Global Scientific JOURNALS

GSJ: Volume 8, Issue 1, January 2020, Online: ISSN 2320-9186

www.globalscientificjournal.com

BIG DATA FOR SECURE PERVASIVE HEALTHCARE SYSTEM
CASE STUDY OF SOCIAL SECTOR OF HEALTH IN RWANDA

NIYONSENGA Aaron[1], TUYISHIME Jean Claude [2], DUSHIMUKIZA Ernest [3], HARUBWIRA Flaubert[4]

Faculty of Computer Science

University of Gitwe

Rwanda/South/ Ruhango

1.1 Background

Pervasive healthcare aims to deliver deinstitutionalized healthcare services to patients anytime and anywhere. Pervasive healthcare involves remote data collection through mobile devices and sensor network which the data is usually in large volume, varied formats and high frequency. The nature of big data such as volume, variety, velocity and veracity, together with its analytical capabilities complements the delivery of pervasive healthcare. Pervasive healthcare, also known as ubiquitous healthcare or mobile healthcare is a concept to deinstitutionalize healthcare services (Ruotsalainen et al., 2012).

Pervasive healthcare improves communication between patients and healthcare professionals by delivering accurate medical information anytime and anywhere. This enables the real time clinical information recording and avoids information duplications, hence leading to a better information sharing and decision making (Drayton, 2012).

Given the nature of pervasive healthcare, data from distributed sources are demanded for the healthcare service provision (Heerden et al., 2012). The wide range of mobile devices, sensor and applications produce data in various formats. Big data is therefore enabling the pervasive healthcare by providing an innovative solution in managing the large, varied and high frequency data sets. Big data is commonly defined with 4Vs: Volume, Velocity, Variety and Veracity (Feldman et al., 2012).

Volume refers to the rapid rate at which data is growing. Velocity represents the increasing frequency with which data is delivered or accumulated. Variety signifies the many forms in which data exist. There are two types of data defined in the big data context: structured data and unstructured data (Hurwitz et al., 2013).

Big data management provides the capabilities in data storage, management, analysis (or known as big data analytics, which is a pivotal element in the big data management) and visualization for handling large (from terabytes to Exabyte) and

complex (from sensor to social media) data sets (Feldman et al., 2012).

However, most of the existing research focuses on the technical context of big data (e.g. developing decision algorithm and decision models) in the healthcare sector. There is limited research in intertwining pervasive healthcare with big data. These two domains are seen as complementing each other to better the healthcare services delivery and to reduce the accelerating operational cost. This in turn causes data related challenges. In addition, most healthcare organisations lose sight of the strategic role of big data which has a direct impact on the quality of healthcare services at the organisational level without considering the big data implementation in a holistic sense. The aim of this study is therefore to propose big data for pervasive healthcare conceptually through an intensive literature review of pervasive healthcare and big data.

1.2 Research Problem

For the health care sector these issues are even amplified, due to fact that health care data are considered privacy sensitive data. However, traditional security and privacy methods to protect privacy health care data seem insufficient or even obsolete.

This is a problem for patients as personal information can unwillingly be derived from these health information systems and end up in wrong hands. Besides that, individuals have some rights against intrusion of their personal information, in wrong hands; personal information can potentially harm individuals.

On the other hand, weak security and privacy methods can hinder the adoption of big data in the pervasive health care. There can be public resistance from individuals or government against the use of big data in pervasive health care, when there is no trust in the protection of their personal information. Hindering in the adoption of big data in the pervasive health care could also hinder potential benefits big data could bring to the pervasive health care, which are for example improved quality of health.

Therefore, the owners of the problem are the hospitals and other organizations in the healthcare that potentially can benefit from an adoption of big data in pervasive health care. These organizations have to deal with hurdles such as privacy legalization and the public perception of privacy before they can successfully adopt big data.

In addition; Tracking, controlling and monitoring data using traditional system, it is difficult to organize traditional system is easily lost or destroyed. All businesses services have to track data in some way, so that system has many weaknesses.

1.3 General Objective

The aim of this study is to examine the role of big data in delivering pervasive healthcare a case study of social sector of health in Rwanda.

1.4 Specific Objectives

The specific objectives of this research are the following:

i. Evaluate pervasive healthcare in social sector of Rwanda.
ii. Evaluate the application of big data to secure health care system.
iii. Examine the effectiveness of the mechanism of big data provision in supporting pervasive healthcare.

