Journal of Research and Development
A Multidisciplinary International Level Referred and Double Blind Peer Reviewed, Open Access
ISSN : 2230-9578 | Website: https://jrdrvb.org Volume-18, Issue-7| July-2026

Original Article

# Big Data Analytics: Recent Trends, Applications, Challenges and Future Directions

**Swati S Hiremath**
Assistant Professor, Department of Computer Science, KLE Society's G H College, Haveri

**Manuscript ID:**
JRD -2026-180748
ISSN: 2230-9578
Volume 18
Issue 7
Pp. 224-233
July- 2026
**Submitted:** 16 June. 2026
**Revised:** 26 June. 2026
**Accepted:** 11 July. 2026
**Published:** 31 July. 2026

***Abstract***

*Big Data Analytics has emerged as an important area of computer science because of the rapid growth of data generated through digital platforms, social media, Internet of Things (IoT) devices, business transactions, mobile applications, sensors, and other digital technologies. The increasing volume, velocity, variety, and complexity of data have created both opportunities and challenges for organizations and researchers. This review paper examines the development of Big Data Analytics and focuses on recent technological trends, applications, challenges, and future directions. The study is based on a systematic review of selected research studies published between 2021 and 2025. The reviewed literature covers Big Data Analytics capabilities, programming frameworks, machine learning, social media analytics, digital twins, Big Data science project management, predictive analytics, testing techniques, and emerging analytical technologies. The review shows that Big Data Analytics is increasingly integrated with Artificial Intelligence, Machine Learning, cloud computing, IoT, edge computing, predictive analytics, and real-time processing. Recent developments are also moving towards Generative AI, automated analytics, privacy-preserving analytics, and federated learning. The review identifies data quality, privacy and security, scalability, computational requirements, interoperability, lack of skilled professionals, and ethical concerns as major challenges. The study concludes that the future of Big Data Analytics will depend on intelligent, real-time, distributed, secure, and responsible analytical systems capable of converting large and complex datasets into meaningful insights and supporting faster decision-making.*

***Keywords:*** *Big Data Analytics, Artificial Intelligence, Machine Learning, Internet of Things, Predictive Analytics, Generative AI, Edge Computing, Data Analytics*

## Introduction

Data has become one of the most important resources in the digital economy. Every day, enormous quantities of data are generated through social media, online transactions, mobile applications, websites, sensors, IoT devices, healthcare systems, banking systems, transportation networks, and business operations. Traditional data processing systems are often unable to efficiently manage such large, diverse, and rapidly generated datasets. This situation has resulted in the development and widespread adoption of Big Data technologies and Big Data Analytics. Big Data generally refers to datasets that are too large, complex, diverse, or rapidly generated to be effectively processed using traditional data management techniques. Big Data Analytics refers to the systematic process of collecting, processing, analyzing, and interpreting such data to identify useful patterns, relationships, trends, and insights. The objective is not simply to store large quantities of information but to transform data into useful knowledge that can support decision-making. The development of Big Data Analytics has been closely associated with advances in distributed computing, cloud computing, machine learning, artificial intelligence, and IoT. Earlier Big Data systems primarily concentrated on large-scale storage and distributed processing using technologies such as Hadoop and MapReduce. Subsequently, technologies such as Apache Spark enabled faster processing and more efficient analytical applications.

***Creative Commons (CC BY-NC-SA 4.0)***
*This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International Public License, which allows others to remix, tweak, and build upon the work noncommercially, as long as appropriate credit is given and the new creations ae licensed under the idential terms.*

**Address for correspondence:**
*Swati S Hiremath Assistant Professor, Department of Computer Science, KLE Society's G H College, Haveri*

**How to cite this article:**
*Hiremath S. S. (2026). Big Data Analytics: Recent Trends, Applications, Challenges and Future Directions. Journal of Research and Development, 18(7), 224–233.*
*https://doi.org/10.5281/zenodo.22027510*

Quick Response Code:
Website: https://jrdrvb.org/
DOI: 10.5281/zenodo.22027510
OPEN ACCESS

Journal of Research and Development
A Multidisciplinary International Level Referred and Double Blind Peer Reviewed, Open Access
ISSN : 2230-9578 | Website: https://jrdrvb.org Volume-18, Issue-7| July-2026

The research literature shows that Big Data Analytics has gradually moved from basic data processing towards predictive, intelligent, automated, and real-time analytics.

