Integrating crowdsourced data into existing system for Preparedness, Immediate response & Reconstruction

Table of Contents

1. Introduction 4

2. Literature Review 4

2.1. Data Collection in crowdsourcing 4

2.2. Role-based user involvement in crowdsourcing 5

2.3. Communication as two-way channel 6

2.4. Different areas in crowdsourcing that can aid Humanitarian Supply Chain 6

2.5. Achieving sustainability with the collected data 6

2.6. Integrating in existing system 8

2.7. Roadblocks/ Challenges 8

2.7.1. Data Quality 8

2.7.2. Data Validation and Verification 8

2.7.3. Infrastructural availability 8

3. WFP – Case Analysis 9

3.1. WFP's Emergency Preparedness and Response Package Data Requirements (WFP, 2012) 9

3.2. Data collection and processing methods used in WFP (WFP, 2014) 10

3.3. WFP's Partnerships – acting as another source of information (WFP, 2014) 11

3.4. Limitations experienced in processing information from previous crowdsourced platforms 11

4. Discussion/ Suggestion 12

4.1. Identified Crowdsourced Areas and Partnerships 12

4.2. Crowdsourced Data Management 13

4.3. Crowdsourced Data Governance 13

4.4. Crowdsourced Analytics Platform 13

5. Further Research areas 14

6. APPENDIX 14

6.1. Tables 15

6.2. Figures 18

7. Bibliography 21

List of Tables

Table 1: Crowdsourcing roles and disaster management cycles (Poblet, et al., 2014) 6
Table 2: Successful crowdsourced platform leading innovation (Kreutz, 2016) 6
Table 3: WFP's potential crowdsourcing areas 10
Table 4: MDCA M&E Assessments (WFP, 2017) 10
Table 5: WFP's Crowdsourcing Application requirement (Richards & Veenendaal, 2014) 12
Table 6: Preparedness Dimensions and Activities (source: (Narvaez, 2012)) 15
Table 7: Disaster-Relevant Information Needs Checklist (Source: (Narvaez, 2012) 15
Table 8: Technology Availability in Countries 16
Table 9: Critical Success Factors of Crowdsourcing Systems (Bott & Young, 2012) 16
Table 10: Future directions in urban sustainability governance through the implementation of crowdsourcing (Certoma, et al., 2014) 17
Table 11: Infrastructure to manage Big Data (Storeya & Song, 2017) 17

List of Figures

Figure 1: Taxonomy of Crowdsourcing Tools (Poblet, et al., 2014) 5
Figure 2: Crowdsourcing roles based on users' involvement and level of data processing (Poblet, et al., 2014) 5
Figure 3: Data Governance Framework (DAMA International, 2017) 7
Figure 4: WFP Overview (WFP, 2018) 9
Figure 5: WFP's Strategic Objective (WFP, 2012) 9
Figure 6: Data and Business Value (Southekal, 2017) 18
Figure 7: Crowdsourcing Management (Grier, 2013) 18
Figure 8: Data Challenges in data analysis (Rashed, et al., 2018) 19
Figure 9: Process Challenges in data analysis (Rashed, et al., 2018) 19
Figure 10: Management Challenges in data analysis (Rashed, et al., 2018) 20

1. Introduction

Data not only provides competitive advantage but is also considered a business in itself. For enterprises, the data is the core of their business. Important drivers behind this data-focused business transformation include improved technologies, increased competition, product and service obsolescence, need for faster time to market, change in customer preferences, new government regulations, and the need to act on emerging new opportunities (Southekal, 2017). Crowdsourcing is an open innovation platform for the direct and indirect volunteers to support empowerment based on the principle of universal participation by communication technologies such as Internet and Mobile. The driving vision behind these phenomena is the philosophy of open-source governance. Crowdsourcing is used to create and increase collective knowledge, community building, collective creativity and innovation, crowdfunding, and civic engagement (Bott & Young, 2012). To crowdsource, one must ask for a piece of information, an idea for a new product, a little bit of work, a large task or even a contribution. In return, the organization offer some kind of compensation. Pay for the worker's services with money, or offer him/her gratitude, or give him/her a gift, or offer him/her membership in a community, for example, Wikipedia (Grier, 2013). Crowdsourcing can be used for emergency preparedness, climate change, to validate against internally collected data or fill in the data gaps. This study used the WFP's crowdsourced case studies performed on Gaza crisis and Typhon Bopha. This is a qualitative study is to understand integrated data platforms from crowdsourcing perspective. A proof of concept must be performed in order to test the concepts in the next stage.

2. Literature Review

After selecting few case studies, crowdsourcing is studied from the data collection level to the governance level with challenges that might come while integrating in to the existing system. On a high level, organizations can start questioning themselves that what value they are generating for the business and how compliant they are with other systems (Refer Figure 6: Data and Business Value). This section has identified few areas that must be assessed to integrate different data platforms from crowdsourced data perspective.

2.1. Data Collection in crowdsourcing

An acute Humanitarian Situation can occur due to man-made or natural reason such as sudden unplanned displacement, new or exacerbated and sustained episodes of armed conflict, impending or already occurred sudden deterioration of nutritional status, natural or industrial disaster, or sudden breakdown of critical administrative and management functions (WHO, 2017). This information depends on many operational objectives where data plays a crucial role in reporting and executing strategies. Crowds can be used to conduct targeted data collection to supplement and add depth to social science research. According to (Brabham, 2013) data can be collected on Crowdsourced using below methods

a) **Crowdsourcing survey response** where an organization uses crowd to find and assemble information
b) **Targeted data collection** efforts where an organization has a set of information in hand and mobilizes a crowd to process or analyze the information
c) **Challenge-based data collection** is used for ideation problems where an organization asks crowd to come up with a solution to a problem that has an objective, provable right answer
d) **Citizen observation networks** production is used for ideation problems, where an organization mobilizes a crowd to come up with a solution to a problem which has an answer that is subjective or dependent on public support

Crowdsourcing mediums such as Internet and Mobile Applications allow applications on these platforms to do data collection, also, perform data tagging and data filtering. In recent years, market for disaster management apps has grown which provide mapping and navigational functionalities. Online Platforms such as OpenIR maps ecological risks to identify vulnerable areas and support its emergency management, Crowdcrafting and ArcGIS focus on mitigation and preparedness, Ushahidi and CrisisTracker utilized the power of social media as their features and OpenStreetMap provides offline and online data. Further, the study (Poblet, et al., 2014) classified the different platforms and mobile apps in four major characteristics illustrated in the Figure 1: Taxonomy of Crowdsourcing Tools below:

i) the phase of the management disaster cycle where it better applies to

ii) the availability of the tool and its source code
iii) the main core functionalities
iv) crowdsourcing role types

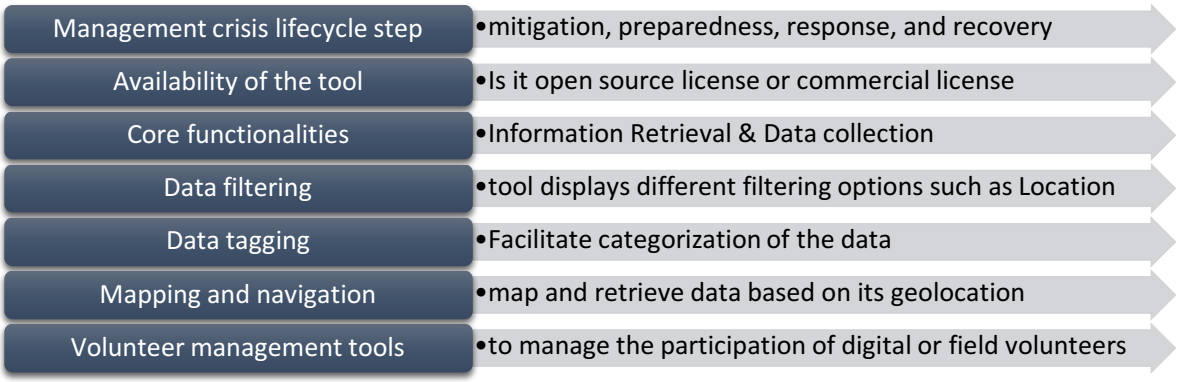

Figure 1: Taxonomy of Crowdsourcing Tools (Poblet, et al., 2014)

