Babylonian Journal of Internet of Things
Vol.2024, **pp**. 10–15
DOI: https://doi.org/10.58496/BJIoT/2024/002; ISSN: 3006-1083
https://mesopotamian.press/journals/index.php/BJIoT

Research Article

A Brief review of big data in healthcare: challenges and issues, recent developments, and future directions

Russel R Majeed [1], Ashour Ali [2], Ahmed Jumaah Yas [3], Saeed Amer Alameri [4], Mohammed M AL-Ani [2], Ahmed Adil Nafea [5*]

[1] *College of Education for Pure Sciences, University of Thi-Qar, Thi-Qar, Iraq*

[2] *Center for Artificial Intelligence Technology (CAIT), Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia (UKM), Bangi, Selangor, Malaysia*

[3] *College of Computer Science and IT, University of Anbar, Ramadi, Iraq*

[4] *Department of Information Technology, Seiyun University, Hadhramout, Yemen*

[5] *Department of Artificial Intelligence, College of Computer Science and IT, University of Anbar, Ramadi, Iraq*

ARTICLE INFO

Article History

Received 17 Dec 2023
Revised 25 Jan 2024
Accepted 01 Feb 2024
Published 26 Feb 2024

Keywords

Artificial Intelligence

Big data

Healthcare

Internet of Medical Things

Machine Learning

ABSTRACT

Big data is improving the healthcare industry and creating opportunities for improved patient care, personalized medicine, and advanced research. This brief review article aims to survey the challenges and issues connected with big data in healthcare, discuss recent developments in the field, and highlight future directions for helping big data to enhance healthcare outcomes by understanding these aspects, participants can make informed decisions and contribute to the advancement of big data analytics in healthcare.

1. INTRODUCTION

Big Data (BD) describes the large volume of information generated, stored, and analyzed by several industries to improve services [1]. It contains structured, semi-structured, and unstructured forms and originates from many sources [2]. BD techniques provide scalability, allowing ontologies and KGs to handle massive volumes of data. This is crucial for applications in diverse domains, such as healthcare, finance, and e-commerce, where the knowledge representation needs to accommodate a vast array of information [3]. Traditional data processing systems are inadequate for handling BD, requiring specialized tools and techniques[4]. BD has become increasingly related in healthcare, improving patient outcomes, increasing efficiency, reducing costs, and facilitating medical research [4]. Big data is revolutionizing healthcare by transforming the way consumer, patient, physiological, and medical data is collected, analyzed, and utilized, surpassing the capabilities of traditional approaches [5]. Health systems need to integrate technology capable of gathering, analyzing, and interpreting healthcare data to tackle challenges including scale, speed, uncertainty, and accuracy [6], [7]. The term "big data" is not new, but its definition is always changing. The data set is characterized by its large size, varied distribution methods, diversity, and need for timely processing, necessitating the use of innovative technology frameworks, analytics, and resources to unlock new market value [8].

*Corresponding author. Email: ahmed.a.n@uoanbar.edu.iq

1.1 Big Data

Developments in knowledge and the adding total of data in organizations necessitate quicker and efficient data analysis [9]. Big data refers to the large and diverse data sets, making it complicated to explain and collect results utilizing current techniques. Big data helps by gathering and combining data from various sources, transforming it into useful data for organizations. This transforms unstructured raw data into valuable insights for decision-making [10] . In healthcare, BD stands for large and intricate electronic health datasets that are challenging to handle with standard software, hardware, data management tools, and procedures. Healthcare professionals find BD intimidating because of its enormous amount, variety of data formats, and rapid processing capabilities. The BD scientist has opportunities inside this vast amount and diversity of data [11].

1.2 Big Data Analytics in Healthcare

In healthcare big data refer large volumes of digital technology utilized to capture medical records and manage hospital results [12]. This data is utilized in several applications, containing Electronic health Records, computer-generated data, health information exchanges, patient registries, portals, genetic databases, and public records [13]. Big data analytics improves address medical challenges by processing large volumes of data on medical servers, clinical databanks, and clinical decision-making systems. This technology-based system transforms treatment to the right patient at the right time, giving useful information for early diagnosis and treatment [14].

