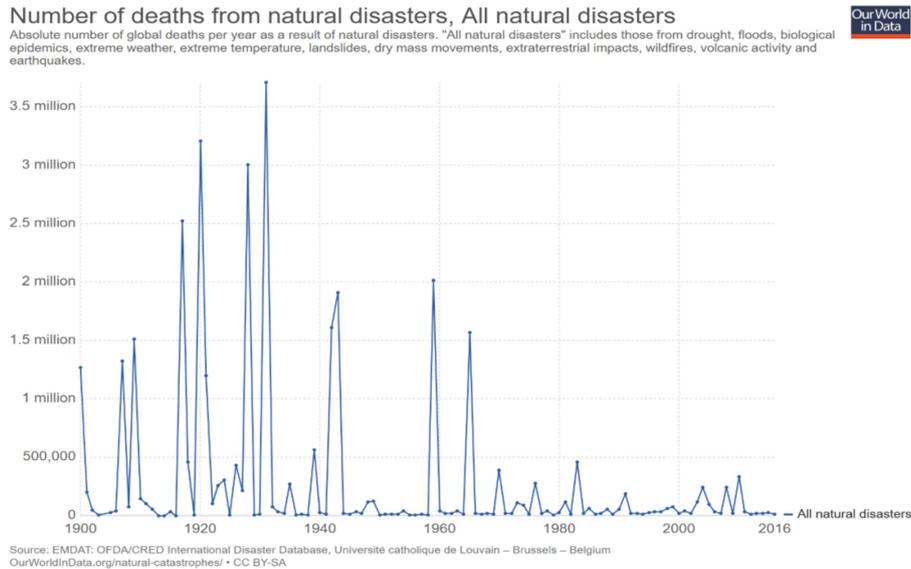


Figure 1b: Number of deaths from natural disaster

Source: EMDAT: OFDA/CRED International Disaster Database, Université Catholique De Louvain-Brussels- Belgium

Figure 1a and 1b can be used to analysis quality of disaster data wherein the first figure shows the damage assessment and next figure shows mortality data. These all parameters can be integrated to get better quality and assurance of disaster data.

The user and system design requirements from the aspect of big disaster data quality control and assurance needs an interface to control big disaster data documentation. The interface document describes the internetworking of two separate platforms in information dissemination and big disaster data quality control and assurance. For example, control room having communication interface in disaster risk reduction i.e. elaborated in terms of data items and messages passed as which rector scale earthquake had come in some which area, messages passed, instruction and SOPs followed, time of occurrence of disaster and streamlining post disaster risk reduction mechanism. The interface document in disaster risk reduction and big disaster data also describes the interaction between two different sources and to gather correct information at one place. This interface is also very similar as other communication interface wherein the role of a user and the system, software component and a hardware device or two software components are interrelated. This class of document is typically used where complex interfaces exist between components for disaster risk reduction that are being developed by different persons sitting at different places. It is jointly prepared by the interfacing groups.

The big disaster data requirements describe what the interface is to achieve together with any hindrance on its design as follows:

- Identification of the disaster interfacing systems/sub-systems
- The reason for the disaster interface's existence including the user requirement that is satisfied
- A description of what the interface does for disaster risk reduction
- Specification of the information to be exchanged for mitigating disaster risk

- Timing and sequencing constraints of pre as well as post disaster
- Capacity and performance requirements for disaster risk reduction
- Requirements for communications protocol standards compliance to mitigate effects of disaster
- Identification of any safety requirements discovered in an Disaster Interface Hazard Analysis

5.1 Disaster data interface between different entities

The design in big disaster database describes how the interface will be implemented for disaster risk reduction. For example in the case of a control room interface the following technical details are provided.