2.2. Role-based user involvement in crowdsourcing

Crowdsourcing data collection is around data which is coming in Media, Cloud, Web, Internet of Things (IoT) and Databases which in turn is classified into structured, semi-structured and unstructured data. As per (Poblet, et al., 2014), by including the citizens ("the crowd") into the platforms, power to bring change is given back to the people and knowledge can be extracted to trigger a global response based on the information. The crowd can be further distinguished based on the role of data processed as shown in Figure 2 : Crowdsourcing roles based on users' involvement and level of data processing which further leads to four types of crowdsourcing roles based on

a. Type of data processed - raw, semi-structured, and structured data
b. Participants' level of involvement – passive or active
c. Skills required – Basic or Specialized

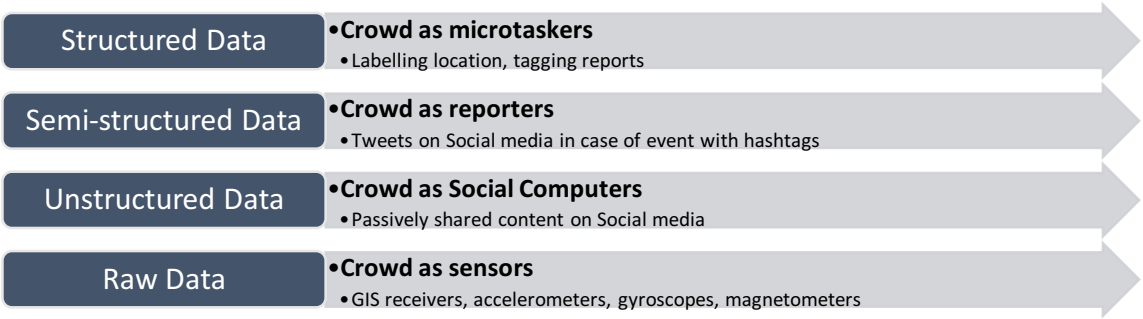

Figure 2: Crowdsourcing roles based on users' involvement and level of data processing (Poblet, et al., 2014)

The four different roles provide data at different level of Disaster Recovery. The data is collected from 'Crowd as a sensors' inadvertently or by explicit consent in the form of GPS location from mobile devices or location sensors such as Accelerometers and Gyroscope. At the level of Crowd as Social Computers and reporters, people and organization (Volunteers, Citizens and Emergency authorities) provide real time update on the events. The amount of information that people share during disaster on social media platform can raise situational awareness during Emergency. The last role of micro-tasker is required to analyzing structured data. The Table 1: Crowdsourcing roles and disaster management cycles shows the roles involvement at different level of the emergency management cycle (Poblet, et al., 2014).

	Crowd as a sensor	**Crowd as a social computer**	**Crowd as a reporter**	**Crowd as a micro-tasker**
Preparedness	X			X
Response		X	X	X
Recovery		X	X	X
Mitigation	X	X	X	

Table 1: Crowdsourcing roles and disaster management cycles (Poblet, et al., 2014)

2.3. Communication as two-way channel

In the last decade, crowdsourcing gained relevance as a mechanism of open innovation. Organizations stage a platform for influencing problem solvers in an open crowdsource challenge. Common way to ask for 'Request for Proposal' where seeker would draft a Problem Statement. To influence the problem solver, it is important that some reward is given in form of points or money (Pollok, et al., 2018). Because the conversations happen when the communication becomes both ways. In this era of virtualization, it is important to engage the participator. It is inevitable that the crowd is managed, informed and if required, rewarded on time (Grier, 2013). The crowd could participate for different reasons such as economic, social recognition, self-esteem, or the development of individual skills. In the past, the successfully implemented crowdsourcing examples are of Linux, Unix demonstrating collective intelligence and OpenStreetMap by Google is another example of Social Mapping by crowd. Many private companies such as Unilever, Heineken and Sony launched open crowdsourcing platforms exploit crowdsourcing as a shortcut for addressing their business challenges in sustainability (Certoma, et al., 2014).

2.4. Different areas in crowdsourcing that can aid Humanitarian Supply Chain

The four types of disasters that can take place: 1 natural, sudden onsets (e.g. earthquakes, hurricanes, tornadoes); 2 human-made, sudden onsets (e.g. terrorist attacks, coups d'e´tat, industrial accidents); 3 natural, slow onsets (e.g. famines, droughts, poverty); and 4 human-made, slow onsets (e.g. political and refugee crises). The vulnerability of a household or community is determined by their ability to cope with their exposure to the risk posed by shocks such as droughts, floods, economic fluctuations, and conflict (O'Connor, et al., 2017). Crowdsourced platform allows the organization to leverage open-source data collection for different areas as listed in the table below:

Table 2: Successful crowdsourced platform leading innovation (Kreutz, 2016)

Brand	Concept
Picnic Green Challenge	Ideas to save the planet.
Open Ideo	Solve big challenges for social good.
Root Cause	One stop shop for social innovation.
Challenge.Gov	Government challenges and crowd sourcing (U.S).
Ushahidi	Crowdsourcing crisis information.
Unilever	Open innovation for social good.
Humanitarian Innovation	Open innovation challenges to solve humanitarian problems.
Start Some Good	Crowdfunding for social good.
Young Foundation	Disruptive social innovation.
Building Change Trust	Resources for the social innovator.
The Social Innovation Partnership	Support for social innovation projects.
CrowdVoice	Tracking voices of protest.
CrowdCrafting	Non-profit crowdsourcing platform with many projects.

Usually in disaster management, a high level of detailed information is needed in order to coordinate the required processes of disaster relief in the imminent time after such an event— i.e., rescue, support, material handling as well as supplied transport processes have to be based on hands-on information on the ground near the disaster site (Klumpp, 2017). The segregated information from different crowdsourced platform can further be aggregated into information and supply flow model for deeper understanding into the crowdsourced system.

2.5. Achieving sustainability with the collected data

Not just in Humanitarian Supply Chain, but in any business operation, data plays a crucial part wherein crowdsourced data is also a data which can be used with organization's data to get insight from the data. In this era of Big data analytics, every big organization wants to leverage Data warehousing capabilities to get more

meaning out of the data. But this assimilation of different types of data at one place requires a framework which can act as a constitution for the data world. In selected three studies (Wang, et al., 2018), (Al-Ruithe, et al., 2016), (Veeneman, et al., 2018), the authors discussed the Data Governance in Big Data from three different perspectives. In (Wang, et al., 2018), the author discussed how Nonprofit organizations (NPO) holding a tremendous quantity of client data, now encountered challenges to managing, analyzing, and applying this acquired data in order to improve service quality and customer satisfaction levels. Thus, for agencies to implement thorough and effective e-governance, they must first plan satisfactory methods and procedures for data governance, so as to ensure that the data they obtain are of high quality and can be well managed and applied. The focus of data governance on data availability, data usability, data consistency, data integrity, and data security. In (Al-Ruithe, et al., 2016), the authors discussed that how cloud computing models are a highly disruptive technology and the adoption of its services requires a more rigorous data governance strategies and programmes. Further, they discussed how organizations need to be aware of the best practices for safeguarding, governing, and operating data in the cloud environment since the loss of control and governance can also lead to the impossibility of complying with the security requirements, a lack of confidentiality, integrity and availability of data, and a deterioration of performance and quality of service, not to mention the introduction of compliance challenges. In (Veeneman, et al., 2018) , big data is discussed from the smart city on a mobility data platform. These platforms combine a wide variety of data sources to support the creation of an overview of network status and flows (historical, current, and future), and allow for use of that information from real-time travel planning to long-term infrastructure planning. The challenge for mobility data platforms is how to get the right data, aggregate that data, model future network states on that data, and make that available to the right people. Challenges exist in creating reliable network flow data from various sources (from induction loop counting via Bluetooth, Wi-Fi, RFID, and number plate tracking, to GPS traces). Also, making the planning effective in an uncertain environment with a wide variety of multimodal options (spanning from private modes like cars, cycling and walking, via shared modes, like cars and cycling, to public modes, like train and bus) in easily usable travel plans that support public goals, has proven challenging. In addition, the role of simulating realities to predict future network states in the management of networks has been a challenge. Tables in the APPENDIX section lists the other factors involved in the Governance. Table 9: Critical Success Factors of Crowdsourcing Systems Table 10: Future directions in urban sustainability governance through the implementation of crowdsourcing, Table 11: Infrastructure to manage Big Data . Before implementing governance, one can argue that it is important to differentiate different areas in the crowdsourcing and it must be managed, (Grier, 2013) identified few areas to manage crowdsourcing as a whole system in Figure 7: Crowdsourcing Management. Further, crowdsourcing data integration can be implemented against proposed a Data Governance framework by (DAMA International, 2017) for performance of policies and processes and conformance to standards and procedures. Refer Figure 3 below

