BIG DATA: APPLICATIONS AND CASE STUDY

Article · June 2024

CITATIONS
0

READS
479

3 authors:

Mubarak Altamimi
Karabük University
21 PUBLICATIONS **24** CITATIONS
SEE PROFILE

Nehad T. A. Ramaha
Karabük University
26 PUBLICATIONS **84** CITATIONS
SEE PROFILE

Ijibadejo Oluwasegun William
Karabük University
22 PUBLICATIONS **18** CITATIONS
SEE PROFILE

3rd International Conference on Engineering, Natural and Social Sciences

May 16-17, 2024 : Konya, Turkey

https://www.icensos.com/

BIG DATA: APPLICATIONS AND CASE STUDY

Mubarak Altamimi [*], Nehad Rammaha[2], Ijibadejo Oluwasegun William [3]

[1] *Department of Computer Engineering/Karabuk University, Karabuk. Almaafer Community College-Taiz, Yemen Karabuk University, Karabuk, Turkey*
[2] *Department of Computer Science and Engineering/ESCAE University, Porto Novo, Republic of Benin*

[*]*(Eng.mobaraktamimi@yahoo.com)*
(nehadramaha@karabuk.edu.tr)
(oluwasegunijibadejo@gmail.com)

Abstract – Big data is the term used to describe extraordinarily huge and intricate datasets that are too big or too complicated for conventional data analysis and processing techniques. Distributed computation, processing in parallel, machine learning, data mining, and other innovative technologies and methods are used in big data analysis to draw out relevant trends, patterns, and knowledge that can inform innovation, decision-making, and optimization across a range of fields and sectors. These applications are made to meet certain demands in the business or sector, streamline processes, promote decision-making, elevate customer loyalty, and provide a competitive edge over adversaries. The goal of this case study is to provide a thorough analysis of the field's applications, challenges, assets, and emerging trends.

Keywords – Big Data, Organization, Application Domains, Classification, Challenges.

I. OVERVIEW

Volume, velocity, and diversity are characteristics of very large and complicated datasets, which are referred to as big data. It includes enormous volumes of data that are unstructured, semi-structured, or organized and that are difficult to handle, process, or analyze with conventional data processing techniques [1]. Big data includes not just the sheer amount of data as well as the value and insights that may be obtained from it. Utilizing innovative techniques like data mining, machine learning, statistical analysis, and natural language processing, big data analysis aims to find patterns, trends, associations, and useful information that can spur innovation, decision-making, and optimization across a range of fields and sectors [2]. Big data is used by organizations to create new goods and services, enhance consumer experiences, increase operational efficiency, acquire a competitive edge, and spur corporate growth. Applications for it may be found in various industries, including manufacturing for supply chain, telecommunications, energy, retail, healthcare, business analytics, financial services, E-commerce, retail, smart cities, social media analytics, energy, governmental, and finance [2].

.

Fig. 1 Big data application fields[2].

Organizations that effectively use big data have an advantage over rivals. Every day, several Terabytes of data are uploaded by Facebook users. Social media data are utilized to create increasingly sophisticated analyses that try to extract additional value from user data [3].

THE MAIN 10 V CHARACTERISTICS OF BIG DATA

Big data is expanding in volume and complexity, particularly from new sources. These enormous data volumes are too big for conventional data processing techniques to manage. Even so, you may be able to use this enormous amount of information to address business problems that you previously could not resolve [4]. Everyone has to grasp the seven more crucial big data qualities, or the 10 V's.
Interestingly, these traits also begin with V, hence the following list of 10 V's of big data is provided:

