Challenges of Big Data and Cyber Security in Health Center

Prof. Sudheer Shetty[1], Monisha N S[2], Moollya Gautami Bhaskar[3], Nandan M R[4], Nayana T[5]

HoD, Department of Information Science & Engineering[1]

Students, Department of Information Science & Engineering[2,3,4,5]

Alva's Institute of Engineering and Technology, Mijar, Karnataka, India

4al20is029@gmail.com, 4al20is030@gmail.com, 4al20is031@gmail.com, 4al20is032@gmail.com

Abstract: *Today, big data analytics are used in many fields. This section introduces the characteristics of big data, which is a visualization tool for big data, main prospects of big data use in healthcare, Benefits of Big Data Analytics in Healthcare and Progress. This document describes big data technologies in healthcare. blockchain and Data protection is becoming more important in the medical field Privacy issues and blockchain-based cyber security are also affected scrutinized closely. Healthcare is changing rapidly, and so is technology. It should be modified to suit the needs expressed by the provider. When Big data is necessary to promote personalized medicine Good luck in the field of healthcare, but there are many challenges Big data in healthcare and cyber security discussed..*

Keywords: Big data, cyber security, deep learning, machine learning, artificial intelligence, blockchain, healthcare, sensing network, IoT

I. INTRODUCTION

Clinical research data are very different from clinical practice data as clinical research data are of outstanding value. physicians typically produce poor quality data. Moderately caused by the enormous stress providers have to manage, It is further affected by the lack of a "data value" culture. Many doctors, nurses, etc. use data logging Bureaucratic effort and lost distractions patient care [1]. In the past, clinical plans were relied upon. static models created by giant analytics, Heterogeneous and multicentric patient datasets. the other day. It's a one-size-fits-all tactic that works for common patients. How satisfying it is for poorly performing patients is questionable. Features differ from normal patients. A cosine similarity-based patient similarity metric (PSM) was organized using: ICU database identifying the largest number of patients Resemblance to other patients and subsequent creation of her 30 days Mortality prediction model [2].for a better understanding of the relevant genetic and environmental factors Underlying it is that researchers are studying disease Different geographic regions and diverse distribution sethnic group. Geoepidemiology provides important evidence for genetics or an environmentally induced mechanism of autoimmunity [3].Data Precision Medicine Initiative announced in the US in 2015Shown: Integrating omics data or information into electronic health records (EHRs) is a necessary step to practice precision medicine [4].New Applications for Personalized mHealth and eHealth Technology has evolved the way people around the world started Using wearable biosensors. As applications emerge, They also bring new challenges, especially when it comes to goals flexible and complex healthcare market, Suitable, safe, consistent and energy efficient solutions [5].Telecommunication systems, nano- and microelectronics, remote sensing, wearable/mobile computing, advanced materials, as well as geographic information systems, etc., deliver big sensing health data ubiquitously and inconspicuously[6]. Various applications of big data and Big Data technology, especially in the sectors of healthcare and business, have seen promising scalable cloud computing infrastructure support [7].Deep learning employs fast computing to swiftly enhance the processing voluminous and multilayered datasets prearranged in various configurations such as filter layers of convolutional neural networks and recursive layers of recurrent neural networks [8].

Deep learning models use numerous samples to extract high-level features and learn hierarchical representations through combining the low-level input of bigdata with the features of high veracity and high variety.

Although various deep learning models have been created fort he feature learning of high-quality data; some approaches were anticipated to perform the feature learning of low-quality data in the past few years [9]. Although diagnostic certainty rarely reaches 100%,New technologies and providers are constantly improving

system speed. By applying natural language processing, Cognitive programs are rapidly impacting medical practice. Browse scientific literature and electronic medical records(EMR) [8]. Ontological and nomenclature is crude Considered essential in most areas of artificial intelligence. The PhD in Biomedical Informatics and Medicine is Knowledge-based organization to big data analysis. Or The category "Semantic Technologies" currently contains the most features. Storage and analysis methods Those created in other application areas are well advanced Entered the biomedical field. including the design of

Temporal Data Warehousing, Temporal Data Mining, or Visual Mining, temporal multidimensional online analytical processing (OLAP) analysis, and integrated mining and analysis Environments are some important application areas [10]. A Multimodal Algorithm Based on Convolutional Neural Networks(CNN) was developed to use structured and unstructured messages Hospital data and disease risk prediction. For example, CNN Used to automatically perform feature extraction We're doing just as well with text (unstructured data)It has attracted attention [11].

