

## Chapter 7

# A Brief Survey on Big Data in Healthcare

**Ebru Aydindag Bayrak**

 <https://orcid.org/0000-0002-2637-9245>

*Istanbul University-Cerrahpaşa, Turkey*

**Pinar Kirci**

*Bursa Uludağ University, Turkey*

### ABSTRACT

*This article presents a brief introduction to big data and big data analytics and also their roles in the healthcare system. A definite range of scientific researches about big data analytics in the healthcare system have been reviewed. The definition of big data, the components of big data, medical big data sources, used big data technologies in present, and big data analytics in healthcare have been examined under the different titles. Also, the historical development process of big data analytics has been mentioned. As a known big data analytics technology, Apache Hadoop technology and its core components with tools have been explained briefly. Moreover, a glance of some of the big data analytics tools or platforms apart from Hadoop eco-system were given. The main goal is to help researchers or specialists with giving an opinion about the rising importance of used big data analytics in healthcare systems.*

### 1. INTRODUCTION

The technological developments helped us in producing more data that cannot be easily processed with currently available technologies. Thus, a new term ‘big data’ is created to describe the data that is large and not processed. Healthcare systems are generating huge amounts of data that present many positive and negative situations at the same time. For this reason, big data management and its analysis in healthcare sector are important (Dash et al., 2019).

Healthcare data increase day by day with the improvement of technology. The correct analysis of this data will increase the quality of maintenance and reduce the costs. This kind of data (big data) have some features such as high volume, variety, high speed production etc. Because of these features, analyz-

DOI: 10.4018/978-1-6684-3662-2.ch007

ing data with traditional hardware and software platforms are pretty hard. Hence, choosing appropriate platform for analyzing and managing big data is very important (Nazari et al., 2019).

Considering the studies related to big data and big data analytics in the field of health in the literature, it is seen that quite a lot of studies have presented. Especially in recent years, it has been seen that there is a great increase in the number of studies on this subject. Some of them can be expressed as follows.

Galetsi et al. (2020) have studied on big data analytics in healthcare. Theoretical frameworks, techniques and prospects about big data analytics have been explained. They have aimed to present a systematic overview of the literature in order to show how much big data analytics has managed to contribute the healthcare system. Shilo et al. (2020) characterized health data by several axes that represent different components of the data. They described the potential and hardship of using big data in healthcare resources. They aimed to contribute to the continued argument of the potential of big data resources to improve the understanding of health and disease. McCall (2020) reviewed the interest in the use of big data in healthcare. Many big data analytics examples were explained for health in Silicon Valley. Alghunaim and Al-Baity (2019) studied on the problem of breast cancer prediction in the big data context. Support Vector Machine, Random Forest and Decision Tree used for machine learning classification by applying on each dataset. Apache Spark was used as a big data framework. Also, big data framework Spark was compared with WEKA traditional data processing environment. Bayrak and Kirci (2019) studied on intelligent big data analytics and machine learning systems for early diagnosis of neurological disorders. Many researches about intelligent big data analysis were reviewed in their study. Also, most used platforms or tools for big data analytics was explained. Carnimeo et al. (2019) was aimed to study a new health care network based on big data analytics for Parkinson's disease (PD). According to healthcare network, collected data during motor examinations of PD patients were analyzed and acquired knowledge was used to create a diagnostic report for patients. Dhayne et al. (2019) searched the topic of big medical data integration solutions. Data integration technologies, tools and applications was examined in the study. They focused on finding the strength and weakness of data integration technologies. Especially, they explained data integration and it is the most important factor for healthcare sector. El Hanafi et al. (2019) studied on characterization of big data platforms for medical data. They described the big data environment with different components based on Hadoop Ecosystem. It was claimed that this system can be useful for helping the doctors to pursue their patients remotely. Nazari et al. (2019) investigated definition of the big data and the big data sources. In concern with big data, the advantages and challenges of big data, big data applications, big data analysis and big data platforms were explained. Palanisamy and Thirunavukarasu (2019) reviewed various healthcare frameworks and summarized their important ideas to learn the impact of big data in healthcare. The implications of big data tools in enhancing healthcare frameworks were considered. The big data tools were grouped data integration tools, machine learning tools, scalable searching and processing tools, visual data analytical tools and real-time and stream data processing tools. Uçar and İlkılıç (2019) worked on epistemological and ethical issues of the big data used in healthcare. Especially rather than whether big data was used in healthcare or not, they asked the question of “how” and in which conditions and in what moral lines should people use big data? Wang and Alexander (2019) explained big health data, big data in healthcare systems, applications, benefits and challenges of Big Data Analytics in healthcare system. They also presented a comparison of tools used for analyzing big data. Bahri et al. (2018) was aimed to explain big data technologies on the performance and outcomes of healthcare system. They explained big data process, technologies, and big data applications in healthcare sector. Big data application on healthcare was classified in five groups as Healthcare monitoring, healthcare prediction, recommenda-