Recent research demonstrates that Big Data Analytics is being applied across healthcare, finance, manufacturing, agriculture, transportation, education, social media, smart cities, cybersecurity, and e-commerce. Tosi et al. (2024), in their review of 15 years of Big Data research, observed the expansion of Big Data applications across numerous sectors and highlighted the increasing integration of AI, ML, cloud computing, and distributed computing. The increasing integration of Artificial Intelligence and Machine Learning has further transformed Big Data Analytics. Machine learning algorithms can identify patterns and relationships in large datasets and support classification, prediction, recommendation, anomaly detection, and decision-making. Nti et al. (2022) found that predictive analytics was a major application area in the reviewed literature and identified healthcare, cybersecurity, IoT, transportation, finance, agriculture, and education among important application domains. At the same time, the growth of Big Data Analytics has created important concerns. Large datasets may contain personal, financial, medical, or organizational information. Therefore, privacy, security, data quality, ethical use, governance, and regulatory compliance have become important considerations. Tosi et al. (2024) identified data quality, scalability, privacy, security, heterogeneous data integration, governance, and shortage of skilled professionals as continuing challenges. Therefore, understanding the recent trends in Big Data Analytics is important for researchers, organizations, technology professionals, and policymakers. This paper reviews the development of Big Data Analytics and examines major recent trends, applications, challenges, and future directions.

**Review of Literature**

**Table 1. Summary of Selected Literature on Big Data Analytics**

| Sl. No. | Author(s) & Year | Major Focus of the Study | Methodology | Major Findings / Contribution |
|---|---|---|---|---|
| 1 | **Sabharwal & Miah (2021)** | Big Data Analytics Capabilities (BDAC) in organizations | Systematic Literature Review and thematic analysis | Identified five major themes: resource management, Big Data proficiency, transformation of data into actionable insights, management attitude, and organizational value assessment. Emphasized the importance of technological, human, and organizational capabilities. |
| 2 | **Abkenar et al. (2021)** | Big Data Analytics in social media | Systematic Literature Review | Classified analytics into content-oriented and network-oriented approaches. Identified applications in sentiment analysis, recommendation systems, cybersecurity, health monitoring, and other areas. Privacy, security, scalability, and real-time processing were major challenges. |
| 3 | **Rathore et al. (2021)** | AI, Machine Learning, Big Data and Digital Twins | Systematic Literature Review | Found that AI, ML, IoT and Big Data form important technological foundations of digital twins. Identified applications in manufacturing, healthcare, transportation, smart cities, and other sectors. |
| 4 | **Abdalla (2022)** | Big Data technologies, terminology and applications | Comprehensive survey | Examined the 5Vs of Big Data and technologies such as Hadoop, Spark, Hive, Cassandra, NoSQL, Storm, HDFS and MapReduce. Identified applications across healthcare, agriculture, education, transportation, smart cities and other areas. |
| 5 | **Belcastro et al. (2022)** | Programming models and frameworks for Big Data Analytics | Technology/ survey review | Examined MapReduce, Workflow, BSP, Message Passing and SQL-like approaches. Reviewed Hadoop, Spark, Storm, Hama, MPI, Hive and Pig. Concluded that framework selection should depend on application and data characteristics. |
| 6 | **Nti et al. (2022)** | Machine Learning in Big Data Analytics | Mini Systematic Literature Review | Identified predictive analytics as a major application area. Discussed SVM, ANN, Deep Learning, Decision Trees, KNN, LSTM, CNN and Random Forest. Apache Spark was identified as an important Big Data platform. |

Journal of Research and Development
A Multidisciplinary International Level Referred and Double Blind Peer Reviewed, Open Access
ISSN : 2230-9578 | Website: https://jrdrvb.org Volume-18, Issue-7| July-2026

| | | | | |
|---|---|---|---|---|
| 7 | **Saltz & Krasteva (2022)** | Execution of Big Data Science projects | Systematic Literature Review | Identified workflows, agility, process adoption, project management, tools and review studies as major themes. CRISP-DM was found to be an important framework, with growing interest in combining it with agile methods. |
| 8 | **Arshad et al. (2023)** | Big Data testing techniques and challenges | Systematic Literature Review | Classified testing into functional, non-functional and combined approaches. Identified challenges related to data volume, performance, scalability, security, data quality, processing speed and node failures. |
| 9 | **Jamarani et al. (2024)** | Big Data Predictive Analytics | Systematic Literature Review | Reviewed applications in industry, e-commerce, healthcare, agriculture, smart cities, ICT and weather. Accuracy was the most frequently used evaluation metric. Identified timeliness, scalability and data complexity as major challenges. |
| 10 | **Tosi et al. (2024)** | Evolution of Big Data research over 15 years | Systematic Literature Review | Demonstrated the expansion of Big Data Analytics across healthcare, finance, manufacturing, education, transportation, smart cities and other domains. Highlighted AI, ML, cloud computing, edge computing, privacy and ethical practices as important future directions. |
| 11 | **Taha (2025)** | ML techniques in IoT, social media, NLP and information security | Survey and experimental benchmarking | Developed taxonomy of ML algorithms for different domains. Examined CNN, GNN, XGBoost, SSL, ELM, KNN, Decision Trees and K-Means. Highlighted scalability, privacy, computational complexity and algorithm selection challenges. |
| 12 | **Shahnawaz & Kumar (2025)** | Big Data characteristics, tools and techniques | Systematic Literature Review | Examined the 7Vs of Big Data and their relationship with ML and IoT analytics. Highlighted challenges involving scalability, real-time processing, data quality, visualization and technology selection. |

