International Journal of Engineering and Information Systems (IJEAIS)
ISSN: 2000-000X
Vol. 2 Issue 9, September – 2018, Pages: 43-56

Leveraging Big Data for Emergencies and Crises - Case Study Gaza Strip

Ihab Zaqout[1], Ibrahim Elhaddad[2]

Department of Information Technology Faculty of Engineering and Information Technology
Alazhar University, Gaza, Palestine
Email: i.zaqout@alazhar.edu.ps[1], ibhadad@gmail.com[2]

Abstract— *Emergencies all over the world have been increasing over the past years due to manmade and natural disasters. Consequently, it has been a challenge to analyze the huge data collected during emergencies and then process these data to formulate a rapid response, especially if in critical times when saving lives is the top priority. With the continuously emerging technologies specifically those that facilitate the process of data collection, processing and analysis, the stockholder's response ability to the affected population will further increase. Using big data tools such as Hadoop or NoSQL if applicable in Gaza context could be leveraged to analyze the large amounts of data produced during emergencies and crises to contribute to changing the situation and providing a proper response for the affected population. Data collection during emergencies has been a main challenge as it's the first stage in the information management lifecycle. Other challenges include how to process the large amounts of data and analyze them to provide the stockholders responding to the emergencies with accurate information and the ability to respond on a timely and professional manner. In this paper, the researchers examined the importance of having clear platforms and tools to be used in case of emergencies regardless their different types in Gaza context and what are the available tools to be used to enhance the response for the affected population.*

Keywords— emergencies; crisis; big data; Gaza; humanitarian**.**

1. INTRODUCTION

In this paper, the problem domain is identified and how it's relevant to Gaza Strip in leveraging of big data for emergencies and crisis. What are the obstacles and challenging in using the technology to respond to the affected population and what's the future plans for using it if it's not available in the current time?

Emergencies are defined as urgent or onset situations in which there is clear evidence that an event or series of events has occurred which causes a humanitarian crisis that affects their lives or livelihoods and which the governments not able to respond to the affected population. The event or series of events may comprise one or a combination of the following: a) sudden calamities such as: floods, earthquakes or unforeseen disasters, which are called natural disasters; b) manmade disasters resulting in an influx of refugees or internal displacement of population; c) slow onset events such as drought, crop failure and diseases that results in increasing the number of vulnerable population caseload; d) severe food access or availability conditions resulting from economic shocks, market failure or economic collapse; and e) complex emergency such as a humanitarian crisis in a country, region or society with more than one generating factor whether internal or external conflict [1].

Emergencies require an international response that goes beyond the mandate or capacities of any single agency and/or the ongoing UN country programme [2].

Big data can be described as massive amounts of public data generated on a daily basis, or it is described by some data scientists as tsunami of data [3]. The data volume is increasing exponentially due to rapid changes. Companies such as Facebook, Google, Twitter, Yahoo, Microsoft and Amazon routinely deal with petabytes of data on a daily basis. It's estimated that we are generating an incredible 2.5 quintillion bytes per day. Big Data is often described as data that are too large to process or analyze on a regular computer with common tools like the popular spreadsheet application Microsoft Excel. Big data is relative to the computing power at our fingertips [4].

Big data refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyse. This definition is intentionally subjective and incorporates a moving definition of how big a dataset needs to be in order to be considered big data—i.e., we don't define big data in terms of being larger than a certain number of terabytes (thousands of gigabytes). We assume that, as technology advances over time, the size of datasets that qualify as big data will also increase. Also note that the definition can vary by sector, depending on what kinds of software tools are commonly available and what sizes of datasets are common in a particular industry. With those caveats, big data in many sectors today will range from a few dozen terabytes to multiple petabytes (thousands of terabytes) [5].

The Gaza Strip has a population of approximately 1.9 million people where they live in an area of 360 km^2. For the last decade, the socioeconomic situation in Gaza has been declining steadily. Years of conflict and blockade have left 80% of the population dependent on international assistance [6]. In Gaza Strip, there are huge amounts of data in need of processing and analysis in order to give the humanitarian actors the ability and knowledge to execute the proper response for the affected population. Over the last decade, Gazans suffered three Israeli wars and attacks leading to major emergencies and increased their vulnerabilities.

The overall crisis life cycle can be divided and analysed into three periods: before, during and after a crisis as depicted in Fig. 1. The first stage is pre-crisis preparedness that deal with the preparation of expected crisis in terms of logistics, planning and warehousing of emergency relief goods in the region or in another location. Using big data techniques in this stage will determine the needed available base line data and identify the gaps that need to be covered. The benefits in this stage is prevention or elimination of the consequences. The second period is during crisis, which is mainly coordinating between humanitarian organizations to address and identify any potential gaps or overlap in responses. Finally, the third period is post-crisis response. This stage deals with reporting, auditing and accounting of the overall humanitarian response. As this stage looks at the crisis response in a more comprehensive, analytical and insightful way in order to reflect the learnt lessons in any future crisis [3].

Currently, benefiting from the data and information collected during crisis and emergencies is a hot research topic yet there is still no clear guidelines or procedures for benefiting from huge volumes of data or any clarifications on the means in which humanitarian organizations or workers can benefit from these data to target the most vulnerable affected population.

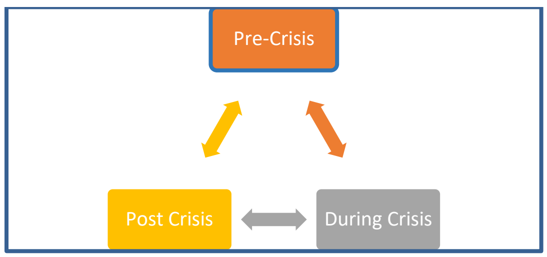


Fig. 1. *Crisis Life Cycle*

The problem is in identifying the best approaches with acceptable accuracy and creditable information from the Gaza Strip ad-hoc emergencies in order to identify the needs of affected population. The humanitarian actors serve in different types of emergencies, and not restricted to only natural disasters. The data collected during emergencies can change the mean of interventions and response of aid organizations. In addition, every emergency has special characteristics for example, (Syrian migrant's crisis as civil wars emergency, Haiti earthquake as natural disasters and conflict such as 2014 Israeli war on the Gaza Strip).