2. CHALLENGES AND ISSUES OF USING BIG DATA IN HEALTHCARE

While there are advantages and benefits of utilizing Big Data in healthcare, there are also problems and concerns associated with its use in this field. Here is a brief discussion of these problems as outlined in Fig. 1.

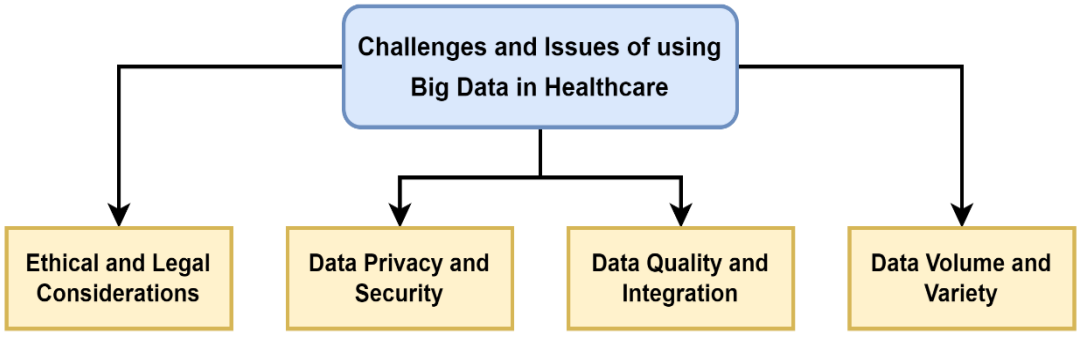

Fig. 1. Common issues of using Big Data in healthcare

2.1 Data Volume and Variety

Healthcare organizations face major challenges in managing and processing large volumes of healthcare data, generated from several sources like electronic health records, medical imaging, wearable devices, and genomic information [15]. Efficient storage, processing, and analysis require accessible structure and advanced data management techniques. The analysis of different data types, including structured and unstructured formats like text, images, and time-series data, can be challenging [16].

2.2 Data Quality and Integration

Data quality is a necessary for accurate analysis and decision-making in healthcare, due to its susceptibility to errors, disputes, and omissions data cleansing and preprocessing techniques are necessary to address these issues. Interoperability challenges, such as data distributed among different systems like electronic health record, laboratory, and billing systems, make it difficult to integrate and exchange data seamlessly [17].

2.3 Data Privacy and Security

Big data in healthcare needs good security methods to protect sensitive patient information, including personal identifiers, medical records, and genomic information [18]. Healthcare organizations must employ encryption, access controls, and secure data storage to ensure data privacy compliance with privacy regulations, such as HIPAA in the US and GDPR in the

EU, is important [19]. Healthcare organizations must achieve informed approval from patients and implement suitable procedures to handle data breaches, ensuring the protection of patient data [20].

2.4 Ethical and Legal Considerations

Big data in healthcare needs to ethical and legal considerations, including balancing data usage with patient rights and consent [21]. Patients should have knowledge when utilized their data, sharing and be informed about potential risks and benefits. Data ownership and sharing are complex legal issues, and clear guidelines and agreements are needed when collaborating with stakeholders like healthcare contributors, researchers, and technology companies transparent and secure data sharing frameworks can ensure responsible data sharing while protecting patient privacy [22].

3. RECENT DEVELOPMENTS

Advancements in healthcare have led to substantial progress in medical technology, patient care, and treatment alternatives. The healthcare field is quickly changing due to the emergence of precision medicine, personalised healthcare, and the incorporation of artificial intelligence (AI) and machine learning into clinical practice. The advancements might transform healthcare delivery, better patient outcomes, and improve the quality of treatment. Fig. 2. shows the recent advancements in healthcare that are discussed in this study.