- Volume: The term "big data" describes vast amounts of data that are larger than what can be processed by conventional data processing methods. Large-scale data management and analysis are required.
- Velocity: The speed at which big data is being created and gathered is unparalleled. Real-time data streams that need quick processing and analysis are included, including financial transactions, sensor data, and social media updates.
- Variety: Unstructured data, semi-structured data, and structured data are among the many forms and types of data that make up big data. It contains recordings of social media postings, text, photos, videos, audio, and more.
- Veracity: Data quality and dependability issues are frequently present with big data. Because of its many sources and extensive collection, it can have mistakes, inconsistencies, and inaccuracies.
- Value: Big Data can yield important insights and information. Organizations can find associations, associations, and trends that can spur creative thinking and well-informed decision-making by examining large and diverse information.
- Variability: Variability may be observed in the volume, velocity, and variety of big data. With variations in the speeds and kinds of data creation, the data flow might be unpredictable.
- Complexity: Because of its many different sources, forms, and structures, big data can be difficult to handle and evaluate. To obtain significant insights, it frequently takes sophisticated instruments, methods, and technology.
- Accessibility: Big data is frequently dispersed over several platforms and sources, such as external sources, cloud computing services, and on-premises systems. Getting access to and combining data from several sources can be difficult.
- Privacy and Security: Big data creates security and privacy issues. Due to the large volume of private and sensitive data involved, data security and privacy protection become crucial factors.
- Scalability: Big data solutions need to be scalable to handle the increasing amount, velocity, and diversity of data. The infrastructure and systems must be able to manage the rising demands for data processing and storage [4].

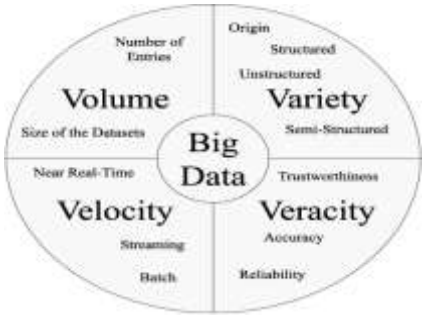

Fig. 2 Big data application Vs[4].

II. BIG DATA APPLICATION FIELDS

Many different sectors are using big data applications to better decision-making, acquire insightful information, and streamline corporate procedures [18]. The following are some noteworthy uses of big data:

Fig. 5 Fields applications of big data [18]

Banking and Securities

In the banking and securities industries, big data is being used more and more for transaction visibility, risk analytics, customer data transformation, and securities fraud detection. It is anticipated that the exponential expansion of data would assist in identifying illegal activity and reducing dangers [18].

Educations Industry

Large volumes of data are being used by the education industry to improve functionality and efficiency. With more than 38,000 students, the University of Alabama has the potential to enhance operations, recruiting, and retention efforts by using analytics and data visualizations to uncover patterns in student changes.

Big Data Application In The Healthcare Industry

Technologies like Industry 4.0 and Big Data emphasize data exchange throughout the whole value chain, necessitating new techniques for processing data. Industry 4.0, which integrates software, technology, and processes, is essential for smart factories and hospitals [20]. By lowering treatment costs, predicting epidemics, detecting illnesses early, assisting in the prevention of avoidable diseases, and enabling effective treatments, big data has greatly benefited healthcare. Patients may keep and access their medical information on their devices thanks to real-time data transmission via wearable technologies and sensors, such as those included in Apple's ResearchKit, CareKit, and HealthKit [20].

APPLICATION OF BIG DATA IN THE GOVERNMENT SECTOR

Every day, governments handle enormous volumes of data, including surveys of the population, growth, energy resources, and geography. Big data has several advantages and is becoming more and more important in the public sector [21].

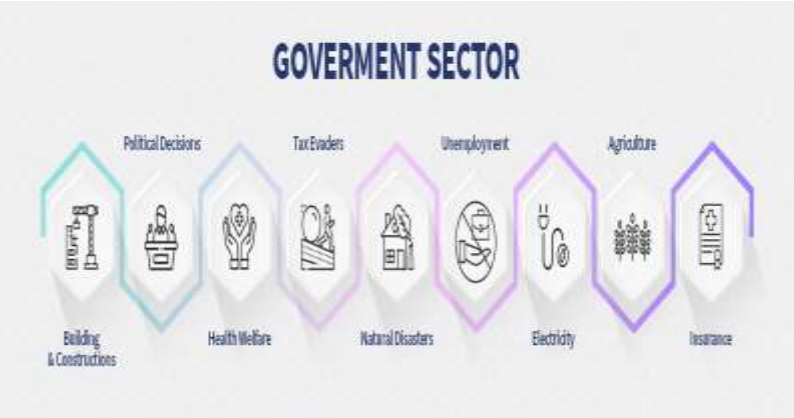

Figure 8. Big data applications in the Government Sector [21].