## 3. Research Gap

The literature shows that Big Data Analytics has developed rapidly and has been examined from several perspectives. Existing studies have provided valuable insights into Big Data technologies, programming frameworks, Machine Learning, social media, digital twins, predictive analytics, testing, and organizational capabilities. However, the reviewed literature also indicates that research is often **technology-specific or application-specific**. Some studies focus mainly on programming frameworks, while others concentrate on machine learning, social media, digital twins, predictive analytics, or testing. Therefore, there is a need for a broader review that brings together these developments and examines the overall direction of Big Data Analytics. Further, traditional Big Data research has increasingly shifted towards **AI-based, real-time, distributed, and privacy-aware analytics**. Recent developments such as Generative AI, edge analytics, federated learning, automated analytics, and privacy-preserving techniques require greater attention in an integrated review. The literature also shows that challenges such as data quality, scalability, privacy, security, interoperability, computational cost, and shortage of skilled professionals continue despite technological advancement. Therefore, understanding the relationship between emerging technologies and these challenges is important.

Hence, this paper attempts to provide an integrated overview of the **recent trends, applications, challenges, and future directions of Big Data Analytics**.

### Objectives of the Study

The study has the following objectives:

1. To examine the development and evolution of Big Data Analytics.
2. To identify the major recent trends in Big Data Analytics.
3. To examine the technologies supporting modern Big Data Analytics.
4. To identify major applications of Big Data Analytics across different sectors.
5. To examine the major challenges associated with Big Data Analytics.

Journal of Research and Development
A Multidisciplinary International Level Referred and Double Blind Peer Reviewed, Open Access
ISSN : 2230-9578 | Website: https://jrdrvb.org Volume-18, Issue-7| July-2026

6. To examine emerging developments such as Generative AI, edge analytics, real-time analytics, and federated learning.
7. To identify future opportunities and research directions in Big Data Analytics.

**Research Methodology**

The study adopts a **secondary research and systematic literature review approach**. The review is based primarily on published research papers dealing with Big Data Analytics, its technologies, applications, challenges, and emerging developments. The literature considered in the study includes research published mainly between **2021 and 2025**, supplemented by relevant recent research where required to understand emerging trends. The literature covers systematic reviews, surveys, technology reviews, and application-oriented studies.

**The major themes considered in the review are:**

1. Big Data Analytics capabilities
2. Big Data technologies
3. Programming frameworks
4. Machine Learning
5. Artificial Intelligence
6. IoT
7. Social Media Analytics
8. Predictive Analytics
9. Real-Time Analytics
10. Digital Twins
11. Data Security and Privacy
12. Edge Computing
13. Generative AI
14. Federated Learning
15. Big Data Testing

