International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

# COPY RIGHT

ELSEVIER SSRN



IJIEMR Transactions, online available on 4th Sept 2019. Link :http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-09

Title **BIG DATA PROCEDURES IN HEALTHCARE SECTOR**

Volume 08, Issue 09, Pages: 152–157.

Paper Authors

**PRANAVI.ADDANKI, PREMANVITHA.KAKARLA, VASANTHI.BHARANI**

Anubose Institute of Technology


USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

# BIG DATA PROCEDURES IN HEALTHCARE SECTOR

[1]PRANAVI.ADDANKI, [2]PREMANVITHA.KAKARLA, [3]VASANTHI.BHARANI
[1,2,3]CSE, AnuboseInsstitute of technology
[1]Pranaviaddhanki15@gmail.com[2]Kakarlaprema9@gmail.com[3]Vasanthiswaroop507@gmail.com

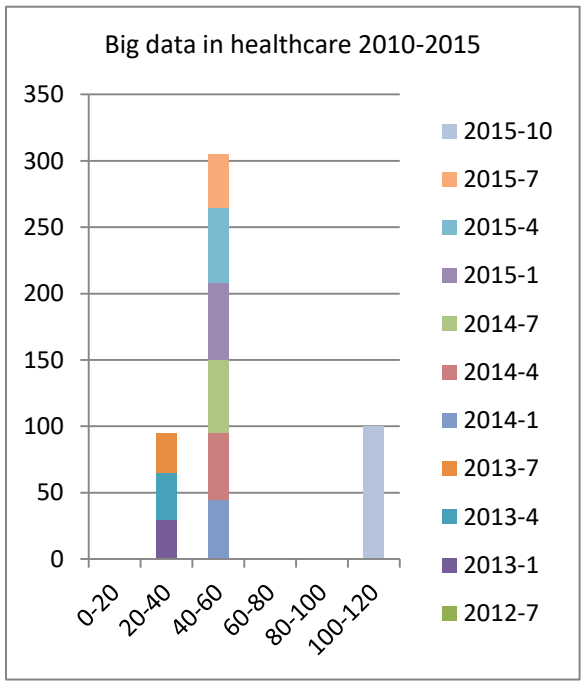
## Abstract

Big data is a fast-growing technology, based on the data collected meaningful insights are generated. The big data extended its wing to the health care sector. The study has explored the current disputes that governments and health care collaborators are facing as well as presented the big data. By using big data there is a lot of benefits sometimes and it is life-saving also. Undoubtedly, for years gathering a massive amount of data for medical use has been expensive and time-consuming. The main theme of this paper is how big data works on the health care sector. The big data tools are used to find these things because most of the time unstructured data approximately of 90% is obtained than the structured/semi-structured data. Methods that are followed by big data are data acquisition, data storage, data management, data analytics, and data visualization. These methods make meaningful insights out of unstructured, semi-structured and structured data. Recent time's big data's data visualization tools are being popular as it is important to understand the situation and results obtained. The main theme of this paper is to understand better big data on health care industries and the application of big data in healthcare services.

## Introduction

One of the world's biggest and widest developing sector is health care industry, world changed around disease-centered to patient centered system. The main advantage of big data era is that it creates new opportunities for health and medical domain as it knows the big data contains massive amount of data. It might be unstructured, semi-structured/structured. A commonly cited statistic volume from EMC(Hopkinton ,MA, USA)said that 4.4 zeta bytes of data existing globally in 2013. And it also predicted, it increases to 44 zeta bytes by 2020. Big data is driving us to value based health care delivery model. Value based health care is educating the superiority of health and decreasing cost. Many potential big data sources are the digital medical records like MRI scan and CT scan. The data might be noisy some times as the users can't use the insights data preprocessing, data modeling, security and visualization are needed to analyze. To make the data analysis efficiently and conviently the data needed to be preprocessed, and typical algorithms, tools based on big data platform are used. Many countries public care systems are providing electronic patient

records.To develop the patient care eventually the big data provides an opportunity for epidemiologists(a person who is an expert in the branch of medicine which deals with the possible control of diseases).in health care big data encounters the upcoming market trends and needs. According to the research the predictions on the world wide big data expenditure in health care business to progress towards compound annual growth rate(CAGR)of 42% during the years 2014-2019.For the future world there is a need for innovative and new big data tools. Technologies that can meet and exceed the ability of managing health care data.

