University of Sindh Journal of Information and Communication Technology (USJICT)
Volume 6, Issue 1, March 2022
ISSN-E: 2523-1235, ISSN-P: 2521-5582
Website: http://sujo.usindh.edu.pk/index.php/USJICT/

Significance and Challenges of Big Data in Healthcare: A Review

Syed Muhammad Nabeel Mustafa[1*], Hassan Zaki[2], Syeda Sundus Zehra[2], Muhammad Shoaib[3]

[1]Department of Computer Science and Information Technology, NED University of Engineering and Technology, Karachi
[2]Department of Software Engineering, Sir Syed University of Engineering and Technology, Karachi
[3]Department of Computer Science, Bahria University, Karachi
[1*]nabeelation@gmail.com, [2]hassanzac@gmail.com, [2]Sundus_zehra@yahoo.com, [3]shoaibxhah386@gmail.com

Abstract: The infiltration of information and technology in all the important sectors of the world has influenced the efflux of data. Healthcare industry has greatly benefited more the blessings of advanced computing. This sector is getting digitized at a fast rate for more precision and accuracy. Hence, more data is produced from the healthcare sector. In this review paper, we discuss Big data's importance in the healthcare industry. We go over the qualities of big data, including volume, veracity, speed, variety, and value. The healthcare industry has been touched by big data in practically every way. The diagnosis, telemedicine, medical research, and enhancing patient care are the remaining key effects. The application of Big data gave rise to many challenges such as accuracy, security and privacy, storage and processing, data complexity, unified format, image preprocessing and data analytics. We also intend to investigate further big data uses in healthcare and to highlight the difficulties this industry faces

Keywords: Healthcare; Big Data; Big Data Analytics; Medical Research; Telemedicine

I. INTRODUCTION

The expansion of data in recent years has been facilitated by the advent of the internet. Due to the extensive use of the internet, more and more data enters the mainstream every day [1]. Many studies on Big Data have been conducted by computer scientists in both the academic and industrial sectors [2]. Diverse data is gathered in bulk and stored as big data in cloud databases in order to automate data processing. By 2024 it expected that, more than 120 zettabytes of data will be generated by different sectors [3]. The amount of data being generated is increasing exponentially. Internet users, social media, e-commerce sites [4], 5G Technology [5], health records, web graphs, devices connected through the Internet of Things (IoT) [6] and all online activities, and more are some of the sources of data that are generated. The emergence of ITS as a data-producing entity is another development [7]. Because the data is so dense, it is classified as big data. Big data refers to the vast volume, diversity, and velocity of the data produced by these sources [8]. Big data has ingested every facet of contemporary business operations and industry, and it has become an essential part of manufacturing. The belief is that people who own big data will be successful in the future.

The definition of big data varies across scientists and scholars. A McKinsey global study has predicted Big data to be the reason of competition and creativity in future. A huge scale, dispersion, diversity, and timeliness of the data, according to the study [9], need the use of an innovative technology framework and exploration to delve deeply and usher in new sources of business value. Big data can be seen as a link that subtly joins and unifies the internet, human culture, and the physical world [10]. Big data from the web, gadgets connected to the internet, and other information technologies reflect the physical world in cyberspace. The components of big data-based mapping in human civilization include mobile internet, brain-machine interfaces, and human-computer interfaces [11]. According to Gartner [12], the support of high-volume, high-velocity, and high-variety information assets necessitates special types of processing to enable decision-making, insight discovery, and process optimization.

The healthcare sector is going through a significant shift. The healthcare sector is producing a lot of data and digitizing medical records [13] [14]. It enables hospitals to have their own digital patient records database. Huge amounts of healthcare data have been produced through record keeping, consent and regulatory requirements, and patient care. It is evident that automation pf healthcare industry is largely been done by Big data [15]. Patients are now participating in telemedicine. Big data has improved the precision and accuracy of medical research. It has improved disease prediction and diagnosis while lowering the cost of care.