1.5 Significance of the Study

This research will contribute in implementation of smart energy saving algorithm to handle big data resources. The smart energy saving algorithm should dynamically decide (context-aware) on closing/pausing background or unessential applications or services, in accordance with continuous metered resources (resource-aware), and scale to the big data resource needs. Also, data reduction/sampling (data-aware) techniques will be incorporated to process and diagnose pervasive health data, for further reduction in resource consumption. The key decision on where to accomplish the analytics which includes the preprocessing, processing and visualization would be our second contribution.

Finally the study is of significance to the academic community by adding to the existing knowledge in the area of big data analytics. Clinicians will refer to this study as a guide to assist in effective big data analytics and also as a benchmark to address the needs of the pervasive healthcare system.

1.6 Justification of the Study

There's a huge need for big data in healthcare as well, due to rising costs in nations like the United States. As Chawla and Davis (2013) state, "After more than 20 years of steady increases, healthcare expenses now represent 17.6 percent of GDP nearly $600 billion more than the expected benchmark for a nation of the United States' size and wealth". In other words, costs are much higher than they should be, and they have been rising for the past 20 years. Clearly, we are in need of some smart, data-driven thinking in this area. And current incentives are changing as well: many insurance companies are switching from fee for service plans (which reward using expensive and sometimes unnecessary treatments and treating large amounts of patients quickly) to plans that prioritize patient outcomes. In the previous scheme, healthcare providers had no direct incentive to share patient information with one another, which had made it harder to utilize the power of analytics. Now that more of them are getting paid based on patient outcomes, they have a financial incentive to share data that can be used to improve the lives of patients while cutting costs for insurance companies.

Finally, physician decisions are becoming more and more evidence-based, meaning that they rely on large swathes of research and clinical data as opposed to solely their schooling and professional opinion. As in many other industries, data gathering and management is getting bigger, and

professionals need help in the matter. This new treatment attitude means there is a greater demand for big data analytics in healthcare facilities than ever before. First of all, it is pleasure to implement the goal of EDPRS3 by using innovation of ICT like this research. The implementation of this research can be effective if they are executed consistently and fairly, this is particularly true for our country that is going in smart for everything typically in ICT.

1.7 Theoretical / Conceptual Framework

1.7.1 Theoretical Framework

The two main themes are pervasive healthcare and big data. Pervasive healthcare is executed by healthcare delivery services such as telemedicine, patient monitoring, location-based medical services, incident detection, emergency response and management, pervasive access to medical data and prevention. Pervasive healthcare is impacted by the trend and challenges. Trend refers to the existing needs or motivations of implementing pervasive healthcare, for example driving the operational costs down, improving patient care quality and reducing medical errors.

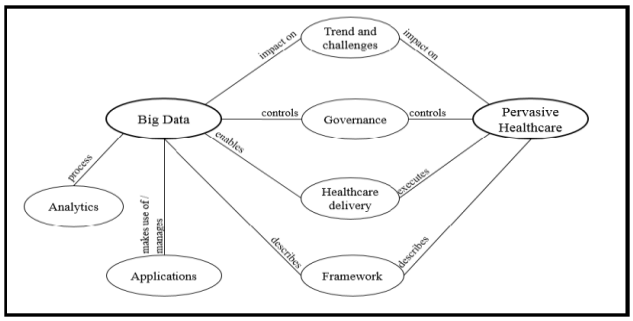
Challenges indicate the data related issues such as data interoperability, privacy and security. Data collection in the healthcare delivery is enabled by big data (Macdonald, 2012).

According to Madden (2012) big data applications make use of or manage big data. Analytics is required for data processing and analysis. The processed data is provisioned to big data, where data is used to enable pervasive healthcare delivery. For example, the patient care quality is improved through a better decision. A better decision can be made via the processed sensor data which usually comes in an unstructured format, together with the structured data captured in the electronic health records (EHR). The challenges in trend and challenges impact big data. Governance controls the data quality in big data for delivering pervasive healthcare. Governance considers the requirements from social layer, where information requirements from related processes are collected before they are facilitated in the technical layer at the data level. Framework describes big data provision for pervasive healthcare, with governance addresses the identified challenges.

1.7.2 Conceptual Framework

In the healthcare sector, government organizations and health insurance agencies, such as the Rwanda Food and Drug Administration, National Healthcare Insurance Corporation, Health Insurance Review and Assessment Service, and National Cancer Center have begun to apply big data analysis towards evaluating services.

In order to provide pervasive healthcare services tailored to a person’s lifecycle, it is imperative to formulate a model for future prediction and policy decision by utilizing big data from various sources of healthcare data. Integrated data management and analysis of big data is necessary to achieve pervasive healthcare system.