II. BIG DATA AND VISUALIZATION TOOLS

TABLE I.   MOBILE BIG DATA [14]

Aspects	Taxonomy
Data types	Application level Network level
Sources	Cellular networks Mobile social networks Vehicular networks Wireless sensor networks Internet of Things (IoT) Wireless body area networks
Characteristics	Personalized Multisensory Multidimensional Volume Variety Velocity Value etc.
Analytics	Text analytics Video analytics Audio analytics Social media analytics Predictive analytics
Security	Availability Privacy Mobile big data integrity
Applications	Social networking Mobile service personalization Mobile service optimization Mobile advertising Urban planning

Big data is often called heterogeneous data, unprecedented complexity and size, structure and Derived from several sources [12]. Raw data collected by bigData source should be cleaned and converted to formatted record through the extraction, transformation, and loading process(ETL) Organize and cleanse data from multiple sources. Many data formats can be used as inputs for big data analytics. A platform characterized by heterogeneity such as structured or unstructured [13]. Mobile big data is often It is produced in various fields including healthcare. mobile big data Classification is based on data type, source, Features, Analytics, Security, and Applications. big data Visualization is very important. Create trends or patterns in raw data and more actionable information. Large scale visualization tool The data are shown in Table II.

TABLE II.   TOOLS FOR BIG DATA VISUALIZATION [15]

Tools	Data Types	Advantage	Disadvantage
R	All kinds of data	Rich documents, thousands of packages, cross platform, powerful	Programming language, difficult to learn
Tableau	Maps and tables	Google map style, beautiful figures,	Commercial software
Circos	Circular	Widely used, beautiful figures, very good for visualizing genome chromosomes	Difficult to learn, difficult to use
GraphViz	Network	Powerful	Command-line, difficult to learn, difficult to use
Cytoscape	Network	Easy to use, graphical user interface, powerful	Time consuming, click button
Gephi	Network	Graphical user interface, powerful, easy to use	Time consuming, click button

III. BIG DATA IN HEALTH CENTER

Big medical data ushered in a critical era in medicine collection or collection, control and interpretation or clarification of information. Complex (with multisystem, systemic manifestations) and uncommon human diseases (affecting less than 1D44 of her population) are of great interest for big data geoepidemiological analysis. Future developments, no of particular relevance to the medical field is the study of ethnogeographic gradients or the identification of possible environmental influences. Factors that explain geographic differences (such as lifestyle)diet, ultraviolet rays, diet, role exposure to environmental toxins or endemic infections).Big data technology is already transforming healthcare

Research used; must continue to make an impact future change. Big data also influences methods Healthcare providers will make the actual medical decisions in the future [3].

Big data goes hand in hand with the EMR and claims databases in use today. Extensive conversion to EMR generated a huge set of quantitative data (e.g. expenditure data) and qualitative data (e.g. demographic data).(experimental value). Nevertheless, most providers still have manyData as a by-product of healthcare, not as a core asset

It is intended to increase the efficiency of processes in healthcare. However, some medical professionals believe it. Which healthcare systems can use data-driven approaches Improving the quality and effectiveness of healthcare It uses collections of patient and practitioner data [16]. Big data holds promise for improving care and treatment. understanding of many diseases. B. Bipolar DisorderThere are some challenges [12]. have an important point of view The use of big data in healthcare is shown in Table III. AdvantageData obtained from big data analytics in healthcare are as follows [17]:

Patient profile analysis: patient identification

You can benefit from a proactive approach.

Security Observation: Using Big Data Analytics

Investigate large volumes of crisp hospital data in real time to aid security monitoring and adverse events weather.

Genome analysis: Efficient protection of this data Analytical Tactics Are Covered in Genomic Analysis ensure that this method becomes a routine medical decision-making process.

Electronic Medical Records (EMR):

This can be standard data (structured or unstructured).Analysis using big data analytics to find patients. Take risks and provide efficient medical services.

Fraud analysis: Analyze large claims Curb fraud, reduce abuse and waste.

Public health: assessing and predicting disease patterns Disease outbreaks and improved public health.