The other sections of this paper are organized as follows: Section II describes the characteristics of big data; Section III examines the use of big data in healthcare; Section III extols the importance; Section IV describes the difficulties; and Section V wraps up the paper.

II. CHARACTERISTICS OF BIG DATA

The term "Big data" describes a sizeable volume of complex, (semi) organized, and unstructured data that is produced in large amounts and enters a system more quickly than usual in order to be analyzed for improved decision-making, strategic planning, and business operations [16]. On the other hand, the approach to managing massive volumes of data is not new. Big data differs from conventional data in five ways: volume, velocity, diversity, truth, and value. As a result, the 5 V's, which are depicted in Figure 1, are known as the characteristics of big data. The main difficulty in dealing with huge data is not merely its size; we may somewhat alleviate this problem by expanding or extending our computing systems appropriately.

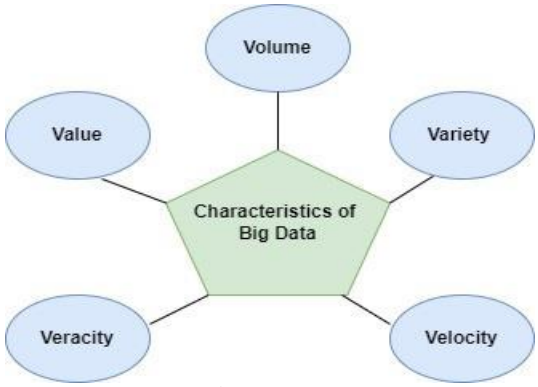

Figure 1. 5 Vs of Big data

A. *Volume*

First, Volume refers to the development of extremely large amounts of data that are acquired and processed and can be expressed in petabytes, exabytes, or zettabytes. Moreover, volume is inversely correlated with data type: a bigger volume of simple data may be equivalent to fewer highly complex data points. This is likely the most straightforward Big Data concept to define [17]. For instance, Twitter receives and processes millions of tweets every day. Similar to this, Facebook manages millions of posts and images every day. Almost a billion searches are made on Google each day. Millions of data recordings are also gathered by sensor technologies utilized in transportation, weather, environmental systems, and other domains.

B. *Velocity*

The velocity of data refers to the rate at which data is produced, processed, and transmitted among various systems and devices. To put it another way, data is accessible at any time, from anywhere and in real time.

C. *Variety*

Variety is the process of combining several data sources to provide the necessary information or results. Big data types and formats include structured, semi-structured, and unstructured data, for instance. Variety is related to the wide number of information/data sources and formats, including social media, video, chat, and mobile devices.

D. *Veracity*

Data's veracity is its consistency, accuracy, reliability, security, and dependability. It implies that data collected should come from a reputable system. It should be less noisy and relevant to the stated aim.

E. *Value*

The value of data is the benefits realized from processing and analyzing vast amounts of data. Value can take many forms, including money, societal worth, research and educational value, and so on.

III. BIG DATA IN HEALTHCARE

The largest and most diverse developing industry in the world is thought to be the healthcare sector. Healthcare management methods in the past were disease- and volume-based. Today, however, it has changed to a value-based healthcare delivery paradigm and a patient-centered strategy globally [18]. Big data is used to predict diseases before they occur based on medical data.

Nowadays, public health systems in many nations use electronic patient records backed by medical imaging media. Businesses in the healthcare sector can foresee future market trends and needs thanks to big data. Epidemiologists, physicians, and health policy experts have a great opportunity to use big data to make data-driven decisions that will eventually improve patient care [14]. Big data in healthcare refers to the collection of enormous volumes of data from several sources, followed by the organization, analysis, visualization, and dissemination of the data for efficient decision-making. Healthcare big data examples include electronic medical records (EMRs), which store patient medical records, doctor notes, clinical reports, and other health-related information, as shown in Figure 2.