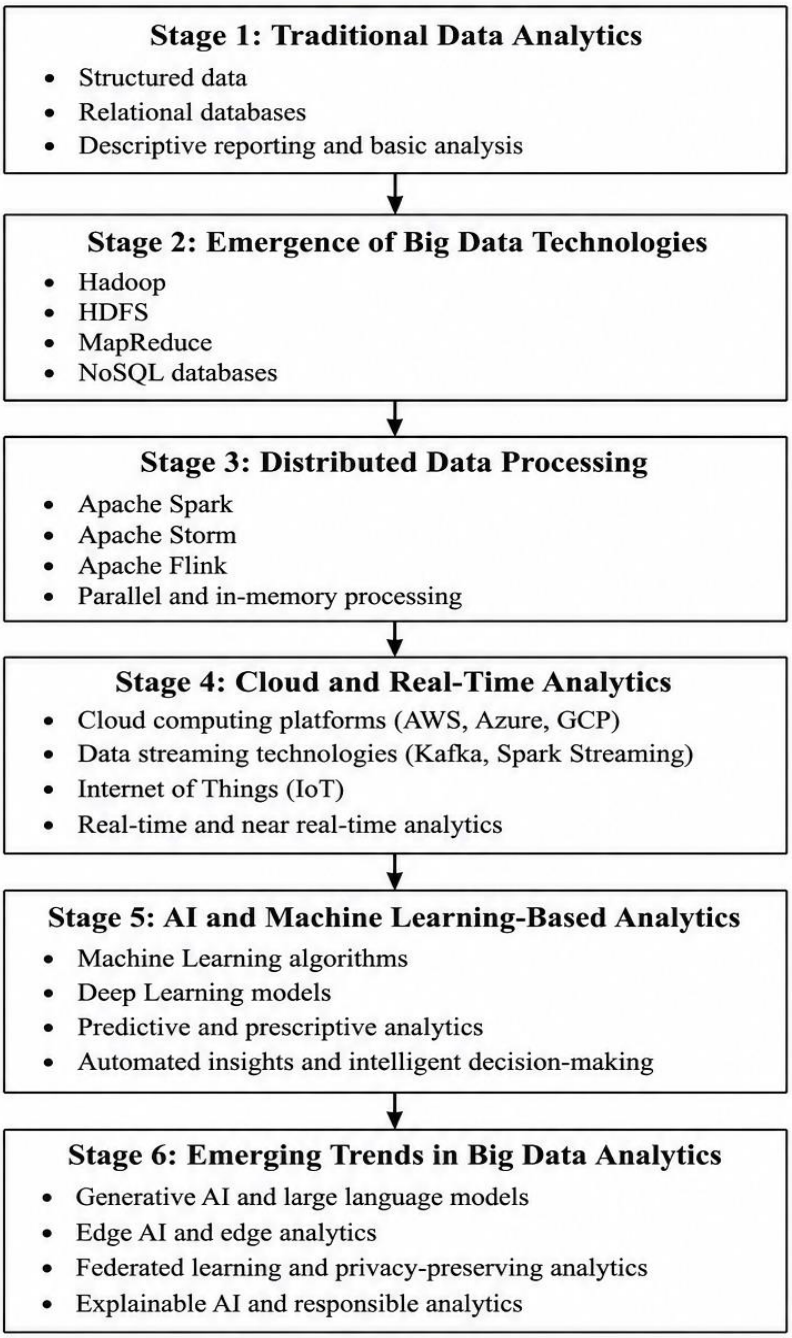
The selected studies were analyzed using a **thematic approach**, in which common concepts, technological developments, applications, challenges, and future directions were identified and compared.

**6. Evolution of Big Data Analytics**

Big Data Analytics has developed through several stages.

Journal of Research and Development
A Multidisciplinary International Level Referred and Double Blind Peer Reviewed, Open Access
ISSN : 2230-9578 | Website: https://jrdrvb.org Volume-18, Issue-7| July-2026

**Figure 1. Evolution of Big Data Analytics**




*Source:* Compiled by the authors based on reviewed literature.

**7. Recent Trends in Big Data Analytics**

| Sl. No. | Recent Trend | Key Technologies / Approaches | Major Applications | Major Benefits | Key Challenges |
|---|---|---|---|---|---|
| 1 | **AI and Machine Learning-based Analytics** | Machine Learning, Deep Learning, Neural Networks, XGBoost, SVM | Healthcare, finance, manufacturing, retail, cybersecurity | Automated pattern recognition, prediction and intelligent decision-making | Computational cost, model complexity, data quality |
| 2 | **Generative AI and Large Language** | Generative AI, LLMs, NLP, foundation models | Data summarization, conversational analytics, report | Natural-language interaction, automation and | Hallucination, bias, privacy, explainability |

Journal of Research and Development
A Multidisciplinary International Level Referred and Double Blind Peer Reviewed, Open Access
ISSN : 2230-9578 | Website: https://jrdrvb.org Volume-18, Issue-7| July-2026