The future of healthcare Leverage big data analytics to analyze and manipulate large amounts of data A set of various unstructured and structured data sources. This is also caused by the exponential growth of medical image sizes. Force computer scientists to drive innovative solutions. To manage large amounts of data in an easy-to-handle manner time scale. There is also some development going on Imaginative and Incredible Computing Systems Physiological signal processing useful in both research areas and exercise, increasing availability, and ultimately save lives [19].

NoSQL databases have CAP (consistency, availability, partitioning) properties, relational databases, etc.B. OLTP, ACID (Atomicity, Consistency, thermal insulation, durability) properties. Applied the CAP theorem OLTP; was an intermediate content-based routing pattern Provided. Parallel data if EHR data is in OLTP mode Processing is based on data distribution and MapReduce mechanism [20].

High-dimensional attributes of healthcare Big data greatly increases the complexity of the anomaly detection process. Researchers proposed an AOR algorithm for separation

Data into small datasets according to attributes Overlap rate (AOR) [21].

TABLE III. OPPORTUNITIES OF BIG DATA IN HEALTHCARE [18]

Category	Description
Improved patient attention	Big Data analytics helps physicians and other medical professionals make optimal decisions and attract patients' personalized and individual attention.
Discovering a remedy of illness	Analyzing data from various areas and subjects helps find a remedy for illness.
Quality care	Physicians can make exact predictions and offer improved healthcare services through Big Data analytics.
Improved operational proficiency	Early analysis leads to better care and attention.
Indication of fraud	Medical tests, results, treatment verdicts, and other items can be analyzed; therefore, fraud and medicine abuse will be probably identified.
Reduced costs and time of waiting	Improved productivity in healthcare due to Big Data analytics reduces or saves costs; the prediction of the number of patients helps make schedules and thus reduce time of waiting.

Researchers designed 'smart clothes' to collect personal databig data in healthcare for sustainable health monitoring, T his allows collection of unobtrusive variables Physiological value of human functions. mobile healthcare cloud platform built with cloud computing, mobile Internet and big data analytics to get persistent information Intelligent clothing system [22]. under consideration size, rapid growth, heterogeneity and complexity, research EHR data will increasingly require tools for big data

Sometimes bigdata analysis requires massive computational power which is unable to be delivered locally; subsequently, specialized data centers take responsibility of handling the data.

There are two possible provisions to big data applications:

big data network traffic that is service network delivery through the packet switched network layer (ensuring the transport of big data traffic); and optical network provision, where big data traffic is directly transmitted over an optical layer by short-lived light paths [23].

Researchers have presented a new method of generating knowledge from the data created by point of care (POC) devices. Additionally, the system suggests advantages to health authorities in the form of optimal resource allocation (from consumables allocation to deciding the best location for newlabs), so supporting more effective and timely

decision-making processes by providers. Utilizing the data created from POC devices (including machine-generated data and test result such as quality control parameters, errors and warnings, the location of the device, and the duration of the test), population-level surveillance data are available without using sensitive data. BigData analysis contributed to this new type of his POC machine Data generated for knowledge extraction. it is also conceivable Plan and execute managed system architecture Batch and real-time processing in unified technologySome big data techniques [24].

Python pandas, scikit-learn, and matplotlib are healthcarebig data tools. Python Pandas has a partial medical purpose However, the industry uses Python Pandas. Python Pandas provides a SQL-like database query interface for drawing andVisualiza tion with matplotlib is highly recommended. The Scikit Learn Python library provides some basic machines learning skills. B. Classification, Clustering, Prediction,etc Packages like Graphviz and Matplotlib can also help Visualize the data [25]. A new pipeline was discovered,Effectively tested with medical imaging data (functional magnetic Resonance Imaging (fMRI) using big data toolsSpark/PySpark on a single node. Data collected by fMRIThe modality is the fourth dimension. Therefore, a huge amount of Image data is stored for a long time. This pipelineHis PySpark on a single node, which can be used to process fMRI Data for extracting brain networks. pipeline saved processing time and proves its effectiveness (about 4 times fasterachieved an accuracy ofsame level. In addition, in-memory data processing, easeuse, and memory consequences of different data structures Key features of the pipeline[26].

IV. PRIVACY AND BLOCKCHAIN-BASED INTERNET SECURITY

Health data, which is often stored and encoded separately, is Laboratory data, billing data, provider data, nurse data, etc .Notes, Medication Records, and Entire Medical Records. Hence the unified and often anonymized Datasets are the first step in predictive analytics in healthcare [27].