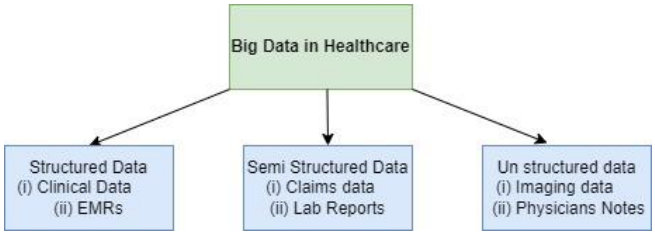

Figure 2. Sources of Heath care Big data

A lot of data is being generated by the healthcare industry, but due to various reasons such as ineffective frameworks, it si not being able to provide better and faster healthcare benefits around the globe [19]. Electronic data is only unavailable, useless, and unsuitable because it is unavailable, useless, and inappropriate. It is challenging to correlate data that can reveal patterns beneficial in the medical industry because to healthcare databases that save health-related information. Big data has been acquired from

medical cell phone applications [20], enabling doctors to practice evidence-based care.

IV. SIGNIFICANE OF BIG DATA IN HEALTHCARE

Public health is only one of several sectors that have been significantly impacted by the rise of big data. The importance of Big data is making a fortune for healthcare systems all around the world. Below is a full explanation of big data's importance in healthcare:

A. Diagnosis

Big data in healthcare is opening up more and more possibilities to investigate novel diagnostic approaches. With more data storage, it will be possible to anticipate disease in every patient by mining disease trends and patterns. Using Big data [21] and various Machine learning (ML) and Deep learning (DL) approaches, predictions about the patients' health are made. The most recent mobile applications aggregate big data to provide patients with medical condition advice.

B. Medical Research

Medical research is an essential aspect. It improves the ability of medical professionals to treat illnesses. Big data enables researchers to analyze extensive datasets in order to identify illness trends buried within the data. It gives researchers the ability to forecast the condition and helps them make decisions about diseases that may have occurred in the past that may have been similar [22]. the investigation of the Big data could provide information about prior pandemics so that people today can plan appropriately. Big data is undoubtedly assisting healthcare researchers in overcoming the new obstacles in the sector.

C. Improving Patient Healthcare

Big data has also been used to improve homes and workplaces while keeping an eye on the wellbeing of patients, the elderly, and workers. Temperature, humidity, and carbon dioxide levels have an impact on employees' comfort and medical wellness. Thus, maintaining these elements at their nominal values is crucial for a positive and productive work environment. Healthcare professionals can identify individuals who are at a high risk of getting specific ailments or diseases and take preventative actions to avert or minimize them by analyzing massive sets of data [23]. Big data can be used to generate patient profiles that include details about each patient's medical history, way of life, and genetic makeup. Personalized treatment programs that are suited to the particular requirements of each patient can then be created using this information. Healthcare professionals can spot patterns and trends in the outcomes of patients' health by studying data on sizable patient populations. With this data, plans for enhancing the health of entire populations can be developed.

D. Real-time Alerting

The most efficient aspect of Big data is the real time monitoring. Big data analytics provides real time monitoring through various tools that keep the doctors updated about the health of patients. Clinical decision support (CDS) software is used in hospitals to assess patient data in real-time and help doctors prescribe treatment regimens [24]. Additionally, by connecting this data to a database on public health, practitioners will be able to analyze it in the context of socioeconomic issues and adjust their delivery methods. With the aid of cutting-edge instruments, institutions and care managers will keep an eye on this large data stream and take swift action if the findings are concerning.

E. Electronic Health Records

Electronic health records are one of the most significant use cases for big data in healthcare (EHRs). By implementing EHRs, it is possible to boost productivity, enhance care coordination, and reduce medical expenses [25]. This is so that testing and medication overlap across doctors and providers can be kept to a minimum. One further important aspect that can result in more individualized treatment is predictive modelling. This plan is based on information from individuals with comparable diseases, inherited factors, and lifestyles.