Major disasters and emergencies bring chaos and confusion. In addition, the lack of availability or applicability of information management standards and guidelines during emergencies with other additional factors affect the response efforts to the affected population. If not processed and analysed in proper ways based on guidelines and standards, it will lead to gaps in responses and duplication of efforts and wasting time of humanitarian workers. In addition, it will restrict decision makers' the ability to provide the proper decision in all areas, which will also result in gaps/delays of response to the affected population.

Lack of information sharing protocols and legal frameworks within the humanitarian organizations, will affect the data quality and the response to the affected population. Another factor is the lack of clear sources of data during emergencies.

All the above mentioned issues will affect the initial phase of information management lifecycle which is data processing and as a result will restrict the ability to use data driven approach to response for emergencies and disasters and leveraging of the big data for emergencies and disasters response. In this study, we tried to achieve the below objectives:

- To describe the promise and potentials of big data analytics in emergencies and disasters.
- To prove that big data will eventually benefit the humanitarian works and respond to emergencies and disasters, as there's so much new data collected, emergency managers will be more confident leveraging analytical information in decision making and responses.
- To reinforce the process of drawing attention to the needs for a higher level of capacity to use it in emergency management community.
- To prove that using big data will increase the accuracy and reliability of response to the affected population.

As the crisis increased in the last years all over the world with continuous growth in the data amounts received from humanitarian responders, and for the purpose of using the collected information in lifesaving, this study appears to be a significant study. This paper will help different stockholders to focus on using the big data in emergencies and crisis, and how the revolution of technology and large amount of data are transforming in all emergencies phases, as parts of preparedness and responding.

The remainder of this paper is organized as follows: Section 2 gives a literature review. Section 3 describes the research methodology. Section 4 presents the data analysis and discussion. Conclusions and recommendations are given in the final section.

2. LITERATURE REVIEW

The researchers took into account other related works done in how to benefit from Big Data in emergencies and how to best use the collected data as response during emergencies and crisis times. As the context of emergencies differ from one crisis to another, the researchers tried in this study to illustrate the most useful tools and guidelines to enhance the response to crises and needs for the most affected population in the Gaza Strip.

Analysing huge amount of data during mass emergencies is not an easy operation, for responders face difficulty in how to assess the conditions of the affected areas and population. This is a hard component to determine how to respond based on priorities, investments and deploying supplies and services. Collecting and processing of the data during the emergency is a fuzzy issue. The importance of data for humanitarian responders has attracted their attention in the recent years, but the challenge is how to transform these huge data into information and then to an effective response or intervention.

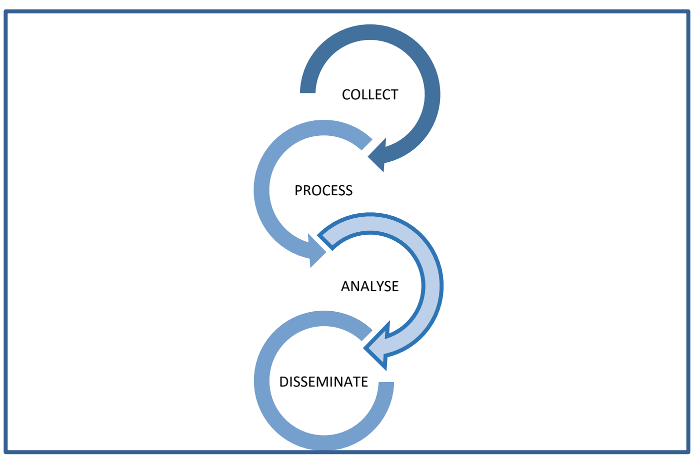
Therefore, we looked deeper on a large scale emergencies and disaster which is Gaza war (2014) and then building a road map or giving guidelines to humanitarian responders for leveraging big data in these situations.

Big Data is one of the trending research topics in science and technology communities, and it possesses a great application potential in every sector of our community/society, such as climate, economy, health, social science, etc. Any failure in the previous sectors will lead for an emergency. Big Data usually includes data sets with sizes beyond the ability of commonly used software tools to capture, curate, and manage. Therefore, the big data specialists can conclude that Big Data is still in her infancy stage, and they will face many unprecedented problems and challenges along the way of this unfolding chapter of human history. It is critical to explore the theoretical perspective of Big Data to efficiently and effectively guide its applications [7].

We consider that there is a significant development in Big Data from various communities including: the mining and learning algorithms from the artificial intelligence community, networking facilities from networking community, and software platforms for software engineering community. However, Big Data applications introduce unprecedented challenges and existing theories and techniques that have to be extended and upgraded to serve the forthcoming real Big Data applications. In addition, the need to invent new tools for Big Data applications is very notable.

Information is the most valuable thing during emergencies or disasters. Information is what everyone needs to make decisions to respond to the humanitarian needs and decisions to gain or lose organization visibility and credibility. In the era of information revolution, information sources are the most important strategic resources. However, to reach a level of satisfaction, understanding the information management stages is very essential. Information has life cycle and steps as any other procedure in our globe. Information management life cycle has main four steps (Collection, Processing, Analysing and Disseminating) as shown in Fig. 2 [8]:

Fig. 2. *Information Management Life Cycle*



1. ***Collection*** refers to those technical and non-technical activities that lead to the establishment of data or information body. In the context of humanitarian operations, the collected information may include aggregated or disaggregated data. The responsibility for primary data collection will in most circumstances rest with specialized agencies. Collection process

should be identified during different stages of emergencies, as for example during pre-emergency stage, humanitarian organizations during preparedness should identify what's the needed information which will be collected as baseline information and update their datasets for unexpected emergencies. During emergencies, the situation will be more challenging as more factors will affect the needs mainly the time factor. Rapid assessment is one of the main methodologies that used during emergencies to identify the needs for the affected population.

