

A Research Insights of Big Data Analytics: Tools and Techniques, Issues and Challenges



D. Bhuvaneswari and J. Gnana Jayanthi

Abstract Big Data Analytics (BDA) supports organizations to recognize new opportunities by analyzing with their existing data which in turn, transforms the business into smarter business efficiently and effectively, with higher profits. Big Data Analytics plays a vital role in perspectives of (i) Data Science, (ii) Business, (iii) Real-time Usability, and (iv) Job Market. Big Data Analytics inspects large amounts of data; and scrutinizes them to obtain certain significant values from the uncovered hidden patterns, correlations and other perceptions. This paper highlights the significances of Big Data, Big Data Analytics, tools and techniques used Data management in Big Data Analytics. This paper also focuses to kindle research interests among the research scholars to march toward the blooming fields in Big Data.

Keywords Big data · Big data analytics · Characteristics · Challenges · Issues · Techniques · Tools

1 Introduction

Big Data is the information collected from the Internet-enabled services, social media, and other similar sources. Kaisler et al. in 2013, stated that Big Data has the volumes in the range of Exabyte, measured in the magnitude of 10^{18} [1]. Cisco, in its executive summary on “Cisco Visual Networking Index: Forecast and Methodology, 2016–2021”, concluded that (i) Annual global IP traffic will reach 3.3 ZB; (1000 Exabytes [EB]) by 2021, i.e., Global Internet traffic in 2021; (ii) Global Internet traffic will cross the range beyond 3.3 ZB annually, and that it will by 2021 [2]. These statistics prove that the generation of digital data is enormously increasing day by day.

Organizations, small or big, are struggling with problems of big data generation, storage, and maintenance. On the contradictory view, with a huge amount of data, those organizations can also find new opportunities and also obtain new insights

D. Bhuvaneswari (✉) · J. Gnana Jayanthi
PG and Research Department of Computer Science, Rajah Serfoji Govt. College (Autonomous),
(Affiliated to Bharathidasan University), Thanjavur, Tamil Nadu, India

by performing some analytics on big data to make efficient and effective decisions, which in turn transforms the business with smart moves. These concerns lead to the development of new techniques, methods and algorithms and Big Data Analytics is one among them. Big Data Analytics is succeeding in Business to Customer (B2C) applications.

However, there exist huge requirements, issues and challenges in the field of Big Data Analytics. This paper is aimed to present an overview of Big Data, Big Data Analytics, architectural framework, types of analytics, storage the mechanism, and tools, techniques, which are used to handle big data.

The remaining portion of this paper is systemized as follows. Section 2 dictates the Big Data overview. Section 3 explains the overview of Big Data Analytics. Section 4 discusses the research challenges and issues. Finally, Sect. 5 determines and tells about future work.

2 Big Data Overview

This section explains the big data, types of big data, and the characteristics of big data.

2.1 Big Data

Utilizing big data as big data services is an innovation. Thomas Davenport et al. in 2011, from their research survey, have concluded that Analytics 3.0 comprises (i) Combination of multiple data types; (ii) Some new methodologies for data integration; (iii) speedy processing of data with new technologies; and (iv) the integration of analytics with operational and decision processes [3].

2.2 Big Data Types

Depending on the sources of available data, Big Data is classified as (i) Structured Data (e.g., relational data), (ii) Semi-Structured Data (e.g., XML data), and (iii) Unstructured Data (e.g., Word, PDF, Text, Media Logs) [4].

2.3 Big Data Characteristics

Based on the characteristics, data are also classified. Richa Gupta in 2014, stated the three major characteristics of Big Data (i) Volume (recently have raised to

Yottabytes); (ii) Velocity (speed of data which is being generated, produced, created, or refreshed); (iii) Variety (different data types collected from sources) [5]. In 2016, Rajkumar et al. mentioned two more attributes (iv) Variability (inconsistencies in data changes and sources); (v) Veracity (reliability of the data source) [6]. According to Pankaj Goel et al. in 2017, Big Data characteristics has 7 V's in the list (vi) Validity (accuracy of data); (vii) Vulnerability (security attacks) [7] (Fig. 1).