Figure 3: Data Governance Framework (DAMA International, 2017)

2.6. Integrating in existing system

Integration of AI, machine learning, big data, and crowdsourcing is still in infancy but here is crowdsourced platform who are living the future such as **CrowdFlower**'s AI allows businesses to perform tasks with algorithms and machine learning, then bring in human judgment if they're not confident. The **Artificial Intelligence for Disaster Response (AIDR)** combines crowdsourcing with real-time machine learning for disaster response. British **WireWax** uses AI and human input to automatically detect countless features in videos. **Republic Systems** allows clients to utilize the best open innovation platforms by monitoring the popularity of over 3,000 existing crowdsourcing platforms (Hershkovitz, 2017).

2.7. Roadblocks/ Challenges

Integrating all the different data sources into one unified view will surely take significant time and effort in the realm of stakeholder alignment, data dictionary rationalization, and data and business rules transformation (Southekal, 2017). Besides integration, the Data Collection can itself pose few challenges since change can be cumbersome. (Rashed, et al., 2018) broadly categorized into Data challenges, Process challenges and Management challenges. Refer Figure 8: Data Challenges in data analysis, Figure 9: Process Challenges in data analysis & Figure 10: Management Challenges in data analysis . Further, in case of emergencies, non-profit organization follows few protocols that should be mentioned to proceed further (Refer Table 6: Preparedness Dimensions and Activities (source:)Table 7: Disaster-Relevant Information Needs Checklist (Source: Table 8: Technology Availability in Countries).

2.7.1. Data Quality

Data is super messy, and data cleanup will always be literally 80 percent of the work. In other words, data is the problem" [Levy, 2015]. According to CrowdFlower, a crowdsourcing company that surveyed data scientists with varying levels of experience, about 60 percent of the time spent by data scientists is on cleaning data and preparing it for analysis (Southekal, 2017). Sometimes,

a. The crowdsourcing process is not effective
b. There can be Clash of Paradigms, when official government and/or donor data is combined with crowdsourced data that usually does not follow the same information management standards (Bott & Young, 2012).

2.7.2. Data Validation and Verification

From data Validation and verification perspective, one must know that the crowd is providing the right data.

a. The crowd is being manipulated, in case of weak governance
b. The crowd is being attacked, when Crowdsourcing contributors can be incriminated through national security

2.7.3. Infrastructural availability

It is necessary to know if in developing and underdeveloped countries, the people have proper tools to collect the data. Since in the case of disaster, it is necessary that the information is received and that is only possible with the help of technology but there are chances that the in some areas, technology infrastructure is destroyed which further can raise questions such as

a. There is no active crowd when crowd is managed centrally. Under authoritarian regimes, it is also more difficult for NGOs and social entrepreneurs to launch a crowdsourcing initiative
b. No data is being shared can be a reason in case of Information Sharing issue such as in China or there is no Technical capability in the country to sustain a hindrance free network.

3. WFP – Case Analysis

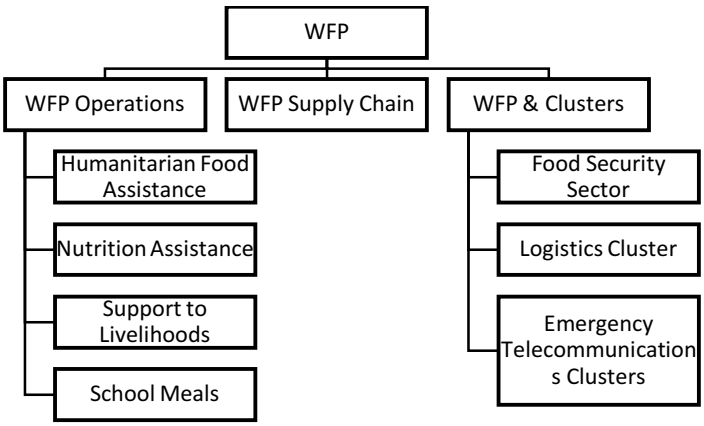

Figure 4: WFP Overview (WFP, 2018)

World Food Programme is the world's largest humanitarian organization addressing hunger and promoting food security by delivering food assistance in emergencies and working with communities to improve nutrition and build resilience. On any given day, WFP has 5,000 trucks, 20 ships and 92 planes on the move, delivering food and other assistance to those in most need. Every year, we distribute more than 15 billion rations at an estimated average cost per ration of US$ 0.31. With strategic objectives so vast and complex in the WFP's ecosystem, it is necessary that the operations are continued with the minimum error and sufficient data. WFP is already having separate platforms for different operations (refer Data collection and processing methods used in WFP) but an integrated platform can help in achieving better results with less errors in next operation and improvised with the gaps experienced in previous cases. WFP's performance is measured against the Strategic Objectives which is high level vision of the organization's aim (Refer Figure 5: WFP's Strategic Objective)

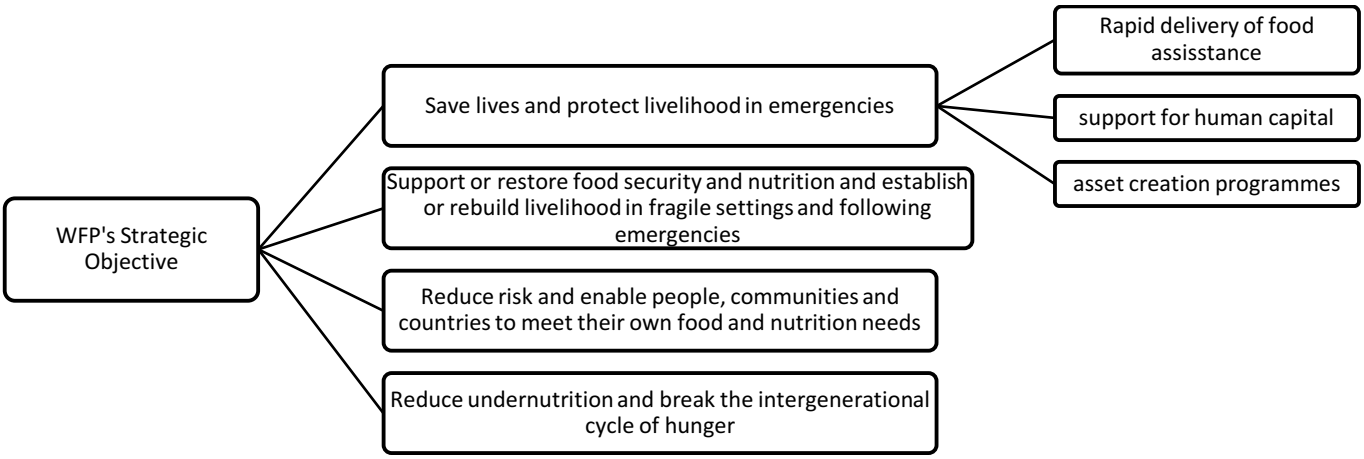

Figure 5: WFP's Strategic Objective (WFP, 2012)

3.1. WFP's Emergency Preparedness and Response Package Data Requirements (WFP, 2012)

The Emergency Preparedness and Response Package builds on the findings and recommendations of the Strategic Evaluation, and follows extensive consultations with Regional Bureaux, Country Offices and relevant technical divisions in WFP Headquarters. The Headquarters-based team can provide tailored analytical services on-demand to Country Offices and Regional Bureaux, and it can collaborate with ODEP's Emergency Preparedness and Response (EPR) and Geographic Information Systems (GIS) teams and other Headquarters units, such as Field Security Division (ODF) and the Vulnerability Analysis and Mapping Unit (VAM), to provide comprehensive and high-quality products. For instance, the team can advise Country Offices and Regional Bureaux as to how to identify, detail and monitor natural hazards and political risks, and help them define indicators and thresholds for emergency readiness actions. In a case of an emergency, the WFP has to respond within first

72 hrs. of the event. In table below, WFP collect data for different purposes where benefits of crowdsourcing can be leveraged.