2. ***Processing*** includes mainly transforming the raw data collected to a format that can be easily understood or manipulated with other collected data for more analysis. This stage includes: cleansing and compiling with other different data collected from various sources. These functions are mainly but not exclusively, carried out by information management staff.
3. ***Analysis*** is the process of presenting the summary of the collected data after processing, and this can come into two formats, one by presenting them through statistics and charts, while the other one is by adding some analysis and narrative based on the context of the humanitarian situation.
4. ***Dissemination*** is the last step of the information management lifecycle, which puts information into the decision and policy makers at different levels.

On daily basis, massive amounts of public data are being generated, the data volume, and there is a constant increasing in data exponentially due to the rapid revolution of mobile phones and increased digitization of all aspects and fields of our modern life due to the availability of technologies such as Internet of Things (IoT), which deploy sensors for almost everything. Big data have more than one definition, one of them is big data to be data that exceeds the processing capacity of conventional database and analytics technology, so the data shape is unstructured, too large or too fast.

In order to understand the relation between big data and crisis analytics, we need to understand what big data is. It can be defined as our emerging ability to collect, process and analyse massive sets of largely unstructured data for example, word documents, blog posts, emails, social media posts/tweets and multimedia data that cannot be stored in an organized relational database [3].

Big data during crisis has many sources such as: data exhaust, online activity (Social Media), sensing technologies, small data/MyData, public/governmental data and crowdsourcing data. Those many sources are explained as follows:

- ***Data Exhaust***: It refers to information that is passively collected along with the use of digital technology. An example of this data source for big crisis data analytics is the Mobile Call Details Records (CDRs), which are generated by mobile companies to capture various important information and details related to any call made over their network. Also, data exhaust may include transitional data such as banking records and credit card history and usage data. An example that is widely used in Africa for the purpose of Big Data in action for development 2014 is that one of the most profitable banking service in Kenya, which is a solution based on the digital currency Bitcoin named BitNation, has been developed to provide banking for the refugees and the crisis affected population (http://refugees.bitnations.co/) [5].
- ***Online Activity***: This includes all types of social media activities on the Internet, in other words all the user generated data on the Internet. Emails, Short Message Service (SMS), blogs, comments, social media platforms (such as Facebook comments and posts, Google+ posts and Twitter tweets), search activities using search engines such as Google/Bing, etc. are examples for data generating sources. Regarding using the online activities in emergencies and crisis, not all of them can be used during this critical situation as the most important issue is apparently saving lives. For more clarifications, SMS is used mostly on the ground by the affected community, while twitter is used mostly by the international aid organizations/communities. One advantage of online data is that it is often publicly available and it is heavily used by the academics in big crisis data researches [9].
- ***Sensing Technologies***: These are mostly gathered information about social behaviors and environmental conditions using a variety of sensing technologies. It is anticipated that sensor data will match or even outgrow social data in the future. There are different types of sensing technologies such as (1) remote sensing in which satellite or high flying aircraft scans the earth in order to obtain information about it, measuring traffic, climate or other environmental conditions; (2) networked sensing in which sensors can perform sensing and can communicate with each other as in wireless sensor networks; (3) sensing technologies like everyday devices, i.e. mobile phones. Different cyber-physical sensing systems can play an important role in disaster management. These devices can be an active source of big data by providing information in the form of images or sensed data. Sensing data is mostly public [3 – 5].
- ***Small data and MyData***: The nature of big data is how to deal with tsunami of data. With big data, the units of sampling and analyses are vastly dissimilar. In small data even if the datasets are too big, the unit of analysis is similarly scoped as the unit of sampling. In terms of MyData, it can help in optimizing individual human lives as there is a lot of work in benefiting of MyData in health system applying this approach and use this source of big data to save people lives [3].
- ***Public-Related Data***: A lot of public related data can be useful during emergencies and crisis, and they are collected in various offices. These data include: census, birth and death certificate, and socio-economic data. By analyzing all of these

available data during emergencies and crisis, the humanitarian organizations and workers' capacities will further get improved for better responding to the affected population and crisis [4].

- ***Crowd Sourced Data***: Applications are actively involving a wide user base to solicit their knowledge about particular topics or events. Crowdsourcing platforms are usually publicly available, and are widely used by big crisis data practitioners. Crowdsourcing is a task is outsourcing. However, it is not to a designated professional or organization but to general public in the form of open call. Crowdsourcing is central to the ways in which this new type of organization makes data. QuakeMap example shows that crowdsourcing of crisis information can result in data sets that reflects existing inequalities, especially digital divides [10].

As described above, big data is unstructured data, so the traditional relational database can’t store such data which require data to be structured; so modern tools and approaches are required to store, save and access the unstructured data. NoSQL or non-relational database have been developed for this purpose, which these types of database have new features such as scalability, speed, and flexibility. Big companies use the NoSQL database such as Amazon, Google.

NOSQL Definition is Next Generation Databases mostly addressing some of the points: being non-relational, distributed, open-source and horizontally scalable. The original intention has been modern web-scale databases. The movement began early 2009 and is growing rapidly. Often more characteristics apply such as: schema-free, easy replication support, simple API, eventually consistent / BASE (not ACID), a huge amount of data and more. So the misleading term "NoSQL" (the community now translates it mostly with "not only sql") should be seen as an alias to something like the definition above [11].

To correctly categorize the data, and instead of using relational database which is very difficult to be used in case of using big data during emergencies and crises, a more flexible management policy is necessary for big data to meet specific needs. Since big data are often collected from heterogeneous resources, applying flexible data management offers a more reliable and instant storage plan for heterogeneous information at different levels, compared to general data management solution. Mass data operations are also suitable for further analysis and statistics [12].

In 2001, “the four Vs.” term was generated to identify the characteristics of big data by providing a comprehensive definition of big data in light of the four Vs. as shown in Fig. 3. Those four Vs. will be used to describe the evolution of the overall change of data we generate, collect and store. They are also responsible for the big data phenomenon [13].

