

# Big Data Challenges in Healthcare

Madina Abdulah<sup>1</sup>, Wael.M.S.Yafooz<sup>2</sup>

<sup>1</sup>Faculty of Computer and Information Technology, Al-Madinah International University, Malaysia, [Madenaj76@gmail.com](mailto:Madenaj76@gmail.com)

<sup>2</sup>Faculty of Computer and Information Technology, Al-Madinah International University, Malaysia, [wael.mohamed@mediu.edu.my.edu](mailto:wael.mohamed@mediu.edu.my.edu)

Received 15 September 2017; accepted 08 December 2017

---

## Abstract

The generation and collection of large data sets have become a necessity in many areas that need to be maintained for a long time. While this provides significant opportunities, the main challenge is to provide economically scalable storage systems for efficient storage and preservation of data, as well as enabling search, accessing data and analysis in the future. Data analysis is the process of inspection, cleaning, conversion, data modelling with a view to highlighting useful information, propose conclusions, and support decision-making. Several business users wish to use data analysis of business data to understand trends, to create predictions and improve their business decisions. This paper is a review study to the importance of Big Data and its challenges in healthcare sector.

*Keywords: Big Data, Healthcare.*

---

## 1. Introduction

We live in the time of the owner of the information is own the power, and the big companies such as Google, Facebook, and others, do not miss any opportunity to collect and store any information, for example Google saves all user search results and even if you write anything in the search box without searching, Facebook also stores all the information you deleted and others. Certainly, this information has great value and just owning it is a fortune, and who can manage this data, treat it and get the results benefits, will achieve a lot of successes.

From the biggest challenges in Big Data management is matching business requirements with the appropriate technology[1]. Many companies have yet to formulate a Big Data strategy, while others relegate it to specific tasks in siloed departments. In addition, hiring employees or consultants with a deep understanding of both the financial services business and data technology. Some firms prefer to hire a team of individuals with the combined skills instead of a single person. And then data privacy is a major concern tied to the implementation of cloud computing technologies. [2]

At the same time, the scrutiny of Big Data has led to valuable information, the discovery of future patterns and trends, and multiple customer choices. Thus, huge data has become knowledge that can be invested in all life areas.

“Big data is high-volume, high-velocity and high variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making”.

## 2. Importance of Big Data

The importance of Big Data consists in working to improve efficiency in the use of an enormous volume of different types of data. When defining Big Data and using it properly, organizations can get a better view of their business and thus lead to efficiency in different areas like sales, improving the manufactured product and so forth. [3, 4]

Health care organizations have a large amount of data available to them and a large part of it is unorganized. This data can be evaluated in various places such as laboratory imaging systems, physician notes, and even CRM systems. More and more healthcare organizations are taking advantage of large data technology to capture all patient information. The aim of this technology is to obtain better insights that can help diagnosis and treatment of patients. And the harnessing Big Data can help achieve the multiple critical healthcare goals. [5]

1. An obvious advantage of storing patient data on-line and utilizing Big Data is that the amount of space saved. Data isn't any longer required to be holding on paper, taking over precious physical space for storing and providing only singular access.

2. The Big Data ability to investigate it. By analysing massive data sets, one will determine clear patterns not only in health but in patient satisfaction or hospital operation. This can assist each doctors and directors in understanding strengths, weaknesses, and more. They can analyze data to predict outcomes and start preventive care, similarly as access past data to create treatment plans.

3. Also, can use this data to analyze trends outside of hospitals, ultimately making an oversized data repository

for population health analysis. Big Data and its capabilities give researchers access to several sources which enable them to reach a better understanding of disease and illness.

4. The ability of all time monitoring and recording varied health factors, and this by using mobile health devices. With this capability comes many data points, which enables the doctor to a lot of easily analyze the recorded data to form a lot of accurate diagnoses. As a result, Big Data may be particularly useful in determinant patient satisfaction and trajectories.

Big Data analytics help organizations take advantage of their data and use it to show their importance in different aspects can be summarised in helping to reduce costs, improve confident decision making immediately, and with developing new products and services. And this leads to smarter business moves, more efficient operations, higher profits and happier customers. [6, 7]

### 3. Challenges for Big Data in Health Care

By the nature of Big Data, It can cause many healthcare challenges as:

Data—especially "big data"—isn't static. Its content and use are constantly evolving as it grows ever larger and more complex. [8]

Data sources are highly variable, which causes inconsistencies, duplications, and inaccuracies that undermine people's trust in its validity.

Working with data can be messy due to a lack of agreed-upon industry standards and definitions. This is a major challenge for healthcare given the wide range of nomenclature, measurement units, and other site-specific standards used by hospitals and networks.

#### Healthcare as a Technology Laggard

Healthcare is notoriously slow to redefine and redesign processes and tends to be a laggard in adopting technology that impacts the healthcare system, outside of some specific areas such as care delivery and research. In addition, the healthcare technology landscape includes vast areas of legacy technology, causing further complications[9].