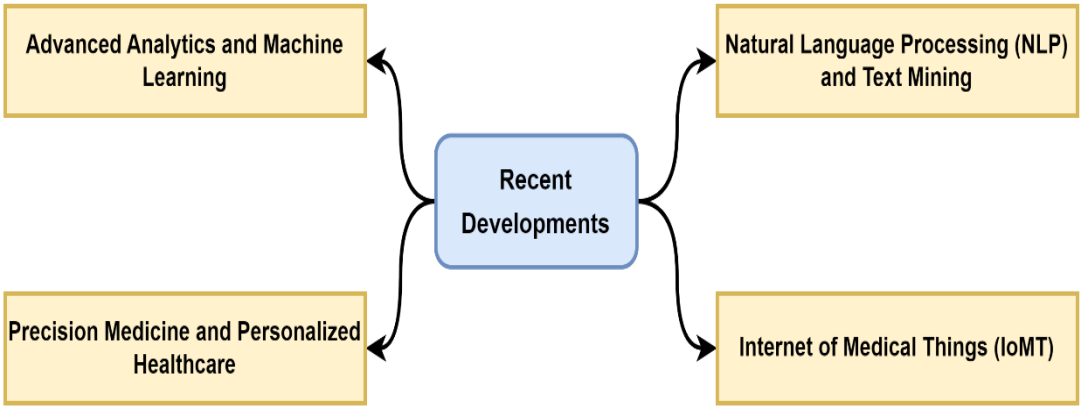

Fig. 2.The recent developments in healthcare.

3.1 Advanced Analytics and Machine Learning

There are a lot of advanced analytics techniques, including machine learning, have supported predictive modeling and decision support systems in healthcare by analyzing large volumes of data, such as electronic health records and medical images, machine learning algorithms can identify patterns, predict outcomes, and support healthcare professionals in making informed decisions. Real-time monitoring and anomaly detection systems help detect and address critical events[23].

3.2 Precision Medicine and Personalized Healthcare

Big data analytics shows a important role in precision medicine, which aims to deliver modified treatment plans based on individual attributes and genetic profiles[24]. Genomic data analysis enables identification of disease markers, genetic mutations, and personalized helpful interventions. Predictive diagnostics and risk assessment models advance in early disease detection, avoidance, and personalized interventions[25].

3.3 Internet of Medical Things (IoMT)

The IoMT shows connecting medical devices, wearables, and sensors to collect real-time patient data [26]. Addition of these devices with big data analytics enables remote patient monitoring, telemedicine consultations, and continuous data tracking. IoMT facilitates proactive and personalized healthcare delivery, improves patient situation, and supports remote care management [27].

3.4 Natural Language Processing (NLP) and Text Mining

Unstructured clinical text, such as physician notes and medical literature, contains valuable information so can utilized NLP and text mining techniques extract insights from this unstructured data, enabling better clinical documentation and

coding[28], [29]. NLP algorithms can analyze large volumes of text to identify disease patterns, extract medication information, and support clinical decision-making [30].

4. FUTURE RESEARCH DIRECTIONS FOR BIG DATA IN HEALTHCARE

The application of BD in healthcare has shown tremendous potential, and as technology continues to evolve, there are several promising future research directions for further advancements in this field. Here are some critical areas for future research in the intersection of big data and healthcare:

4.1 Data Governance and Standardization

To optimize big data in healthcare, it's important to create data governance frameworks, standardization guidelines, and data quality standards. Encouraging data sharing and collaboration can enhance research outcomes and comprehensive datasets [31].

4.2 Artificial Intelligence and Machine Learning

The integration of AI and machine learning algorithms in healthcare offers important potential for enhancing big data analytics, enabling accurate predictions, pattern recognition, and anomaly detection, and fostering intelligent decision support systems for informed healthcare professionals [32].

4.3 Blockchain Technology

Blockchain technology can improve healthcare data security and privacy by utilizing decentralized systems, ensuring transparency, and empowering patients with greater control over their data through secure exchange and consent management[33].

4.4 Real-World Evidence and Comparative Effectiveness Research

It is important in healthcare for decisions that are based on evidence. RWE generates safety, effectiveness, and value insights, while CER compares treatments for specific patient populations, improving clinical decision-making [34].