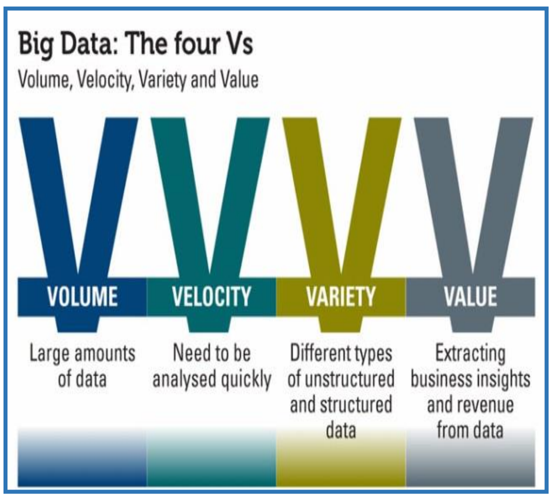


Fig. 3. *Big Data: The four Vs* [14]

- ***Volume:*** The volume of data organizations and individuals own grows rapidly. This growth can be quantitative measured by looking at the amount of bytes’ organizations collect and store, the number of records they have stored in their databases, the number of transactions being run or by counting the number of tables, files and other forms of data.
- ***Variety:*** Although the increasing amount of data is a main driver for big data analytics, the variety of data being generated, stored and made public, plays an important role in describing the big data phenomenon. Data can be divided into two groups: structured and unstructured data as explained. In order to store and analyze unstructured data, the need to dynamically add new attributes to support new forms of data without changing the structure of the database is increasing. Since storing unstructured data is becoming easier, analyzing this type of data will be more feasible too and this is important since in 2003, more than 85 percent of all business information exists as unstructured data.

International Journal of Engineering and Information Systems (IJEAIS)
ISSN: 2000-000X
Vol. 2 Issue 9, September – 2018, Pages: 43-56

- ***Velocity:*** It refers to the speed at which new data is being generated. A source of big data that generates data at a high speed would naturally produce large volumes of data. Social media such as: Facebook and Twitter are examples of such sources of data. Facebook users upload enormous numbers of postings and photographs with around 1.6 billion active Facebook users in 2016, the speed requirements for big data is clearly being fulfilled. The use of social media posting as a source of data has become popular [15].
- ***Veracity:*** It refers to the accuracy of database. Based on the "garbage in garbage out" principle, if a database contains too much inaccurate data, the analysis conducted using this database will generate inaccurate results. In the content of Big Data, a certain degree of inaccuracy is acceptable because the volume of data is much larger than in traditional small datasets, offsetting the potential errors [15].

Although 4Vs definition of big data is popular, it is still difficult for scholars to determine whether a dataset has fulfilled the 4Vs definition or not. The definition is also difficult to use. For example, the term volume is the concept of big which means that the data is difficult to be quantified, and it is difficult to come up with an objective threshold value to classify whether a dataset is big or small. The concept depends on the times and discipline of domain.

The big data crisis analytics is mainly about the aim to leverage big data analytics techniques, along with digital tools and platforms (such as mobiles, internet) for effective humanitarian response to different crisis. There are many thematic applications of big crisis data including:

1. Data-driven digital, an example of this is digital epidemiology in which public health researchers are conducted using social media. In humanitarian crisis, the challenges include how to analyze the collected data related to such crisis.
2. Population surveillance and urban analytics: in which big crisis data is used for tracking the movement of the affected population during the crisis.
3. Crisis information and sociology. Here, data along with participatory mapping and crowdsourcing technology is used for analyzing the sociological behavior of the affected community through behavioral inference and reality mining that is data mining to extract and study social patterns of an individual or a group of individuals.

Also, we can consider that the various big data crisis analytics tasks that can be performed on big crisis data such as various discovery tasks, e.g. clustering the dataset to obtain natural grouping, outlier detection (to detect any anomalies).

Unfortunately, in the Gaza Strip, the applications of big data in crisis and emergencies is not applicable in a way or in another, despite of the huge data collected during the three crisis phases. Based on the United Nations Office for Coordination of Humanitarian Affairs (OCHA) as mandated to the information management during the crisis, the techniques is still used is depending on data collection mechanisms and no big data analytics used due to many other challenges. During this paper, the researcher examined the best and suitable ways of big data analysis that fit the Gaza Strip context.

3. RESEARCH METHODOLOGY

The study tools aimed at collecting information about the following: the research design, research population, questionnaire design, statistical data analysis, content validity and pilot study. The data were collected from the primary and secondary resources. The secondary resources include the use of books, journals, statistics and web pages. On the other hand, the primary data were collected by using questionnaires developed specifically for this research. Many of measurement tools "questionnaires" used by other researchers were adapted, translated, combined and modified to fit the purpose of this research ended up in developing one questionnaire distributed to specialized Information management focal points in their respect organization to collect the primary data, the researchers retrieved 25 out of them. In addition, the researcher conducted seven unstructured interviews to understand more about the challenges and opportunities.

The methodology flowchart which has been followed by the researchers and which lead to achieve the research objective is shown in Fig. 4.