APPLICATION OF BIG DATA IN MEDIA AND ENTERTAINMENT

Personalized music suggestions and stream scheduling are made possible by the growing prevalence of big data in the media and entertainment industry. With the use of optimized digital media distribution methods and insights from client testimonials, Spotify leverages big data analytics to collect consumer data [22].

BIG DATA APPLICATION IN WEATHER PATTERNS

IBM Deep Thunder enhances natural catastrophe forecasts, anticipates weather patterns, and predicts power outages using big data and high-performance computing. Numerous applications of this data exist, such as forecasting water supply, disaster patterns, weather prediction, studies on global warming, and emergency planning[23].

BIG DATA IN THE TRANSPORTATION INDUSTRY

Big data's assistance with route planning, traffic management, and control has greatly increased the efficiency of transportation. Users may locate routes with the least amount of traffic by using its real-time assessment of traffic patterns and congestion. To lower accident rates and raise traffic safety, predictive analysis, and real-time big data processing may also be used to pinpoint accident-prone locations[24].

BIG DATA APPLICATION IN THE BANKING SECTOR

Data volume in the financial sector is increasing quickly; by year's end, it is expected to have expanded by 700 percent, which is good news for anti-money laundering initiatives like SAS AML [25].

Fig.10 Big data applications in the Banking Sector [25].

BIG DATA IN MARKETING

While traditional marketing strategies relied on one-on-one interactions with clients and surveys, digital marketing arose with the advent of big data and the internet. Big data helps organisations manage campaigns, increase click-through rates, and improve products by rapidly learning the preferences of millions of customers. Based on worldwide customer purchase habits, Amazon leverages this data to offer new promotions and specials [26].

Fig.11 Big data applications in Marketing [26].

BIG DATA IN BUSINESS INSIGHTS

Sixty percent of the data collected by companies and social media platforms is unstructured, and big data is being used to produce business insights. Revenue, customer happiness, and product development are all addressed by this data. Organizations such as Netflix are utilizing technology to effectively manage and examine this information, augmenting their comprehension of user conduct [27].

BIG DATA IN THE SPACE SECTOR

Every day, space organisations get enormous amounts of data from rovers, satellites, probes, and observations. Before rocket launches, this data is analysed to model flight trajectories taking into account weather, payload, and orbital position [28].

BIG DATA APPLICATION CHALLENGES

Big Data's incompleteness, fuzziness, complexity, softness, speed of arrival, and vastness pose obstacles despite its magnitude. It is distinct in that it is relational, varied, qualitative, and incomplete. For instance, precise tsunami motion models combined with well-structured data can yield reliable information when

anticipating a tsunami. However, there are serious mathematical and scientific issues when attempting to combine this enormous amount of disorganised data with more precise data. Handling, visualising, predicting, modelling, comprehending, and regulating systems with care are necessary to meet these problems [29]. The goal of the research is to find patterns in massive data sets, a human inclination to make sense of the confusing reality. Automation needs sophisticated mathematical algorithms because of the abundance of data. The primary obstacles are gathering, storing, finding, distributing, analysing, and displaying data [30].

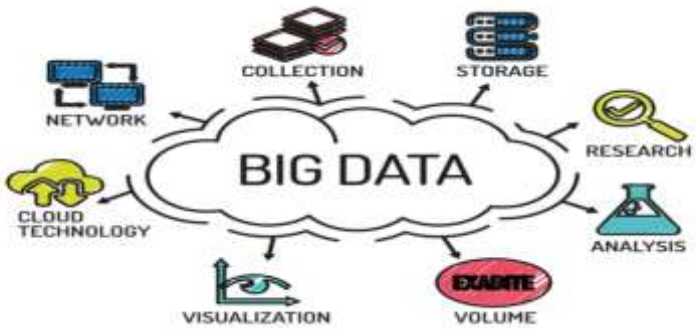

Fig. 12 A main challenge of big data application [30].