13 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:  
[www.igi-global.com/chapter/a-brief-survey-on-big-data-in-healthcare/290980](http://www.igi-global.com/chapter/a-brief-survey-on-big-data-in-healthcare/290980)

## Related Content

---

### Comprehensive Analysis of State-of-the-Art CAD Tools and Techniques for Chronic Kidney Disease (CKD)

Mynapati Lakshmi Prasudha, Rakesh Kasumolla and Deepak Sukheja (2021). *International Journal of Big Data and Analytics in Healthcare* (pp. 1-12).

[www.irma-international.org/article/comprehensive-analysis-of-state-of-the-art-cad-tools-and-techniques-for-chronic-kidney-disease-ckd/287605](http://www.irma-international.org/article/comprehensive-analysis-of-state-of-the-art-cad-tools-and-techniques-for-chronic-kidney-disease-ckd/287605)

### Enhancing the Employee Experience Through Psychological Safety

Rosa Isabel Rodrigues and Paula Cristina Nunes Figueiredo (2025). *Approaching Employee Experience Management With Data Science* (pp. 95-114).

[www.irma-international.org/chapter/enhancing-the-employee-experience-through-psychological-safety/370200](http://www.irma-international.org/chapter/enhancing-the-employee-experience-through-psychological-safety/370200)

### Data Visualization of Big Data for Predictive and Descriptive Analytics for Stroke, COVID-19, and Diabetes

Richard S. Segall and Soichiro Takashashi (2023). *International Journal of Big Data and Analytics in Healthcare* (pp. 1-31).

[www.irma-international.org/article/data-visualization-of-big-data-for-predictive-and-descriptive-analytics-for-stroke-covid-19-and-diabetes/331996](http://www.irma-international.org/article/data-visualization-of-big-data-for-predictive-and-descriptive-analytics-for-stroke-covid-19-and-diabetes/331996)

### Real Time Analysis Based on Intelligent Applications of Big Data and IoT in Smart Health Care Systems

Mamata Rath (2018). *International Journal of Big Data and Analytics in Healthcare* (pp. 45-61).

[www.irma-international.org/article/real-time-analysis-based-on-intelligent-applications-of-big-data-and-iot-in-smart-health-care-systems/223166](http://www.irma-international.org/article/real-time-analysis-based-on-intelligent-applications-of-big-data-and-iot-in-smart-health-care-systems/223166)

### A Review of Text Analytics Using Machine Learning

Ajaypradeep Natarajsivam and Sasikala R. (2022). *Handbook of Research on Opinion Mining and Text Analytics on Literary Works and Social Media* (pp. 54-69).

[www.irma-international.org/chapter/a-review-of-text-analytics-using-machine-learning/298866](http://www.irma-international.org/chapter/a-review-of-text-analytics-using-machine-learning/298866)