A description of the control room communications protocol may include:

- Message format for hazards, risk, vulnerability and description including user error messages, user information messages and inter process messages
- Message component names (like earthquake, tsunami etc.)
- Message initiation (Heat wave or Cold wave started)
- The processing of message interruptions. Fragmentation and reassembly of messages
- Error detection, control and recovery procedures
- Disaster Data Synchronization, including connection establishment, maintenance, termination and timing and sequencing for disaster risk reduction mechanism
- Disaster risk reduction data transfer rate
- Data transmission services including priority and grade
- Disaster Data Security including encryption, user authentication and auditing
- Error codes in Disaster Data

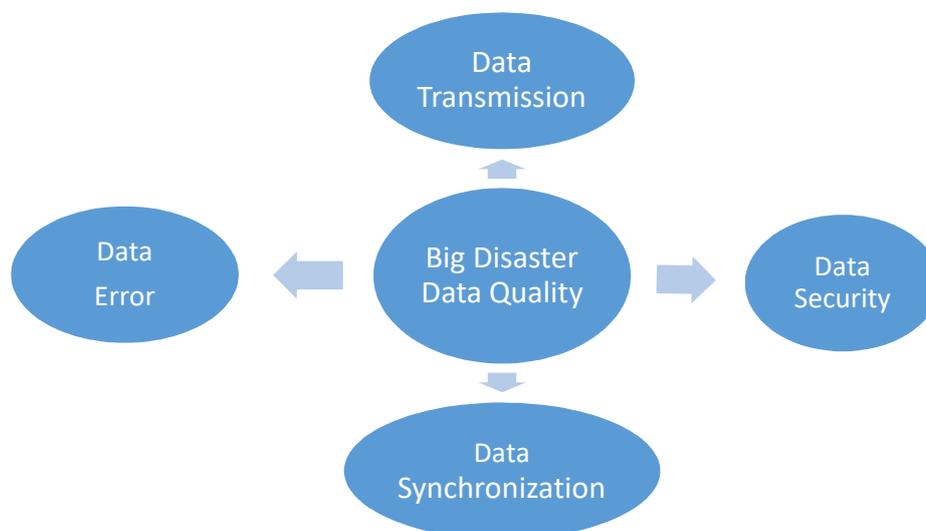


Figure 2: Interrelation of disaster data quality

Above figure 2 shows the impact of transmission, errors, security and synchronization in maintaining big disaster data quality .The same can be processed in terms of other parameters. Data quality is a perception or an assessment of data's fitness to serve its purpose in a given context. Various aspects of big disaster data quality include certain other parameters of considerations: Accuracy, Completeness, Update status, Relevance, Consistency across data sources, Reliability, Appropriate presentation, Accessibility. In the following section, we will further discuss the issues related to disaster data visualization.

6. Big disaster data visualization

Chronically, dealing with disasters focused on emergency response, but from the beginning of the 21st century it was increasingly recognized that disasters are not natural (even if the associated hazard is) and that it is only by reducing and managing conditions of hazard, exposure and vulnerability that we can prevent losses and alleviate the impacts of disasters. Since it is not possible to minimize the severity of natural hazards, the main task for reducing risk lies in mitigating vulnerability and exposure. Reducing these two components of risk requires identifying and reducing the underlying drivers of risk, which are particularly related to poor economic and urban development choices and practice, degradation of the environment, poverty and inequality and climate change, which create and exacerbate conditions of hazard, exposure and vulnerability. Addressing these underlying risk drivers will reduce disaster risk, lessen the impacts of climate change and, consequently, maintain the sustainability of development. DRR is a part of sustainable development, so it must involve every part of society, government, non-governmental organizations and the professional and private sector. Big Disaster Data visualization enables to create a culture of prevention and resilience for implementing disaster risk reduction mechanism. Big Disaster Data visualization will also provide decision makers to build strategy for multiple, cascading and interacting hazard.

6.1 Data visualization for disaster management

A picture is worth a thousand words – especially when you're trying to find relationships and understand risks & hazards data, which could include thousands or even millions of variables. Big Disaster Data Visualization has become the de facto standard for Disaster Risk Reduction. Data visualization tools in Disaster Risk Reduction have been important in democratizing data & analytics and making data-driven insights available to disaster professionals throughout an organization. They are typically easier to operate than traditional statistical analysis which was not software based and rather we had to depend on written data which partially correct. This has led to a rise in lines of hazard mapping, vulnerability analysis & risk reduction implementing data visualization on their own, without support from IT. Data visualization software also plays an important role in big data disaster visualization and advanced analytics projects. Big Disaster Risk reduction measures accumulated massive troves of data during the early years of the big data trend about

various hazards, related risks & disasters; they needed a way to quickly and easily get an overview of their data. Visualization tools were a natural fit.