| | | | | | |
|---|---|---|---|---|---|
| | **Models** | | generation, code generation | faster data interpretation | |
| 3 | **Real-Time Big Data Analytics** | Stream processing, Apache Kafka, Apache Storm, Spark Streaming | Fraud detection, IoT, stock markets, healthcare, cybersecurity | Faster response and immediate decision-making | Low latency, high data velocity, scalability |
| 4 | **Edge Analytics and Edge Computing** | Edge AI, IoT, distributed processing | Smart cities, autonomous systems, healthcare, manufacturing | Reduced latency, lower bandwidth use and faster local processing | Limited computing resources, security and interoperability |
| 5 | **Cloud-Based Big Data Analytics** | Cloud computing, distributed storage, cloud AI/ML services | Business analytics, e-commerce, healthcare, finance | Scalability, flexibility and reduced infrastructure cost | Data security, privacy, vendor dependency |
| 6 | **IoT and Big Data Integration** | IoT sensors, streaming analytics, cloud/edge platforms | Smart agriculture, smart homes, smart cities, manufacturing | Continuous data collection, monitoring and optimization | Data volume, heterogeneous devices, security |
| 7 | **Predictive and Prescriptive Analytics** | Predictive models, ML, statistical analytics, optimization | Demand forecasting, risk management, healthcare, agriculture | Future prediction and improved decision-making | Data quality, accuracy, scalability and model reliability |
| 8 | **Automated Machine Learning (AutoML)** | Automated feature selection, model selection, hyperparameter optimization | Business analytics, healthcare, finance, marketing | Reduces manual effort and supports wider use of ML | Limited interpretability, computational cost and model selection |
| 9 | **Federated Learning and Privacy-Preserving Analytics** | Federated Learning, differential privacy, secure aggregation | Healthcare, finance, smart cities, IoT | Collaborative analytics while reducing the need to share raw data | Privacy–accuracy trade-off, communication cost, scalability |
| 10 | **Explainable and Responsible AI** | Explainable AI, interpretable ML, responsible AI | Banking, healthcare, insurance, public services | Transparency, trust and accountability | Model complexity, bias and explainability |
| 11 | **Digital Twins and Big Data** | AI, ML, IoT, simulation, real-time analytics | Manufacturing, healthcare, transportation, smart cities | Real-time monitoring, prediction and optimization | Interoperability, standardization, security and computational requirements |
| 12 | **Data Governance and Data Quality** | Data governance frameworks, data quality management, metadata management | Banking, healthcare, government, business | Reliable data, improved compliance and better analytics | Data integration, governance complexity and regulatory requirements |
| 13 | **Big Data Security and Privacy** | Encryption, access control, privacy-preserving techniques | Banking, healthcare, e-commerce, government | Protection of sensitive data and reduction of security risks | Cyberattacks, privacy threats and complex distributed environments |
| 14 | **Green and Sustainable Big Data Analytics** | Energy-efficient computing, optimized algorithms, sustainable data centres | Data centres, cloud computing, AI systems | Reduced energy consumption and environmental impact | High computational demand and infrastructure cost |

**Applications of Big Data Analytics**

Big Data Analytics is being used in a wide range of sectors because organizations generate large quantities of structured, semi-structured, and unstructured data. Analytics helps convert this data into useful information for

Journal of Research and Development
A Multidisciplinary International Level Referred and Double Blind Peer Reviewed, Open Access
ISSN : 2230-9578 | Website: https://jrdrvb.org Volume-18, Issue-7| July-2026

planning, prediction, monitoring, and decision-making. The reviewed literature shows that the application of Big Data Analytics has expanded from traditional business applications to healthcare, agriculture, smart cities, manufacturing, transportation, cybersecurity, social media, and digital twins.

**1. Healthcare**

Big Data Analytics is increasingly used in healthcare for disease prediction, patient monitoring, clinical decision-making, medical image analysis, and personalized treatment. Healthcare organizations generate large amounts of information through electronic health records, medical devices, laboratory reports, and wearable devices. Analytics can help identify patterns and support early prediction of diseases. Machine Learning and Big Data can also support predictive healthcare by analyzing historical and real-time patient information. The literature reviewed by Nti et al. (2022) identifies healthcare as one of the major application areas of Big Data Analytics. Similarly, the integration of AI, IoT, and Big Data has created opportunities for intelligent healthcare systems.

**2. Banking and Finance**

Banking and financial institutions generate large volumes of transaction and customer data. Big Data Analytics can be used for fraud detection, credit risk assessment, customer segmentation, investment analysis, and financial forecasting. Machine Learning algorithms can identify unusual transaction patterns and support fraud detection. Customer data can also be analyzed to understand customer behaviour and provide personalized financial services.

**3. E-Commerce and Retail**

E-commerce platforms generate data through customer searches, purchases, product reviews, clicks, and browsing behaviour. Big Data Analytics helps organizations understand customer preferences and improve marketing strategies.
Major applications include:

1. Recommendation systems
2. Customer segmentation
3. Demand forecasting
4. Personalized advertising
5. Inventory management
6. Sentiment analysis

Predictive analytics can help retailers estimate future demand and improve inventory planning.

