Global Scientific JOURNALS

GSJ: Volume 10, Issue 1, January 2022, Online: ISSN 2320-9186

www.globalscientificjournal.com

Jimma University

Institute of Technology

Faculty of Computing and Informatics

Department of Information Science

Graduate Seminar for PhD program in Information Science

Title: Application of Big Data Analytics in Natural Disaster Management in African Countries

Prepared By: - Tefera Sime (PhD. Students)

Advised By: - Dr. Getache Mamo (PhD)

Tefera Sime

Jimma, Ethiopia

December, 2021

Approval Sheet

Application of Big Data Analytics in Natural Disaster Management in African Countries

Sign of Candidate: -

Name: - _____Tefera Sime (PhD. Student) Sign _____Date_____

Advisor signature

Name: - _____Dr. Getachew Mamo (PhD Sign _____ Date _____

Examiner signature

Name: - _____ Dr. Worku Jimma (PhD) Sign _____Date _____

Contents

List of figures

Executive Summary

Disaster is the disorder in reliable situation happening in a level of suffer that exceeds volume of change and disturb society. The disaster management phases, 'Mitigation', 'Preparedness', 'Responses' and 'Recovery'. Big data uses Hadoop, Apache Spark & extra outlines for distributing and saving data. Whatever, application of big data is endless, in Africa still infancy age. Objectives, to analysis applications of BD in natural disaster management, to explore consequence of natural disaster and narrow knowledge gap on big data analytics developed and developing globe. Method used comprehensive literature review. Sources database; Scopus, IEEE Xplore, Elsevier, ScienceDirect and keywords used 'disaster', 'disaster management', 'big data', 'evolution of big data', 'big data application on disaster management'… By Nima Yaghmaei and Regina,2019, rapid population growth and increasing impact of climate change likely increase in coming decades and over 50 African countries and 30 million square kilometers land threat of natural hazards. (2000-2019*), through cases of drought, extreme temperature, flood and storm in Africa, 36,704 people died from ten countries. In Somalia 20,739, Algeria 3,777, Mozambique 2,291, Nigeria 1,696, Madagascar 1,644, Ethiopia 1,639, Kenya 1,572, Sierra Leone 1,289, D.R. Congo 1,072 and Malawi 985. These countries need to improve big data infrastructure, produce educated human power. Because, evolution of digital data, social media, satellite remote sensing and sensor networks has contributed to a data deluge. Properly use technology key to solve problems by natural disaster. Security and privacy issue in data transmission and storage need to investigation to ensure validity of disaster data.

Keywords: Big Data, Big Data Analytics, Natural Disaster Management, Big Data Tools, Big Data Value Chain

Section One
Introduction

The term disaster is the existences of disorder in regular circumstance which happening a level of suffer and volume of change which affects society in some area. By William Danner in 2011, disaster occurs when the hazard attacks the vulnerable community with low capacity; resulting in damages, loss and serious distraction of community and through outlook of international disaster database, natural disasters kill around 60,000 people per year processes. Such disasters comprise; famine, drought, earthquakes, flooding, wildfires, mudslides, hurricane, epidemics and pandemics like COVID-19 can push even the most strong people into a crisis (Chaudhary & Piracha, 2021). The 'disaster management' encompasses both natural and man-made, highly infused with various categories of sensors. Consequently, big data applications in field of disaster management should adopt a segmental view, going from a component to nation scale (Arslan et al., 2018). Big data have been widely used to process data, improve disaster prediction and decision-making processes through utilizing community information linking victims with emergency responders (Iglesias & Favenza, 2020). Briefly, disasters tend to be boring and disaster prediction usually contains several activities that actions to be taken before, during & after disaster in order to reduce its impact. According to (Yu et al., 2018) and (Iglesias & Favenza, 2020), disaster management encompasses through four separate phases; 'Mitigation', 'Preparedness', 'Responses' and 'Recovery'. But other experts include these attributes 'Prevention' and 'Rehabilitation' phases. For instance, Training Package held in 'Pan-African Emergency Training Centre' (PAETC) held in 2002 included these two phases. The goal of the 'Mitigation' phase is to minimize the effects of a disaster (building warning codes and risk zones, risk analysis, public instruction). Emphasis of 'Preparedness' phase is on planning how to respond to a disaster; includes preparedness plans, emergency exercises and training, also, Early Warning System (EWS) advance & implementation.

The 'Response' phase pertains to providing the required disaster management services to save lives and safeguard property and protect the environment during disaster management situations. Finally, the 'Recovery' phase is the process of returning systems to normal levels after a disaster happened. In disaster management situations quickly generated big volume of data extreme

weather simulation TB/day (Tera bytes) needs to be analyzed in a real-time fashion for immediate action and archives of historical data needs to be shared through researchers online and services of evaluation and validation of analysis methods that need to be standardized and made publicly available(Yu et al., 2018). The main purpose of disaster management is the incorporation of the interrelated processes that can provide efficient means to analyze, monitor or predict disasters oppo (Shah et al., 2019). The probable & utility of big data paradigm is growing for disaster management as the number and access to different datasets is expanding rapidly (Khalid et al., 2017). The diverse companies have been storing and using large amounts of information or data with technical progress that is visible in field of information technology requirements and emphasis on collecting, storing, analyzing and processing information has increased significantly. The big data is an opportunity to gain new insights into emerging types of data and content, in order to create more interesting businesses and answer questions which were before considered unanswerable new generation of technologies, architectures & variety of data sources, quality of data to be integrated and data visualization are some of challenges for big data integration (Alexandru et al., 2016).

Nowadays, big data has improved human life through universities, research organizations and colleges increased their research and development centers. In this digital and computing world, information is generated and collected with the fastest rate, which can rapidly exceed the edge range. When the size is outside the capability of representative database apparatuses to capture, control, store, manipulate, organized, manage and analyses. The big data technology enhanced to develop Hadoop framework with various eco-system. For the scientific and economic growth of developing countries using big data on disaster management how essential is? that policymakers understand the importance of investment in big data (Serra, 2016), and choose one of the 17 UN Sustainable Development Goals (UNSDG) and explain why big data will encourage developing countries to achieve their goal by 2030. These 17 UN Sustainable Development Goals (UNSDG) are by 2030, no-poverty, zero-hunger, good-health, education, gender-equality, clean-water, clean-energy, economic-growth, industry and infrastructure, no-inequality, sustainability, responsible-consumption, climate-action, life-underwater, peace and justice, and partnership. Unquestionably, in age of big data has opened new options for natural disaster management, principally because of the varied possibilities it providing to visualizing, analyzing, and predicting natural disasters(Yu et al., 2018). Based on this perspective, big data has radically

changed ways through which human societies adopt disaster management strategies to reduce human suffering and economic losses.