#### Nature of Data Structure

The biggest obstacle to effective use of Big Data is the nature of healthcare information. The majority of data in health care is unstructured, like from language processing. It is commonly fragmented, dispersed, and barely standardized. These are fundamentally difficult to integrate because of concerns about privacy and propriety, the complex and fragmented nature of the data, as well as the different schemas and standards underlying the data and lack of metadata within each silo. Even if everyone shared their data, there would be enough challenges integrating it within the silo, much less outside it. It isn't any secret that the EHRs don't share well across organizational lines, however with unstructured data, even inside the same organization, it is not easy to aggregate and analyze unstructured data. [10]

#### Data Security

There are privacy concerns regarding the use of large data analyses, specifically in healthcare by the enactment of Health Insurance Portability and Accountability Act (HIPAA) legislation. [10] The health data is the most personal data, it is freely available and, hence, highly vulnerable. And protection and security are expected to data owners. The healthcare industry needs to walk the line between patient privacy and security, complying with the increasing numbers. Sensitive, personally identifiable information includes, patient names, diagnosis, medications, hospital records, insurance IDs, social security numbers, billing information, genetic history, and it is at risk. The entire healthcare system can realize benefits from democratizing Big Data access; for example, researchers can more easily collaborate, engage in peer review and eliminate duplication of efforts. Researchers will also be able to more readily identify opportunities where they can contribute and collaborate.

The cloud makes exposing and sharing Big Data easy and relatively inexpensive. However, significant security and privacy concerns exist, including the Health Insurance Portability and Accountability Act (HIPAA). A credentialing process could facilitate and automate this access, but there are complexities and challenges. Since providers, patients and other interested parties such as researchers need secure access; data access should be controlled by group, role, and function. Finally, the security of the data once it leaves the cloud also needs to be assured. Big Data can be used to identify patterns and irregularities indicating and preventing security threats, as well as other types of fraud.

Data Standardization Limited interoperability poses a major challenge to large data, with data rarely being standardized. This leaves Big Data to face issues related to the acquisition and cleansing of data into a standardized format to enable analysis and global transmission and sharing.

With the globalization of data, Big Data has to deal with a variety of standards, barriers in language, different terminologies. [10] Big Data solution architectures have to be flexible enough to cope with not only the additional sources but also the evolution of schemas and structures used for transporting and storing data. To ensure analytics are meaningful, accurate and suitable, metadata and semantic layers are needed that accurately define the data and provide business context and guidance, including appropriate and inappropriate uses of the data. This evolution of standards will eventually improve data quality.

#### Timeliness

One of the key requirements in health care is to be able to utilize Big Data in real time. Real-time is defined by enabling the use of applications such as cloud computing to view said data in real time. The use of these technologies leads to issues of security and privacy within patient information. [11]

Careful attention to data and query structure, scope and execution are needed to ensure that the constraints of the processing windows are observed while still obtaining the best possible answer.

#### 4. Conclusion

Big Data and the use of advanced analytics have the potential to advance the means in which providers leverage technology to create informed clinical decisions. The area of health care is made in data, that leads to the increasing demand of Big Data analyses, and therefore the knowledge of the complexities of "Big information" leads to the development of opportunities for Big Data analyses, which might offer immediate results for patients and reduce the costs of care, and this leads to many ideas about creating the correct decisions Time for patients.

The literature review discovered the challenges that Big Data offers to the healthcare industry. The literature additionally mentioned the importance of Big Data in healthcare sector which might be formed in increased service quality, better management of population health, early detection of disease, and the real time confident decisions.

#### References

- [1] Wael M.S.Yafooz., Abidin, S. Z., & Omar, N. (2011, November). Challenges and issues on online news management. In Control System, Computing and Engineering (ICCSCE), 2011 IEEE International Conference on (pp. 482-487). IEEE.
- [2] Debra Walton, "Big data challenges in financial services", Thomson Reuters, Sep 2014.
- [3] Elena Geanina ULARU.. et al., "Perspectives on Big Data and Big Data Analytics", Database Systems Journal vol. III, no. 4, 2012.
- [4] Yafooz, W. M., Abidin, S. Z., Omar, N., & Hilles, S. (2016, September). Interactive Big Data Visualization Model Based on Hot Issues (Online News Articles). In International Conference on Soft Computing in Data Science (pp. 89-99). Springer Singapore.
- [5.] Vignesh Subramanyan, "Why is Big Data Important in Healthcare?", B2C, 2013, <http://www.business2community.com/big-data/why-is-big-data-important-in-healthcare-0576823#vbS1FICEq5zp238f.97>.
- [6.] Thomas H. Davenport and Jill Dyché, "Big Data in Big Companies", Thomas H. Davenport and SAS Institute Inc., 2013.
- [7] Tom Davenport, "How Big Data will Change Business", Thomas H. Davenport, 2014.
- [8] McKinsey Global Institute. "Overcoming Healthcare Big Data Challenges", 2015.
- [9] Cognizant 20-20 insights, "Big Data is the Future of Healthcare", September, 2012.
- [10.] Clemens Scott Kruse et al., "Challenges and Opportunities of Big Data in Health Care: A Systematic Review", JMIR Publications, Vol 4, No 4, 2016
- [11]Jiang Bian et al., "Challenges and Opportunities of Big Data in Health Care: A Systematic Review", JMIR Med Inform. 2016 Oct-Dec; 4(4): e38. Published online 2016 Nov 21. doi: 10.2196/medinform.5359.