EHR enables quick data retrieval and reporting of significant healthcare quality indicators to companies. Public health surveillance is improved by the ease with which viral illnesses and epidemics can be reported [26]. Finally, billing and claims administration delays and misconceptions can be greatly reduced or eliminated with the use of electronic health records (EHRs). Patients now have access to millions of health-related medical records that are essential to their life because to EHRs and the internet. Big data in healthcare allows doctors and other healthcare professionals to find connections between various diseases and improve treatments using data from actual clinical experiences.

F. Telemedicine

In the contemporary era, the excessive use of portable wearable and devices is helping people to monitor their health. As our society becomes increasingly mobile in practically every part of life, the healthcare infrastructure needs to be modernized to embrace mobile devices. Mobile applications can be useful for healthcare operations and services. These apps let hospital administrators and doctors keep tabs on their patients and transmit and receive health reports instantly. In other words, mobile apps enhance patient-provider communication and service quality [19]. With the help of these apps, data may be seamlessly integrated with a wide range of consumer electronics and embedded sensors. Clinicians can instantly access patient health information thanks to telemedicine technologies. The patient's current condition is always communicated to the doctor. These smart devices and apps can help us plan for

our wellness and motivate us to lead healthier lives [27]. Users or patients may decide to promote their own health

G. Reduction in Cost

Based on a patient's medical history, lifestyle, and genetic information, big data analytics can assist in creating individualized treatment regimens for them. This strategy may result in more efficient therapies, fewer problems, and shorter hospital stays, all of which may aid in lowering treatment costs. As a result, money won't be squandered on pointless procedures or hospital stays, which will cut costs and better utilize resources. These data can also help clinicians manage patients more effectively, which could, in some cases, result in fewer hospitalizations and re-admissions. By bringing down the price of medical care, this helps patients as well as medical institutions [28].

Big data analytics provides extremely helpful therapeutic insights that cannot be discovered using conventional methods. It improves patient care and lowers costs by enabling healthcare professionals to prescribe medications and make clinical decisions more accurately. Also, it is possible to calculate each patient's specific therapy expenses, enabling meticulous treatment planning to significantly improve healthcare efficiency.

V. CHALLENGES OF BIG DATA IN HEALTCARE

Today, maximizing the use of big data involves a variety of challenges, including the creation of analytical tools at the higher layer, the design of processing systems at the lower layer, and a number of open scientific issues. Big data's unique properties contribute to some of these issues, while the shortcomings of the available analytical models and techniques contribute to others. The key issues and difficulties are briefly covered in this section.

A. Storage and Processing

For big data scientific research, powerful big data processing systems that can process complex structured and un structured medical are the need of industry. High computer complexity, a lengthy duty cycle, and real-time requirements are necessary for processing data that is large in volume, complex in structure, and sparse in value [29]. These requirements not only present fresh challenges for the development of system architectures, computing frameworks, and processing systems, but they also place strict restrictions on their working efficacy and energy usage. Such a large amount of data requires a lot of processing time. The complete data set would need to be inspected in order to find all acceptable parts, which is not practical. So, it is a wise practice that considerably improves performance to create indexes as soon as the data is collected and saved.

B. Accuracy

Due to inadequate EHR usability, complicated workflows, and data acquisition, clinical data accuracy is hazy. All the aforementioned factors contribute to a rise in Big data difficulties over the entire lifespan. the quality and communication of data in clinical workflows is enhanced by EHRs, notwithstanding the problem of data accuracy. Using patient self-report questionnaires for their symptoms may help to enhance the standard of documentation.

C. Image pre-processing

The medical imaging is a tough field to analyze. There are many issues data quality and incorrect interpretations of recent medical records. Medical photography is often hampered by technical issues including noise and artefacts [30]. When medical images are inappropriately processed, for as by defining anatomical components like veins in a way that is at odds with actual environments, image tampering can take place. The issue of images can be done by reducing the noise, image quality repair, and artefact removal.