DATA CAPTURE

Efficient collection and integration of data is a difficulty when dealing with Big Data that originates from several sources. Efficient data collecting, pre-processing, data transmission, and automated metadata synthesis can provide businesses with a competitive edge [32].

DATA STORAGE

To handle Big Data, large-scale distributed systems need creative data management strategies. Because of its schema-free architecture, simplicity of replication, and ability to store enormous amounts of data, NoSQL databases like Apache Cassandra, Apache HBase, and Project Voldemort are at the forefront of Big Data technology [38].

DATA SEARCH

Timeliness, comprehensiveness, and correctness of data are essential for making decisions. Complex analytical SQL queries require query optimization, and physical design and query optimization are crucial infrastructure elements even in parallel systems[33].

DATA SHARING

Due to legal paperwork requirements and public repositories, big data endeavors frequently encounter difficulties when attempting to obtain data from other sources. For prompt decision-making, businesses must supply thorough, accurate, and timely data [34].

DATA ANALYSES

In government initiatives, the scientific and technological domains, and business endeavours, big data analytics is essential. Distributed methods such as cloud computing, MapReduce, and MPI may tackle problems like software framework performance optimisation and distributed data management ecosystems [35]. Big data poses analytical difficulties, such as handling an enormous amount of data and locating

important data points. Semi-structured, unstructured, and structured data are the three forms of data that are examined. One can choose to incorporate a lot of data or decide in advance which big data is relevant [36].

DATA VISUALIZATION

The new language of business is data visualisation, which makes it possible to communicate complicated data accurately and efficiently, helps companies make wise decisions, and shows the connections between daily operations.

Fig.13 Data visualized benefits [37].

Visual trends improve learning, accessibility, and clarity of data, which facilitates comprehension and memory of knowledge. They also offer a novel viewpoint and support the creation of successful tactics[37].

Fig. 14 A data properties[37].

Given that every sort of data has a different pace, size, and variety, visualization is essential for processing large data. Word clouds and symbol maps are two examples of modern data visualization techniques that are intended for unstructured and semistructured data [37].

CONCLUSION

We have seen a few real-world uses for big data in this blog. Given all of its applications, it makes sense why big data is the subject of so much hype. Big data significance is not determined by an organization's ability to gather vast amounts of data, but rather by how it uses the information it does collect. Big data can be easily and effectively analyzed with the help of big data solutions. Big data is becoming more and more important in several industries, and this trend is predicted to continue. This study examined big data research articles from prior years and categorized the works into twelve application areas: Biochemistry, Genetics, Molecular Biology, Computer Science, Engineering, Mathematics, Business, Management, and Accounting, Physics and Astronomy, Social Sciences, Material Sciences, Medicine, Decision Sciences, Multidisciplinary, and Arts and Humanities. Additionally, Understanding the value of big data is being utilized in many fields of knowledge, making analysis the most significant problem for big data research. Most industrial operations find it extremely difficult to implement big data analytics in real-time and in the

Internet of Things process, even though several research papers in the field of big data applications have attempted to recognize and clarify the challenges in industrial or supply chain big data applications.

REFERENCES

[1]. Rodríguez-Mazahua, L., Rodríguez-Enríquez, C. A., Sánchez-Cervantes, J. L., Cervantes, J., García-Alcaraz, J. L., & Alor-Hernández, G. (2019). A general perspective of Big Data: applications, tools, challenges and trends. *The Journal of Supercomputing*, *72*, 3073-3113.

[2]. Yu, S., & Guo, S. (Eds.). (2016). *Big data concepts, theories, and applications*. Springer.

[3]. Khine, P. P., & Shun, W. Z. (2017). Big data for organizations: a review. *Journal of Computer and Communications*, *5*(3), 40-48.