Panimalar et al. in [8], included additional three characteristics (viii) Volatility (data currency and availability); (ix) Visualization (readability, understandability, accessibility); (x) Value (importance of data).

In the same manner, Big Data analytics consists of 6 C's for factory work and Cyber-physical systems. Data management can become a very complex process where data comes from multiple sources. These data need the following steps (i) Connection (sensor and networks); (ii) Cloud (computing and data on demand); (iii) Cyber (model and memory); (iv) Content/context (meaning and correlation); (v) Community (sharing and collaboration) (vi) Customization (personalization and value) [9].

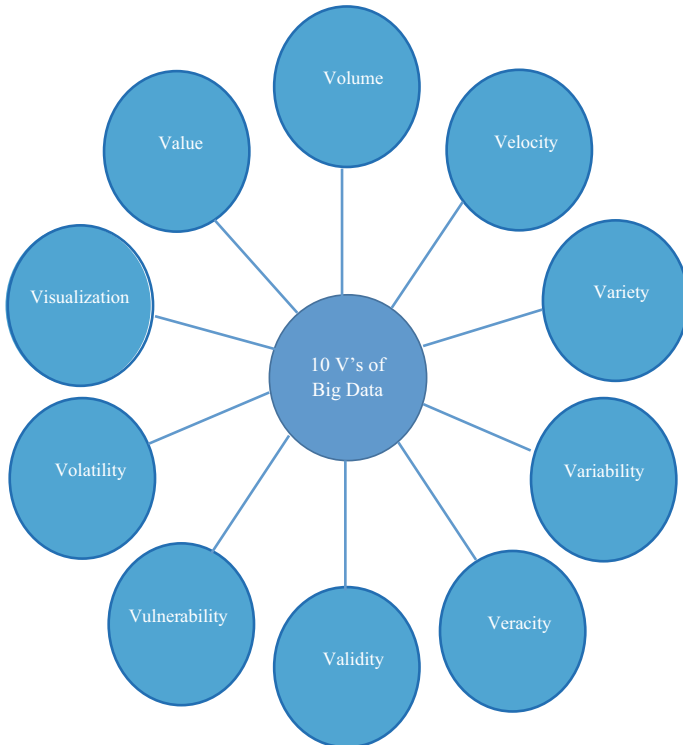
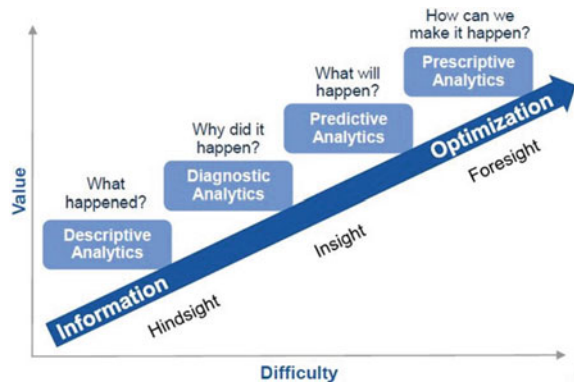


Fig. 1 Characteristics

Fig. 2 Big data analytics

3 Big Data Analytics Overview

This section summarizes the importance of Big Data Analytics, types of Big Data Analytics, its architecture, some techniques which are used to make decisions, tools to implement, and storage mechanisms, etc.

3.1 Importance of Big Data Analytics

Big data analytics assists organizations to assume their data and involve it to invent new opportunities for intelligent business. Many organizations can gain high profits and also customer satisfaction with the help of analytical tools. Big Data Analytics also affect government sectors in education, transport, and health care. Unemployment can be reduced by predicting the desired task. World Health Organization uses Big Data Analytics to get a detailed report about the children who get quality education from their state [10] (Fig. 2).

3.2 Types of Big Data Analytics

Big Data Analytics has the facility to make an accurate and quick decision in lively conditions and ambiguous environment. Significant benefits of Big Data Analytics are Cost reduction, Faster and better decision making, New products and services. There are four categories of analytics, (i) Descriptive analytics, (ii) Diagnostic analytics, (iii) Predictive analytics, and (iv) Prescriptive analytics [11].