D. Data Analytics

Big Data presents a wide range of analytical challenges. Large-scale data analysis is done using a technique called big data analytics. This will help in the discovery of obscure trends, puzzling connections, and other valuable information that could be utilized to make better judgements. A significant level of technical expertise is needed to perform these studies on vast amounts of unstructured, semi-structured, and ordered data. Furthermore, it is still unclear how the best analytics architecture should be able to manage both real-time and historical data simultaneously.

E. Data Complexity

These days, the most widely used formats for storing digital data are unstructured databases and discrete text files. Moreover, as data volume grows daily, it becomes more difficult to organize data from multiple sources and formats [31]. Big data is characterized by a vast variety of data quality levels, complex interrelationships, and a wide variety of data types and patterns. Due to the intrinsic complexity of Big Data, it is more challenging to recognize, describe, comprehend, and calculate it, which raises the level of computational complexity. Working with huge data sets makes standard data analysis and mining tasks, such as topic identification, semantic analysis, sentiment analysis, and others, more difficult [32]. The reasons for data complexity include the variety of data formats, the number of data sources, and the rate of data expansion.

F. Unified format

Patients produce an enormous amount of data, which is challenging to acquire in a traditional EHR format due to its complexity and handling issues [33]. It is difficult for healthcare providers to manage vast amounts of data when there is poor data organization. Several codes, including Current Procedural Terminology (CPT) and International Classification of Diseases (ICD) code sets, were created to create medical reports in a uniform format.

G. Security and Privacy

Big data analytics research and implementation within enterprises are typically hampered by privacy issues. Big data analytics usually makes use of personal data—

information gathered with an individual's agreement for a particular purpose—that was initially gathered for other purposes. Integrating personal data with other sources of data can raise a number of moral and legal questions, including the risk of exposing and disclosing people's private information [34]. Social stratification would be another important result of Big Data predictive analysis, with the knowing person benefiting from it while the underprivileged would be more easily identified and subjected to harsher treatment. Users' security and privacy concerns are increased by moving not only data but also processing to external infrastructures, leading to data loss, data breach, and data theft.

H. Skills and Expertise

Big Data in healthcare industry is still growing. Due to complexity of medical data, the skill set needed by professionals is high level. Lack of personnel skills with big data or general analytics experience is one of the main challenges facing businesses looking to utilize big data. Even worse, a lot of organizations lack the technology necessary to find and fix missing or inaccurate data. These skills should not only be technical; they should also be investigative, analytical, interpretive, and creative in nature. These skills must be acquired by individuals; hence businesses must offer training programs. Moreover, Big Data programs at institutions are necessary to produce knowledgeable professionals [35].

VI. CONCLUSION

Due to a growth in information and technology, big data has had reverberations. Big data is a product of the growing amount of data. Big data is now required by many industries due to its traits. In this essay, we looked at the value and difficulties of big data in healthcare. Big data generation has facilitated disease diagnosis and, as a result, enhanced patient treatment. Big data has helped to lower the cost of treatment by enabling the construction of electronic health records. The patient's health can now be easily tracked in real time thanks to telemedicine. has Big data has some difficulties when its significance is of this magnitude. Big data is difficult to manage because of its size, processing, and security. Since big data is still in its infancy, people lack the necessary knowledge and abilities to manage it. Big data will eventually rule everything, though, as its use spreads. This will make it simple to manage.

REFERENCES

[1] H.-N. Dai, H. Wang, G. Xu, J. Wan, and M. Imran, "Big data analytics for manufacturing internet of things: opportunities, challenges and enabling technologies," *Enterprise Information Systems*, vol. 14, no. 910, pp. 1279–1303, 2020.

[2] W. A. Gunther, M. H. R. Mehrizi, M. Huysman, and F. Feldberg,¨ "Debating big data: A literature review on realizing value from big data," *The Journal of Strategic Information Systems*, vol. 26, no. 3, pp. 191–209, 2017.