Table 3: WFP's potential crowdsourcing areas

Data area	Purpose
Monitoring analysis and early warning	Droughts, hurricanes or conflicts in the area
Sound analysis and extensive monitoring	Preparedness for slow-onset disasters
Post-harvest handling and quality control	Post-harvest food loss
Nutrition Monitoring	Nutrition-specific and nutrition-sensitive activities
Seasonal climate forecasts	To trigger funding and action to build resilience before shocks occur
Gender capacity assessment	To track resources allocated to both genders

3.2. Data collection and processing methods used in WFP (WFP, 2014)

Over the years, WFP has incorporated different strategies to enhance data collection methods to serve their beneficiaries better. WFP's vulnerability analysis and mapping (VAM) uses a mobile system using mobile telephones and other remote data-collection approaches such as voice calls, text messages and interactive voice response technology to collect real time information on household food security and markets. Performance and risk-management systems, SRF and MRF, and approach to monitoring and review are improving WFP's ability to design, implement and report on its activities. COMET is used to design, implement and monitor programmes, improve organizational performance and enable more systematic measurement of food-assistance outcomes, Optimus for Supply Chain Optimization, and Mobile data collection and analytics (MDCA) for monitoring and evaluation purposes. Refer below for MDCA's assessment areas. In the end, Palantir and Tableau serving Data warehousing and Business Intelligence capabilities (Tableau, 2017) (Palantir, 2017).

Table 4: MDCA M&E Assessments (WFP, 2017)

Business Area	Process Area	Assessment
CBT	Emergency Assessment	72 HRS Assessment
		Checklist of Protection
		Emergency Micro Finance Assessment
		Financial Strength Assessment
		Micro IT Assessments
		Macro Supply Chain Assessments
	Process Monitoring	Distribution Monitoring Checklists
		Post Distribution Monitoring Checklists
		Retailer Monitoring
		Non-participating Retailer
Monitoring (Cross-sectoral)	Activity Implementation Monitoring	Activity Implementation Monitoring
	Distribution Monitoring	Food Transfer - Distribution Monitoring
	Post Distribution Monitoring (PDM)	Food Transfer - Post Distribution Monitoring Gender, Protection and Assistance to Affected Populations
		Food Security and Livelihood
		Nutrition module
		School Feeding
Nutrition		Emergency Nutrition Assessment
		Rapid Nutrition Assessment
		PDM - Nutrition module
VAM	Vulnerability Analysis	72 HRS Assessment
		Rapid Food and Nutrition Security Assessment
		Emergency Food Security Assessment (EFSA)
		Food Security Assessment Comprehensive
		Food Security and Vulnerability Analysis (CFSVA) Survey
		Crop and Food Security Assessment
		Market Assessment
		Trader Assessment
		Transporter Assessment

3.3. WFP's Partnerships – acting as another source of information (WFP, 2014)

To achieve Zero Hunger, WFP works closely with a variety of partners to help fight global hunger and poverty, within the framework of the 2014-2017 Corporate Partnership Strategy. Government partners are involved in all levels of WFP's work, including resource mobilization, implementation of activities, as well as at the policy and advocacy level. NGO partners for strategic collaboration in the fields of policy, advocacy and programme design. The three Rome-based UN agencies - WFP, the Food and Agriculture Organization (FAO) and the International Fund for Agricultural Development (IFAD) - work together with the common vision of promoting world food security. WFP works with other agencies across the UN system to ensure that the humanitarian and development community continues to provide efficient and effective assistance with greater beneficial impact on the people we serve. Private partners/donors are classed as 'for-profit' corporations or business associations, foundations, trusts, educational or service organizations, or an individual (WFP, 2018).

3.4. Limitations experienced in processing information from previous crowdsourced platforms

In a (Richards & Veenendaal, 2014) study, the authors studied Gaza crisis'12 and Philippine Typhoon Bopha'12 for crowdsourced platform for timely response. After weighing the pros and cons of crowdsourcing technology, it appeared that this technology is in its infancy at present and will need to develop further before it will be adopted as part of WFPs' existing crisis mapping operations. Further, it is argued that this research will lead to technological development in terms of a crowdsourcing application that meets; critical time frames for both slow and sudden-onset crisis situations, sustained and consistent crowd participation, affected population participation (which is not dependent on a countries economic development), practical and accurate attributes, accurate positional accuracy, effective spatial data handling capability, an ease in extraction of actionable information and an application which meets the geospatial data needs required for humanitarian crisis mapping operations. However, in a study (Riccardi, 2016), social media and crowdsourced information proved useful in Colorado Wildfires'13 as it provided

information to other inhabitants in the area. Further, this study discussed the voice of the 'crowd' that in Haitian earthquake, volunteers connected through Twitter arranged and delivered resources to victims on the ground. Crowdsourcing has shown to be an effective tool in providing situational awareness for both victims and disaster managers and is an excellent method to recruit resources such as volunteers, donations, imagery analysts, etc., to help control the chaos that occurs during recovery operations but for that to happen, it is necessary that the place have seamless access to technology in order to report the event on Internet. Further, the author suggested a table below which identifies the WFP's crowdsourcing application requirements

Table 5: WFP's Crowdsourcing Application requirement (Richards & Veenendaal, 2014)

Criteria	WFP Crowdsourcing Application Requirement
Criteria 1: Time	The application would need to be able to support the continual and rapid crisis mapping timeframes for the ER to both slow and sudden-onset crisis situations respectively.
Criteria 2: Collaborative Knowledge (Crowd Participation)	The application would need to have a high level of crowd participation to ensure that crisis mapping tasks are completed within typically limited timeframes. Crowd participation would need to be constant and sustained to support both the slow-onset and sudden-onset crisis situations. The resultant collaborative knowledge needs to be a true representation of the whole affected population.
Criteria 3: Local knowledge	The application would need to be accessible and usable by the affected population in all countries in which WFP has a presence, despite the countries' level of economic development.
Criteria 4: Attribute Accuracy	Not only would the application need to achieve a high degree of attribute accuracy, but it would also need to have attributes which do not delay crisis mapping operations. E.g. long textual comments associated with each observation or attributes containing uncertainty will delay critical crisis mapping operations and are not practical.
Criteria 5: Positional Accuracy	The application would need to achieve a high degree of positional accuracy and need to generate reliable crisis data. Million-dollar food aid decisions need accurate and reliable location information.
Criteria 6: Spatial Data Handling Capability (Big Data)	In the typically time-constrained crisis mapping operations, the application would need to support ease in spatial data handling in order to permit further manipulation of the geospatial data.
Criteria 7: Actionable Data	The application would need to define and extract actionable geospatial data with ease to permit timely assessment of food security and factors affecting food aid decision making.
Criteria 8: Meets WFP Spatial Data Needs	The application would need to be suited to support WFP's geospatial data needs as outlined within this chapter.