- **Descriptive Analytics:** This is a simple method of analytics converting the huge size of data into small bytes. It collects and analyzes the data to realize what

happened in the past. This helps us to understand past behavior and learn how it can affect future outcomes. The outcome is monitored through emails or dashboard.

- **Diagnostic Analytics:** Diagnostic Analytics is based on the questions that framed on the existing data such as how what and why it happened. This type of analytics is used to discover any hidden patterns and identify the factors which are causing effects directly or indirectly. This type of analytics is used in social media to find out user's opinion from history.
- **Predictive Analytics:** This is the best advanced analytics which studies the historical data, identifying insights, trends and patterns and predict "what might happen next" by using statistical models and forecasts techniques with the probabilistic nature. It discloses activities and recommendations of what step should be taken in future.
- **Prescriptive analytics:** This is the most advanced and promising analytics which helps to make truly data-driven decisions through simulation and optimization. Prescriptive analytics reveals the action and make the situation in a focused way.

3.3 Architectural Framework

The need for architecture is to store, process, and transform the unstructured data for analytics and to produce reporting. Big Data Analytics handle unbounded streams of data to capture, process and analyze. Figure 3, represents the logical components of the Big Data Analytics process architecture. The components of Big Data Analytics architecture comprises of (i) Data Source, (ii) Data Storage and Transformation, (iii) Data Processing, (iv) Analytical data storage, (v) Analysis and reporting.

- **Data Source:** Web and public data: Data from Blogs, Twitter, Facebook, LinkedIn, i.e., Clickstream data.
- **Machine to machine data:** Data from wearable and sensor devices. Big transaction data: Billing records in health care and insurance. Biometric data: Fingerprints, calligraphy, x-ray, medical images and other data. Human-generated data: Electronic Medical Records, physician notes, email, and paper documents are unstructured and semi-structured data. Data from various sources are transmitted at high speed and also in different formats. It is tedious to rank and characterize these data for immediate use. Data ingestion process identified whether it is structured or unstructured [12]. The 'raw' data needs to be processed or transformed.
- **Data Storage and Transformation:** Data storage is responsible for the management of large-scale storage system. The distributed file system is used to store high volumes of large files in various formats and serve as a repository. Commonly, this method of storage is called Data Lake [13]. Data are cleaned and are ready to use stage by using Intelligent Engineering tools, Sqoop, and Log management tools [14].
- **Data Processing:** Data process in two ways (i) Batch processing and (ii) Stream processing. (i) Batch Processing: In batch processing model, data are stored first and then analyzed. Batch processing is the high-volume nature of big data process.