5. CONCLUSION

Big data could improve healthcare by offering important insights, enhancing patient outcomes, and revolutionizing delivery methods. Challenges in this situation include handling substantial volumes of data, ensuring data accuracy, safeguarding patient confidentiality, and dealing with ethical and legal issues. Despite obstacles, progress in precision medicine, real-time monitoring, predictive analytics, and population health management illustrate the influence of big data analytics. Blockchain technology can improve data security and integrity, while real-world evidence can bolster evidence-based decision-making. Future studies, cooperation, and creativity are necessary to connect the potential of big data in the healthcare sector.

Conflicts Of Interest

No conflicts of interest.

Funding

This work is funded by Universiti Kebangsaan Malaysia.

Acknowledgment

The author extends appreciation to the Universiti Kebangsaan Malaysia. for their support.

References

[1] S. Dash, S. K. Shakyawar, M. Sharma, and S. Kaushik, “Big data in healthcare: management, analysis and future prospects,” *J. big data*, vol. 6, no. 1, pp. 1–25, 2019.

[2] U. Sivarajah, M. M. Kamal, Z. Irani, and V. Weerakkody, “Critical analysis of Big Data challenges and analytical methods,” *J. Bus. Res.*, vol. 70, pp. 263–286, 2017.

[3] A. Ali, S. A. M. Noah, and L. Q. Zakaria, “Representation of Event-Based Ontology Models: A Comparative Study,” *IJCSNS*, vol. 22, no. 7, p. 147, 2022.

[4] L. Baloch *et al.*, “A Review of Big Data Trends and Challenges in Healthcare,” *Int. J. Technol.*, vol. 14, no. 6, pp. 1320–1333, 2023.

[5] L. Bote-Curiel, S. Munoz-Romero, A. Gerrero-Curieses, and J. L. Rojo-Álvarez, “Deep learning and big data in