4. Discussion/ Suggestion

WFP is already working with crowdsourcing platforms such as Ushahidi and Maano but there are few other few data areas identified in Table 3: WFP's potential crowdsourcing areas which can be outsourced to the data intelligence and data M&E companies in identified emergency prone areas and leverage the power of crowdsourced data to empower the executives with the data. Based on the limitations in the case analysis in previous section, crowdsourced data management and governance framework could help in maintaining sustainable data quality which can further help in migrating data later to other systems. Further, the crowdsourced analytics platform can be implemented for open innovation and generate insight from the data using the voice of crowd. From literature review and analysis, authors also discussed another limitation from the big data and data gap perspective which can also be solved with the data tools, further discussed in sub section below.

4.1. Identified Crowdsourced Areas and Partnerships

The data was always around since no information can be processed without the data. In last few years, the independent data collection companies identified some gap areas, collected data and started collaborating with the other organizations for greater good, companies such as Ushahidi. With the advent of fourth industrial revolution, the world wants to jump onto new state-of-the-art technologies but limitation from the data side is

visible. In humanitarian supply chain, the partnerships are already present but new partnership can be fostered with the successful start-up who are working to collect data in the identified risk area. Table 2: Successful crowdsourced platform leading innovation & Table 3: WFP's potential crowdsourcing areas mentioned few platforms and areas respectively which can further be studied for more in-depth knowledge and partnerships.

4.2. Crowdsourced Data Management

There has been a research activity in studying the potential of big data in decision making. This interest has been further renewed by the explosion of massive data sets, social media streams, and the emergence of high performance computing. Data management and utilization in the policy formulation has to be accompanied by a set of specifications that safeguard the compliance with ethical, legal, data privacy rules and norms. As separate Data Layer can be implemented which will be responsible for management of data and knowledge resources coming from the Information Systems. Various tools used at this level are listed below (Androutsopoulou & Charalabidis, 2018)

a. **Data Discovery Tools** to find out data from numerous data sources and extract data
b. **Data Cleansing Tools** ensure appropriate data quality and enable their re-use
c. **Data Transformation Tools** transforms external cleansed data (e.g. Open Data, IoT Data, Government Data) to use standard formats in order to facilitate data interoperability
d. **Data Linking Tools** maintaining compatibility evaluation is based on the structure of the data and the desired type of join

The study proposed a framework aims to facilitates cooperation among all societal actors, managing information flows across organizational boundaries and among the participants and to improve understanding of the social problem under investigation and its parameters, revealing hidden Patterns and, consolidate governmental, scientific and behavioral data into a knowledge base.

4.3. Crowdsourced Data Governance

“Crowdsourcing is viewed as a core mechanism of new systemic approaches to governance addressing the highly complex, global, and dynamic challenges of climate change, poverty, armed conflict, and other crises” (Bott & Young, 2012). Though this study did not talk much about big data analytics but the massive data will eventually be needed to process information. So, it is imperative that the proper framework such as Data Governance Framework by (DAMA International, 2017) is in place to integrate with other systems. Or any Governance tools are used as in one study, according to (Meers, et al., 2017), the data governance team is not to be staffed with a full-time resource for the sole effort of administration as managing user’s logins and access is not time-consuming, but performing quality checks, running ETL or database scans, doing database backups, and other infrastructure tasks can be a full-time job. Many aspects of data governance technologies, such as data modeling, metadata management, and data quality, have been addressed by technologies for three decades, the concepts of data governance management are an emerging market. The selection of data governance technology is a complex activity for many organizations that are not experienced in technology selections. So, the authors recommended a formal evaluation process and suggested few governance tools such as SAP which offers **a.** Data Services Enterprise, **b.** Information Lifecycle Management and, **c.** SAP Access Control.

4.4. Crowdsourced Analytics Platform

This study is an overview at a macroscopic level, which could be further broken to country or regional level for more insight into the Preparedness, Immediate response & Reconstruction by executing a proof of concept on Big Data Analytics platform in a small area. Also, this study talks about mostly from the crowdsourced Data collection perspective since effective analytics results can only be obtained if the organization have sufficient cleansed data. Once the data has been obtained and datasets can be formed using that data, the crowdsourced analytics platform could be implemented for the open source innovation and rewarding crowd in exchange with points or money such as Amazon Mechanical Turk or PepsiCo. Platforms like these can also be used for spreading the word as an informal advertisement channel.

Usually, the partners are government agencies and they have their own protocols around Data Privacy. With so many initiatives already in the countries and recent discussion on Data Protection Laws, it is very important that the organization maintains a framework or a checklist to keep the data secure and only use for the mentioned purpose.

5. Further Research areas

The scope of this study is wide and present study can be further studied for areas listed below

1. An Integrated Data Analytics process to Optimize Data Governance of Non-Profit Organizations
2. Attracting individuals on a crowdsourcing platform for competition such as Amazon Mechanical Turk
3. Gamification for Crowdsourcing Data Collection such as Fitbit and Freeletics. Setting levels to engage volunteers.
4. Detection of privacy leaks on an integrated platform
5. Sentiment Analysis and data analysis of crowdsourcing analysis
6. Machine Learning algorithms for Data Quality

6. APPENDIX

6.1. Tables

Table 6: Preparedness Dimensions and Activities (source: (Narvaez, 2012))

Dimensions of Preparedness	Goals Sought	Activities
Hazard Knowledge	To provide activities a knowledge-base about hazards, the likelihood of different types of disaster events, and likely impacts	Conducting hazard, impact, and vulnerability assessments;
		Using loss estimation software, scenarios, census data;
		Understanding potential impacts on facilities, structures, infrastructure, populations;
		Providing hazard information to diverse stakeholders
Management, Direction and Coordination	To formulate *strategies* make it possible for households, organizations, and other units of analysis to manage both preparatory activity and response processes	Assigning responsibilities;
		Developing a division of labor and a common vision of response-related roles and responsibilities;
		Forming preparedness committees, networks;
		Adopting required and recommended management procedures (e.g., National Incident Management System);
		Providing training experiences, conducting drills, educating the public
Formal and Informal Response Plans and Agreements	To target the development of disaster plans and other agreements	Developing disaster plans, evacuation plans, memoranda of understanding, mutual aid agreements, collaborative partnerships, resource sharing agreements;
		Participating in broader and more general planning arrangements (e.g., neighborhood and community preparedness groups, Urban Area Security Initiative regional plans, industry-wide preparedness initiatives)
Supportive Resources	To identify and establish internal and external resources necessary for disaster response and recovery	Acquiring equipment and supplies to support response activities;
		Ensuring coping capacity;
		Recruiting staff;

Table 7: Disaster-Relevant Information Needs Checklist (Source: (Narvaez, 2012)

Information	Preparedness	Response	Recovery	Reconstruction
Local community, regional and national levels				
Local disaster plans, procedures and policies				
City and Housing Plans				
Phone management databases				
Emergency Centers: where, who, what				
Telecom: infrastructure, laws, organizations				
Social, demographic and economic data				
Safety and environmental standards and codes				
Cultural asset information: significance, age, construction material and condition				
Land use plans				
Critical infrastructure inventories				
Building inventories				
Property ownership records				
Birth and medical records				
Hazard maps: nature, site and real-time changes				
Vulnerability data: who, where, how				
Loss/damage data				
Weather data: short and mid-term				
Available resources: what, how much, who				
Formal key decision-making: who makes what decisions, when, how				
Informal local authority structures in the family, in the community				

Table 8: Technology Availability in Countries

	Mobile Usage (apprx)	Internet Usage (apprx)
	millions	millions
Benin	**5.1**	**3.8**
Burkina Faso	**7.7**	**3.7**
Cote d'Ivoire	**12.5**	**6.3**
DRC	**28.29**	**5**
Ethiopia	**34.7**	**16**
Guinea	**5.9**	**1.6**
India	**442**	**462**
Kenya	**28.3**	**43**
Mali	**11.4**	**12.4**
Mauritania		**0.81**
Niger	**5.5**	**0.95**
Nigeria	**86**	**98**
Pakistan	**140**	**44**
Senegal	**8.4**	**9**
Tanzania	**23.7**	**23**
Togo	**2.9**	**0.9**
Uganda	**17**	**19**