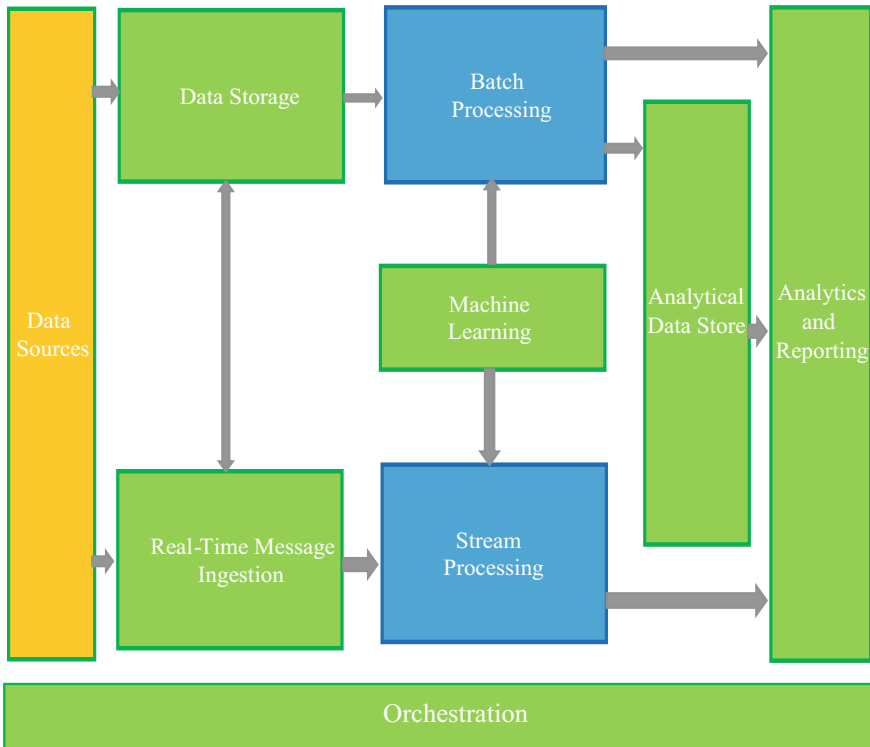


Fig. 3 Conceptual architecture of big data analytics

Hive, Pig and Spark are used in batch processing. Stream Processing: The stream processing model analyzes the data immediately possible to derive its results. There are unceasing arrival of data with high speed and carries enormous volume. The stream processing paradigm is widely used for online applications. Real-time message ingestion is used to capture and store real-time messages for stream processing. This message ingestion works as a buffer for messages, and support scale-out processing, consistent delivery, and also message queuing semantics. This process of streaming is generally called as stream buffering. A major challenge in stream processing is to work on the data quickly, presenting the data in a real-time dashboard and also generating alerts [15].

- Analytical Data Storage: Big Data Analytics is aimed to organize the data for analysis and report. The processed data is organized in a specific format and are queried by analytical tools [16]. Hadoop and MapReduce are the current popular choice for implementation in big data infrastructure.
- Analysis and Reporting: Analyzed data are presented in the form of Query reports. That is readable, accessible and Visualization charts. Visualization is a predominant theme so that many techniques and technologies are developed and adapted for reporting. It is a presentation of pictorial or graphical format [17].

3.4 *Techniques*

Here, are some techniques to overcome the challenges of big data and make decision making [18–21].

- **Classification:** Classification classifies the dependent variable from one or more predictor variables based on measurements. It derives a new model to determine the data class or concept. It predicts the class label for the objects of the unknown class.
- **Regression:** Regression derives a model that maps values from predictors. The dependent variable is predicted by estimating the relationship among variable.
- **Nearest Neighbor:** Nearest neighbor is used to predict the values from the predicted values of the record which is the nearest record that needs to predict.
- **Clustering:** Clustering is used even if the training data is not available and also when the class data are not known previously. From the data, collection clustering is used to find a grouping of different data.

3.5 *Big Data Analytics Tools*

Many open-source tools are available to analyze Big Data with emphasis, which focuses on both batch processing and stream processing. Conversely, Big Data Analytics tools are complex, programming intensive, and requires a variety of skills. Some of the most used tools are overviewed here in this literature.

- Hadoop is a most conventional product from Apache Software Foundation for batch processing. Many concerns use Hadoop because, it supports voluminous data sets, best handing out ability and offers updates and enhancements [22].
- MapReduce tool is used for handling large datasets. It is implemented with two steps, i.e., Map and Reduce like divide and conquer method. High throughput and fault-tolerant storage are the main advantages of MapReduce [23].
- A mahout is a tool for elevating big challenges. It provides machine learning techniques for decision making in Big Data Analytics. Clustering, classification, pattern mining, regression, and dimensionality reduction are the algorithms used in mahout [24].
- The storm has a superior stream processing capability in real time. It combines with Apache Sliders tools to manage and secure the data. It is a very fast processing tool which is used in dashboards, cyber analytics [25].
- Spark is an open-source processing framework used for fast and classy analytics. It has a machine learning library. Resilient Distributed Dataset (RDD) stores the data and also responsible for fault tolerance and eliminates duplication. In Spark, programs can run in Java, R, Python, or Scala [26].
- Dryad is a suitable programming design to handle huge context bases on dataflow graph by implementing parallel and distributed programs. It provides many

functionalities, i.e., generating job graph, machine scheduling, transition failure handling, and invoking user-defined policies [27].