**4. Manufacturing**

Manufacturing industries use sensors and IoT devices to continuously generate information from machines and production processes. Big Data Analytics can analyze this information to identify abnormal conditions and predict equipment failures. The combination of IoT, Big Data, AI and digital twins is particularly important in modern manufacturing. Rathore et al. (2021) identified manufacturing as an important application area for AI, Machine Learning, Big Data and digital twin technologies.

**5. Agriculture**

Big Data Analytics is increasingly used in agriculture through sensors, satellites, weather information, drones, and IoT devices. Analytics can support:

1. Crop prediction
2. Soil analysis
3. Weather forecasting
4. Irrigation management
5. Pest detection
6. Yield prediction

This can contribute to more efficient use of water, fertilizers, land, and other agricultural resources.

**6. Smart Cities**

Smart cities generate data from traffic systems, public transportation, energy systems, sensors, surveillance systems, and public services. Big Data Analytics can be used to improve traffic management, energy consumption, waste management, public safety, and urban planning. The combination of IoT, cloud computing, edge computing, and AI is particularly useful for developing intelligent urban services.

**7. Transportation**

Transportation systems generate data through GPS devices, vehicles, traffic cameras, sensors, and mobile applications. Big Data Analytics can support traffic prediction, route optimization, demand forecasting, accident analysis, and intelligent transportation systems. Real-time analytics is particularly important because transportation decisions often need to be made immediately.

Journal of Research and Development
A Multidisciplinary International Level Referred and Double Blind Peer Reviewed, Open Access
ISSN : 2230-9578 | Website: https://jrdrvb.org Volume-18, Issue-7| July-2026

**8. Social Media**

Social media platforms generate enormous quantities of text, images, videos, comments, and network information. Big Data Analytics can be used for sentiment analysis, trend identification, customer opinion analysis, recommendation systems, and social network analysis. Abkenar et al. (2021) highlighted content-oriented and network-oriented analytical approaches in social media and identified several applications of Big Data Analytics in this area.

**9. Cybersecurity**

Big Data Analytics can analyze large volumes of network traffic, system logs, user behaviour, and security events. Machine Learning can identify unusual patterns and potential cyber threats.

Applications include:

1. Intrusion detection
2. Malware detection
3. Fraud detection
4. Anomaly detection
5. Threat intelligence

**10. Education**

Educational institutions generate data through student records, attendance, examination results, learning management systems, and online learning platforms. Big Data Analytics can be used for learning analytics, student performance prediction, identifying students who require support, and improving educational planning.

**Major Technologies Used in Big Data Analytics**

Big Data Analytics depends on several technologies that support data collection, storage, processing, analysis, visualization, and decision-making.

**Table 3. Major Technologies Used in Big Data Analytics**

| Technology | Main Purpose | Major Use |
|---|---|---|
| Hadoop | Distributed storage and processing | Large-scale batch processing |
| HDFS | Distributed data storage | Storing massive datasets |
| MapReduce | Parallel processing | Processing large datasets |
| Apache Spark | Fast distributed processing | Real-time and batch analytics |
| Apache Storm | Stream processing | Real-time analytics |
| Apache Kafka | Data streaming | Continuous data transmission |
| NoSQL Databases | Flexible large-scale storage | Unstructured and semi-structured data |
| Hive | Data querying | Data warehousing and analysis |
| Cloud Computing | Scalable computing and storage | Big Data infrastructure |
| Machine Learning | Pattern recognition and prediction | Predictive analytics |
| Deep Learning | Complex pattern analysis | Image, text and speech analytics |
| IoT | Data generation and collection | Sensor-based analytics |
| Edge Computing | Local data processing | Real-time IoT applications |
| Generative AI | Content and insight generation | Conversational and automated analytics |
| Federated Learning | Distributed model training | Privacy-preserving analytics |

Belcastro et al. (2022) emphasized that different programming models and frameworks are appropriate for different Big Data applications. Hadoop, Spark, Storm, Hive, Pig and other frameworks differ in their processing models, scalability and application suitability. Therefore, technology selection should depend on the characteristics of the data and the requirements of the application.

**Challenges of Big Data Analytics**

Although Big Data Analytics provides significant benefits, organizations face several challenges while implementing and managing Big Data systems. Large quantities of data do not necessarily mean high-quality data. Missing values, duplicate records, inaccurate information, inconsistent formats, and outdated data can affect analytical results. Big Data may contain sensitive personal, financial, medical, and organizational information. Protecting such information is a major concern. Large distributed systems can create additional security risks. Unauthorized access, cyberattacks, malware, and data breaches can affect organizations and individuals. Big Data systems must be able to handle continuously increasing quantities of data. Analytical systems therefore require scalable storage and processing infrastructure. Advanced Machine Learning and Deep Learning models require considerable computing power, memory, and storage. This can increase the cost of implementing Big Data Analytics. Data comes from multiple sources such as databases, websites, social media, sensors, mobile devices, and IoT systems. Combining these heterogeneous datasets is technically challenging.