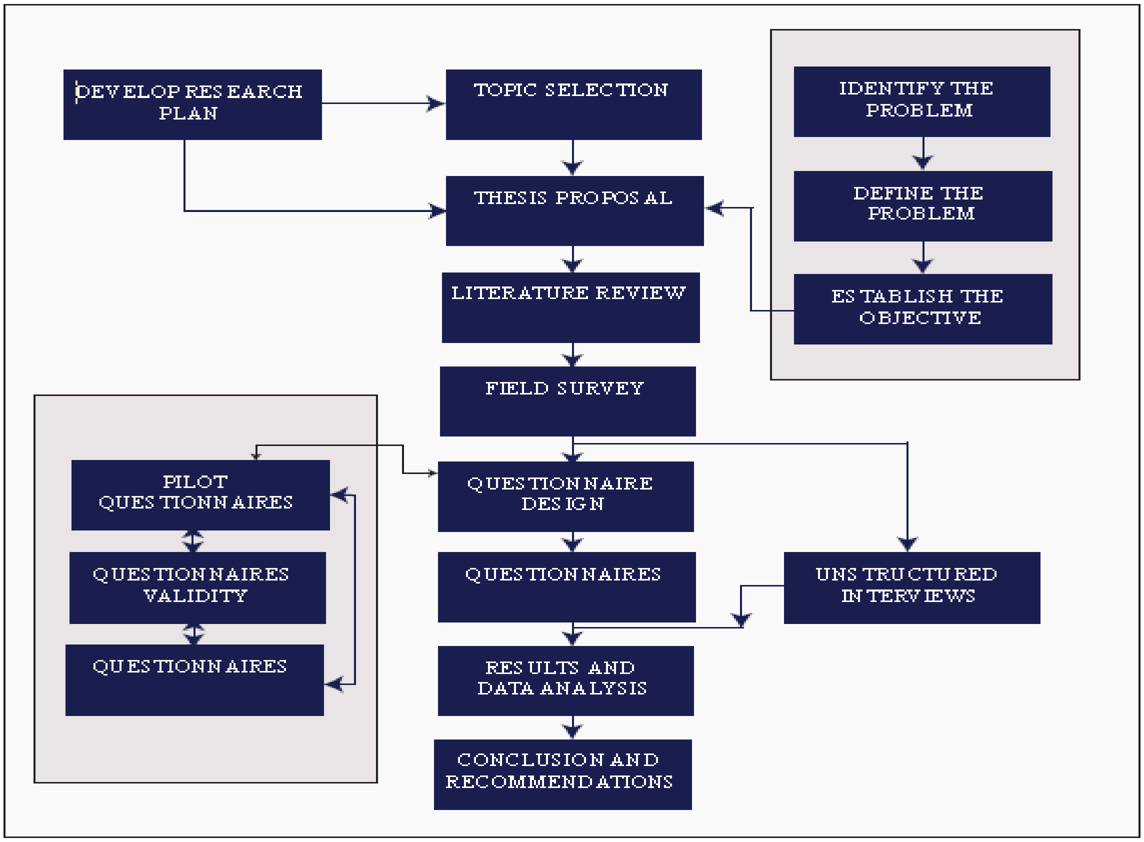


Fig. 4. *Methodology Flow Chart*

3.1 Population and Sample

The population of research consisted of international and local senior staff members who are engaged in emergencies in the Gaza Strip working in different bodies. The total number of those staff members is (25) staff representing six main counterparts. Those staff members are distributed among six operations bodies. The researchers inspected them both physically or by e-mails.

Fig. 5, displays the number of responders distributed according to the different organizations they are working at including: United Nations agencies, international non-governmental organizations, Palestinian non-governmental organizations, governmental bodies and other private sector organizations. The majority of the sampled which constitutes about 44% of the responders are working in the United Nations agencies. This percentage is followed by 32% who are working with international non-governmental organizations. Only 12% of the responders are working with governmental bodies, followed by 8% are working with local non-governmental bodies, and only 4% of the responders are working with private sector.

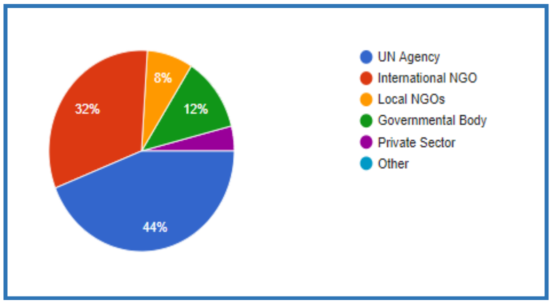


Fig. 5. *Number of Responders Disaggregated by Organization Type*

The percentages of responders disaggregated in accordance to their position types whether international or local ones are illustrated in Fig. 6. It appeared that 84% of the responders are working in local positions while only 16% are working in international positions.

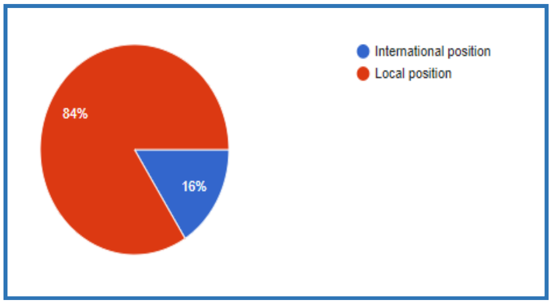


Fig. 6. *Number of Responders Disaggregated by Occupation Type*

The below chart (Fig. 7) shows the number of responders according to the academic qualifications. While 64% of the responders are bachelor degree holders while 36% are master degree holders.

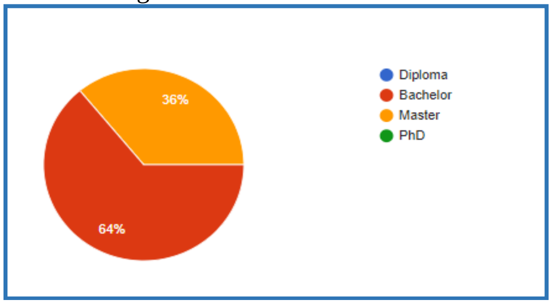


Fig. 7. *Number of Responders Disaggregated by Qualification*

Fig. 8, presents information on the percentage of the responders according to their gender. While 88% of the responders are females and 12% are males.

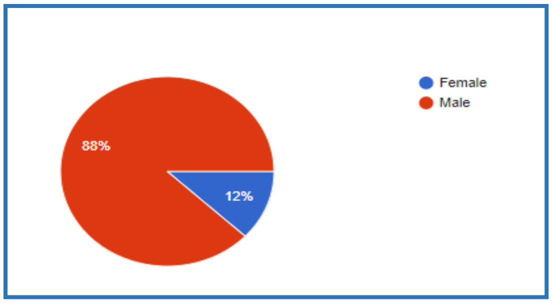


Fig. 8. *Number of responders disaggregated by Gender*

As shown in the below Fig. 9, 48% of the responders are occupying “officer” positions, followed by 24% are occupying “senior officer” positions, followed by 16% are occupying “head of department positions”, followed by 8% are occupying “director” positions, and 4% are occupying “deputy head of department” positions.