- Apache Drill uses a distributed file system to store and also uses MapReduce to perform batch processing. The drill has the ability to support different data sources, data formats and query languages. It has the facility to work from 10,000 and more servers and can process petabytes of data and trillions of records in seconds [28].

3.6 Storage Mechanism

Traditional data handling techniques may not support big data due to its gargantuan bulk of data. It cannot handle the scalability and challenges of Big Data [29–31].

- Relational database management system uses ACID (Atomicity, Consistency, Isolation, Durability) to ensure data consistency. Relational databases cannot classify unstructured data. PostgreSQL is an open-source relational database management system.
- The non-relational database uses BASE mechanism. The BASE stands for Basically Available Soft State and Eventual Consistency. NoSQL systems provide data consistency. NoSQL is the most popular in Big Data storage.
- Key-value pair databases are created by a basic data model. In this, data are stored with the key values. Each key has a unique value. Customers can input the query with the key value. Key-value pair databases do not require any schema (like RDBMS). It employs key-value pair model. Ex. Cassandra, Azure, LevelDB and Riak.
- Document database has two types, (i) Collection and (ii) Description. The collection is composed of documents which consist of fields. A description is composed of documents which consist of fields, attachments and description of the document in the form of metadata. Ex. MongoDB, CouchDB.
- In the column-oriented database, the data are stored across rows. Data also process according to column. Rows and columns are split into several nodes to understand the expandability. HBase is a columnar database. It is a cloned version of Google's Bigtable programmed in Java.
- Graph database contains the data in a graph structure and allows queries on such data and the queries are implemented efficiently by using graph algorithms. Neo4J is a widely used graph database. It is scalable and simple to design because of node-relationship properties.
- Everyone interacts with spatial data in day to day life. Global Positioning System (GPS) is used for directions and to know the locations. Geographical data are stored in a geodatabase. Point, line and polygon are the characteristics of the object. Objects of Geographical Information System (GIS) are identified by the latitude and longitude values. PostGIS is the spatial data database that contains mapping information, i.e., GIS.

4 Research Challenges and Issues

This section discusses the challenges and issues faced by Big Data Analytics. Big Data Analytics is gaining so much attention these days but there are a number of research problems that still need to be addressed.

- **Scalability:** Most important characteristics of big data that describes the capacity to cope. Datasets are increasing in size day by day because of mobility trends and Internet of Things (IoT). Data can be scale and/or extend in future. The growth of data is much faster than the processor speeds. This issue leads to Big Data toward cloud computing and parallel computing for storage and computing [32].
- **Data heterogeneity:** In recent, 80% of data are in an unstructured format. It embraced various types of data which has been generated daily through social media, sharing of files, transfers of fax, chats, emails, messages and a lot more [33].
- **Data usability:** It has two main features. Increase the availability of dataset and offering the user a clean understanding of data. Provide a mapping between the data and corresponding analytics.
- **Data inconsistency:** Data changes from time to time, and these underlying data can reduce performance. This is the major issue faced by analytics. Big Data Analytics tools are calibrated to avoid replications in data sources.
- **Accuracy:** Analytics has the responsibility to give accurate result without missing the goal. Before processing the level of data can be measured. Big Data accuracy strongly depends on the requirements of analytics. If the requirements are detailed then the accuracy is higher.
- **Trust and provenance:** In Big Data management data provenance is fundamental for the trustworthy of Big Data Analytics. Data quality and data origin provide better accuracy of the result and increase the trust of the end-user [34].
- **Processing:** Processing of Exa and Zeta bytes of data is still a ridiculous task. Now many advanced indexing schemes and processing methods are used to improve the processing speed. Even though, more efficient parallel processing algorithms are needed for high response and actionable information [35].
- **Management:** Digital representation of data uses the modified methodology for data collection. Big Data needs special multidimensional tools to provide high performance else the outcomes may be undesirable. Organizing of data includes methods like cleaning, transforming, elucidation, dimension reduction, justification etc. It is very hard for business organizations to analyze the big data in unmanaged form. New research techniques must be used to fill the crack [36].
- **Security:** Security plays a major role in hindering patterns depends on data outsourcing. Data integrity, availability, and confidentiality are the major issues related to the protection of data. Now, these issues are related to the effect of an entirely outsourced structure. Speed of technology and volume of information makes it more difficult. Big Data security is quite complicated and still not thoroughly known. Preventing users against attacks is major challenge of insecurity [37].