Intel has developed an open architecture into which anonymization fits. It has great advantages provided by cloud storage. A tool for anonymizing and re-identifying weblog entries. Corporate data properties differ from the average example. Literature on anonymization during architecture implementation process. Intel similarly found that the data was anonymized No protection against correlation attacks regardless of masking Obvious personally identifiable information (PID) such as IPaddress and username. We then explored the trade-offs in fixing these vulnerabilities and discovered users.Agent information is strongly correlated with individual users [28].Widely used and practical method for data protection Privacy is the anonymization of data through generalization To appease certain privacy models. local recording Big Data and the Difficulties of Proximity Anonymization Investigate data breaches and find scalable solutions for the problem. A scalable two-phase clustering method Consists of environment-aware agglomerative clustering

Algorithms and t-ancestor clustering (such as k-means)Algorithms were expected to address the above issues. Extensive experiments with real datasets showed thisapproach greatly improved scalability capabilities.

Countermeasures against information leaks in the surrounding area and time efficiencyAnonymization of local records compared to existing anonymizationmethod [7].Blockchain technology has privacy potentialprotection and security, parallel and distributed computing andbackup capability. Can be used on effective datecollection, storage, and analysis mechanisms for retrieving themA more efficient data processing model. BlockchainMedical data and genetic data can be safely stored as personal "DNA" "Wallet". Blockchain technology can provide high quality

personalized health care since the personal history ofPersonal health, doctor visits, etc. may be saferecorded and stored. It can also be used to combatPharmaceutical counterfeiting by affixing our own certification labelReduces drug packaging and long processing timesHealth Insurance Claim. Researchers Developed and ResearchedA method of parallel computing for big data analytics.New Blockchain Parallel Computing ComponentGeneral Parallel and Distributed Computing ExploredParadigm components can be created based on blockchain.Blockchain programmable and time-stamped ledgerEnables intelligent control of record access. DecentralizedRecord keeping system developed based on blockchainBeth Israel Deaconess Medical Center and MIT Media Lab

For EPA administration. data sharing, accountability,Authentication and Confidentiality. IBM Watson Health andFood and Drug Administration (FDA) collaborated and focusedToward blockchain frameworks and oncology-related dataImproving public health. Blockchain enables data collection fromCollect various sources of information and maintain an audit trail of transactions.

Therefore, both transparency and accountability are guaranteedin all data exchanges. data is genomic data,Clinical Trials, EHRs and Other Information Provided by New

Sources such as IoT, wearable devices, and mobile devices. [30]. Programmable blockchain has potential as a solutionhealth care issues. Problems include inefficienciesDelivery of clinical reports, awkward communication,Fragmented health data. Metric createdEvaluate blockchain-based decentralized apps (DApps) accordinglyFor projected capacity, feasibility and compliance Healthcare[31].

A digital health application has been announced to make it possibleRecruitment of clinical trials utilizing IoT data. blockchain Technologies are used to ensure that no personal data is stored It will remain open until one person is registered. CombinationIoT and blockchain technologies that can provide high-quality digital serviceshealth [32]. blockchain is EHR Accessibility and Privacy. framework based onBlockchain was designed for interoperability, security, and efficiencyaccess to medical records used by healthcare providers;Patients and Third Parties. The framework uses smartContracts and uses on Ethereum-based blockchainsEncryption methods for additional security [33].

V. BIG DATA CHALLENGES AND CYBER SECURITY IN HEALTHCARE

People have resorted to "dataism" because it exists There is a lot of data everywhere, A very powerful technology required to process data.

However, the potential of big data lies in healthcare and big data Analytics can reach its limits as personalization increases Medicine, but it's just theory. Comparison with other companies There is a clear contradiction between big data focusesOn Population Health and the Impulse of Health Care Providers to PracticePersonalized medicine; this has slowed things down considerably Adoption of big data in healthcare is limited. and it's still possibleBarriers to big data in healthcareand healthcare [34]is becoming increasingly important.