Visualizing big disaster data is crucial when the disaster risk comes between big and impersonal, and small & intimate data will blur as we've never seen before. The greater the quantity and kinds of disaster data collected, the more we need to experiment with how to make it unique. The requirement in mitigating disaster risk to start with a plain sheet of paper and then experiment with custom visualization. Even very minute details from your process which everyone think is of no use can play and enhance basic charts to reveal more about disaster risk reduction measures.

The ideal process to create big disaster data visualization is to equip all the disaster related data in a specific tool , pick it from unconventional sources and frame it so that the complete task can be completed within a couple of clicks. Simplified solutions are much required to solve big disaster data collection and transmission methodologies. To get clarity it is much required that all the big disaster data should be compared, synthesized and concluded, then transmitted in a single piece of information to the users. Big disaster data-driven doesn't mean unmistakably true because data and the tools that collect it are human-made. DRR data is not pure fact, but evidence that filters reality in a very subjective way. The intention is not to get a black and white paper rather a good visualized data that delivers a meaning to people reading that information.

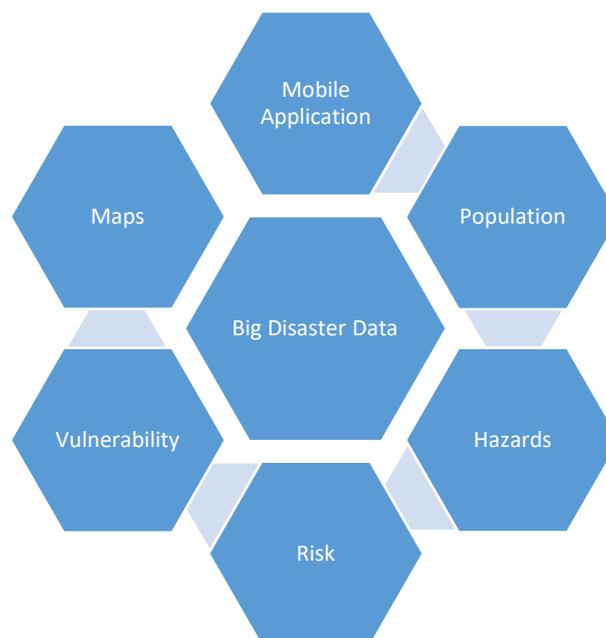


Figure 3: List that is required to fetch big disaster data

As shown in Figure 3 the approach of data visualization is very much analyzed and compiled by certain important parameters like the usage of android, population data, hazard analysis, risk mapping, vulnerability studies, maps etc. and finally combining all the parameters to form a visualized big disaster data. The same approach can be prepare for traffic congestion data as number of vehicles on road, population travelling as well as nearby population as shown in Figure 4. The

important parameters always start with hazard analysis and risk mapping along with vulnerability study.

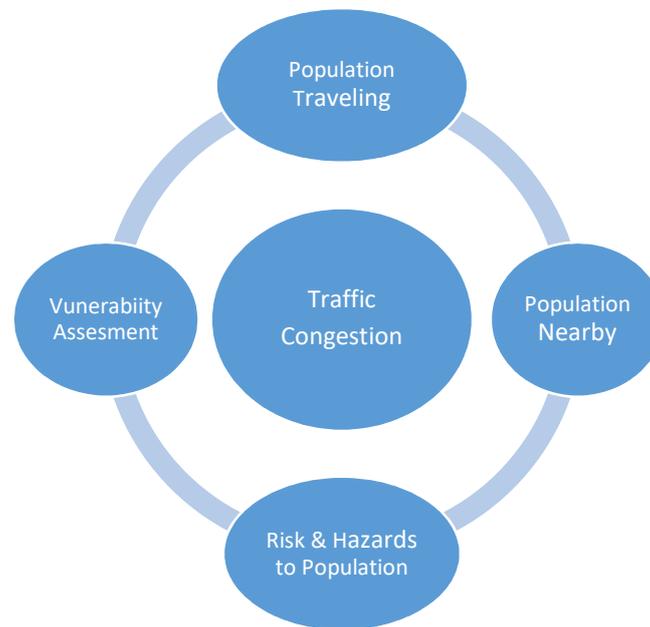


Figure 4: Data visualization for traffic congestion

Big Disaster Data visualization technologies can be as powerful as they are easy to use, allowing decision makers to quickly and easily understand, articulate and share the insights across the organization to others. The main objective will be easy sharing of information and channelize the response mechanism as rapid as possible.