[4]. Bansal, B., Jenipher, V. N., Jain, R., Dilip, R., Kumbhkar, M., Pramanik, S., ... & Gupta, A. (2022). Big Data Architecture for Network Security. *Cyber Security and Network Security*, 233-267.

[5]. Boyapati, S., Swarna, S. R., Dutt, V., & Vyas, N. (2020, December). Big data approach for medical data classification: A review study. In *2020 3rd international conference on intelligent sustainable systems (ICISS)* (pp. 762-766). IEEE.

[6]. Jeong, S., & Law, K. H. (2022). Data management technologies for infrastructure monitoring. In *Sensor Technologies for Civil Infrastructures* (pp. 397-426). Woodhead Publishing.

[7]. Dass, S., & Prabhu, J. (2022). Amelioration of Big Data Analytics by Employing Big Data Tools and Techniques. In *Research Anthology on Big Data Analytics, Architectures, and Applications* (pp. 1527-1548). IGI Global.

[8]. Pentzold, C., Brantner, C., & Foelsche, L. (2019). Imagining big data: Illustrations of "big data" in US news articles, 2010–2016. *New Media & Society*, *21*(1), 139-167.

[9]. Hendawi, A., Gupta, J., Liu, J., Teredesai, A., Ramakrishnan, N., Shah, M., ... & Ali, M. (2019). Benchmarking large-scale data management for Internet of Things. *The Journal of Supercomputing*, *75*, 8207-8230.

[10]. Pijanowski, B. C., Tayyebi, A., Doucette, J., Pekin, B. K., Braun, D., & Plourde, J. (2014). A big data urban growth simulation at a national scale: configuring the GIS and neural network based land transformation model to run in a high performance computing (HPC) environment. *Environmental Modelling & Software*, *51*, 250-268.

[11]. Sharma, D., & Hasija, Y. (2021, July). Big Data: A Boom in biomedical sciences. In *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1-5). IEEE.

[12]. Wang, B., Wang, Y., Yan, F., & Zhao, W. (2022). Safety intelligence toward safety management in a big-data environment: A general model and its application in urban safety management. *Safety science*, *154*, 105840.

[13]. Tripathi, M. K., Kumar, R., & Tripathi, R. (2020). Big-data driven approaches in materials science: A survey. *Materials Today: Proceedings*, *26*, 1245-1249.

[14]. Kraus, S., Schiavone, F., Pluzhnikova, A., & Invernizzi, A. C. (2021). Digital transformation in healthcare: Analyzing the current state-of-research. *Journal of Business Research*, *123*, 557-567.

[15]. Demirkan, H., & Delen, D. (2013). Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. *Decision Support Systems*, *55*(1), 412-421.

[16]. Sánchez-Corcuera, R., Nuñez-Marcos, A., Sesma-Solance, J., Bilbao-Jayo, A., Mulero, R., Zulaika, U., ... & Almeida, A. (2019). Smart cities survey: Technologies, application domains and challenges for the cities of the future. *International Journal of Distributed Sensor Networks*, *15*(6), 1550147719853984.

[17]. Demchenko, Y., de Laat, C., & Los, W. (2021). Future scientific data infrastructure: towards platform research infrastructure as a service (PRIAAS). In *The 2020 International Conference on High Performance Computing Simulation (HPCS 2020), The 18th Annual Meeting, Barcelona, Spain*.

[18]. Lu, K., Liu, J., Zhou, X., & Han, B. (2020). A review of big data applications in urban transit systems. *IEEE Transactions on Intelligent Transportation Systems*, *22*(5), 2535-2552.

[19]. Baig, M. I., Shuib, L., & Yadegaridehkordi, E. (2020). Big data in education: a state of the art, limitations, and future research directions. *International Journal of Educational Technology in Higher Education*, *17*(1), 1-23.

[20]. Karatas, M., Eriskin, L., Deveci, M., Pamucar, D., & Garg, H. (2022). Big Data for Healthcare Industry 4.0: Applications, challenges and future perspectives. *Expert Systems with Applications*, *200*, 116912.