healthcare: a double review for critical beginners," *Appl. Sci.*, vol. 9, no. 11, p. 2331, 2019.
[6] A. A. Nafea, M. S. Ibrahim, M. M. Shwaysh, K. Abdul-Kadhim, H. R. Almamoori, and M. M. AL-Ani, "A Deep Learning Algorithm for Lung Cancer Detection Using EfficientNet-B3," *Wasit J. Comput. Math. Sci.*, vol. 2, no. 4, pp. 68–76, 2023.
[7] Y. Wang, L. Kung, and T. A. Byrd, "Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations," *Technol. Forecast. Soc. Change*, vol. 126, pp. 3–13, 2018.
[8] M. J. Mazzei and D. Noble, "Big data dreams: A framework for corporate strategy," *Bus. Horiz.*, vol. 60, no. 3, pp. 405–414, 2017.
[9] H. Tao *et al.*, "Development of integrative data intelligence models for thermo-economic performances prediction of hybrid organic rankine plants," *Energy*, p. 130503, 2024.
[10] R. Kitchin and G. McArdle, "What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets," *Big Data Soc.*, vol. 3, no. 1, p. 2053951716631130, 2016.
[11] E. A. A. Ghaleb, P. D. D. Dominic, A. Muneer, A. A. Almohammedi, and IEEE, "Big Data in Healthcare Transformation: A Short Review," *2022 INTERNATIONAL CONFERENCE ON DECISION AID SCIENCES AND APPLICATIONS (DASA)*, no. International Conference on Decision Aid Sciences and Applications (DASA). Univ Teknol PETRONAS, Dept Comp & Informat Sci, Seri Iskandar 32160, Perak, Malaysia, pp. 265–269, 2022.
[12] C. Guo and J. Chen, "Big data analytics in healthcare," in *Knowledge Technology and Systems: Toward Establishing Knowledge Systems Science*, Springer, 2023, pp. 27–70.
[13] M. F. Collen, W. V Slack, and H. L. Bleich, "Medical databases and patient record systems," *Hist. Med. Informatics United States*, pp. 207–288, 2015.
[14] T. Nguyen, Z. Li, V. Spiegler, P. Ieromonachou, and Y. Lin, "Big data analytics in supply chain management: A state-of-the-art literature review," *Comput. Oper. Res.*, vol. 98, pp. 254–264, 2018.
[15] M. Chen, S. Mao, Y. Zhang, and V. C. M. Leung, *Big data: related technologies, challenges and future prospects*, vol. 100. Springer, 2014.
[16] T. V Nguyen, L. ZHOU, V. Spiegler, P. Ieromonachou, and Y. Lin, "Big data analytics in supply chain management in supply-chain management: A state-of-the-art literature review," 2017.
[17] S. Bhartiya and D. Mehrotra, "Challenges and recommendations to healthcare data exchange in an interoperable environment," *Electron. J. Heal. Informatics*, vol. 8, no. 2, p. 16, 2014.
[18] M. Kayaalp, "Patient privacy in the era of big data," *Balkan Med. J.*, vol. 35, no. 1, pp. 8–17, 2018.
[19] E. S. Dove and M. Phillips, "Privacy law, data sharing policies, and medical data: a comparative perspective," *Med. data Priv. Handb.*, pp. 639–678, 2015.
[20] L. Gostin, "Health care information and the protection of personal privacy: ethical and legal considerations," *Ann. Intern. Med.*, vol. 127, no. 8_Part_2, pp. 683–690, 1997.
[21] J. P. Woolley, "Trust and justice in big data analytics: Bringing the philosophical literature on trust to bear on the ethics of consent," *Philos. Technol.*, vol. 32, pp. 111–134, 2019.
[22] R. Reed-Berendt, E. S. Dove, M. Pareek, and U. S. C. Group, "The Ethical Implications of Big Data Research in Public Health:'Big Data Ethics by Design' in the UK-REACH Study," *Ethics Hum. Res.*, vol. 44, no. 1, pp. 2–17, 2022.
[23] A. A. Nafea, S. A. Alameri, R. R. Majeed, M. A. Khalaf, and M. M. AL-Ani, "A Short Review on Supervised Machine Learning and Deep Learning Techniques in Computer Vision," *Babylonian J. Mach. Learn.*, vol. 2024, pp. 48–55, 2024.
[24] T. Hulsen *et al.*, "From big data to precision medicine," *Front. Med.*, vol. 6, p. 34, 2019.
[25] V. Gligorijević, N. Malod-Dognin, and N. Pržulj, "Integrative methods for analyzing big data in precision medicine," *Proteomics*, vol. 16, no. 5, pp. 741–758, 2016.
[26] O. AlShorman, B. AlShorman, M. Alkhassaweneh, and F. Alkahtani, "A review of internet of medical things (IoMT)-based remote health monitoring through wearable sensors: a case study for diabetic patients," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 20, no. 1, pp. 414–422, 2020.
[27] R. Dwivedi, D. Mehrotra, and S. Chandra, "Potential of Internet of Medical Things (IoMT) applications in building a smart healthcare system: A systematic review," *J. oral Biol. craniofacial Res.*, vol. 12, no. 2, pp. 302–318, 2022.
[28] C. Dreisbach, T. A. Koleck, P. E. Bourne, and S. Bakken, "A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data," *Int. J. Med. Inform.*, vol. 125, pp. 37–46, 2019.
[29] T. Hao, Z. Huang, L. Liang, H. Weng, and B. Tang, "Health natural language processing: methodology development and applications," *JMIR Med. informatics*, vol. 9, no. 10, p. e23898, 2021.
[30] A. A. Nafea, N. Omar, and Z. M. Al-qfail, "Artificial Neural Network and Latent Semantic Analysis for Adverse Drug Reaction Detection," *Baghdad Sci. J.*, 2023.

[31] P. Saranya and P. Asha, "Survey on big data analytics in health care," in *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 2019, pp. 46–51.
[32] Z. Ahmed, K. Mohamed, S. Zeeshan, and X. Dong, "Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine," *Database*, vol. 2020, p. baaa010, 2020.
[33] R. Vatambeti, E. S. P. Krishna, M. G. Karthik, and V. K. Damera, "Securing the medical data using enhanced privacy preserving based blockchain technology in Internet of Things," *Cluster Comput.*, pp. 1–13, 2023.
[34] Z. K. Lu, X. Xiong, T. Lee, J. Wu, J. Yuan, and B. Jiang, "Big data and real-world data based cost-effectiveness studies and decision-making models: a systematic review and analysis," *Front. Pharmacol.*, vol. 12, p. 700012, 2021.