Big Disaster Data Visualization is central to advanced analytics for similar reasons. When a data scientist is writing advanced predictive analytics or machine learning algorithms on hazards, vulnerability, risks & disasters as shown in Figure 5, it becomes important to visualize the outputs to monitor results and ensure that models are performing as intended. The main reason is because visualizations or understanding of complex algorithms are generally easier to interpret than numerical outputs. Most of today's data visualization tools come with connectors to popular data sources, including the most common relational databases, Hadoop and a variety of cloud storage platforms. The visualization software pulls in data from these sources in case of normal data management and applies a graphic type to the data or in broader term disaster data.

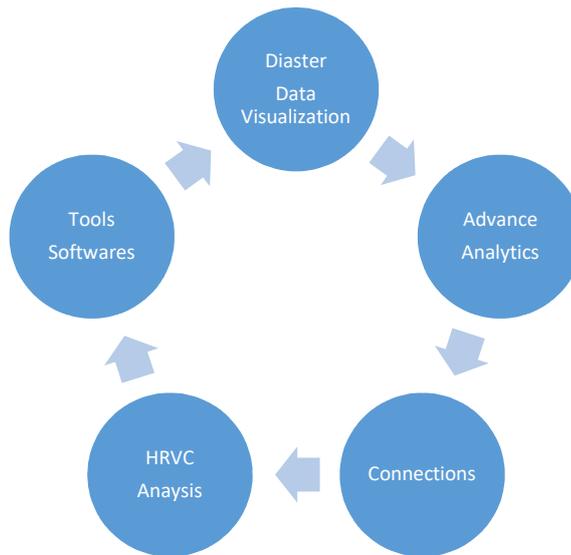


Figure 5: Big disaster data visualization cycle (HRVC* – Hazard, Risk, Vulnerability & Capacity)

The same technology can be used to visualize the big disaster database .One disaster several times gets converted into another disaster so get unique integration it is very much required to use the tools for data visualization specially for disaster risk reduction. Big Disaster Data Visualization will provide the best of both worlds: specific, precise, uniform analytics you can rely on and the ease-of-use and speed you need to change it. Disaster data visualization will explore parameters like

- Search for response, recovery & rehabilitation points and get to data quickly thanks to guided navigation,
- Analyze disaster data, anywhere with instant mobile,
- Highlight one visual to automatically see disaster risk related information in the others
- Add new data and see intelligent updates as you go
- Use pre-built data connections to load and integrate data from a wide variety of sources
- Combine information to uncover new insights
- Capture insights and add comments to create visual stories
- Collaborate with coworkers simply by sharing disaster risk reduction stories

Big Disaster Data Visualization is very important step to mitigate the disaster risk because of the way the human brain processes information, using charts or graphs to visualize large amounts of complex data is easier than poring over spreadsheets or reports. For example if there is an earthquake then the probability of landslide also comes into picture and to develop an understanding of what is required & how to handle this situation data visualization can become very efficient tool to any person who is not even present at that particular place. Big Disaster Data visualization will be a fast, feasible, easy way to deliver concepts in a universal manner – and we can

experiment with different disaster related scenarios either natural or man-made by making slight adjustments.

Sense-making (also called data analysis) and communication is the graphical display of abstract information for two purposes in data visualization & understanding. Important stories live in our data and data visualization is a powerful means to discover and understand these stories, and then to present them to others. The information is abridgement in that it describes things that are not physical. Statistical information is mostly theoretical. Whether it concerns hazards findings, vulnerability, risks or gap analysis, capacity building, or anything else, even though it doesn't pertain to the physical world, we can still display it visually, but to generate such outcome we must find a new way to give form to that which has none. This translation of the abstract into physical attributes of vision (length, position, size, shape, and color, to name a few) can only succeed if we understand a bit about visual perception and cognition. In other words, to visualize disaster data effectively, we must follow design principles that are derived from an understanding of human perception in disaster risk reduction.