[3] N. Forgo, S. H´ anold, and B. Sch¨ utze, "The principle of purpose¨ limitation and big data," in *New technology, big data and the law*. Springer, 2017, pp. 17–42.

[4] M. Tahir, R. N. Enam, and S. M. N. Mustafa, "E-commerce platform based on machine learning recommendation system," in *2021 6th International Multi-Topic ICT Conference (IMTIC)*. IEEE, 2021, pp. 1–4.

[5] S. S. Zehra, M. Magarini, R. Qureshi, S. M. N. Mustafa, and F. Farooq, "Proactive approach for preamble detection in 5g-nr prach using supervised machine learning and ensemble model," *Scientific reports*, vol. 12, no. 1, p. 8378, 2022.

[6] M. Tahir, R. N. Enam, S. M. N. Mustafa, and R. Javed, "Emas: Environment monitoring and smart alert system for internet of things (iot)," in *2020 Global Conference on Wireless and Optical Technologies (GCWOT)*. IEEE, 2020, pp. 1–4.

[7] S. S. Zehra, S. M. N. Mustafa, and R. Qureshi, "Comparing artificial bees colony algorithm and firefly algorithm to achieve optimization in route selection processing time in vanets," *Pakistan Journal of Engineering and Technology*, vol. 4, no. 2, pp. 159–164, 2021.

[8] J. Liu, J. Li, W. Li, and J. Wu, "Rethinking big data: A review on the data quality and usage issues," *ISPRS journal of photogrammetry and remote sensing*, vol. 115, pp. 134–142, 2016.

[9] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, A. Hung Byers *et al.*, *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute, 2011.

[10] X. Jin, B. W. Wah, X. Cheng, and Y. Wang, "Significance and challenges of big data research," *Big Data Research*, vol. 2, no. 2, pp. 59–64, 2015.

[11] Y.-Z. Wang, X. Jin, and X. Cheng, "Network big data: present and future," *Chinese Journal of Computers*, vol. 36, no. 6, pp. 1125–1138, 2013.

[12] C. P. Chen and C.-Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on big data," *Information sciences*, vol. 275, pp. 314–347, 2014.

[13] K. Benke and G. Benke, "Artificial intelligence and big data in public health," *International journal of environmental research and public health*, vol. 15, no. 12, p. 2796, 2018.

[14] S. Senthilkumar, B. K. Rai, A. A. Meshram, A. Gunasekaran, and S. Chandrakumarmangalam, "Big data in healthcare management: a review of literature," *American Journal of Theoretical and Applied Business*, vol. 4, no. 2, pp. 57–69, 2018.

[15] P. Saranya and P. Asha, "Survey on big data analytics in health care," in *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*. IEEE, 2019, pp. 46–51.

[16] B. Gartner and M. R. Hiebl, "Issues with big data," in¨ *The Routledge companion to accounting information systems*. Routledge, 2017, pp. 161–172.

[17] C. J. Cremin, S. Dash, and X. Huang, "Big data: Historic advances and emerging trends in biomedical research," Current Research in Biotechnology, 2022.

[18] M. Naeem, T. Jamal, J. Diaz-Martinez, S. A. Butt, N. Montesano, M. I. Tariq, E. De-la Hoz-Franco, and E. De-La-Hoz-Valdiris, "Trends and future perspective challenges in big data," in *Advances in Intelligent Data Analysis and Applications*. Springer, 2022, pp. 309–325.

[19] H. Zafar, M. Shoaib, S. M. N. Mustafa, M. S. Khan, and S. Ijaz, "Application of data mining of clinical data," in *2020 International Conference on Information Science and Communication Technology (ICISCT)*. IEEE, 2020, pp. 1–8.