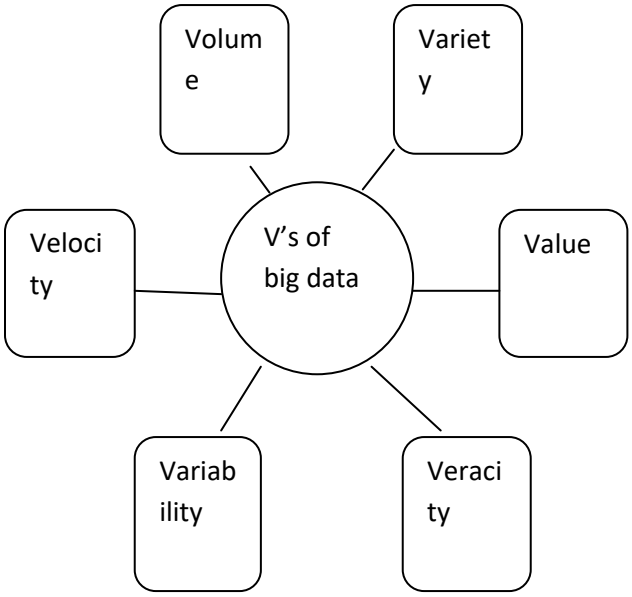



Google inclines an unreservedly accessible online entryway of Google Inc. which is supreme. That enables clients to help with Internet pursue information, which may give savvy dreams into individuals' activities and wellbeing associated potential outcomes. Google dengue and Google influenza patterns are utilized for throwing the compass of sicknesses like influenza and dengue. Google patterns are utilized in many research distributions [9]. It very well may be seen from the above diagram on enormous information's medicinal services around 2013. Increment in enthusiasm for this term can be identified with a famous report by McKinsey and Company that turned out in mid 2013 [10]. The report features that human services working expense contributes about 17.6% of GDP and have a conceivable to diminish medicinal services spending by $300 billion to $450 billion.

*Big Data Meaning:*

There are different definitions that are provided for research of big data A huge research on the definition of big data has done by Baro et Al and he also proposed that a data set can also be qualified as big data set only if log(n*p)is higher-up or equal to 7.The u. s. congress in August 2012has submitted a report on big data it explains about “massive volume of high complex, velocity and variable data that need advance technologies and techniques to sanction the storage, management, capture, analysis of a information and distribution. Health care of bid data involve collection of large data from diverse health care foundations like managing, visualizing, storing, convey information for effectual decision making and analyzing. The bid data in healthcare is related with 6diffeerent attributes veracity, volume, variety, velocity, value and variability.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org




Volume indicates the massive quantity of data fabricates by the organization. Nowadays health care data are in peta bytes (10^15 bytes),terabytes(10^12 bytes)or Exabyte(10^18 bytes).in future the broad gross of records on clinical data will be increased to yotta byte(10^24 bytes)or zetta byte(10^21 bytes) .the various patterns of data in healthcare can be categorized as semi-organized, unorganized and organized. Organized data involves clinical,data from relational database,laboratory data and sensor data. Unorganized data are text free data that generally does not have a particular design such as radiological images, manually written data from x-ray images, or medical images,graphics,physiological measures. Semi-organized data involves the stored data in extensible markup languages format (xml). Velocity refers to the large frequency in the period of current date is supplied,created and managed. The generated data can be either in batch or real -time data. The data contents are often altering through the attentiveness of corresponding data assemblies. Example of velocity, there is increasing growth rate of patients with the increase in growth rate of data by55-60% every year. Velocity mentions to the precision and correctness of information. There is low veracity for big data it can't be 100 % accurate,it's validity is also difficult. Variability is related to data altering throughout the lifecycle and handling.Mckinsey and company trust that modify of data assign with what is worthy or right for patients or right for health care environment. They follow the five ways of significance based on the concept of patient outcomes and healthcare cost Right care- patient must take the mostly time and suitable treatment offered Right value- for maintaining and improving health care the customers must often look for suppliers and customers for the way to expand

Right living- patients must be restorative to be inspire by taking an announcement part in there health care

Right innovation- for health care delivery the investors must concentrate on the new approaches and therapies. Right provider- for succeeding best results the professionals must have powerful performance records to treat with the patients

*Methods*:

This article with complete analysis deals with the big data in health care .during 2014-2015 a thorough search of about big data in health care is used for review. Below table involves the large publishers' article of peer-reviewed journal. pubMed, Taylor, science direct and Francis. After detaching unrelated approaches and duplications 573 papers where involved for title review. after going through the titles there are 459 papers

on big data which didn't include the word health care in their title where disbar. There are 36 papers which were disbar because they were not directly to health care.78 papers applicable to health care with text full standards were reviewed. Finally the endure 76 papers met all the additions standards were maintained.