Generally data visualization features key relationships between quantitative values; it can also display certain relationships that are not quantitative in nature & can also derive some unique features. For instance, the connections between people on a social networking site such as Facebook, applications like Whatsapp which is near the disaster site or between suspected terrorists who can create man-made disaster can be displayed using a node and link visualization. Even the mobile technology provides a gateway to get instant disaster data only by using the application which is inside that mobile device. So big disaster data visualization will create a great platform in helping the life and property of human beings by using such an efficient and effective technology.

The bigger challenges of big disaster data collection is that we rely on grass root workers for data sharing, who are never sensitized and educated for that particular hazard or risk data we desire. Intellectuals want to analyze disaster risk data and simulate modes but have no control over policies and mechanisms that exists to collect them. United Nations Office for Disaster Risk Reduction (UNISDR), KoboToolbox (A suite of tools for field data collection for use in challenging environments), EM-DAT (Platform for maintain international disaster database) and several country specific mechanisms exists for post disaster need assessment for estimating several parameters .All these platforms are used for short & long term loss estimation but has a limited scope on futuristic predictive role & recovery.

6.2 Information systems design for disaster management

All four phases of disaster management, readiness, response, recovery and reduction (MCDEM, 2007) [26], involve heavy use of data and information that belongs to all three categories of data, viz. structured data, semi-structured data and unstructured data (Batini and Scannapieco 2006) [27]. Further the data used during disaster situations can be classified as *Big Data*. As per Gartner, Big data is defined when the volume velocity and verity of data is in high scale (Kailser et al, 2013) [28].

For instance data supporting the response phase of a disaster generates significantly large volumes of data from sensor networks, satellites, social media (Prasanna and Huggins, 2016) [29] and photos, videos and GPS signals from cell phones and other multimedia devices. Also the generation of such data occurs relatively in a shorter period of time (Yang, Prasanna and King, 2009) [30] and thus creating a high velocity, further they focus on numerous varieties like weather, medical related, supplies, relief, warnings, traffic, transport etc. and thus belongs to multiple varieties.

Creating situational awareness (SA) is the foremost task in an emergency response phase. Situational awareness is defined as “*The perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future*”(Ref). In order to develop SA data should be visualised in an effective and efficient manner so that it supports three levels of knowledge viz. (1) perception (2) comprehension and (3) projection of the incident so that the incident commanders can plan and carry out the rescue operations. In order to develop each level of knowledge system interfaces are required for data visualization.

- Interfaces for Level 1 SA – Perception

The first step in achieving SA is to perceive the status, attributes and dynamics of relevant elements in the environment. Lack of basic perception on important information can easily lead to form an inaccurate picture of the situation. These interfaces will support an end user to maintain a global picture relevant to a particular job role at any given time of the incident (Yang et al, 2009) [30]. Thus, as shown in the Figure 6, with this type of interfaces, incident commanders will be able to have a high-level, summarized overview of the situation.

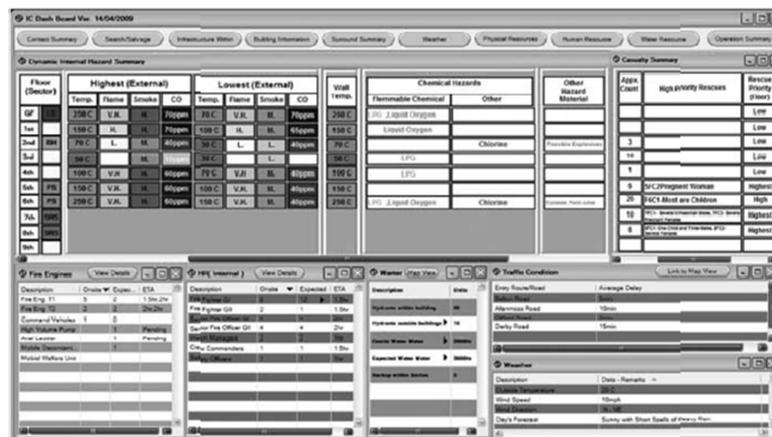


Figure 6: Interface supporting Level 1 SA of a IC: ‘Dash Board’. [30]