- Individual's privacy details include name, signature, address and telephone number, date of birth and commentary or opinion about a person. To protect personal privacy, everyone should limit the amount of information revealed to others. Organizations and government accumulate individual's information and they should prohibit others to access such data. Information privacy includes national or traditional origin, political attitudes, membership and association, spiritual beliefs, business affiliation, trade union, criminal record, and health information. Privacy issues breaches in the following areas (i) Social Media, (ii) Location-Based Service or Application, (iii) Cloud Storage, (iv) Government Web sites, (v) Science and Technology Research. Big data facilitates organizations, to analyze, accurate, manage and process the data with high speed. Need protection of privacy either at an individual level and/or at a corporate/organization level [38].

In recent years, Big Data Analytics has crossed the threshold in several domains like recommender systems, business, retail, optimizing cities, health care, and telecommunication. Many organizations also improve the performance of the company. Even though there exist some issues in these fields.

- Recommender Systems work based on the concept of preference analytics. E-commerce web sites are the basic component of recommender systems. The goal is to provide the right items to the right users from their past likes and dislikes. Give suggestions to buy products/to watch movie depends on ratings and opinions. User's opinions may change from time to time that makes more items to a user. Data sparsity is the issue in recommender system when a user has the least number of friends/contacts/likes [39].
- Big Data Analytics extracts business intelligence by analyzing the data and make a better decision for good business growth [40]. It supports to recognize the customer needs, products demand and also provide better business choice.
- Smart devices and machines receive the innumerable amount of data through the Internet which generates Internet of Things (IoT). In future, the creation of data, communication technology, and network are the promising issue. Knowledge acquisition from streaming is the significant issue and data extraction can be done by machine learning techniques in Big Data Analytics [41].
- Big Data analytics are perpetually recognized by worldwide organizations. Customer purchase behavior patterns are predicted by most companies. To achieve this, the data is to be visualized efficiently, i.e., like third eye [42].
- In order to make optimizing cities, real-time traffic information, social media, and weather data have been used to progress the aspects of our cities and countries. Recently, most of the cities are currently piloting with Big Data Analytics [43]. This leads our city to become a Smart City.
- In the health care arena, the condition of the patient is analyzed with respect to various medicines and drugs for a specialist to make choices with the historical data for future health care [44]. Use of Big Data Analytics in health care has many benefits like prevention of diseases, cost reduction, better decision making, and quality of care. Data produced from intensive care units for a patient are large.

Health care system records patients streaming data from the ICU and make a decision and gives alert to the caretaker.

- In telecommunication huge amount of high quality, data are generated frequently and these customer base data are operated by a highly competitive environment. Prevents impact from fraud, and knowledgeable customer location and travel patterns to support real-time promotions, advertising and up-sell of services [45].

5 Conclusion and Future Work

Big Data Analytics gets a lot of attention in the marketing realm since it transforms the market space forever, and helps to make large-scale data-driven decisions. A conceptual view of Big Data Analytics tools, techniques and applications are briefed. This paper presents an eye-opener to the researchers to focus their interest in Big Data Analytics. Big Data Analytics is studied presented in this literature. A full-fledged Health Care System is not yet framed, even though several researchers have proposed an enhanced one. Our future research is to study the research issues and challenges in e-Health Care systems.