The Seminar Leading Questions

The seminar needs to answer following key questions to attain the application of big data for natural disaster Management in Africa: -

- What is disaster means?
- What disaster Management means?
- What are opportunities and challenges of big data analytics in natural disaster management and why it is evolving significant?
- How do we address these challenges and take opportunities offered by big data within the reach of African countries?
- Who are the main stakeholders that can help in obtaining the benefits of big data for disaster management? And
- What tools need to analyzing the big data in natural disaster management?

General Objective of the seminar

The main objective of this seminar is to identify the application of big data analytics in natural disaster management, in of African countries.

The Specific Objectives of the Seminar

Besides the above general objective, the seminar has the following specific objectives: -

- To analysis applications of big data on natural disaster management,
- To create knowledge on disaster management in African countries,
- To identify big data analysis tools,
- To enhance the state-of-the-art aid and insights for precise and timely decision-making in African countries.
- To close the gap between traditional data warehouses and big data technology.

- To display and notice natural disaster, alleviate their effects, assist in relief efforts and contribute to recovery and reconstruction processes.

Section Two
Methodology

In order to accomplish this seminar used various mechanisms, such as a comprehensive literature review was carried out using virous key words on the search engines, such as ‘big data’, ‘evolution of big data’, ‘big data analytics’, ‘disaster management’, ‘big data application on disaster management’, ‘big data analytics tools’, ‘challenges of big data’, ‘benefits of big data on disaster prediction’ and used some common and popular academic research sources, published papers, high-impact-factor journals, articles and conferences journals. These survey sources comprise from Google Scholar, Scopus, IEEE Xplore, Elsevier, ScienceDirect and ResearchGate. The document analysis methods followed to select relevant literature related to big data and how to manage the natural disaster, consuming access to reliable information during emergencies is essential for effective emergency management.

Evolution of Big Data

What is big data? As good-looking, it is hard to define and its platforms are undertaking a rapid and continuous transformation the new technologies being developed called big data. Currently, with uprising of fourth web (Web 4.0) technology utilization, such as big data has involved rapidly and researchers started more thought. The evolution of social media, satellite remote sensing, sensor networks and connected devices has contributed to the data deluge beyond what can be captured, processed, stored and understood traditional tools, known as a big data problem(Iglesias & Favenza, 2020). Big data is the perception of enormous amounts of data being generated daily in diverse fields due to the increased use of technology and internet sources, and despite several advancements and the expectations of better understanding, big data management and analysis remain a challenge, calling for more rigorous and detailed research, as well as the identifications of methods and ways in which big data could be tackled and put to good use (Munawar et al., 2020). The massive amount of data is accessible through internet and

as part of daily operations, and most companies are previously trying to process it in order to extract specific information that will increase their value. However, the amount of data that can potentially be processed is huge, leading to information overload for company managers, decision makers and administrators, the reason why this big data become essential(Guhr et al., 2016).Therefore, the fast rising of data in every organization enhance to the formation of big data as research area. Big data and artificial intelligence are two words that are widely used when discussing the future of business and ability to adventure granularity of data brings can potentially enable insights into a variety of predictable behaviors and incidents (Arishee & Al-Ghamdi, 2020). As a result, researchers well-defined big data differently but they do agree on what it involves. as tried to mentioned above, nowadays the term big data is the special claim of data science, in which the data sets are enormous and require overcoming logistical challenges to deal with them. The main concern is efficiently capturing, storing, extracting, processing & analyzing information from these enormous datasets (Chaudhary & Piracha, 2021). The big data mentions the data requires more dispensation resources to produce worth than are available within the restrictions of the current business problem (Fast & Ph, 2016). The big data technologies are occasionally considered as destructive technologies as they revolutionized the traditional ways of doing things in this data intensive era(Phyu & Shun, 2018), and concepts from distributed and parallel system are reapplied as the foundation of big data such as MapReduce paradigms for handling the big Vs characteristics such as, volume, velocity, variety, value and value. Due to advancing technological infrastructure and arrival of 'Internet of Things' companies continue to innovate, finding new ways to capture & leverage ever-expanding amounts of data called big data (Ivanova et al., 2016). The two major contributing issues to the emergence of big data era, rapid advances in computing technologies and resulting explosion of data; the former including hardware technologies such as CPU speeds and network bandwidths, as well as software technologies which beginning of distributed parallel processing frameworks (MapReduce &Hadoop); plus increasing popularity of web-based software (e.g. search engines, social media networks & e-commerce systems or widespread usages of various sensors (Song & Zhu, 2017). According to (Bhadani & Jothimani, 2016), over the past two decades, incredible growth of data and digitization of most of the processes, emergence of different social network platforms, blogs, placement of different kind of sensors, adoption of hand-held digital devices, wearable devices and explosion in usage of internet, huge amount of data are being generated on

continuous basis. Due to this the storage and retrieval of massive amount of structured as well as unstructured data at a wanted time lag is a challenge. Some of these limitations to handle and process vast amount of data with the traditional storage techniques led to the emergence of the term big data in this era. The term big data is the concept of massive quantities of data being generated daily in different fields due to increased use of technology & internet sources(Munawar et al., 2020). Big data (BD) are a data tendency present around us mainly through internet, social networks, smart devices and being aware of them terms gained popularity in the last few years due to data tendencies developing in its direction (Ptiček & Vrdoljak, 2021). At present, big data is infancy stage and inspire for the creation of data scientist in the world. This skill is the creation of combined through new generation of modern expertise and architecture that can bind extremely huge volumes of different data with real time processing and analysis. Sources of big data can be numerous eco-systems; educational settings, governmental agencies, banking society, tele-communications, transportation agencies, web-technologies and social media, shopping centers, manufacturing organizations, national or international revenues bureaus, meteorological agencies, military agencies, agricultural and farming organizations, airport, health-related organizations, research and developments agencies. Therefore, due to in ability of traditional SQL database and data warehouse tools handling huge and variety of data with generation of multitudinous data by fields of science, business, art, technology and social science is a global problem. Based on these problem big data knowledge realized to spread the current data warehouse with Hadoop eco-system, that deliver as platform for data analysis tools and change of 7vs such, (1) data variety, (2) data value, (3) data volume, (4) data veracity, (5) data velocity (accumulation time), (6) data visualization, and (7) data variability discussed the shortcomings advantages of varied mechanisms and challenges that emphasized need to develop well-organized big data processing techniques in modern using cloud computing science. On other hand, internet of things, block chain along with big data analytics, is measured one of the growing technologies in this world and gives rise to numerous medical applications; including medical staff and resource allocation, remote health monitoring, diagnosis and prediction of diseases at early stages, emergency care services, elderly care and others (Ahmed et al., 2021).