- Interfaces for Level 2 SA – Comprehension

At this level information obtained through observation is combined and interpreted. Rather than presenting a set of isolated information, mostly via numbers and text as in the perception level, as shown in the Figure 7, with this type of interface dynamic information is meaningfully integrated with static information using

graphical presentations. It provides the appropriate level of comprehension of the situation at any given moment of time to further improve the SA.

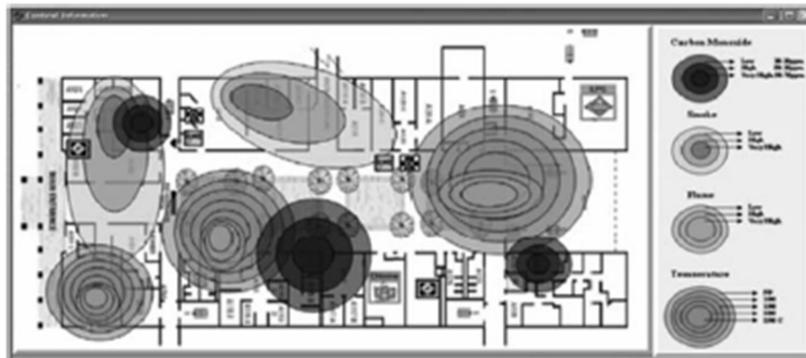


Figure 7: Interface supporting Level 2 SA of a IC. [30]

- Interfaces for Level 3 SA – Projection

By developing these interfaces, projection of future events is supported by providing the incident commanders with information on current and past trends on various situational parameters. Together with the above described Level 1 – Perception and Level 2 –Comprehension interfaces, as shown in Figure 8, this Level 3 – Projection interfaces are capable of providing a higher level of SA for incident commanders in making difficult predictions with confidence at any given moment of time of the emergency.

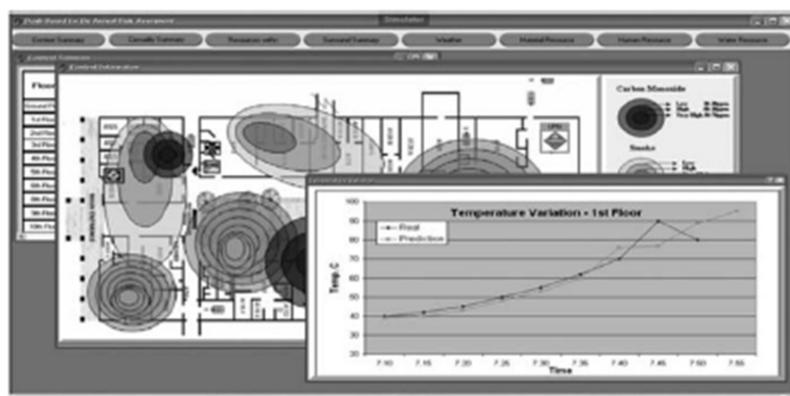


Figure 8: Interface supporting Level 3 SA of a IC. [30]

The system design requirements for data visualization in each level explained above will be different depending on the nature of the emergency or disaster. For example, in case of a fire fighting scenario nearly 350 different information interfaces have been proposed for the use of four fire-fighter job roles (Prasanna et. al 2013) [31]. Thus depending on the nature of the disaster, ground level data visualization requirements may change.

Visualization of disaster related data positioned on geographical data (area maps, building maps etc.) is an important aspect for all phases of emergency management.

GIS technology provides the capability to map and analyse hazards of all types and visualize their potential impacts. Thus, the emergency management systems should be capable of producing interactive maps such as vulnerability maps, operations maps, logistics maps, tactical maps, air deployment maps, transportation maps and incident prediction maps should be used in all walks of emergency management phases. Some examples of these information visualizations in maps are information about trapped persons, medical resources, damaged buildings, closed roads, and the availability and whereabouts of specific needs such as food, water and shelter (Beatson et. al 2014) [20].