*Process of Big Data Analysis in Healthcare Industry*

This big data is introduced to health care just to change the perspective of healthcare practice and equipment. This really helps in increase of awareness about the data and leads to make a clear, declared conclusion

Big data analytics has 5 steps:

- Data acquisition
- Data storage
- Data management
- Data analytics
- Data visualization and report

Data acquisition:

In data acquisition, the data is collected first. It might be structured/unstructured or semi-structured. Primary sources are CPDE, clinical decision, support systems, etc. And secondary sources are laboratories, pharmacies.

The important sources in big data health care:

- Electronic health records these are set of data for personalized medicine is required. so, they can collect all the data about the patient and give a beneficial treatment.

- Image processing often in hospitals scanning, MRIs, X-rays, images are produced. They analyze the data.

- Social media and smartphones health care data can be collected through social media like facebook, twitter, LinkedIn, etc.

- Applications of smartphones are doing the same thing. They collect the data about symptoms and then they predict the diseases accurately.

- Web base data is also the source of data acquisition; the popular web sites are 2andme which provides the study of DNA .where asUBiome is a microbiome sequencing service.

Data storage:

Usually, the collected data has to be stored properly. As the platform is increasing, there should be an efficient & effective place to store all the data. Nowadays it is done with the help of the cloud, cloud computing has become prominent for storing the data. Cloud provides elasticity and proficiency to get into data. For using cloud we no need to have costly hardware with numerous space. It can be accessed by the remote clients also, From anywhere we can access the cloud.The main reason that we use the cloud in big data is in the investigation we need efficient and cost-effective infrastructure

Data management:

Storing the data doesn't mean all things are done. The data management in health care

includes organizing, cleaning, method of validating whether there are any null/scrap values.Managing the data for each patient gives a personalized discharge plan. Major tools for managing data are Apache Ambary and HCatalog. In the management, data gets retrieved and management of security, integrity, availability.Recording each patient's information confidentially is also an important thing in healthcare management. So, government regulations intended to address healthcare data privacy.Some of the important acts in the data governance act are HITECH, HIPPA, HDI, GINA& FOIA.HITECH encourage the acceptance and significant use of health information knowledge.

Data analytics:

It is the process of changing raw data into information or make meaningful insights into it. It is divided into 4 ways in big data healthcare analytics:

- Descriptive analytics
- Diagnostic analytics
- Predictive analytics
- Prospective analytics

Descriptive analytics is based on past data/historical data. Also, known as unsupervised learning. It simply tells the thing happened in the health care management & impact of it on the system.Diagnostic analytics, it is based on the past data sets and it analyzes the root cause of the problem. It simply explains why it happened like that.Predictive analytics is based upon both the real-time and historical data predicts the future. Also, known as supervised learning. This analytics is beneficial because it gives an idea about the future trends and makes us take decisions according to it.

Perspective analytics synthesizes the big data and provides advice on the number of possible outcomes. It is also called an advanced version of perspective analytics.

Data visualization:

Data visualization is nothing but presenting the (data) results which were analyzed. It can be represented as a pictorial or graphical format for better understanding of complex data based on this we can make a better decision.

It can be used to understand the pattern and correlation among the data.

*Applications of Big Data in Health Care:*

Huge information is utilized during all periods of pharmaceutical development, for the most part for medication identification [57]. Pfizer has as of late started exactness Medicine Analytics Environment program that partners the dabs among electronic therapeutic record information, clinical preliminary, and genomic to perceive opportunities to rapidly pass on imaginative drugs for specific patient Populations.

Huge Data will make liable to bring best and redid patient consideration. In adjacent future, new enormous information determined impacts will convenient reasonable updates of diagnostic help, clinical system and patient triage to enable progressively specific and changed treatment to propel therapeutic outcome for patients (Yang et al., 2014).

## Conclusion:

This study evaluates information about big data in health care. Data privacy, data leakage, efficient handling of the huge capacity of medical imaging data, data security are the challenging parts of big data in health care. There are some more challenging parts of big data in health care miscarriage to safeguard the health care information and unordered clinical notes in context. After the review there are some limitations that are identified, firstly the publication is limited availability of health care in big data management. This reading tells the major tests of health care in data governance. This study of the article will be gaining the practitioners, health care academicians, testing who are unavailable in the area of health care management

## References:

http://www.ajtab.org/article/328/10.11648.j.ajtab.20180402.14

http://article.sciencepublishinggroup.com/pdf/10.11648.j.ajtab.20180402.14.pdf

https://ieeexplore.ieee.org/document/8345663/