[20] S. M. N. Mustafa, H. Zaki, S. Alam, M. S. Khan, and H. Zafar, "Anti depression and anti suicidal application," in *2020 International Conference on Information Science and Communication Technology (ICISCT)*. IEEE, 2020, pp. 1–6.

[21] M. Saadoon, S. H. A. Hamid, H. Sofian, H. H. Altarturi, Z. H. Azizul, and N. Nasuha, "Fault tolerance in big data storage and processing systems: A review on challenges and solutions," Ain Shams Engineering Journal, 2021.

[22] P M. Karatas, L. Eriskin, M. Deveci, D. Pamucar, and H. Garg, "Big data for healthcare industry 4.0: Applications, challenges and future perspectives," *Expert Systems with Applications*, p. 116912, 2022.

[23] S. Leonelli, "Scientific research and big data," 2020.

[24] G. Sheares *et al.*, "Internet of things-enabled smart devices, biomedical big data, and real-time clinical monitoring in covid-19 patient health prediction," *American Journal of Medical Research*, vol. 7, no. 2, pp. 64–70, 2020.

[25] S. M. S.-F. Rose, K. Contrepois, K. J. Moneghetti, W. Zhou, T. Mishra, S. Mataraso, O. Dagan-Rosenfeld, A. B. Ganz, J. Dunn, D. Hornburg *et al.*, "A longitudinal big data approach for precision health," *Nature medicine*, vol. 25, no. 5, pp. 792–804, 2019.

[26] G. Aceto, V. Persico, and A. Pescape, "Industry 4.0 and health: Internet´ of things, big data, and cloud computing for healthcare 4.0," *Journal of Industrial Information Integration*, vol. 18, p. 100129, 2020.

[27] S. Dash, S. K. Shakyawar, M. Sharma, and S. Kaushik, "Big data in healthcare: management, analysis and future prospects," Journal of Big Data, vol. 6, no. 1, pp. 1–25, 2019.

[28] M. El Khatib, S. Hamidi, I. Al Ameeri, H. Al Zaabi, and R. Al Marqab, "Digital disruption and big data in healthcare-opportunities and challenges," ClinicoEconomics and Outcomes Research, pp. 563–574, 2022. [40] S. J. Miah, E. Camilleri, and H. Q. Vu, "Big data in healthcare research: a survey study," Journal of Computer Information Systems, vol. 62, no. 3, pp. 480–492, 2022.

[29] M. Marwan, A. Kartit, and H. Ouahmane, "Security enhancement in healthcare cloud using machine learning," *Procedia Computer Science*, vol. 127, pp. 388–397, 2018.

[30] A. Alharthi, V. Krotov, and M. Bowman, "Addressing barriers to big data," *Business Horizons*, vol. 60, no. 3, pp. 285–292, 2017.

[31] J. Stefanowski, K. Krawiec, and R. Wrembel, "Exploring complex and big data," *International Journal of Applied Mathematics and Computer Science*, vol. 27, no. 4, 2017.

[32] A. Katal, M. Wazid, and R. H. Goudar, "Big data: issues, challenges, tools and good practices," in *2013 Sixth international conference on contemporary computing (IC3)*. IEEE, 2013, pp. 404–409.

[33] C. A. Ardagna, P. Ceravolo, and E. Damiani, "Big data analytics as-aservice: Issues and challenges," in *2016 IEEE international conference on big data (big data)*. IEEE, 2016, pp. 3638–3644.

[34] A. Ashabi, S. B. Sahibuddin, and M. S. Haghighi, "Big data: Current challenges and future scope," in *2020 IEEE 10th Symposium on Computer Applications & Industrial Electronics (ISCAIE)*. IEEE, 2020, pp. 131–134.

[35] M. Saadoon, S. H. A. Hamid, H. Sofian, H. H. Altarturi, Z. H. Azizul, and N. Nasuha, "Fault tolerance in big data storage and processing systems: A review on challenges and solutions," *Ain Shams Engineering Journal*, 2021.