Moreover, it is necessary to establish an organization to analyze health risk and services, establish policies, and evaluate on the individual, community, and country levels. We named the assumed organization the 'Health Risk Analysis Center' (above Figure). The health risk analysis center is expected to enable early response to diseases at the national level. It will predict the distribution and trends of major diseases through monitoring health behaviors and home environments, the utilization of medical services, and other health-related data and population statistics.

1.8 Scope of the Study

As far as geographical scope is concerned, this study is going to be conducted in Nyanza District, southern province of Rwanda at Gitwe Hospital. This area has been chosen mainly because it is the home district for the researcher and therefore both the interest and benefit are high.

As for time scope, the study will be carried for a period of one year (from January to December 2019).

The study has been limited to this period of time because it is a period during which the researcher can get relevant data on big data and pervasive healthcare system.

As for content scope, the information will be obtained by using both primary and secondary data collection methods. The primary data will be obtained through the engagement of the respondents in an exercise of questions and answers while seeking to obtain their view in application of big data in pervasive healthcare. Secondary data will be gathered using the documented sources and journals.

1.9 Definition of Terms

Big Data

Big Data is defined as a collection of data elements whose size, speed, type and/or complexity require one to seek, adopt and invent new hardware and software mechanisms in order to successfully store, analyze and visualize the data.

Pervasive healthcare

Pervasive healthcare aims to deliver deinstitutionalized healthcare services to patients anytime and anywhere. Pervasive healthcare involves remote data collection through mobile devices and sensor network which the data is usually in large volume, varied formats and high frequency.

1.10 Literature Review

Pervasive healthcare is defined as the use of portable devices, sensors and wireless technologies, where data is created, stored, retrieved and transmitted in real time for the purpose of improving patient care (Akter and Ray, 2010).

Unlike the conventional delivery of healthcare services, pervasive healthcare extends the provision of healthcare services outside a hospital by heavily utilizing the mobile and wireless technologies for enabling real-time data collection, monitoring, and even interactive intervention of individual patient's activities. It leads to a dependency of powerful technical infrastructure which is capable of handling big data in high frequency and delivering context-dependent personalized service.

According to Gaggioli and Riva (2012) the benefits of pervasive healthcare are: enhancing continuous care provision, where patient's treatment is moved from hospitalization to home-based treatment, encouraging patient centric care provision, where patients can monitor their health condition with the assistive decision support systems, improve communication among clinicians and patients, avoiding data duplication as clinicians can view and share data of the clinical services involved in patient care, and reducing operational costs as the needs for hospitalization is decreased.

Big data, together with its analytic technology revolutionizes how healthcare works (Hardin, 2013). It has been applied in various healthcare areas such as: optimizing the clinical pathways, using predictive modeling to enhance patient care, enhancing personalized care, where medicine is developed through understanding the patterns, commonalities and correlations with patients which leads to embracing the pharmaceutical industry (e.g. inventing drugs for HIV with big data), improving healthcare services delivery performance, enhancing public health surveillance where real time data is collected through mobile devices and sensors to provide insight on how people move and behave which helps in understanding the spread of disease and contributing to genomics and proteomics research (e.g. the ribonucleic acid (RNA) research requires analysis on large scale of genomic data in order to search for a new RNA gene sequence and predict their target.

Big data solution is incorporated with big data analytics where certain algorithms are built for processing the data in order to gain certain insights of a specific application area (Hrickiewicz, 2012).

There are other co-products developed on top of the big data analytics process, such as the MOLMeth monitors the process of data capture, data analysis and publications, Modular API (MAPI) integrates the large data sets in different formats and protocols prior to the big data processing, and Sagace, a web based search engine that retrieves a wide range of biological data such as gene expression profiles and proteomics data. The research in cloud-based big data services helps promoting pervasive healthcare.

More importantly, the reviewed literature did not seem to provide sufficient emphasis on how to achieve the "Value" of big data, which is a

potential add on to the "V" series when describing the characteristics of big data. Making sense of the health data is imperative for delivering the value of pervasive healthcare such as to increase patient safety and to reduce the operational costs. Forming a hypothesis is a departure point for any sense-making process and it articulates the process of how data is collected, processed, analyzed and disseminated in the pervasive healthcare setting. In hindsight, sense-making of health data is hypothesis-driven. Despite of the existing technological platforms that enable the features such as finding patterns, trends and relationship through the collected health data, it is still short of literature in postulating the methodology of making sense of these data.

1.11 Hypotheses of the Study

Hypothesis 1: Pervasive healthcare have more positive impact in social sector of Rwanda.

Hypothesis 2: Big data evokes a secured healthcare system in Rwanda.

Hypothesis3: There is a significant impact of big data provision in supporting pervasive healthcare.