Over the past two decades huge data were generated by different organizations and number of scholars and researchers mentioned the reason for upcoming of big data from different

disciplines. The quantity of data produced in a globe is increasing rapidly, from 33 zettabytes in 2018 to predictable 175 zettabytes in 2025 (Reinsel et al., 2018), due to the declining costs of sensors and data storage, swift progress in advanced analytics and computing capabilities and greater connectivity with faster and cheaper data transmission. In 2010 Eric Schmidt, decision-making chairman of Google, expresses in a conference that as much data is now being created every two days, as was created from the beginning of human civilization to the year 2003 (Zennaro, 2016), and the big data is a broad term for data sets so large or complex that traditional data processing applications are inadequate. According to (Ivanova et al., 2016),the digitization of products, processes and business models through the corresponding explosion of big data has led to an evolution within business organizations. Accomplishment far beyond information technology's traditional role in business strategy, the implications of this big data phenomenon are considered through an exploration into what big data is, how it is currently being used by existing firms, and how it factors into strategic thinking. More than 2.5 quintillion bytes of data are generated every day, which expected that 1.7 MB of data be created by each person every second in 2020 & almost 95% of businesses are producing unstructured data and they spent $187 billion dollars in 2019 for big data management and analytics(Munawar et al., 2020). Big data analysis needs large dataset and formation of advanced analytic algorithms that require a great deal of processing power and resources. Besides the spatial Big Data Analysis (BDA) is very significant in disaster management scenarios to realize distribution patterns of the situation Smart Cities (Amovi, 2021). According to (Guhr et al., 2016), the big data is growing rapidly in its importance over the past few years', academics and practitioners have been considering the means through which they can incorporate the shifts these technologies bring into their competitive strategies. And the growing interest in big data requires a focused discussion on how it can be observed empirically and what theories can be used to understand the critical success factors as well as the business value. On other hand the storage and retrieval of massive volume of structured as well as unstructured data at a desirable time interval is a challenge and some of these limitations to handle and process massive amount of data with the traditional storage techniques led to the emergence of the term big data in this era (Bhadani & Jothimani, 2016).

What Big Data Analytics is?

Big data analytics chiefly involves logical approaches of systematic architecture and mining of big data and software for analysis, which enhance to extract meaningful information from huge dataset in order to decision making. Data investigation is the most important step in big data, for exploring meaningful values, giving suggestions and decisions (Rajaraman, 2016). On other article Rajaraman described big data analytics is procedure of inspecting, distinguishing and transforming big data through the goal of identifying valuable information, suggesting conclusion and helping to take precise decisions (Rajaraman, 2018). The storage and retrieval of vast quantity of structured as well as unstructured data at a desirable time lag is a challenge. Therefore, some of these limitations to handle and process vast amount of data with the traditional storage techniques led to the emergence of the term big data analytics (Bhadani & Jothimani, 2016). The big data analytics enhance to know about how these data can be stored, processed, and understood, such it can be used for predicting the future course of action with a great precision and acceptable time delay. The over-all, explanation of big data analytics is procedure of analyzing massive amount of data in order to extract expressive data from a given data set. According to (Mirjana Stankovic, 2021), big data can complement traditional statistics and inform to public policy-makers and regulators, if extra targeted enquiry is necessary, or prompt immediate response by analyzing it.

The Big Data Tools

At present-day, big data technology is actual complicated sub-sets of science that can be hard to implement through easy practices, and it needs the tool in order to implement in business. For this practice Hadoop and other interrelated environment used as open source and actuality the most prominent tools that planned as a data storage and batch processing engine. According to (Ajah & Nweke, 2019), definition the “Hadoop” is an open source project lead by Apache and originally designed to handle massive amounts of data, rapidly, efficiently and inexpensively improve to handles together structured, semi-structured and unstructured data. The Hadoop Distributed File System (HDFS) is the storage layer that is responsible for creating a distributed repository. In general, the following are the most valuable tools of the big data technology.

Hadoop extended definition: - The best reasonable approach for analyzing massive volumes of big data is to distribute and process it in parallel on multiple nodes. But, the size of data is

usually so large that thousands of computing machines are required to distribute and finish processing in a reasonable amount of time. When working with hundreds, thousands or millions of nodes, one has to handle issues like how to parallelize the computation, distribute the data and handle failures (Dash et al., 2019). Therefore, one of most prevalent open-source distributed application for big data analysis is Hadoop. This Hadoop technology enhance to implements MapReduce algorithm for processing and generating large datasets. MapReduce uses map and reduce primitives to map each logical record' in the input into a set of intermediate key/value pairs, and reduce operation combines all the values that shared the same key.

Apache Spark: - While, it is another open-source substitute to Hadoop and unified engine for distributed data processing that comprises higher-level libraries for supporting SQL(Spark SQL), streaming data (Spark Streaming), machine learning and graph processing (Dash et al., 2019).

Grid Computing: - Grid computing is a means of assigning the computing power in a distributed manner to solve problems that are typically vast and requires lots of computational time and power (Bhadani & Jothimani, 2016). In this concept every computer that is associated to the internet and wants to become a part of the grid is measured to be a node in very large computing machine.

Apache Hive: - an open-source platform that provides facilities for querying and managing large datasets residing in distributed storage & similar to SQL, called as HiveQL (Bhadani & Jothimani, 2016).