Visualization of crowd-sourced information is another important aspect that should be considered in disaster response phase. With the emergence of web 2.0 tools like Twitter, Facebook and YouTube a massive amount of data is exchanged during disasters (Bruns and Burgess 2012) [32]. Majority of such crowd-sourced information is generated by the general public (Harvard Humanitarian Initiative 2011) [23] and are vital in developing real time live maps to produce and visualise a bird's eye perspective of what usually is a complex and often rapidly changing environment (Beatson et. al 2014) [20]. Organisations such as Volunteer & Technology Communities (V&TCs) needs and visualising hot-spots of activity within timeframe to mobilise large numbers of internationally dispersed volunteers in order to collaboratively solve informational and logistical management issues.

7. Conclusion

This white paper discusses the next generation disaster data infrastructure from four different aspects: disaster data collection, disaster data processing, disaster data quality control and disaster data visualization. In disaster data collection, sensor data and crowd-sourced should be considered together in order to have a comprehensive view of threatened areas. In addition, the availability of real-time location-aware information, as well as the capabilities to effectively integrate and utilize available information with different autonomous agencies is the key to effective decision making and resource deployment to respond to crises. Predictive data analytics can also play a wider role in disaster readiness and reduction phases. It is necessary to use new approaches such as recognizing relevant patterns of natural disasters through automated machine learning techniques etc. to make predictions. The government should also consider a cross reference platform for capturing and sharing disaster related terms, data element definitions and semantics. Also, when appropriate quality assurance and quality control measures are implemented on big disaster data, we can be confident that competent authority's decisions are based on sound and reliable data. Visualization of disaster related data positioned on geographical data is important all phases of emergency management. GIS technology provides the capability to map and analyses hazards of all types and visualize their potential impacts. Thus, the emergency management systems should be capable of producing interactive maps such as vulnerability maps, operations maps, logistics maps, tactical maps, air deployment maps, transportation maps and incident prediction maps should be used in

all walks of emergency management phases. Finally, governments should take into account all the above mentioned aspects and the needs of different categories of users and data dissemination in order to reduce the impact of natural disasters.

Reference

1. Liu, J. and E. Chu. *Dependability of active emergency response systems*. in Proc. DEPEND. 2015.
2. FEMA. *Integrated Public Alert & Warning System Open Platform for Emergency Networks* | FEMA.gov. 2017 Dec Available from: <https://www.fema.gov/integrated-public-alert-warning-system-open-platform-emergency-networks>.
3. Suri Vinay Sachidananda Abdelmajid Khelil, N., *Quality of Information in Wireless Sensor Networks: A Survey*. 2010.
4. Qin, Z., et al., *Quality-Aware Sensor Data Management, in The Art of Wireless Sensor Networks: Volume 1: Fundamentals*, H.M. Ammari, Editor. 2014, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 429-464.
5. OASIS. *EDXL-DE: Emergency Data Exchange Language Distribution Element, V1.0*. 2006; Available from: http://docs.oasis-open.org/emergency/edxl-de/v1.0/EDXL-DE_Spec_v1.0.pdf.
6. OASIS. *CAP: Common Alerting Protocol Version 1.2*. 2010; Available from: <http://docs.oasis-open.org/emergency/cap/v1.2/CAP-v1.2-os.html>.
7. Waidyanatha, N., G. Gow, and P. Anderson, *Hazard Warnings in Sri Lanka: Challenges of Internetworking with Common Alerting Protocol*. 2007.
8. Vladimir, B. *Virtual Building Environments - Applying Information Modeling to Buildings. in European Conference on Product and Process Modeling in the Building and Construction Industry (ECPPM) 2004*. Istanbul, Turkey.
9. Chan, M., et al., *A review of smart homes-Present state and future challenges* %J *Comput. Methods Prog. Biomed.* 2008. 91(1): p. 55-81.
10. Caragliu, A., C. Del Bo, and P. Nijkamp, *Smart Cities in Europe*. *Journal of Urban Technology*, 2011. 18(2): p. 65-82.
11. Uhler, P.F., et al., *Open Data for Global Science*. *Data Science Journal*, 2007. 6: p. OD36-OD53.
12. Marulanda, M., O. Cardona, and A. Barbat, *Revealing the socioeconomic impact of small disasters in Colombia using the DesInventar database*. Vol. 34. 2010. 552-570.
13. Tschoegl, L., *An Analytical Review of Selected Data Sets on Natural Disasters and Impacts*. 2006.