1.12 Materials and Methods

This study used descriptive and correlational research design. It described independent variable and dependent variable. It also looked for the relationship between independent variable and dependent variable to stress on the contribution of big data on the pervasive healthcare in terms of percentages. To determine the extent to which the above variables are related, the researcher will made recourse to a survey to record data.

To characterize respondents and get primary data from them at one time, a written survey made of partially open-ended or restricted items will be designed and administered to sampled respondents. The types of data sought from them were numeric or nominal but likely to be turned into quantifiable data. Probability sampling and non probability sampling techniques were used namely systematic sampling methods and purposive sampling method. To get participants respond effectively, the researcher was physically present while participants were completing the survey.

The population of this research work is made of 230 employees of Gitwe hospital. The sample size was derived from population 230 employees using the Sloven's formula at a confidence interval of 90% and margin of error of 10% as described below:

$$n = \frac{N}{1 + N(e)^2}$$

Where:

n is the minimum sample size of Gitwe hospital employees

N is the population from which the sample was drawn estimated at 230 employees

e is the margin of error estimated at 10%.

Substituting in the above formula, the sample size is determined as;

$$n = \frac{230}{1 + 230(0.1)^2} = \frac{230}{1 + 230(0.01)}$$

$$n = 69.697 \simeq 70$$

This means that 70 employees out of 230 were the sample size of the respondents from the whole population in this study.

1.13 Ethical Consideration

In dealing with the human beings who participated in this study, it was incumbent on me to follow ethical guidelines expressed by Tagliamonte (2006). First, I have guaranteed anonymity of respondents. Then, I have made sure their participation in this research was voluntary and finally, I have assured them that they will have free access to me and to the research findings even after the completion of this research.

1.14 Limitations

The limitations are the literature gap and the technical issues affecting the proposed architecture. From the literature perspective, there is limited published work about big data in pervasive healthcare as an integrated solution. From the technical perspective, the informatics experts find that it remains challenging to manage and process the unstructured data in the pervasive healthcare context.

They emphasize explicitly that the data provision process should focus on the semantics of the data, i.e. the right content in a right context.

Furthermore, the health socio sectors are subjected to the feedback loop from the users to revise the meta-data standards in order to improve the health outcome. They suggest that the health socio sectors to be developed as a distributed model. This practice will reduce the implementation costs and at the same time the stakeholders involved will hold the same accountability. Then again, when the health socio sector is a live system, it should deliver a timeliness and accurate data for increasing the pervasiveness of the health services delivery.

1.15 References

Ruotsalainen, P. S., Blobel, B. G., Seppälä, A. V., Sorvari, H. O. & Nykänen, P. A. (2012). A conceptual framework and principles for trusted pervasive health. *Journal of medical Internet research*, pp.14.

Drayton, K. (2012). *How mobile technology can improve healthcare*. Nursing times, pp.18-16, 18.

Heerden, A. V., Tomlinson, M. & Swartz, L. (2012). *Point of care in your pocket: a research agenda for the field of m-health*. Bulletin of the World Health Organization, pp.393-394

Feldman, B., Martin, E. M. & Skotnes, T. (2012). *Big Data in Healthcare Hype and Hope*. US: Dr. Bonnie 3600.

Hurwitz, J., Kaufman, M., Halper, F. & Kirsch, D. (2013). *Big Data for Dummies, For Dummies*.

Akter, S. & Ray, P. (2010). *M-Health-an ultimate platform to serve the unserved*. Faculty of Commerce Papers, pp.94-100.

Gaggioli, A. & Riva, G. (2012). *From mobile mental health to mobile wellbeing: opportunities and challenges*. Studies in health technology and informatics, pp.141-147.

Hardin, S. (2013). *ASIS&T annual meeting plenary speaker: Edward chang: Mobile opportunities.* Bulletin of the American Society for Information Science and Technology, pp. 46-48.

Hrickiewicz, M. (2012). *Thinking big about data. Health Facilities Management.*

Chawla N, Davis D (2013) Bringing big data to personalized healthcare: a patient-centered framework. J Gen Intern Med, pp.660–665.

Macdonald, C. (2012). *Using Big Data to Improve Health*. E–The Environmental Magazine, pp.15-16.

Madden, S. (2012). *From databases to big data. Internet Computing,* IEEE, 16, 4-6.

Authors:

1. **NIYONSENGA Aaron,** email: 7niyonsenga@gmail.com
2. **TUYISHIME Jean Claude,** email: tjclaudemvuyekure@gmail.com
3. **DUSHIMUMUKIZA Ernest ,** email: dushimukiza@yahoo.fr
4. **HARUBWIRA Flaubert** , email: daflaube@gmail.com