Table 9: Critical Success Factors of Crowdsourcing Systems (Bott & Young, 2012)

Dimension	Success criteria
Infrastructure	Ease of accessibility, reliability, and quality of communication technologies and infrastructure
Vision	Well-defined set of ideals, goals, and objectives that is sensitive to the dynamics of its environment
Human Capital	Language skills, managerial skills, national orientation, traditions, and level of education
Financial Capital	Not very capital intensive if crowd is relying on Mobile devices and network
Linkages and Trust	Geographic, cultural, linguistic, or ethnic linkages between individuals, work groups, or organizations can be used to minimize costs of doing business
External Environment	Favorable regulatory environment and ease of doing business can encourage crowdsourcing initiatives
Motivation	Performance expectancy, effort expectancy and social influence can aid in collection
Criteria of Governance	Possibility of anonymous participation via a central registrar, public key infrastructure, and a trusted central authority; decentralization of authority, thus minimizing the principal-agent problem; centralization of information via one platform and interoperability of interfaces and applications with this platform; open and equal opportunity of participation in deliberations or peer reviews

Table 10: Future directions in urban sustainability governance through the implementation of crowdsourcing (Certoma, et al., 2014)

Urban sustainability governance requires	**Crowdsourcing will impact the future by**	**This prefigures a future characterized by**
Participation: people involvement in knowledge production and batch decision-making	Allowing technology-mediated forms of public participation in science and policy production	Ad hoc design of freely accessible crowdsourcing platforms issued by research and policy-making institutions (ex. Citizen Sense, PublicLab)
Multidimensionality: balance of the ecological/environmental measures with social measures	Helping social actors to establish pro-active relationships with the environment	Use of sensing tools and procedures for improving the perception of environmental conditions (ex. Extreme Citizen Science, Communitymaps)
Purposiveness: operationalize principles of ecosystem protection, environmental management and the fulfilment of basic human needs	Feeding problem-solving through direct contact between scientists, citizens and administrators with standardized on-line based processes	Diffusion of interactive platforms for sharing solutions and implementation tools between citizens and institutions (ex. Citizen Cyberlab, openIDEO)
Equality: mechanism to promote social and environmental justice (recognition, redistribution, empowerment)	Distributing agency in data production, collection and interpretation, and in policy-proposal elaboration	Web-platform for creating brand-new sets of data (or feeding existing ones) (ex. Every aware project, Mapping for Change)
Transparency: advancing effective democratic institutions	Granting access to meaningful information, data and software	Increase in open-access knowledge and open-data provided by research, political and administrative institutions (ex. Open Data Aarhus, Environmental Agency Data Share)
Collaboration(internal): substantive agenda that is sharable by institutions, businesses, civil society organizations, NGOs and people on the basis of reciprocity	Allowing institutions, businesses, civil society organizations, NGOs and people to negotiate in the virtual space the fate of an issue on an on-discriminatory base	City-based and issue-oriented collaborative platforms (ex. Sustainable City Network, The Community Planning Website, Cidade Democratica)
Cooperation(external): links between international and transnational actors in order to global environmental goals	Developing across-border public space without the need for institutional mediation and the physical presence of all actors	Interconnectedness of local issues in the global arena, shared practices and creation of distance-actions (ex. Global Voices, Hacking the city)
Adaptability: progressive and dynamic goal-changing agenda	Undertaking tasks of variable complexity and modularity in which the crowd contributes work, money, knowledge and/ or experience	Software to collect the voice of the general public provided by institutions or business (ex. Idea Connection)

Table 11: Infrastructure to manage Big Data (Storeya & Song, 2017)

Data stack	Consists of structured, as well as unstructured, data.
Big data ecosystems	Consists of search and visualization, data orchestration, and data access. Professional services for big data include the hardware, such as the computer, storage and networks. Specific database-related services are SQL-based systems, NoSQL systems, and Hadoop and its ecosystems.
Enterprise information management	Focuses on issues related to data governance, data integration, data quality, data visualization, and master data management.
Data Science platforms and tools	Have a variety of tools for pattern extraction and visualization of results. These include: machine learning algorithms; predictive analytics [7]; prescription techniques (e.g., simulation with alternative variables and subsets of data); descriptive techniques (e.g., statistics and assessment techniques for historical reporting); and reports (e.g., represented by scorecards or dashboards).

6.2. Figures

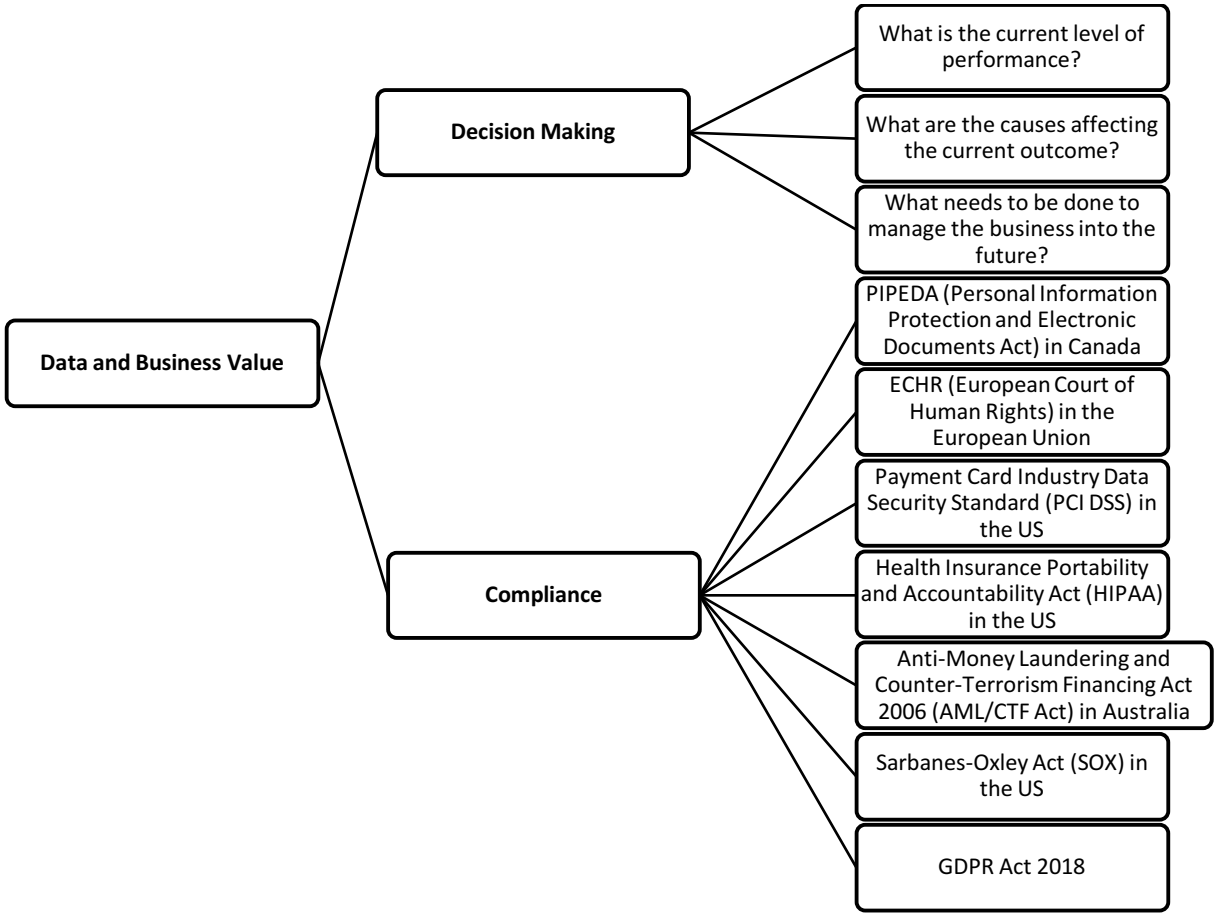

Figure 6: Data and Business Value (Southekal, 2017)