First, expected benefits of big data, second, all expected benefits Risks of polysomnography (PSG) phenotyping as big dataAnalytics strives for more thrust. leaving a huge recordstatistical misunderstandingsPitfalls Encountered in the Routine Clinical Research Context Highlighted, but can be expressed more complexlyTask. highlight some important considerations When analyzing sleep medicine big data, researchersAnalyzing Simulated Data Using Analyze Large Samples

Diagnostic PSG data obtained from their center [35]. Although the current solution seems to be all-encompassing, Promising Solutions, Big Data, Cyber SecurityHealthcare still faces many challenges. ThisIssues are categorized into four types, including:

skills, organization and culture, technical competenceAbout Big Data Capabilities and Security, and Ethical and Legal Issuesfor privacy. They are shown in Table IV [36] [37].

Additionally, some healthcare providers have experienceChallenges in migrating to blockchain-based systems.Mental health challenges stem from health care issuesData Integrity, Security and Privacy. cultural issues

The health sector can be difficult to overcome [30].

VI. Conclusion and future work:

Big data analytics can be used for patient profile analysis.Security monitoring, genome analysis, fraud analysis, EMRanalysis to identify at-risk patients, analysis of clinical presentation,Predict disease outbreaks and improve public health. BigData analytics has numerous applications supporting providers

Enabling patients to receive personalized care

Big data challenges. big data analysisfamily and personal history tests, clinical information,

By understanding the medical history, etc., we provide new care that is distinct from medical care.We will establish our position as a base in the future.according toAbility to analyze large amounts of big dataof medical data, privacy and cybersecurity based onBig data has become a very important issue in healthcare.

Blockchain technology, a novel emerging technology,It is very useful in providing security and privacy protection.Blockchain is helping drive interoperability in healthcareTheir immutability, transparency and decentralization.Blockchain-based EHR management can provide patientsEHR Ownership and Ultimate Control, EHR Secure ManagementEnsure document accessibility and track records and approvalsrecord transfer. However, there are some challenges from BigData analytics and cybersecurity in healthcare, and theseThe challenges lie in technology,

organization and culture,Ethical and legal issues due to skill competencies and data protection. Psychological and cultural challenges when moving toBlockchain-based medical systems have been a big problem.

processing large-capacity, high-speed, and diverse big data,Mutability and veracity are particularly difficult. Real-time processing of big data from sensor networks and IoTFacilitating telemedicine and MHealth, one of the future research direction. Blockchain-based cybersecurityHealthcare, data mining with large scale healthcare dataSocial networks (such as Facebook), social microblogging (such as)Twitter) may be a future work.

TABLE IV. CHALLENGES OF BIG DATA AND CYBERSECURITY HINDERING HEALTHCARE INNOVATION

Challenges	Description
Challenge in competence	
Skills	Big Data analytics and technologies need new and advanced skills; however, only a very limited number of people have the skills.
Cultural and organizational challenges	
Access to data	Restricted access to various datasets leads to difficulty in data utilization and integration.
Data sharing	Limited data sharing due to lack of trust or business competition
Conservatism	It is still rather cautious to perform data-driven decision-making in healthcare.
Value	People are not sure how big data are used to solve healthcare problems and how much value will be achieved.
Data retrieval, processing, and management	How to retrieve, process, and analyze data well is still a challenge; this limits reproducibility.
Data features and technological challenges	
Data volume	High data volume makes it difficult to capture and store data.
Data velocity	Difficulty in performing the real-time processing of big data with a high velocity.
Data variety	Data in healthcare can be structured, semi- structured, or unstructured due to heterogeneity in data sources.
Data variability	Data can be varied with time and instable.
Data veracity	Data can be poor in integrity and quality due to inconsistency, inaccuracy, and untruthfulness.
Data limitations	There are limitations of observational data.
Validation	Proving or checking the accuracy and validity of data as well as data analytics methods.
Re-identification risk	Big data permit the re-identification of anonymized individuals through the combination of various datasets.
Data security	Abuse or misuse of big data should be avoided.
Challenges in ethics and legal issues	
Debate on privacy	The public debate on the healthcare data sensitivity.
Privacy regulation	Privacy regulations can be a barrier to Big Data analytics and data utilization.
Informed consent	Not easy in collecting informed consent for every individual related to datasets.
Healthcare data gaps	Big data capture is often only for the purpose of reimbursement, leading to big gaps in available data.