Apache Pig:- a platform allows analysts to analyzing large data sets and a high-level programming language, called as Pig Latin for creating MapReduce that requires Hadoop for data storage.

The Big Data Value Chain

The term value chain first presented by **Porter** in 1980, that remarks the set of activities accomplished by a firm to add value at each step of delivering a product or service to its customers. Based on this (Bhadani & Jothimani, 2016), declared big data value chain in a comparable way, that refers to the framework that deals with a set of activities to create value from available data. these activities counting seven phases: data generation, data collection, data transmission, data pre-processing, data storage, data analysis and decision making.

Data Generation: - The first and leading step in big data value chain is the generation of data. As deliberated in the preceding section, data is generated from various sources that include data from education, industry, government agencies, private or public organization, transportation, military, finance, social media and others.

Data Collection: - In this stage, the big data is obtained from all possible data sources.

Data Transmission: - After the data is collected, it can transfer to the data storage and processing infrastructure for further processing and analysis and apart from storage of data, data center helps in collecting, organizing and managing data.

Data Pre-processing: - Data that collected from several data sources may be redundant, noisy and inconsistent, hence, in this phase; the data is pre-processed to improve the data quality required for analysis.

Data Storage: - The big data storage systems should deliver reliable storage space and powerful access to the data, and the distributed storage systems for big data should consider factors like consistency, availability, efficiency and partition tolerance.

Data Analysis: after the data is collected, converted and deposited, the next procedure is data exploitation or data analysis, which is computed using the following steps. Such as, define metrics, select architecture based on analysis type, selection of appropriate algorithms and tools and data visualization.

Decision Making: - Based on the analysis and visualized results phase, the decision makers can decide whether and how to reward a positive behavior and change a negative one of big data.

Benefits of Big Data Technology

The big data is one of the most principal fields of knowledge and research that has generated high impact in the process of digital transformation of the organizations in recent centuries and main goal is to improve work processes through analysis and interpretation of large amounts of data, therefore, knowing how big data works, its benefits, challenges & tools, are essential elements for business success (Almeida, 2017). Big data have a positive impact in nearly every sphere of life, in military intelligence, space science, aviation, banking and health (Akinnagbe et al., 2018). Big data as the generation of data has received attention since scholars, organizations,

and individuals over the past few years, in the belief that data helps with making decisions that are supported by facts (Seseni & Mbohwa, 2021). The big data from social media sources become progressively available for disaster management, due to advances of various kinds of capable sensors and their significant growth (Dev et al., 2018), with the combination of powerful computer systems and networks sensors, smart phones, and cyber-physical provisions is creating massive data streams that can help decision makers during disasters, the data estimated to have been created in 2020 as the whole is 40 zettabytes (Marston, 2019). This means the data created per individual was 5200 gigabytes, enhances to individuals for business creation, increase business efficiencies, business innovation, increase individual skills, provide to countries growth, rapid decision making, increase industrial revolution, educational succusses & technology innovation. In general, the problems of developing countries on bib data analysis are related to lack of capacity, lack of appropriate infrastructure for big data initiatives &inadequate investment on skill evolution. Therefore, big data analytics has the following benefits of the globe: -

- Data buildup from several sources, counting internet, aviation, social media platforms, online shopping sites, company databases and outside third-party sources,
- Real-time prediction and intensive care of business as well as natural disaster,
- Recognize critical points hidden inside large datasets to influence disaster management,
- Prompt on mitigation of disaster risks by improving complex decisions for unexpected events
- Enhance on identifying issues in schemes and analytic processes in real-time,
- Big data analytics and apparatuses can dig into massive datasets to extract valuable insights facilitate innovation in a continent.

Section Three

Application of Big Data Analytics on Natural Disaster Management in Africa Countries

African countries are affected by several disaster in each year. The big data has fundamentally transformed the ways through which human societies adopt natural disaster management strategies to reduce suffering and economic losses(Yu et al., 2018). In Africa, there have been experiences of natural disasters (Adjei, 2019), as Adjei mentioned in August 2017, with three days of heavy rainfall led to flooding and mudslide in Sierra Leone in which over 1,000 people were died and many were unaccounted for, the 2011 drought that hit the East African countries of Kenya, Eritrea, Ethiopia, Somalia and Djibouti has been labeled as the worst drought to have been experienced in those countries in over 60 years. Volcanic eruption in Eritrea in 2011 and severe flooding suffered in Southern Africa in the same year are just a few examples of many natural disasters has recorded in its recent history (Adjei, 2019). Through the rapid growing population and increasing impact of climate change, the disaster influence in Africa will likely increase in coming decades (Nima Yaghmaei & Regina Below, 2019) and over 50 countries and around 30 million square kilometers of land threat of natural hazards in Africa. From these disaster types drought, extreme temperature, flood and storm are vital. Based on the report within 20 years (2000-2019*), in Africa completely 36,704 people are died from 10 countries. Countries are listed orderly, in Somalia 20,739, Algeria 3,777, Mozambique 2,291, Nigeria 1,696,

Madagascar 1,644, Ethiopia 1,639, Kenya 1,572, Sierra Leone 1,289, D.R. Congo 1,072 & Malawi 985. Big data has increased academic relevance over the last decade and interest to other role-players such as governments, businesses and general public. Based on Big Data Readiness Index (BDRI), applied in Africa contains the topic of pleasure within the digital wellbeing driver, need these two further investigations (1) what is the relation between pleasure and big data? and (2) how does Africa perform in digital wellbeing? (Jouberst et al., 2020). Along the international organization big data involved attention of many industrialized countries of the world on disaster management. According to (Sh et al., 2017) and (Dhawan et al., 2018), based on research output on big data since 2007-2016, the key initiative in the field has been put forward by US &later expanded to several countries, orderly Chian, India, U.K., Germany, South Korea, Australia, Japan, Italy, Canada, France &Spain have estimated big data as a strategic resource as oil, in addition, the great importance, it gives information to the problems area and many relevant documents have been adopted. While, developing countries have low levels of 'datafication. Therefore, making issue of data creation and digitization's are significant to the prevention of disaster. This contain digitization of existing files, knowledge, information and creating new digital data by digitalizing facilities in health care, education, social security (Mirjana Stankovic, 2021). Whatever, its advantage is high, in Africa the application of big data is in infancy age (Akinnagbe et al., 2018). South African organizations are struggle to find the right skills to employ for big data project, because of a new innovation, that require diverse combination of skill set associated to the traditional data warehouse and data analytics (Mneney & Belle, 2016).The influence of big data applications will convey its own complexities to natural disaster management in Africa. Therefore, these countries will encourage working on big data management that enhance to their economy growth through preventing natural as well as man - made disaster by use of internet of things and big data technology as a tool. The big data bring us the data-driven paradigm and enlighten us to challenge new classes of problems, that we were not able to solve in earlier (Song & Zhu, 2017). Disasters tend to be boring, disaster management usually defines cycles that comprise all activities and measures to be taken before, during and after the disaster enhance to reduce its impact (Iglesias & Favenza, 2020). Currently, effective disaster management is a global challenge, therefore, the potential and utility of big data paradigm is growing for disaster management as the number and access to different datasets is expanding rapidly (Khalid et al., 2017). And mentioned an overview of what kind of data is used