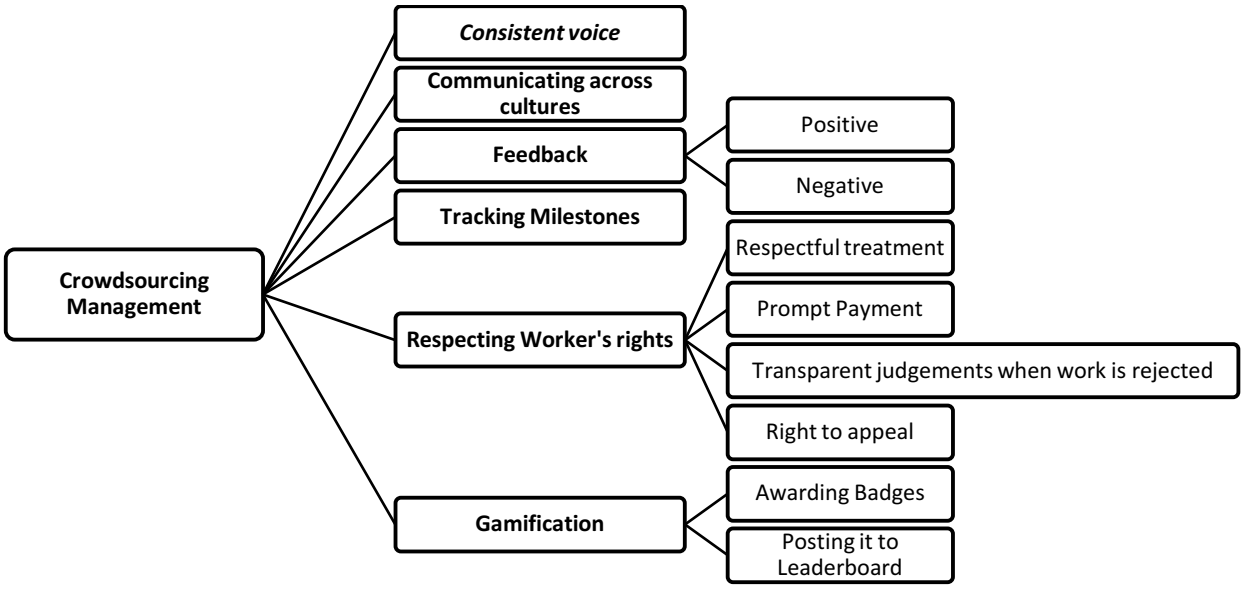

Figure 7: Crowdsourcing Management (Grier, 2013)

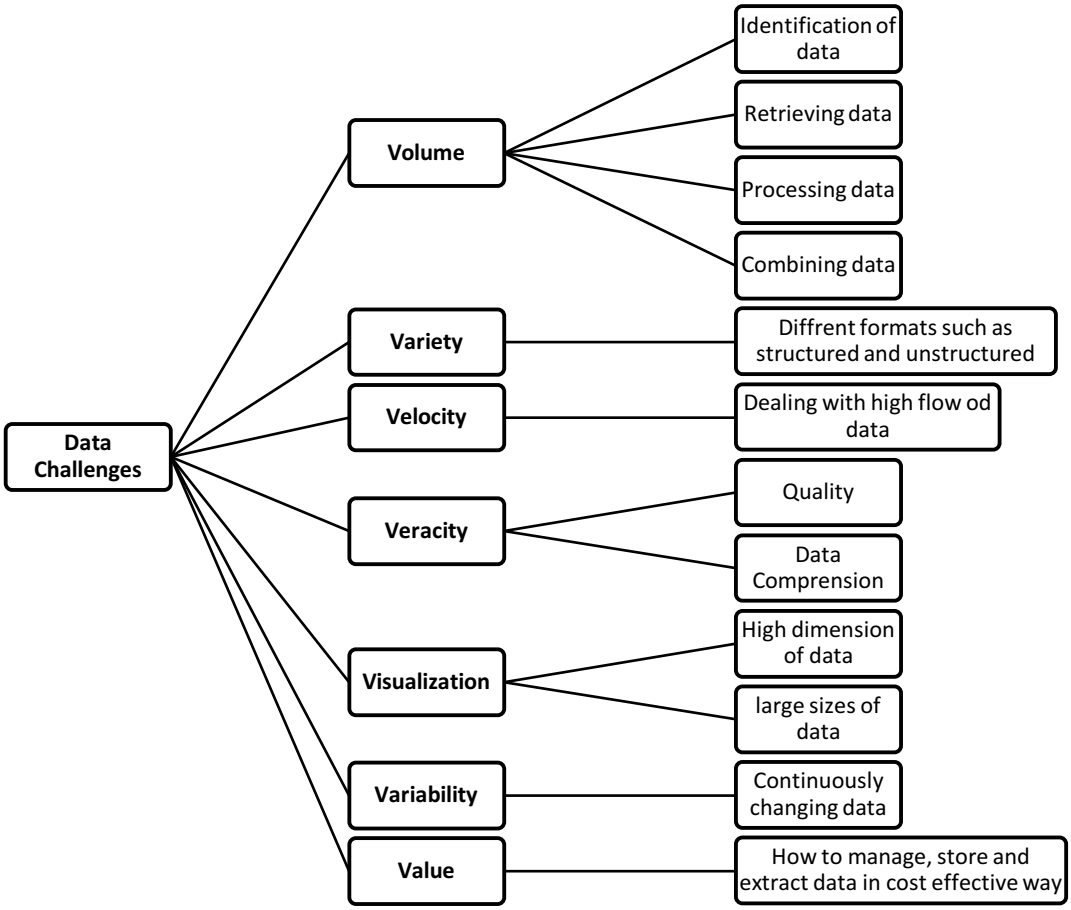

Figure 8: Data Challenges in data analysis (Rashed, et al., 2018)

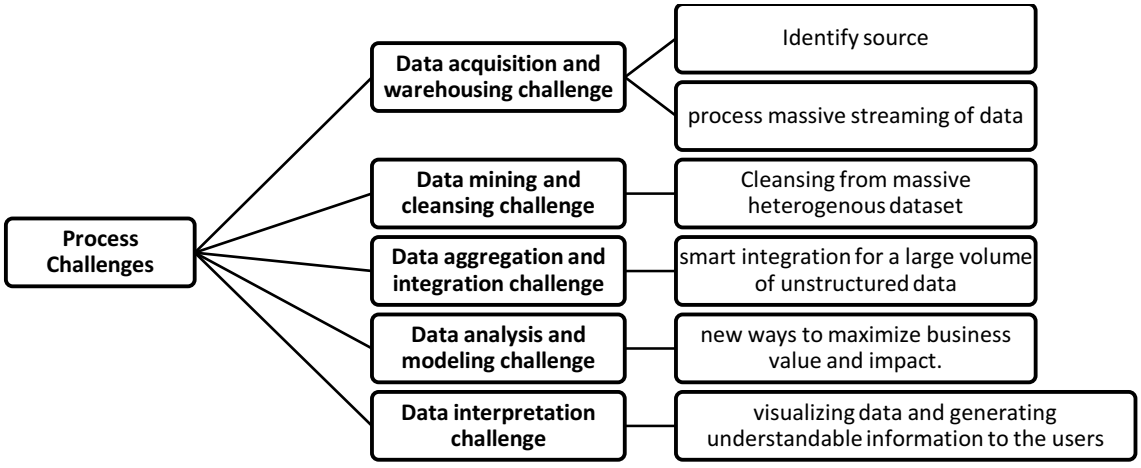

Figure 9: Process Challenges in data analysis (Rashed, et al., 2018)