International Journal of Engineering and Information Systems (IJEAIS)
ISSN: 2000-000X
Vol. 2 Issue 9, September – 2018, Pages: 43-56

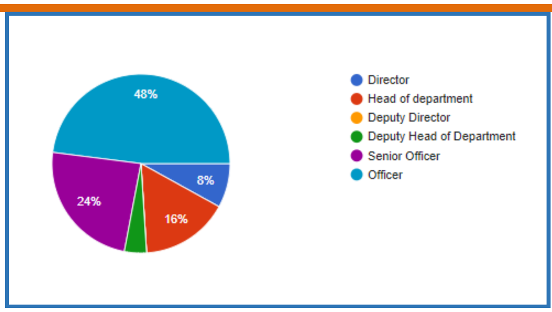


Fig. 9. *Number of Responders Disaggregated By Occupation*

The below Fig. 10, gives information about the number of the responders disaggregated by the organization size. 48% of the responders are working in organization size between 10-00 staff, followed by 24% are working in organization size more than 1000 staff member, followed by 16% of the responders are working in organization size between 100 – 999 staff member and 12% are working in organization size between 1-10 staff members.

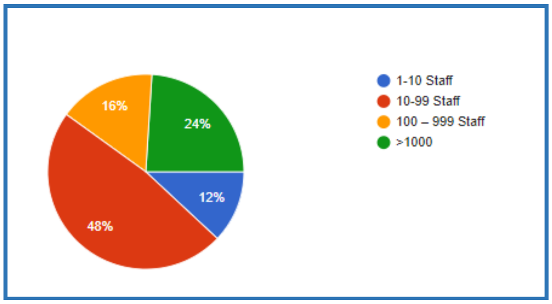


Fig. 10. *Number of responders disaggregated by Organization Size*

As shown in the below Fig. 11, 68% of responders have information management experience in both types of emergencies, followed by 24% of the responders have experience in only man-made emergencies and only 8% have experience in natural disasters.

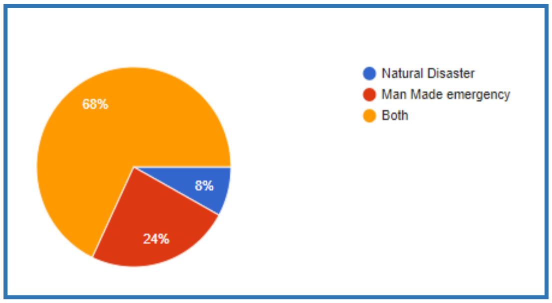


Fig. 11. *Number of Responders Disaggregated by Emergency Type*

3.2 Pilot Study

A pilot study of five respondents for the questionnaire was conducted before collecting the results of the sample. It provided a trial run for the questionnaire, which involves testing the wordings of questions, identifying ambiguous questions, testing the techniques that used to collect data, and measuring the effectiveness of standard invitation to respondents. Thereby, it can be said that the researchers proved that the questionnaire is valid, reliable and ready for distribution for the population samples.

International Journal of Engineering and Information Systems (IJEAIS)
ISSN: 2000-000X
Vol. 2 Issue 9, September – 2018, Pages: 43-56

3.3 Interviews

In addition to the questionnaire, the researchers used unstructured interviews with the stockholders. Five unstructured interviews conducted with key different stockholders to collect additional information and discuss the challenges and opportunities to understand more about how to leverage from the big data in case of emergencies and crises. The interviews provided the researchers with additional and clearer ideas about the scope of work needed to be done and this will be included in the next section about the data analysis and discussion.

4. DATA ANALYSIS AND DISCUSSION

This section represents the research findings and the analysis of the data collected as part of this study. The purpose of this section is to provide a comprehensive overview of the entire data set collected and the characteristics of the respondents.

4.1 Data-Driven innovation in the organization

The majority of the respondents clarify that there are no experiences in their organization with big data and data-driven innovation. Their organizations don't have a strategy to work on big data or data analytics. The below charts and analysis illustrates the findings.

Seven responders (28%) responds that there is no experience with big data and data-driven innovation, (28%) considering the data-driven innovation in their organization is applied. However, 36% responded that either they are planning DDI or piloting it. While just two received responses (8%) answered that they are effectively using DDI. More details available in Fig. 12.

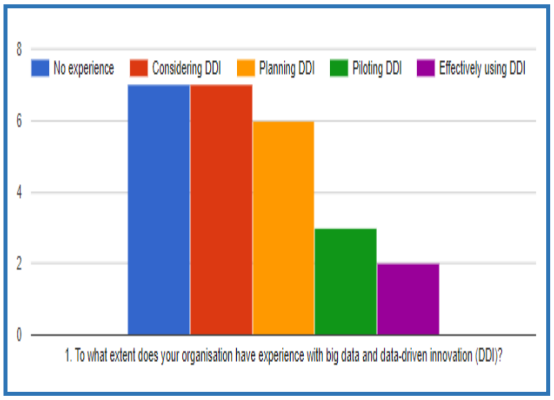


Fig. 12. *Experience with Big Data and Data-Driven Innovation (DDI)*

Concerning the availability of strategies on big data or data analytics, the research results found out that (52%) of the received responses agreed that their organizations didn't have strategies on big data or even data analytics. More details available in Fig. 13.

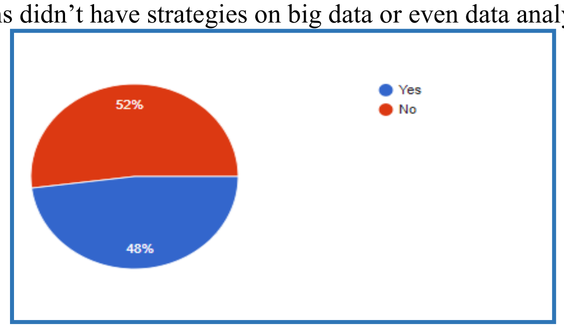


Fig. 13. *Strategy on Big Data or Data Analytics*

In terms of the sources of data collection, (Fig. 14) analyses the received responses which illustrate that data collection move forward and there are some useful resources currently and others are expected in next five years.