References

1. Kaisler S, Armour F, Alberto Espinosa J, Money W, Washigton G (2013) Big data: issues and challenges moving forward. In: The proceedings of international conference on system sciences, pp 995–1004. ISBN: 978-1-4673-5933-7
2. Verma A (2018) Why is big data analytics so important? <https://www.whizlabs.com/blog/big-data-analytics-importance/>
3. Davenport TH, Dyché J (2011) Big data in big companies: executive summary. White paper, SAS Institute Inc.
4. Agrahari A, Dharmaji Rao DTV (2017) A review paper on big data: technologies, tools and trends. *Int Res J Eng Technol* 4:640–649. ISSN: 2395-0072
5. Gupta R (2014) Journey from data mining to web mining to big data. *Int J Comput Trends Technol* 10:18–20. ISSN: 2231-2803
6. Rajkumar D, Usha S (2016) A survey on big data mining platforms, algorithms and handling techniques. *Int J Res Emerg Sci Technol* 3:50–55. ISSN: 2349-7610
7. Goel P, Datta A, Sam Mannan M (2017) Application of big data analytics in process safety and risk management. In: The proceedings of international conference on big data, pp 1143–1152. ISBN: 978-1-5386-2715-0
8. Arockia Panimalar S, Varnekha Shree S, Veneshia Kathrine A (2017) The 17 V's of big data. *Int Res J Eng Technol* 04:329–333. ISSN: 2395-0072
9. Singh S, Firdaus T, Sharma AK (2015) Survey on big data using data mining. *Int J Eng Dev Res* 3:135–141. ISSN: 2321-9939
10. Dinesh. R, Selva Kumar K, Manibharathi R (2018) Survey: importance and research issues in big data. *J Netw Commun Emerging Technol* 8:56–63. ISSN: 2395-5317
11. Gandomi A, Haider M (2015) Beyond the hype: big data concepts, methods, and analytics. *Int J Inf Manage* 35:137–144. ISSN: 0268-4012

12. Mavaluru D, Shriram R, Sugumaran V (2014) Big data analytics in information retrieval: promise and potential. In: The proceedings of international conference on international research forum for engineers and researchers, pp 41–46. ISBN: 978-93-84209-33-9
13. Gao J, Xie C, Tao C (2016) Big data validation and quality assurance—issues, challenges, and needs. In: The proceedings of the international conference on IEEE international symposium on service-oriented system engineering
14. Choudhury T, Chabra AS, Kumar P, Sharma S (2016) A recent trends on big data analytics. In: The proceedings of the SMART—2016, IEEE international conference on system modeling and advancement in research trends, pp 225–231. ISBN: 978-1-5090-3543
15. Joseph Charles P (2018) Big data—concepts, analytics, architectures—overview. *Int Res J Eng Technol* 05:125–129. ISSN: 2395-0072
16. Al-Kahtani MS (2017) Security and privacy in big data. *Int J Comput Eng Inf Technol* 9:24–29. ISSN: 2412-8856
17. Vijayarani S, Sharmila S (2016) Research in big data—an overview. *Inform Eng Int J* 4:1–19
18. Chaudhary P (2015) Data mining system, functionalities and applications: a radical review. *Int J Innov Eng Technol* 5:449–455. ISSN: 2319-1058
19. Kaur N, Singh G (2017) A review paper on data mining and big data. *Int J Adv Res Comput Sci* 8:407–409. ISSN: 0976-5697
20. Gorade SM, Deo A, Purohit P (2017) A study of some data mining classification techniques. *Int Res J Eng Technol* 04:3112–3115. ISSN: 2395-0072
21. Sin K, Muthu L (2015) Application of big data in education data mining and learning analytics—a literature review. *ICTACT J Soft Comput* 05:1035–1049. ISSN: 2229-6956
22. Lakshmi C, Nagendra Kumar VV (2016) Survey paper on big data. *Int J Adv Res Comput Sci Softw Eng* 6:368 – 381. ISSN: 2277 128X
23. Naganathan V (2018) Comparative analysis of big data, big data analytics: challenges and trends. *Int Res J Eng Technol* 05:1948–1964. ISSN: 2395-0072
24. Leelavathy B, Rsa TA, Singh H (2017) Big data analytics: challenges, research issues, tools and applications a survey. *Int J Eng Sci Res Technol* 3:548–554. ISSN: 2277-9655
25. Vijayaraj J, Saravanan R, Victor Paul P, Raju R (2016) A comprehensive survey on big data analytics tools. In: The proceedings of international conference on green engineering and technologies. ISBN: 978-1-5090-4556-3
26. Althaf Rahaman SK, Sai Rajesh K, Girija Rani K (2018) Challenging tools on research issues in big data analytics. *Int J Eng Dev Res* 6:637–644. ISSN: 2321-9939
27. Shaw R, Mukherjee S, Haldar N, Changdar S (2016) Big data-concepts, tools and foresight. *Int J Comput Sci Inf Technol* 7:285–291. ISSN: 0975-9646
28. Akhil S, Uma K (2017) Survey on the challenges and issues on big data analytics. *Int J Mech Eng Technol* 8:138–149. ISSN: 0976-6340
29. Aggarwal D, Sonika R (2016) Emerging technologies for big data processing: NOSQL and NEWSQL data stores. *Int J Eng Comput Sci* 5:15598–15604. ISSN: 2319-7242
30. Hammad KAI, Fakhraldien MAI (2015) Big data analysis and storage. In: The proceedings of the international conference on operations excellence and service engineering, pp 648–659. ISBN: 978-0-9855497-3-2
31. Goyal H, Sharma C, Joshi N (2017) An integrated approach of GIS and spatial data mining in big data. *Int J Comput Appl* 169:1–6. ISSN: 0975–8887
32. Koti S, Seeri SV (2017) A survey on big data issues and challenges. *IOSR J Comput Eng* 19:75–78. ISSN: 2278-8727
33. Tarekegn GB, Munaye YY Big data: security issues, challenges and future scope. *Int J Comput Eng Technol* 7:12 -24. 0976-6367
34. Raziuddin M, Venkata Ramana T (2015) Literature survey in data mining with big data. *Int J Adv Eng Glob Technol* 03:521–527. ISSN: 2309-4893
35. Katal A, Wazid M, Goudar RH (2013) Big data: issues, challenges, tools and good practices. In: The proceedings of international conference on contemporary computing, pp 404–409. ISBN: 978-1-4799-0192-0