REFERENCES

[1]. M. Viceconti, P. J. Hunter, R. D. Hose, "Big data, big knowledge: big data for personalized healthcare," IEEE J. Biomedical and Health Informatics, vol. 19, no. 4, pp. 1209-1215, 2015.

[2]. J. Lee, D. M. Maslove, J. A. Dubin, "Personalized mortality prediction driven by electronic medical data and a patient similarity metric," PloS One, vol. 10, no. 5, e0127428, pp. 1-13, 2015.

[3]. M. Ramos-Casals, P. Brito-Zerón, B. Kostov et al., "Google-driven search for big data in autoimmune geoepidemiology: analysis of 394,827 patients with systemic autoimmune diseases," Autoimmunity

Reviews, vol. 14, no. 8, pp. 670-679, 2015.

[4]. P. Y. Wu, C. W. Cheng, C. D. Kaddi et al. "–omic and electronic health record big data analytics for precision medicine," IEEE Transactions on Biomedical Engineering, vol. 64, no. 2, pp. 263-273, 2017.

[5]. F. Firouzi, A. M. Rahmani, K. Mankodiya, M. Badaroglu et al., "Internetof-Things and big data for smarter healthcare: from device to architecture, applications and analytics," Future Generation Computer Systems, vol. 78, pp. 583-586, 2018.

[6]. J. Andreu-Perez, C. C. Y. Poon, R. D. Merrifield et al. "Big data for health," IEEE J Biomed Health Inform, vol. 19, no. 4, pp. 1193-1208, 2015.

[7]. X. Zhang, W. Dou, J. Pei, S. Nepal, C. Yang et al., "Proximity-aware local-recoding anonymization with mapreduce for scalable big data privacy preservation in cloud," IEEE Transactions on Computers, vol. 64, no. 8, pp. 2293-2307, 2015.

[8]. D. D. Miller, E. W. Brown, "Artificial intelligence in medical practice: the question to the answer?" The American Journal of Medicine, vol. 131, pp. 129–133, 2018.

[9]. Q. Zhang, L. T. Yang, Z. Chen, P. Li, "A survey on deep learning for big data," Information Fusion, vol. 42, pp. 146-157, 2018.

[10]. N. Peek, C. Combi, R. Marin, R. Bellazzi, "Thirty years of artificial intelligence in medicine (AIME) conferences: A review of research themes," Artificial Intelligence in Medicine, vol. 65, no. 1, pp. 61-73, 2015.

[11]. M. Chen, Y. Hao, K. Hwang, L. Wang et al. "Disease prediction by machine learning over big data from healthcare communities," IEEE Access, vol. 5, pp. 8869-8879, 2017.

[12]. S. Monteith, T. Glenn, J. Geddes, P. C. Whybrow et al., "Big data for bipolar disorder," International Journal of Bipolar Disorders, vol. 4, no. 1, pp. 10-27, 2016.

[13]. M. I. Pramanik, R. Y. Lau, H. Demirkan& M. A. K. Azad, "Smart health: Big data enabled health paradigm within smart cities," Expert Systems with Applications, vol. 87, pp. 370-383, 2017.

[14]. E. Ahmed, I. Yaqoob, I. A. T. Hashem et al., "Recent Advances and Challenges in Mobile Big Data," IEEE Communications Magazine, vol. 56, no. 2, pp. 102-108, 2018.

[15]. T. Huang, L. Lan, X. Fang, P. An, J. Min, F. Wang, "Promises and challenges of big data computing in health sciences," Big Data Research, vol. 2, no. 1, pp. 2-11, 2015.

[16]. Y. Chen, G. F. Guzauskas, C. Gu, B. Wang et al., "Precision health economics and outcomes research to support precision medicine: Big data meets patient heterogeneity on the road to value," Journal of Personalized Medicine, vol. 6, no. 4, pp. 20-27, 2016.

[17]. N. Sharma, A. Panwar, U. Sugandh, "Big Data Analytics in Health Care: A Literature Survey," International Journal of Recent Research Aspects, vol. 5, no. 1, pp. 2349-7688, 2018.

[18]. A. G. Alexandru, I. M. Radu, M. – L. Bizon, "Big Data in Healthcare - Opportunities and Challenges," Informatica Economic, vol.22, no. 2, pp. 43-54, 2018.

[19]. A. Belle, R. Thiagarajan, S. M. Soroushmehr et al., "Big data analytics in healthcare," BioMed Research International, vol. 2015, Article ID 370194, pp. 1-16, 2015.