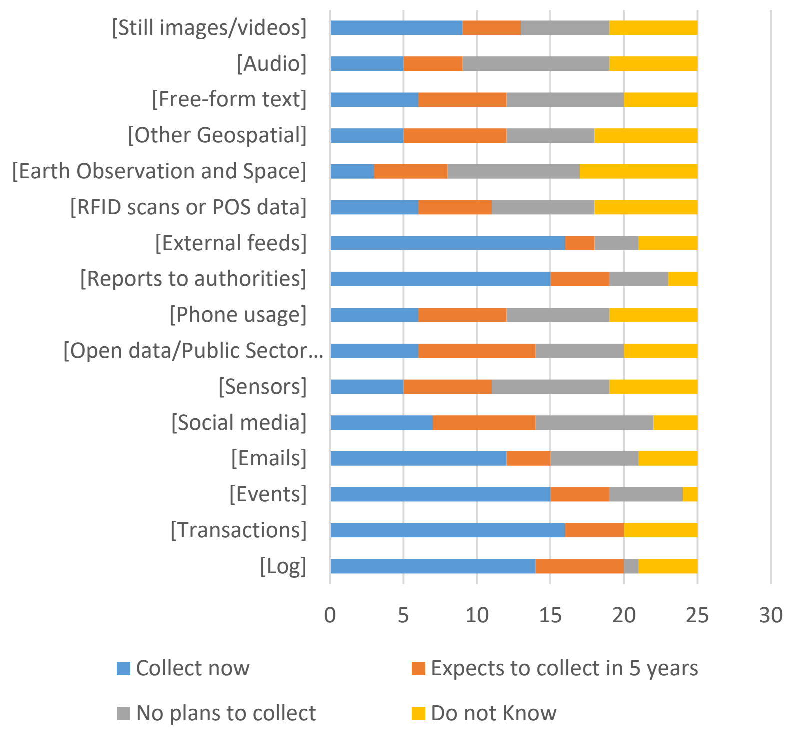


Fig. 14. *Sources of collected data or expected to be collected in 5 years*

In terms of data collection and percentage of data used to improve the process in the respected field. The researchers received 10 responses which means that (40%) of the responders are dependent partially (10-40%) of the collected data based on external sources of data (outside their organization). However, the data is not used in a systematic way to be used in the improvement of the organization responses or activities. More details available in Fig. 15.

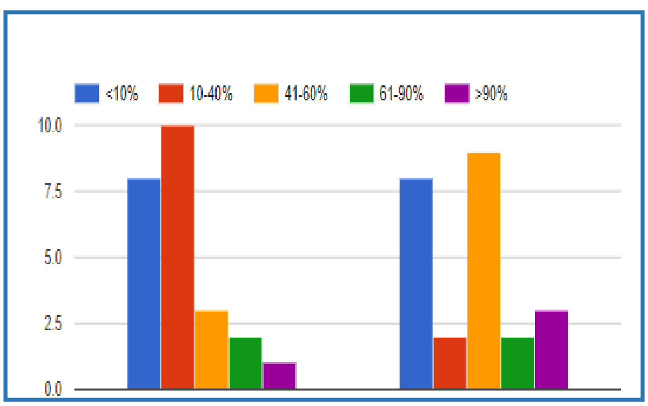


Fig. 15. *Data Collection and data usage*

The main obstacles that prevent accessing or sharing the related data are identified based on the received responses by the organization rules and policies, privacy and confidentiality and not clear information sharing protocol between the stockholders.

4.2 Use of data analytics

Based on the received responses, the majority of the collected data used currently from 41-60%, but it's expected in the next five years, that the data collected usage percentage reached more than 90%. More details in Fig. 16. In another area, the organization that they don't have right now or they will have in the future the right analytical tools to handle data represented by 68%, and this will be one of the recommendation in section five. Fig. 17 provides more details.

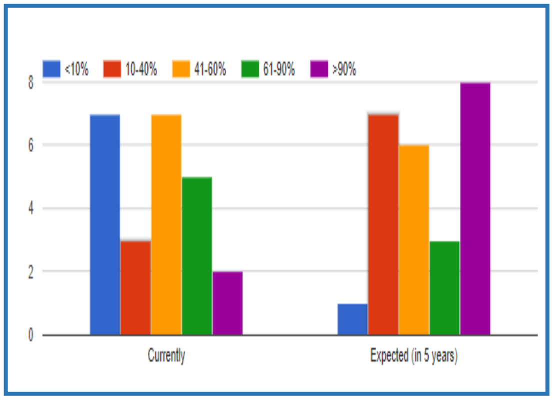


Fig. 16. *Processing of the collected data for value generation*

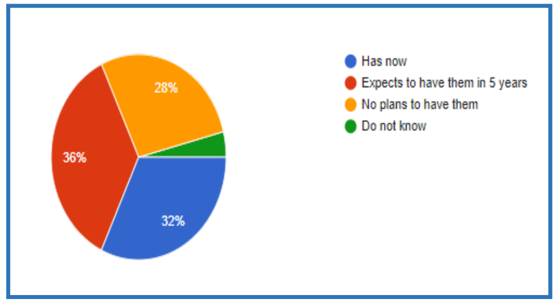


Fig. 17. *Availability of Right Analytical Tools to Handle Data*

Regarding the right tools to handle unstructured data expressed in a natural language, the received responses illustrates that only 12% have the tools and remain either they will have in the next five years or there are no plans to have such a tools. More details in Fig. 18.

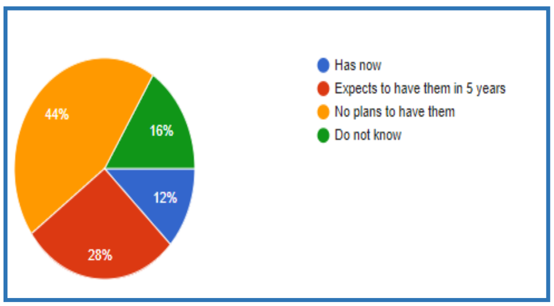


Fig. 18. *Right Tools Availability to Handle Unstructured Data*

5. Conclusions and Recommendations

This section provides an overview of the conclusions and recommendations. Section 5.1 concludes the work of this research. Section 5.2 describes the contributions to the literature. Lastly, section 5.3 elaborates on the limitations of this study and the possibilities for future research.