[21]. Aggarwal, A. K. (2019). Opportunities and challenges of big data in public sector. *Web services: Concepts, methodologies, tools, and applications*, 1749-1761.

[22]. Smith, M. D., & Telang, R. (2016). *Streaming, sharing, stealing: Big data and the future of entertainment*. Mit Press.

[23]. Fathi, M., Haghi Kashani, M., Jameii, S. M., & Mahdipour, E. (2022). Big data analytics in weather forecasting: A systematic review. *Archives of Computational Methods in Engineering*, *29*(2), 1247-1275.

[24]. Welch, T. F., & Widita, A. (2019). Big data in public transportation: a review of sources and methods. *Transport reviews*, *39*(6), 795-818.

[25]. Nobanee, H., Dilshad, M. N., Al Dhanhani, M., Al Neyadi, M., Al Qubaisi, S., & Al Shamsi, S. (2021). Big data applications the banking sector: A bibliometric analysis approach. *Sage Open*, *11*(4), 21582440211067234.

[26]. Lies, J. (2019). Marketing intelligence and big data: Digital marketing techniques on their way to becoming social engineering techniques in marketing.

[27]. Saidali, J., Rahich, H., Tabaa, Y., & Medouri, A. (2019). The combination between big data and marketing strategies to gain valuable business insights for better production success. *Procedia Manufacturing*, *32*, 1017-1023.

[28]. Gawankar, S. A., Gunasekaran, A., & Kamble, S. (2020). A study on investments in the big data-driven supply chain, performance measures and organisational performance in Indian retail 4.0 context. *International journal of production research*, *58*(5), 1574-1593.

[29]. Talebkhah, M., Sali, A., Marjani, M., Gordan, M., Hashim, S. J., & Rokhani, F. Z. (2021). IoT and big data applications in smart cities: recent advances, challenges, and critical issues. *IEEE Access*, *9*, 55465-55484.

[30]. Patel, A. B., Birla, M., & Nair, U. (2012, December). Addressing big data problem using Hadoop and Map Reduce. In *2012 Nirma University International Conference on Engineering (NUiCONE)* (pp. 1-5). IEEE.

[31]. Wüest, R. O., Zimmermann, N. E., Zurell, D., Alexander, J. M., Fritz, S. A., Hof, C., ... & Karger, D. N. (2020). Macroecology in the age of Big Data–Where to go from here?. *Journal of Biogeography*, *47*(1), 1-12.

[32]. ur Rehman, M. H., Yaqoob, I., Salah, K., Imran, M., Jayaraman, P. P., & Perera, C. (2019). The role of big data analytics in industrial Internet of Things. *Future Generation Computer Systems*, *99*, 247-259.

[33]. Arfat, Y., Usman, S., Mehmood, R., & Katib, I. (2020). Big data for smart infrastructure design: Opportunities and challenges. *Smart Infrastructure and Applications: Foundations for Smarter Cities and Societies*, 491-518.

[34]. Qi, W., Sun, M., & Hosseini, S. R. A. (2023). Facilitating big-data management in modern business and organizations using cloud computing: a comprehensive study. *Journal of Management & Organization*, *29*(4), 697-723.

[35]. Pevec, D., Vdovic, H., Gace, I., Sabolic, M., Babic, J., & Podobnik, V. (2019, July). Distributed data platform for automotive industry: A robust solution for tackling big challenges of big data in transportation science. In *2019 15th International Conference on Telecommunications (ConTEL)* (pp. 1-8). IEEE.

[36]. Zgurovsky, M. Z., & Zaychenko, Y. P. (2020). *Big data: conceptual analysis and applications*. Springer International Publishing.

[37]. Lv, Z., & Qiao, L. (2020). Analysis of healthcare big data. *Future Generation Computer Systems*, *109*, 103-110.

[38]. Gandhi, P., & Pruthi, J. (2020). Data visualization techniques: traditional data to big data. *Data Visualization: Trends and Challenges Toward Multidisciplinary Perception*, 53-74.