[20]. H. Salavati, T. J. Gandomani, R. Sadeghi, "A robust software architecture based on distributed systems in big data healthcare," In Advances in Computing, Communications and Informatics (ICACCI), 2017 International Conference on, IEEE, 2017, pp. 1701-1705.

[21]. K. Yan, X. You, X. Ji, G. Yin, F. Yang, "A hybrid outlier detection method for health care big data," In Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom)(BDCloudSocialComSustainCom), 2016 IEEE International Conferences on, 2016, pp. 157- 162.

[22]. M. Chen, Y. Ma, J. Song, C. F. Lai, B. Hu, "Smart clothing: Connecting human with clouds and big data for sustainable health monitoring," Mobile Networks and Applications, vol. 21, no. 5, pp. 825-845, 2016.

[23]. B. Cyganek, M. Graña, B. Krawczyk et al., "A survey of big data issues in electronic health record analysis," Applied Artificial Intelligence, vol. 30, no. 6, pp. 497-520, 2016.

[24]. P. Amirian, F. van Loggerenberg, T. Lang et al., "Using big data analytics to extract disease surveillance

information from point of care diagnostic machines," Pervasive and Mobile Computing, vol. 42, pp. 470-486, 2017.

[25]. A. R. Rao, D. Clarke, "A fully integrated open-source toolkit for mining healthcare big-data: architecture and applications," In Healthcare Informatics (ICHI), 2016 IEEE International Conference on, 2016, pp. 255-261.

[26]. S. Sarraf, M. Ostadhashem, "Big data application in functional magnetic resonance imaging using apache spark," In Future Technologies Conference (FTC), IEEE, 2016, pp. 281-284.

[27]. I. Hernandez, Y. Zhang, "Using predictive analytics and big data to optimize pharmaceutical outcomes," American Journal of Health-System Pharmacy, vol. 74, no. 18, pp. 1494-1500, 2017.

[28]. K. Abouelmehdi, A. Beni-Hssane, H. Khaloufi, et al., "Big Data security and privacy in healthcare: a review," Procedia Computer Science, vol. 113, pp. 73-80, 2017.

[29]. Z. Shae, J. J. P. Tsai, "On the design of a blockchain platform for clinical trial and precision medicine," In 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS), 2017, pp. 1972-1980.

[30]. N. Kshetri, "Blockchain's roles in strengthening cybersecurity and protecting privacy," Telecommunications Policy, vol. 41, no. 10, pp.1027-1038, 2017.

[31]. P. Zhang, M. A. Walker, J. White, D. C. Schmidt and G. Lenz, "Metrics for assessing blockchain-based healthcare decentralized apps," In 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom), IEEE, pp. 1-4, October 2017.

[32]. F. Angeletti, I. Chatzigiannakis and A. Vitaletti, "The role of blockchain and IoT in recruiting participants for digital clinical trials," In 2017 25th International Conference on Software, Telecommunications and Computer Networks (SoftCOM), IEEE, pp. 1-5, September 2017.

[33]. G. G. Dagher, J. Mohler, M. Milojkovic and P. B. Marella, "Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology," Sustainable Cities and Society, vol. 39, pp.283-297, 2018.

[34]. J. A. Sacristán, T. Dilla, "No big data without small data: learning health care systems begin and end with the individual patient," Journal of Evaluation in Clinical Practice, vol. 21, no. 6, pp. 1014-1017, 2015.

[35]. M. T. Bianchi, K. Russo, H. Gabbidon, T. Smith et al., "Big data in sleep medicine: prospects and pitfalls in phenotyping," Nature and Science of Sleep, vol. 9, pp. 11-29, 2017.

[36]. T., Iyamu& S. Mgudlwa, "Transformation of healthcare big data through the lens of actor network theory," International Journal of Healthcare Management, vol. 11, no. 3, pp.182-192, 2018.

[37]. E. R. Witjas-Paalberends, L. P. M. van Laarhoven, L. H. M. van de Burgwal, J. Feilzer, J. de Swart, E. Claassen & W. T. M. Jansen, "Challenges and best practices for big data-driven healthcare innovations conducted by profit–non-profit partnerships–a quantitative prioritization," International Journal of Healthcare Management, vol. 11, no. 3, 171-181, 2018.