5.1 Conclusion

This research conducted to check what are the available tools in the context of Gaza Strip to leverage the available data in decision making especially in emergencies and disasters. However, the topic is still new and classified as hot research area. The researcher achieves the research objectives and beyond based on the results and try to create a road map for the future researcher.

In Gaza Strip, there are willing to use the big data for the sake of response and many of the stockholders and responders have plans to use the data driven innovation within the coming five years and building capacities, but this might be affected by the general situation inside the strip.

Also, those organizations that are interested in investing in the big data should use clear road map for their own operations. This roadmap should start from capacity building for the staff, identifying the proper tools that fits the purpose of the organizations and also, and not only focusing on building smarter big data system.

Visualization, is a very important tool to reflect the data and clarify it in a simple way. Therefore, selecting the right visualization tool would help in presenting the data in a more understandable and clear manner.

The last point, the last V "Value" should be considered in any initiative of big data leveraging in organizations, as this will extract the business value from data and will reinforce the process of drawing attention to the need for a higher level of capacity to use it in emergencies.

5.2 Contribution to literature

This study has several contributions to the literature. The first contribution would be the performed literature review. The paper has shown the current state of knowledge on big data in the Gaza Strip. From a scientific perspective, the performed literature review on Big Data will provide other researchers with a clear overview of the available articles, papers, and other scientific material especially related to the context of the Gaza Strip.

The second contribution are the results of both the conducted interviews and the distributed questionnaires. Big Data are still a relatively novice research topic; hence, the results of this research can contribute widely to the Big Data literature. The conducted interviews can provide others and interested researchers with some empirical evidences on how people think and feel about Big Data. In addition, these interviews can also help to confirm or deny interesting patterns between Big Data especially in time of emergencies.

The last and most important contribution of this study is the development of a set of recommendations for other organizations. The aim of these recommendations is to help other organizations to start with Big Data in a more systematic and steady manner. These recommendations are based on literature, the implemented interviews, and on experiment. The main difference with other studies is that this research goes beyond the "buzzword" and has a more practical approach to Big Data.

5.3 Limitations and future research

The study is limited by the constraints of participants' number and availability. As this paper is restricted to only dedicated stakeholders which resulted in that multiple scenarios are not covered. In addition, the study limits itself to the humanitarian workers inside the Gaza Strip which open the door for future researchers to work on other areas or add other areas in terms of the geographic scope of their works like adding the West Bank cities or other cities inside 1948 Palestinian lands.

Another limitation would be applied on the implementation of these study results. These results and recommendations with its totality as they were produced with specific conditions and contexts. Therefore, the researchers recommends the organization to modify these results in a way that suit the operational context.

An important aspect of the level of generalization of a paper is the sample size. Since this paper only made use of dedicated stakeholders, for example one questionnaire filled by UNRWA reflected their approach in dealing with big data, there is a chance that responders do not cover all the variety of scenarios. In addition, all responders were from a single country. Therefore, future research should be conducted with more cases, a bigger sample size, and cases from all occupied Palestinian territories.

The other limitation is that the generalization of these results are somewhat limited. Due to the nature of this study and the points mentioned above, the results are not generalizable for all organizations. Though, most organizations could tweak the set of recommendations of this study and will therefore still benefit from this research. Therefore, future research should focus on improving these recommendations with the help of a bigger sample and more cases from different countries and branches.

REFERENCES

[1] World Food Programme. (2005). Definition of Emergencies. Available on: www.wfp.org.

[2] UNHCR. (2007). Handbook for Emergencies. Genève, Swizerland.

[3] Qadir, J., Ali, A., Raihan, R., Andrej, Z., Arjuna, S. and Jon, C. (2016). Crisis analytics: big data-driven crisis. Journal of International Humanitarian Action, vol. 1, no. 12, pp. 1 – 8.

[4] Meier, P. (2015). Digital Humanitarians - How BIG DATA is Changing the Face of Humanitarian Response. CRC Press is an imprint of Taylor & Francis Group, an Information business.

[5] Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Byers, A.H. (2011). Big data: The next frontier for innovation, competition, and productivity, McKinsey Global Institute.

[6] UNRWA. (2017). Where we Work. Available on: https://www.unrwa.org/where-we-work/gaza-strip.

[7] Yu, D.S. (2016). Theoretical Foundations for Big Data Applications: Challenges and Opportunities. IEEE Advancing Technology for Humanity, vol. 4, pp. 5730 – 5732.

[8] OCHA IM Guidelines. (2016). Information Management and the Humanitarian Context.

[9] Chu, W. (2014). Data Mining and Knowledge Discovery for Big Data: Methodologies, Challenge and Opportunities. Studies in Big Data, Springer-Verlag Berlin Heidelberg.

[10] Mulder, F., Ferguson, J., Groenewegen, P., Boersma K. and Wolbers, J. (2016). Questioning Big Data: Crowdsourcing crisis data towards an inclusive humanitarian response", Big Data & Society, vol. 4.

[11] Your Ultimate Guide to the Non-Relational Universe. (2017). [Online]. Available: http://nosql-database.org/.

[12] Wang, J., Wu, Y., Yen, N., Guo, S. and Cheng, Z. (2016). Big Data Analytics for Emergency Communication Networks: A Survey", IEEE Communications Surveys & Tutorials, vol. 18, no. 3, pp. 1758 – 1778.

[13] Doug, L. (2011). 3D Data Management: Controlling Data Volume, Velocity, and Variety", META Group Inc.

[14] The world news media network. (2017). Big Data: The four Vs. Available on: https://newsbizblog.org/2013/11/14/big-data-the-four-vs/.

[15] Wong, H.T., Chiang, V.C., Choia, K.S. and Loke, A.Y. (2016). The Need for a Definition of Big Data for Nursing Science: A Case Study of Disaster Preparedness", International Journal of Environmental Research and Public Health, vol. 13, no. 10, pp. 1015.

Authors

Ihab Zaqout, Associate Professor of Faculty of Engineering and Information Technology in Alazhar University, interested in digital image processing and data mining applications.

Ibrahim Elhaddad, Post-graduated student for master degree for computing and information systems in Alazhar University.