36. Muthulakshmi P, Udhayapriya S (2018) A survey on big data issues and challenges. *Int J Comput Sci Eng* 6. ISSN: 2347-2693
37. Ali GH, Nithya A (2017) Challenges and open research issues and tools on big data analytics. *Int J Adv Res Comput Eng Technol* 6:1690–1703. ISSN:2278-1323
38. Kumar K, Patel K (2017) Big data, its issues and challenges. *Int J Eng Dev Res* 5:123–127. ISSN: 2321-9939
39. Farhin VP, Patel M (2017) A review on recommender systems. In: *The proceedings of international conference on innovations in information embedded and communication systems*. ISBN: 978-1-5090-3294-5
40. Kumar K, Dwivedi A (2017) Big data issues and challenges in 21st century. *Int J Emerging Technol* 8:72–77. ISSN: 0975-8364
41. Acharjya DP, Kauser Ahmed P (2016) A survey on big data analytics: challenges, open research issues and tools. *Int J Adv Comput Sci Appl* 7(2)
42. Mukherjee S, Shaw R (2016) Big data—concepts, applications, challenges and future scope. *Int J Adv Res Comput Commun Eng* 5:66–74. ISSN: 2319-5940
43. Kale DR, Todmal SR (2014) A survey on big data mining applications and different challenges. *Int J Adv Res Comput Eng Technol* 3:3835–3838. ISSN: 2278-1323
44. Gupta S, Kataria A, Rathore S, Rajput DS (2016) Information security issues in big data: solution using PPDm (privacy preserving data mining). *Int J Pharm Technol*. ISSN: 0975-766X
45. Santhiya K, Bhuvanewari V (2016) Big data insight: data management technologies, applications and challenges. *Int J Control Theor Appl* 9:77–91. ISSN: 0974-5572