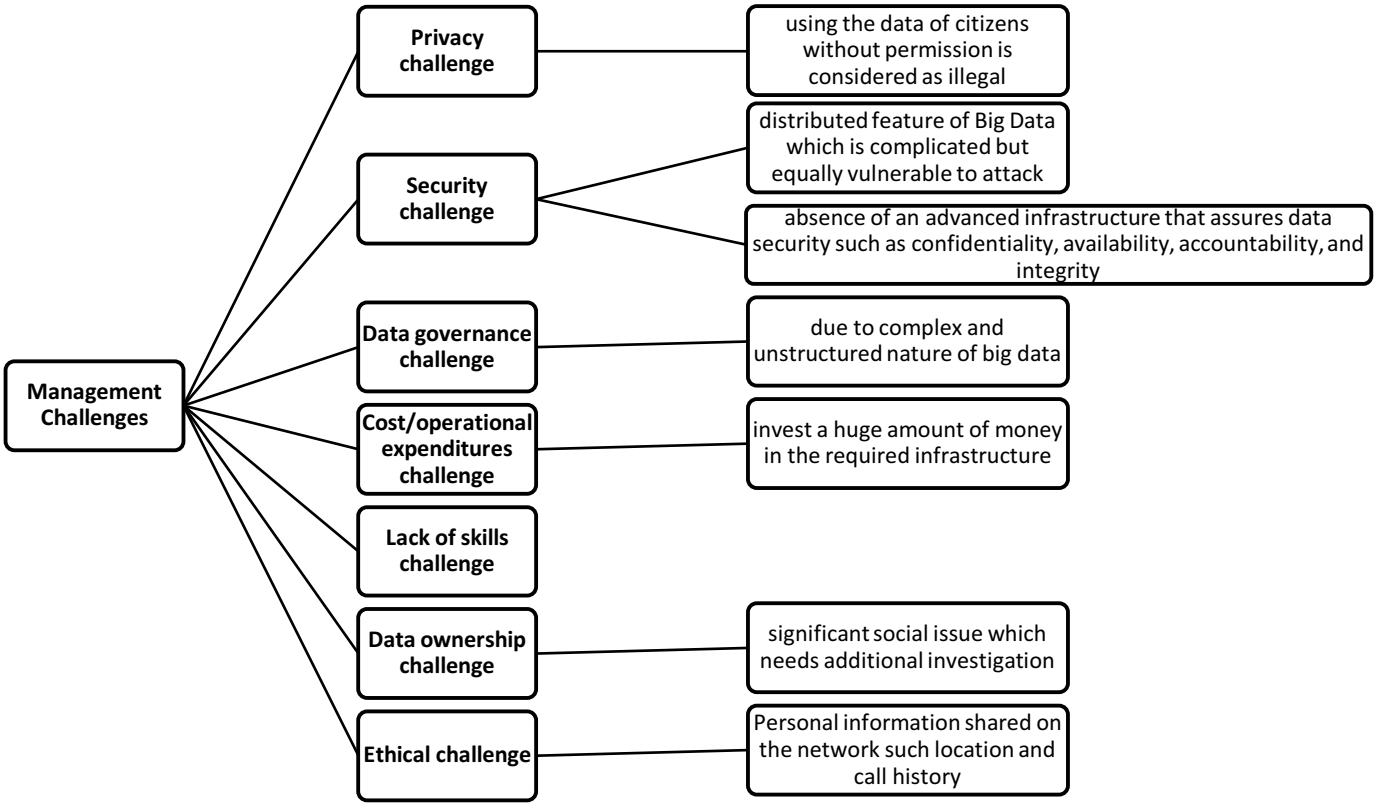

Figure 10: Management Challenges in data analysis (Rashed, et al., 2018)

7. Bibliography

Al-Ruithe, M., Benkhelifa, E. & Hameed, K., 2016. *A Conceptual Framework for Designing Data Governance for Cloud Computing.* s.l., Elsevier.

Androutsopoulou, A. & Charalabidis, Y., 2018. A framework for evidence based policy making combining big data, dynamic modelling and machine intelligence. *Association for Computing Machinery.*

Bott, M. & Young, G., 2012. *The Role of Crowdsourcing for Better Governance in International Development.* s.l.:s.n.

Brabham, D. C., 2013. *Crowdsourcing.* s.l.:s.n.

Certoma, C., Corsini, F. & Rizzi, F., 2014. Crowdsourcing urban sustainability. Data, people and technologies in participatory governance. *Futures.*

DAMA International, 2017. *DAMA-DMBOK: Data Management Body of Knowledge.* s.l.:Technics Publications.

Grier, D. A., 2013. *Crowdsourcing For Dummies.* s.l.:ForDummies.

Hershkovitz, S., 2017. *The future of crowdsourcing: Integrating humans with machines.* [Online] Available at: https://thehill.com/blogs/pundits-blog/technology/324807-the-future-of-crowdsourcing-integrating-humans-with-machines

Klumpp, M., 2017. Crowdsourcing in Logistics: An Evaluation Scheme. *Dynamics in Logistics.*

Kreutz, C., 2016. *We Thinq.* [Online] Available at: https://www.wethinq.com/en/blog/2014/08/12/39-Great-Crowdsourcing-Examples.html

Meers, D., Lampshire, G. & Fryman, L., 2017. *The Data and Analytics Playbook.* s.l.:Elsevier.

Murphy, J., Hill, C. A. & Dean, E., 2013. *Social Media, Sociality, and Survey Research.* s.l.:John Wiley & Sons.

Narvaez, R. W. M., 2012. *Crowdsourcing for Disaster Preparedness: Realities and Opportunities.* Geneva: s.n.

O'Connor, D., Boyle, P. & Ilcan, S., 2017. Living with insecurity: Food security, resilience, and the World Food Programme (WFP). *Global Social Policy.*

Palantir, 2017. *Palantir.* [Online] Available at: https://www.palantir.com/philanthropy-engineering/annual-report/2017/wfp.html

Poblet, M., García-Cuesta, E. & Casanovas, P., 2014. Crowdsourcing Tools for Disaster Management: A Review of Platforms and Methods.

Pollok, P., Lüttgens, D. & Piller, F. T., 2018. Attracting solutions in crowdsourcing contests: The role of knowledge distance, identity disclosure, and seeker status. *Research Policy.*

Rashed, A. H., Karakaya, Z. & Yazici, A., 2018. *Big Data on Cloud for Government Agencies: Benefits, Challenges, and Solutions.* [Online].

Riccardi, M. T., 2016. The power of crowdsourcing in disaster response operations. *International Journal of Disaster Risk Reduction.*

Richards, S. E. & Veenendaal, B., 2014. Understanding the Gap between the United Nations World Food Programme Crisis Mapping Operations and Crowdsourcing Technology.

Southekal, P. H., 2017. *Data for Business Performance: The Goal-Question-Metric (GQM) Model to Transform Business Data into an Enterprise Asset.* s.l.:Technica Publication.

Storeya, V. C. & Song, I.-Y., 2017. Big data technologies and Management: What conceptual modeling can do. *Data & Knowledge Engineering.*

Tableau, 2017. *Tableau.* [Online] Available at: https://www.tableau.com/foundation/featured-projects/wfp

Veeneman, W. et al., 2018. PETRA: Governance as a key success factor for big data solutions in mobility. *Research in Transportation Economics.*

Wang, C.-S., Lin, S.-L., Chou, T.-H. & Li, B.-Y., 2018. An Integrated Data Analytics Process to Optimize Data Governance of Non-Profit Organizations. *Computers in Human Behavior.*

WFP, 2012. *Emergency Preparedness and Response Package,* s.l.: s.n.

WFP, 2014. *ANNUAL PERFORMANCE REPORT FOR 2014,* s.l.: s.n.

WFP, 2014. *WFP Corporate Partnership Strategy (2014–2017),* s.l.: s.n.

WFP, 2017. *MDCA Mobile Data Collection and Analytics.* [Online] Available at: https://documents.wfp.org/stellent/groups/public/documents/ena/wfp291140.pdf

WFP, 2018. *Partners.* [Online] Available at: https://www.wfp.org/partners/

WFP, 2018. *World Food Programme in Syria - Year in Review 2017,* s.l.: World Food Programme.

WHO, 2017. *Vaccination in Humanitarian Emergencies,* s.l.: s.n.