Big Data Analytics requires knowledge of databases, statistics, programming, Machine Learning, cloud computing, AI, and domain-specific applications. The shortage of professionals with these combined skills can affect

Journal of Research and Development
A Multidisciplinary International Level Referred and Double Blind Peer Reviewed, Open Access
ISSN : 2230-9578 | Website: https://jrdrvb.org Volume-18, Issue-7| July-2026

successful implementation. The use of large datasets and AI-based analytical systems raises concerns about privacy, discrimination, surveillance, bias, and responsible use of information. Some advanced AI and Machine Learning models are difficult to interpret. Users may obtain predictions without understanding how the system reached a particular result. Big Data systems require specialized testing because they involve distributed architectures, large datasets, multiple data sources, and complex processing. Arshad et al. (2023) identified challenges related to performance, scalability, data quality, security, processing speed, and node failures. In applications such as healthcare, cybersecurity, transportation, and financial transactions, information must be processed immediately. Maintaining low latency while handling large volumes of continuously generated data remains a challenge.

**Future Directions**

The future of Big Data Analytics is expected to focus on the integration of advanced technologies that can make data processing more intelligent, faster, scalable, secure, and accessible. The increasing use of Artificial Intelligence and Machine Learning is likely to further strengthen predictive and prescriptive analytics, while Generative AI and Large Language Models may enable users to interact with large datasets through natural-language queries and support automated data interpretation and report generation. Edge computing and Edge AI are expected to become more important for real-time applications by allowing data to be processed closer to its source and reducing latency. At the same time, federated learning and privacy-preserving techniques may provide opportunities for collaborative analytics while protecting sensitive data. The future is also likely to emphasize Explainable AI and responsible analytics to improve transparency, fairness, accountability, and user trust. In addition, automated machine learning, cloud-edge integration, digital twins, and energy-efficient Big Data systems are expected to receive greater research attention. Overall, future research should focus on developing Big Data Analytics systems that can combine intelligence, real-time processing, scalability, privacy, security, explainability, and sustainability while addressing the practical challenges identified in existing research.

**Discussion**

The review demonstrates that Big Data Analytics has undergone a significant transformation. Initially, the primary concern was the storage and processing of large datasets. Technologies such as Hadoop, HDFS and MapReduce addressed the requirements of distributed storage and processing. Subsequently, the development of platforms such as Spark and Storm improved the speed and flexibility of Big Data processing. Cloud computing further increased the availability of scalable computing and storage resources. The next major development was the integration of Machine Learning and Artificial Intelligence. Big Data was no longer used only to understand past events; it increasingly became a foundation for prediction, classification, recommendation, anomaly detection and automated decision-making. Nti et al. (2022) and Taha (2025) demonstrate this increasing integration of ML techniques with Big Data Analytics.The literature also shows increasing integration between Big Data, IoT and digital twins. IoT devices continuously generate data, while Big Data platforms process the information and AI systems interpret it. Digital twins further use this information for real-time monitoring, prediction and optimization.Another important development is the movement towards real-time and edge analytics. Traditional centralized systems may not be sufficient for applications requiring immediate responses. Processing data closer to its source can reduce latency and support faster decision-making. Recent developments also indicate a movement towards Generative AI, federated learning, explainable AI and privacy-preserving analytics. These technologies can make Big Data Analytics more accessible, intelligent and secure. However, they also create new concerns relating to privacy, bias, explainability, computational requirements and responsible use.

Overall, the evolution of Big Data Analytics can be represented as:

Traditional Analytics → Distributed Big Data → Cloud Analytics → Machine Learning → AI and Predictive Analytics → Real-Time and Edge Analytics → Generative and Privacy-Preserving Analytics

The reviewed studies therefore indicate that Big Data Analytics is no longer a standalone technological area. It has become an integrated ecosystem involving AI, ML, IoT, cloud computing, edge computing, distributed systems, cybersecurity and data governance.