in existing systems for managing disasters? which the specific phases of disaster management system targeting? and what are enabling technologies have been used along with big data technology to supplement disaster management processes? .The description via (Giest & Samuels, 2020), the big data movement joint through more sophisticated methods in recent years has opened up new opportunities to use existing data in different ways as well as fill data knowledge gaps through techniques that make predictions based on lookalike data. According to the 'panel discussion of African business day concept note' held (Note, 2021), the most African countries have problems on big data. Since, they don't produce on time, frequently produced but lacks proper handling, generated poor quality data, data are not accurate, lacks proper management & handling, not shared among sectors, are not even given due consideration and difficult to make data driven decisions and not contribute for socio-economic development. Where, improvements in internet connectivity are consider an important driver for big data in Africa, since the connectivity in Africa is still not up to the global standards seen in emerging and more developed markers, this also acts as a limit to the full realization of big data's benefits. From the panel discussion the following are challenges that need to addressed to realize the development and adoption of big data in Africa: -

- **Lack of Government support and awareness**: - The government has little awareness on the use of big data operation and providing no or little support for some of the startup big data projects in Africa.
- **Partnership**: -There are no compulsory guidelines and rules among the private sectors, government & institutions to work in collaboration for sharing big data for common goals.
- **Data protection, privacy, confidentiality** and **cybersecurity**: - big data facilitates the tracking of people's movements, behaviors and preferences and personal information that may be exposed to privacy, confidentiality, and cybersecurity risks.
- **Access to big data**: - To access, experiment with, and use big data effectively, users need to enter into agreements with private data owners, while maintaining their independence, and ensure a legal environment that addresses both privacy and confidentiality.
- **Building big data capability**: - That big data needs to be clean, accurate and transparent that needs to be stored, analyzed and shared appropriately which all require the capabilities to do that in turn need to make the required investment.

- **Skilled manpower and technologies**: - The lack of human resources and expertise represents another major barrier to the adoption of big data across the continent.
- **Standards & big data**: - The interworking of multiple systems and technologies, legacy and new are key issues in relation to technology integration that demand for standards to facilitate interoperability among the components of the big data value chain.
- **Platform interoperability**: - As African private sector struggle to accelerate adoption of big data, the issue of platform interoperability should be considered in order to address concerns related to communications protocols, gateways and data integration platforms.
- **Enhancing access to open data**: - African private sector use of open data available from the public sector has been a source of innovation and business opportunity.
- **Infrastructure and services** – availability and access to high-speed and high-bandwidth connectivity which is a key challenge in Africa is of paramount importance for effective use of big data.

Application of Big Data on Natural Disaster Prediction

The disaster incidence is out of human regulator; but, over the disposition of several state-of-the-art smart technologies, we can predict, mitigate and even prevent loss of human lives, materials and infrastructure using big data and IOT by the use of sensor. According to Amir Elichai (2018), around 10,000 people died and more than 95 million affected, in 2017 sadly noticeable record-breaking year for natural disasters as a worldwide, ranging from Hurricane Harvey and massive earthquake in Mexico to Hurricane Irma and mudslides in Colombia. Therefore, the expansion of big data and evolution of internet of things technologies have played an important role in feasibility of smart city creativities. The big data offer the potential for cities to obtain valuable insights from a large amount of data collected through various sources and Internet of Things allows to control disaster through integration of sensors, radio frequency identification and Bluetooth in real-world environment using highly networked services (Abaker et al., 2016). In order to minimize the possibilities of casualties and environmental destruction, disaster prediction measures need to be both preventive and reactive. The key roles of disaster prediction are enhance to early warnings, collect information in real-time, quickly figure out the routes and well manage emergency resource (Shah et al., 2019). In order to predict the natural disaster occurrence accurately, the following mechanism is used as global: -

Synthetic-Aperture Radar (SAR), an energetic microwave remote sensing technology that measures the phase difference between a radar wave emitted from an antenna attached to a satellite or aircraft to generate high-resolution images of a surface in disaster setting.

- **Aerial Imagery Sensor** (AIS), can be captured with a high spatial resolution and processed much faster in comparison to satellite imagery. and
- **Unmanned Aerial Vehicles** (UAVs), which is during a disaster event, could bring saviors the bird view of the disaster environment, which is imperative for inclusive disaster management system that involves data collection, victim localization and lifesaving optimization. Let's see the following Universal Network Architecture for UAV-based Disaster Management.