14. R., G.-S.D.B., *The Quality and accuracy of disaster data: A comparative analyses of 3 global datasets Working paper prepared for the Disaster Management facility*, World Bank, Brussels CRED, 2002.
15. EM- DAT. *EM- DAT criteria and definition [EB/OL]*. 2006; Available from: <https://www.emdat.be/>.
16. Integrated Research on Disaster Risk. (2014). *Peril Classification and Hazard Glossary (IRDR DATA Publication No. 1)*. Beijing: Integrated Research on Disaster Risk.
17. Haghghat, M., M. Abdel-Mottaleb, and W. Alhalabi, *Discriminant Correlation Analysis: Real-Time Feature Level Fusion for Multimodal Biometric Recognition*. IEEE Transactions on Information Forensics and Security, 2016. 11(9): p. 1984-1996.
18. Saran, J.F., et al. *Data and information fusion in the context of emergency management: The DF100Fogo project. in 2017 12th Iberian Conference on Information Systems and Technologies (CISTI)*. 2017.
19. Laylavi, F., *A framework for adopting Twitter data in emergency response*. 2016.
20. Beatson, A., A. Buettner, and T. Schirato, *Social Media, Crisis Mapping and the Christchurch Earthquakes of 2011*. Vol. 14. 2014.
21. Fosso Wamba, S., A. Edwards, and R. Sharma, *'Big Data' as a Strategic Enabler of Superior Emergency Service Management: Lessons from the New South Wales State Emergency Service*. 2012.
22. Cai, L. and Y.J.D.S.J. Zhu, *The challenges of data quality and data quality assessment in the big data era*. 2015. **14**.
23. Initiative, H.H., *Disaster relief 2.0: The future of information sharing in humanitarian emergencies*, in *Disaster Relief 2.0: The future of information sharing in humanitarian emergencies*. 2010, HHI; United Nations Foundation; OCHA; The Vodafone Foundation.
24. Tapia, A.H., K.A. Moore, and N.J. Johnson. *Beyond the trustworthy tweet: A deeper understanding of microblogged data use by disaster response and humanitarian relief organizations*. in *ISCRAM*. 2013.
25. Goodchild, M.F. and J.A. Glennon, *Crowdsourcing geographic information for disaster response: a research frontier*. International Journal of Digital Earth, 2010. **3**(3): p. 231-241.
26. MCDEM (Ministry of Civil Defence and Emergency Management). *The 4Rs*. 2017; Available from: <http://www.civildefence.govt.nz/cdem-sector/cdem-framework/the-4rs/>.
27. Scannapieco, M., *Data Quality: Concepts, Methodologies and Techniques. Data-Centric Systems and Applications*. 2006: Springer.
28. Kaisler, S., et al. *Big Data: Issues and Challenges Moving Forward*. in *2013 46th Hawaii International Conference on System Sciences*. 2013.

29. Prasanna, R. and T.J. Huggins, *Factors affecting the acceptance of information systems supporting emergency operations centres*. *Computers in Human Behavior*, 2016. **57**: p. 168-181.
30. Yang, L., R. Prasanna, and M. King, *On-Site Information Systems Design for Emergency First Responders*. Vol. 10. 2009.
31. Prasanna, R., L. Yang, and M.J.R.M. King, *Guidance for developing human-computer interfaces for supporting fire emergency response*. 2013. **15**(3): p. 155-179.
32. Bruns, A. and J.E. Burgess. *Local and global responses to disaster: #eqnz and the Christchurch earthquake*. in *Disaster and emergency management conference, conference proceedings*. 2012. AST Management Pty Ltd.