**Conclusion**

Big Data Analytics has emerged as an important technological area because of the continuous growth of data generated through digital systems, IoT devices, social media, online transactions, sensors, and other sources. The present review examined the development of Big Data Analytics, its recent trends, applications, technologies, challenges and future directions. The review of twelve selected studies shows that Big Data Analytics has developed from traditional data processing and distributed storage towards intelligent, predictive, real-time and automated analytics. Hadoop, MapReduce and other distributed technologies established the foundation for large-scale data processing, while Spark, cloud computing and streaming technologies improved processing speed and scalability. The integration of Machine Learning and Artificial Intelligence has further expanded the capabilities of Big Data Analytics. Predictive analytics, anomaly detection, recommendation systems, digital twins and intelligent decision-making are now important applications. At the same time, IoT and edge computing are enabling data to be collected and processed closer to its source. Recent developments such as Generative AI, Edge AI, federated learning, privacy-preserving

analytics, explainable AI and automated analytics indicate that the field is moving towards more intelligent and decentralized systems. However, challenges related to data quality, privacy, security, scalability, computational requirements, ethical issues and shortage of skilled professionals continue to require attention. Therefore, the future of Big Data Analytics will depend not only on the ability to process large quantities of data but also on the ability to generate reliable insights, protect information, explain analytical results and support responsible decision-making. The integration of Big Data with AI, ML, IoT, cloud and edge computing is expected to create new opportunities across healthcare, banking, manufacturing, agriculture, education, transportation, smart cities and other sectors.

In conclusion, Big Data Analytics is moving from a traditional data-processing approach towards an intelligent, real-time, automated and responsible analytics ecosystem. Continued research and technological development will be necessary to ensure that these systems are scalable, secure, explainable, sustainable and beneficial to organizations and society.

**References**

1. Abdalla, H. B. (2022). A brief survey on big data: Technologies, terminologies and data-intensive applications. *Journal of Big Data, 9*, 107. https://doi.org/10.1186/s40537-022-00659-3
2. Abkenar, S. B., Haghi Kashani, M., Mahdipour, E., & Jameii, S. M. (2021). Big data analytics meets social media: A systematic review of techniques, open issues, and future directions. *Telematics and Informatics, 57*, 101517. https://doi.org/10.1016/j.tele.2020.101517
3. Arshad, I., Alsamhi, S. H., & Afzal, W. (2023). Big data testing techniques: Taxonomy, challenges and future trends. *Computers, Materials & Continua, 74*(2), 2739–2770. https://doi.org/10.32604/cmc.2023.030266
4. Belcastro, L., Cantini, R., Marozzo, F., Orsino, A., Talia, D., & Trunfio, P. (2022). Programming big data analysis: Principles and solutions. *Journal of Big Data, 9*, 4. https://doi.org/10.1186/s40537-021-00555-2
5. Jamarani, A., Haddadi, S., Sarvizadeh, R., Haghi Kashani, M., Akbari, M., & Moradi, S. (2024). Big data and predictive analytics: A systematic review of applications. *Artificial Intelligence Review, 57*, 176. https://doi.org/10.1007/s10462-024-10811-5
6. Nti, I. K., Quarcoo, J. A., Aning, J., & Fosu, G. K. (2022). A mini-review of machine learning in big data analytics: Applications, challenges, and prospects. *Big Data Mining and Analytics, 5*(2), 81–97. https://doi.org/10.26599/BDMA.2021.9020028
7. Rathore, M. M., Shah, S. A., Shukla, D., Bentafat, E., & Bakiras, S. (2021). The role of AI, machine learning, and big data in digital twinning: A systematic literature review, challenges, and opportunities. *IEEE Access, 9*, 32030–32052. https://doi.org/10.1109/ACCESS.2021.3060863
8. Sabharwal, R., & Miah, S. J. (2021). A new theoretical understanding of big data analytics capabilities in organizations: A thematic analysis. *Journal of Big Data, 8*, 159. https://doi.org/10.1186/s40537-021-00543-6
9. Saltz, J. S., & Krasteva, I. (2022). Current approaches for executing big data science projects—A systematic literature review. *PeerJ Computer Science, 8*, e862. https://doi.org/10.7717/peerj-cs.862
10. Shahnawaz, M., & Kumar, M. (2025). A comprehensive survey on Big Data Analytics: Characteristics, tools and techniques. *ACM Computing Surveys, 57*(8), Article 196. https://doi.org/10.1145/3718364
11. Taha, K. (2025). Big Data Analytics in IoT, social media, NLP, and information security: Trends, challenges, and applications. *Journal of Big Data, 12*, 150. https://doi.org/10.1186/s40537-025-01192-9
12. Tosi, D., Kokaj, R., & Roccetti, M. (2024). 15 years of Big Data: A systematic literature review. *Journal of Big Data, 11*, 73. https://doi.org/10.1186/s40537-024-00914-9