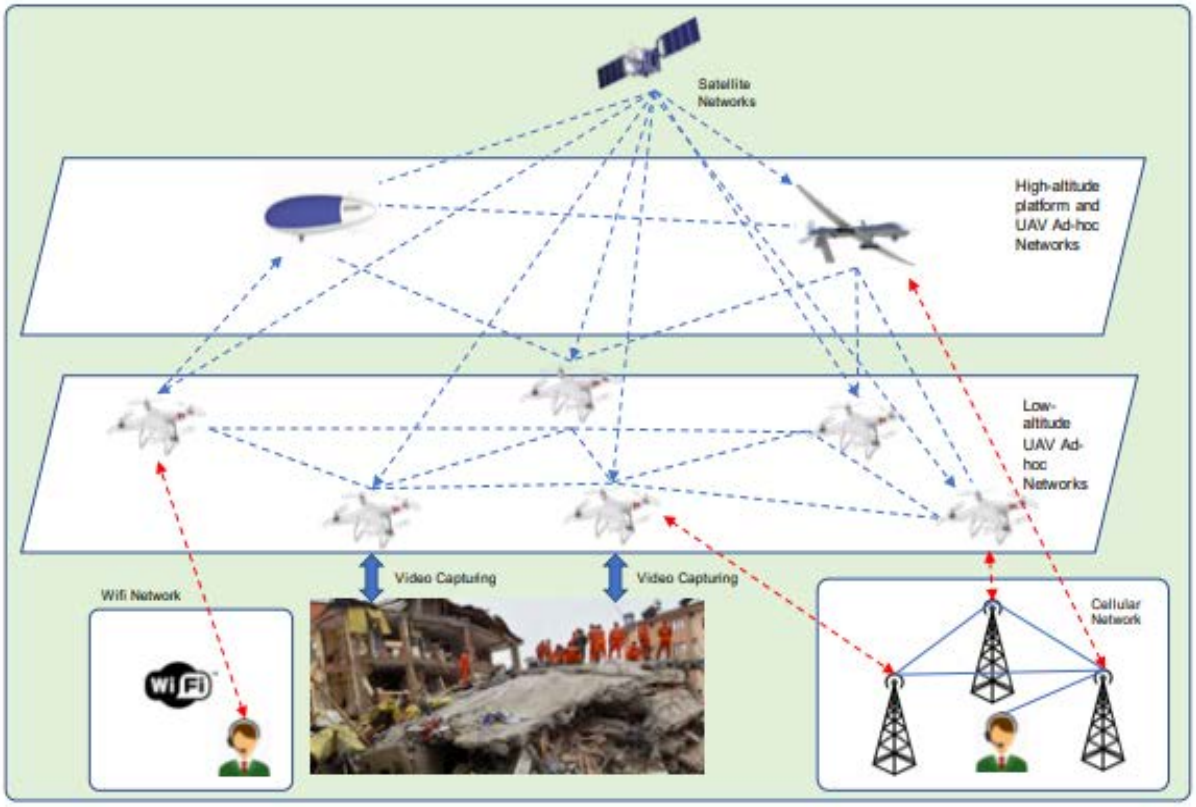

Figure.1. Universal Network Architecture for Unmanned Aerial Vehicles for Disaster Management adopted from (Margolis & Walsh, 2019).

From the perspective of the above architecture (figure 1), at the top layer of the UAVs is connect to the Global Positioning System (GPS) satellite network by equipping a GPS receiver on-board,

which occasionally delivers the geolocation and time information. The information is critically important for UAVs to accurately and safely accomplish the disaster response tasks.

Challenges of Big Data on Natural disaster management

Nowadays the massive amounts of data are produced from various environments (government agencies, city administrators, community health organizations, educational institutions, shopping, social media, government or non-governments agencies, agricultural and farming, transportation, mining departments, construction agencies, manufacturing industries and fire management departments). Thus, knowing and managing the data from these various sources is the challenging situation for individuals in each department. Because, currently we are in the world of technology and huge volume of data flowing and collected from many sources such as from our personal computers, laptop, mobile phone, TV channels, social media site, web site and many other devices and sources. These data are categorized as structured, semi-structured and unstructured. On other hand on perspective of (Seseni & Mbohwa, 2021), the primary impairments to the investigation of dynamic phenomena has been challenges associated with collecting data at a sufficient frequency and duration to accurately model such changes. Such challenges including data quality and reliability, data theft, storage need, requires sophisticated tool &professional skills, difficulty to analysis, difficulty to collect and complex tool to use.

Analysis

The up and coming of big data enhance to increase economic strength and resolve several social problems counting natural disaster management and prediction. The big data bring us the data-driven paradigm and enlighten us to challenge new classes of problems, that we were not able to solve in earlier (Song & Zhu, 2017). Whatever, big data have various benefits the African countries still did not use properly. Due to lack of infrastructure, like Synthetic-Aperture Radar, Aerial Imagery Sensor, Unmanned Aerial Vehicles, lack of skill on big data tools and lack of awareness on big data analytics within 20 years (2000-2019*), by cases of natural disaster in Africa 36,704 people died from ten countries, by drought, extreme temperature, flood and storm are worse. Also, 16.72% world's population lived in Africa, but in terms of academic research output generates in Africa less 1% of world research. But, only South African organizations are struggle to find the right skills to employ for big data project, since of a new innovation, that require diverse combination of skill set associated to the traditional data warehouse & data

analytics (Mneney & Belle, 2016). From analysis, national and international technologies, strategies and infrastructure for future research are beginning to have positive effects on African research output, that makes formation of big data, but issues such as discrimination, academic access, publisher relationship and financial investment are affecting this progress. For the coming era what expect from African countries are now the globe is evolving with expansion of up-to-date technologies in diverse situations, such industrialized countries use big data as energetic oil for problematic cracking, comparable disaster prediction and management. Therefore, the big data is evolving as a dramatic shift through interaction of human life and enhances to solve global challenges and deliver considerable benefits. Besides, the novel sources of data, internet of thingies, artificial intelligence skills and new analytical methods, applying successfully and sensibly on big data can allow more agile, well-organized and evidence-based decision-making in African continent.

Section Four
Conclusion

Disaster is the disorder in consistent situation in which happening in the level of hurt which surpasses the volume of change that disturb society in some area. Such disasters contain; famine, drought, earthquakes, flooding, tsunamis, wildfires, mudslides, hurricanes, epidemics &pandemics plus human-made disasters. Nowadays effective disaster management is a global challenge. The probable and usefulness of big data paradigm is growing for disaster administration as the amount and access to different datasets is expanding rapidly in Africa. Big data is the comprehensive term for data sets which affords to solve problem of large or traditional data processing applications that are inadequate. This big data is new science and massive collections of structured, semi-structured and unstructured data that come from different sources. The methodology used in this seminar is comprehensive literature review with different keywords and various online databases. Whatever, the 16.72% world's population lived in Africa, but in terms of academic research output generates in Africa less than 1% of the world research. Over-all, big data has the competence to improve disaster recovery by operating through public information and connecting victims with emergency and emergency personnel can minimize their search time and maximize their recovery time when they have access to real-time information highlighting the areas most affected. In order to address problems of natural

disaster the big data technology uses Hadoop, Spark and other frameworks for distributing and retrieving data for the support of disaster prediction. Hadoop is focused on disk-based data and basic map-reduce scheme, while, Spark advances in several directions. These together can accommodate multiple types of databases and accomplish their performance using parallel workers. Hadoop framework provides to advance programming platform, distributed filesystem, parallel execution and software ecosystem. In general, Africans need to change their attitude towards disaster management. since, disaster is a global challenge. The probable and usefulness of big data paradigm is growing for disaster management as the amount and access to diverse datasets is expanding rapidly. Knowing the big data skill gives an overview of what kind of disaster happened in existing systems and enhance to identify and predict the precise phases of disaster management that targeting to and what are the enabling technologies that have been used along with big data technology to supplement disaster management. Main issue on application of big data on disaster management need more emphasis that identify causes and ensure data consistency, accuracy and completeness for decision making.

Recommendations

The present and evolving emphasis of big data analytics is to discover and update the outdated techniques of data analysis such as rule-based system, pattern mining, decision trees & data mining systems should develop among each country of Africa. Therefore, to attain global sustainable development program African nation will be participate on following list of big data achievement program to the future.

- The African countries, rarely have educated human resources with necessary big data analytics skills, so, need to effectively deploy on building adequate big data analytics skills and capacity.
- Need to create big data governance: Because, there is lack of adequate data governance standards and skill on how data are captured and stored for liability.
- Should participate on reduction of digital dissimilarities from developed world.
- Need to grow plans for data sharing agendas respect to privacy and intellectual property rights.

- Need to develop the sensible big data infrastructure: through contributing on investment into software, hardware and broadband connectivity for widespread data access and use.
- Guarantee to the adequate levels of privacy, security and management of data: example adjust the program on use of data without agreement, reduce the risk of identification of individuals through data selection bias, the resulting discrimination of big data models.
- Big data tools should be robust and reliable: the data produce must be analysis by reliable algorithms at least as robust and reliable as the traditional data warehouse processes.
- Big data technology should be costumer centered: As big data is used to increase human abilities, the protection of consumer interests, counting well-being and security, should be the primary considerations in the design, development and deployment of big data.
- The African union ought to enhance to develop ethical guidelines for trustworthy on big data by high-level expert group on big data analytics.
- Generate an enabling setting in (public as well as private sector) for effective decision making over big data. Which including an appropriate strategies and regulatory measures, establishing data protection frameworks and sectoral regulatory frameworks.

The Line of Future Work

In Africa, there are various problems in public as well as private sectors grounded on data generating, data storing, data quality and data knowledge among specific areas. For the coming period this category of challenges should be address in way big data can influence policies and need specific attention on producing quality data, improve big data analytic skill, need to have capacity on data analytics and to avoid digital divide. Therefore, Africans should inspire on this large dataset of structured, semi-structured and unstructured data professional activities, in order to reduce the above-mentioned problem of natural disaster occurrence in each year using big data analytics applications and other components, for identification and finding a solution. For the future what needs from African countries are; currently the massive volumes of digital data being produced by a vast and growing number of platforms and plans to offer extraordinary opportunities to advance a better understanding of multipart connections, and obtain real-time feedback on how well strategy responses are working. Finally, security as well as privacy issues in data transmission, data analysis and

storage also need to be under constant investigation to ensure the validity of disaster management and prediction.

The publication Tracking ID: 57566

Funding: there is no funding agencies for this research, it was self-funded by authors.

References

Abaker, I., Hashem, T., Chang, V., & Anuar, N. B. (2016). *The Role of Big Data in Smart City The Role of Big Data in Smart City*. *October*. https://doi.org/10.1016/j.ijinfomgt.2016.05.002

Adjei, E. a. (2019). *Munich Personal RePEc Archive Natural disasters and economic growth in Africa*. *95588*.

Ahmed, I., Ahmad, M., Jeon, G., & Piccialli, F. (2021). A Framework for Pandemic Prediction Using Big Data Analytics. *Big Data Research*, *25*, 100190. https://doi.org/10.1016/j.bdr.2021.100190

Ajah, I. A., & Nweke, H. F. (2019). *Big Data and Business Analytics : Trends , Platforms , Success Factors and Applications*.

Akinnagbe, A., Peiris, K. D. A., & Akinloye, O. (2018). *Prospects of Big Data Analytics in Africa Healthcare System*. *10*(6), 114–122. https://doi.org/10.5539/gjhs.v10n6p114

Alexandru, A., Alexandru, C. A., Coardos, D., Tudora, E., & Definition, A. B. D. (2016). *Big Data : Concepts , Technologies and Applications in the Public Sector*. *June 2017*.

Almeida, F. L. F. (2017). *Benefits , Challenges and Tools of Big Data Management Benefits , Challenges and Tools of Big Data Management*. *October*. https://doi.org/10.20470/jsi.v8i4.311

Amovi, M. (2021). *applied sciences Big Data in Smart City : Management Challenges*. 1–27.

Arishee, J. A., & Al-Ghamdi, F. A. (2020). The Use of Big Data and Artificial Intelligence in Confronting the COVID-19 Epidemic. *Arab Journal for Security Studies*, *36*(2), 249–264. https://doi.org/10.26735/gycx5740

Arslan, M., Roxin, A., Cruz, C., Ginhac, D., Arslan, M., Roxin, A., Cruz, C., Ginhac, D., Review, A., & Arslan, M. (2018). *A Review on Applications of Big Data for Disaster Management To cite this version : HAL Id : hal-01678238 A Review on Applications of Big Data for Disaster Management*.

Bhadani, A. K., & Jothimani, D. (2016). Big data: Challenges, opportunities, and realities. *Effective Big Data Management and Opportunities for Implementation*, 1–24. https://doi.org/10.4018/978-1-5225-0182-4.ch001

Blake, A. (2019). *Dynamics of data science skills How can all sectors benefit*. https://www.burning-glass.com/wp-content/uploads/dynamics-of-data-science-skills-report.pdf

Chaudhary, M. T., & Piracha, A. (2021). *Natural Disasters — Origins , Impacts , Management*. 1101–1131.

Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. *Journal of Big Data*, *6*(1). https://doi.org/10.1186/s40537-019-0217-0

Dev, K. A., Pradeepkumar, A. P., & Mohan, M. (2018). *Integrating Disaster Science and Management*. *April*. https://doi.org/10.1016/B978-0-12-812056-9.00016-6

Dhawan, S. M., Gupta, B. M., & Gupta, R. (2018). Mapping global research output in big data during 2007-16. *Emerging Science Journal*, *2*(3), 117–129. https://doi.org/10.28991/esj-2018-01135

Engineering, F. (2017). *D w b d*. *9*(2), 1–17. https://doi.org/10.5121/ijcsit.2017.9201

Fast, A., & Ph, D. (2016). *The Evolution of " Big Data " The Basics • This is a computer . advanced algorithms for processing data can produce*.

Giest, S., & Samuels, A. (2020). 'For good measure': data gaps in a big data world. *Policy Sciences*, *53*(3), 559–569. https://doi.org/10.1007/s11077-020-09384-1

Guhr, N., Loi, T., Wiegard, R., Breitner, M. H., & Guhr, N. ; (2016). Big Data and Strategy: A research Framework. *ACIS 2010 Proceedings*, *50*, 76–87. http://aisel.aisnet.org/mcis2016http://aisel.aisnet.org/mcis2016/50%0Ahttp://aisel.aisnet.org/ecis2016_rphttp://aisel.aisnet.org/ecis2016_rp/99%0Ahttps://aisel.aisnet.org/ecis2016_rp/99%0Ahttp://www.ecis2016.eu/papers/2024.pdf%0Ahttp://aisel.aisnet.org/ec

Iglesias, C. A., & Favenza, A. (2020). *A Big Data Reference Architecture for Emergency*

Management. 1–24. https://doi.org/10.3390/info11120569

Ivanova, N., Gugleva, V., Dobreva, M., Pehlivanov, I., Stefanov, S., & Andonova, V. (2016). We are IntechOpen , the world ' s leading publisher of Open Access books Built by scientists , for scientists TOP 1 %. *Intech*, *i*(tourism), 13.

Joubert, A., Murawski, M., Julian, B., & Bick, M. (2020). *Happiness and Big Data – Theoretical Foundation and Empirical Insights for Africa*. 443–455. https://doi.org/10.1007/978-3-030-44999-5

Khalid, M. A., Roxin, A., Cruz, C., & Ginhac, D. (2017). *A REVIEW ON APPLICATIONS OF BIG DATA FOR DISASTER MANAGEMENT*. *December*. https://doi.org/10.1109/SITIS.2017.67

Margolis, J. D., & Walsh, J. P. (2019). Section 1: Introduction. *People and Profits?*, *August*, 13–13. https://doi.org/10.4324/9781410600622-9

Marston, C. K. S. (2019). *No Title*. *16*. https://doi.org/10.4301/S1807-1775201916006

Mirjana Stankovic, et al. (2021). *Emerging technology trends: Artificial intelligence and big data for development 4.0*.

Mneney, J., & Belle, J. Van. (2016). *Big Data Capabilities and Readiness of South African Retail Organisations*. *January*. https://doi.org/10.1109/CONFLUENCE.2016.7508129

Munawar, H. S., Qayyum, S., & Ullah, F. (2020). *Big Data and Its Applications in Smart Real Estate and the Disaster Management Life Cycle : A Systematic Analysis*.

Nima Yaghmaei & Regina Below. (2019). Disasters in Africa: 20 Year Review 2000-2019. *Cred & Undrr*, *56*, 1–2.

Note, C. (2021). *Panel Discussion – African Business Day On Fostering the African Private Sector in Harnessing Big Data in Health and Energy Sectors*. 1–7.

Phyu, K. P., & Shun, W. Z. (2018). *Data lake : a new ideology in big data era*. *03025*, 1–11.

Ptiček, M., & Vrdoljak, B. (2021). *Big Data and New Data Warehousing Approaches*. *February*. https://doi.org/10.1145/3141128.3141139

Rajaraman, V. (2016). Big data analytics. *Resonance*, *21*(8), 695–716. https://doi.org/10.1007/s12045-016-0376-7

Rajaraman, V. (2018). Big data analytics. *Resonance*, *21*(8), 695–716. https://doi.org/10.1007/s12045-016-0376-7

Reinsel, D., Gantz, J., & Rydning, J. (2018). The Digitization of the World - From Edge to Core. *Framingham: International Data Corporation*, *November*, US44413318. https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf

Rupnawar, M. A. (2016). *Comparative Study On Data Warehouse And Big Data*. *January*, 258–261.

Santoso, L. W. (2018). ScienceDirect Data Warehouse with Big Data Technology for Higher

Education. *Procedia Computer Science*, *124*, 93–99. https://doi.org/10.1016/j.procs.2017.12.134

Serra, C. (2016). *Big data : a road to development*. *28*(2), 12–14.

Seseni, L., & Mbohwa, C. (2021). The significance of big data in the success of SMES in emerging markets: A case of South Africa. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, *Smit 2020*, 1986–1996.

Sh, M., Technology, I., & Academy, A. N. (2017). *Big data initiatives of developed countries*. *January*. https://doi.org/10.25045/jpis.v08.i1.02

Shah, S. A., Seker, D. Z., Rathore, M. M., Hameed, S., Yahia, S. B. E. N., & Draheim, D. (2019). Towards Disaster Resilient Smart Cities : Can Internet of Things and Big Data Analytics Be the Game Changers ? *IEEE Access*, *7*, 91885–91903. https://doi.org/10.1109/ACCESS.2019.2928233

Song, I., & Zhu, Y. (2017). *A rticle*. *October 2015*. https://doi.org/10.1111/exsy.12130

Venkata, L., & Enka. (2015). *A Survey on Challenges and Advantages in Big Data*. *8491*, 115–119.

Yu, M., Yang, C., & Li, Y. (2018). *Big Data in Natural Disaster Management : A Review*. https://doi.org/10.3390/geosciences8050165

Zennaro, M. (2016). *Intro to Big Data & Analytics*. *December*, 34. https://www.itu.int/en/ITU-D/Regional-Presence/AsiaPacific/SiteAssets/Pages/Events/2016/Dec-2016-IoT/IoTtraining/BigData -Zennaro.pdf