Golden Sun-Rise
International Journal of Multidisciplinary on Science and Management
ISSN: 3048-5037/ Volume 1 Issue 2 Apr-Jun 2024 / Page No: 30-40
Paper Id: IJMSM-V1I2P104 / Doi:10.64137/30485037/V1I2P104

Research Article

The Integration of Big Data and Business Intelligence: Challenges and Future Directions

Mr. Rahul Cherekar

Independent Researcher, USA.

Received: 18 April 2025 Revised: 29 April 2025 Accepted: 07 May 2025 Published: 17 May 2025

Abstract*: BI and big data provide insights for the business which are helping to re-form the procedures of business. Several industries are now using Big Data analysis to improve performance, customer satisfaction, and operations. This paper seeks to discuss the integration of big data and BI focusing on the opportunities that come with it and the problems that may be encountered while at it. It explains the main topics, including information obtainment, data storage and analysis, and data visualization. The other subtopic discusses technological support like cloud computing, AI, and ML in developing BI-enhanced features. Nonetheless, several issues arise when applying Big Data with BI systems, such as data quality, governance, scalability, and real-time processing. The paper also explores possible areas of development like edge computing, blockchain and the enormous potential that quantum computing holds in Big Data analysis. Thus, this research aims to offer a rich understanding of how organizations can use Big Data with BI to gain competitive superiority. In this paper, based on the literature review, methodological framework, and survey findings, we suggest the recommendation for BI to maximize Big Data's value for enterprises.*

Keywords*: Big Data, Business Intelligence, Data Analytics, Cloud Computing, Machine Learning, Artificial Intelligence, Data Governance.*

I. INTRODUCTION

Business intelligence and big data can be named the vital tools for implementing the data-driven approach in the organization. [1-4] The difference between the contemporary BI system is the data coming from internal sources is structured, while the Big Data BI involves semi-structured and unstructured data from such sources as social media, smart devices, and streaming services.

A. Importance of Integration

The merge of Big Data with BI allows organizations to enhance how to analyze, interpret, and utilize data for decision making. When Big Data is integrated with the BI tools, businesses can get instant information, increase predictability, and have better control over organizational functions. The following are some of the areas of concern that advocate for this integration:

- **Enhance Real-Time Decision-Making:** Businesses of the present world are considerably dynamic and need information inputs in real-time. Combined, Big Data and BI help organizations to analyze large volumes of raw data feed and make decisions based on end-user interactions with those data feeds so that the organizations can proactively work through change in the market, operations, and consumers. Therefore, business intelligence plays a critical role in today's industries, especially in finance for fraud detection, health care for remote monitoring of patients, and e-commerce for product recommendations. Some technologies like Apache Kafka, Spark Streaming, and BI on the cloud help smoothly process real-time data and help organizations be in a better position.
- **Improve Predictive Analytics:** It analyzes data to make future estimations about potential risks and opportunities. With the help of big data and BI combined, companies can use machine learning and AI for analytics and learn about patterns and probable outcomes that could be used in strategic planning. For instance in the retail industry, variables based on predictive BI are used in determining customers' demanding trends, while in the finance industry, it is used in assessing credit risks. The idea refers to the ability to utilize sophisticated algorithms

in analyzing the data, making it possible for organizations to be more proactive rather than reactive in decision-making, hence eradicating risks that may hinder high profitability.

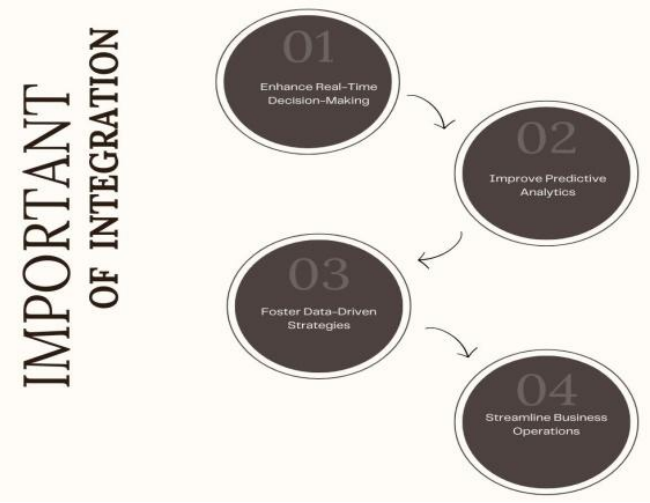

Figure 1. Importance of Integration

- **Foster Data-Driven Strategies:** This approach means that decision-making is not based on guesswork but depends on actual data collected from the business environment. Therefore, the Big Data integration with BI ensures that organizations have a rich variety of quality data that can be used to produce such insights. Today, marketing, manufacturing, and logistics sectors incorporate BI with business real-value goals to manage supply chains, improve customer relations, and optimize success. Besides, BI dashboards and automated reporting tools allow managers to have detailed checks on business performance through visual analytics.
- **Streamline Business Operations:** The two concepts are very important as they contribute to the success of an organization. The combination of Big Data and BI leads to better control of the business processes, allocation and distribution of resources, and the overall work processes. From large data analysis, large amounts of money, general productivity, and expense ratios, organizations can find problem areas they can work to eliminate. For instance, in the healthcare sector, analytics provide for the more effective management of patients; in the manufacturing sector, predictive maintenance helps maintain low levels of downtime and production stoppage. In this case, using cloud-based business intelligence and artificial intelligence based process automation, businesses can take additional capability to expand, be more flexible, and gain additional business advantage.

B. Definition of Big Data and Business Intelligence

a) Big Data

Big Data can thus be described as a concept that relates to bulky volumes of data that cannot be processed in a normal way by conventional tools. Hadoop is characterized by four Vs, namely Volume, Velocity, Variety, and Value that define the Big data. [5,6] Volume pertains to the real sense of data magnitude originating from numerous sources such as social media, connected devices, and business operations. Velocity stresses the rapid generating of data that require immediate action or analysis, thus, real-time data analysis is core. Big data: Variety refers to the type of data, structured, semi-structured, and unstructured, consisting of databases, XML/JSON, and videos, images, and social media posts, respectively. Veracity brings into the limelight that data should always be correct and of good quality since wrong data leads to wrong conclusions. Last, Value relates to the capability of decision-making and utilizing the big data flow to develop business strategies, make new inventions, and improve efficiency. The advancement in big data technologies like Hadoop, spark, and cloud analytics in organizations is used in arriving at the right decisions by analyzing a multitude of data. AI and ML integrated with Big Data provide more effectiveness and efficiency by enhancing the facility of predictive and prescriptive analytics which again assures bettering business and customer satisfaction.

b) Business Intelligence (BI)

Therefore, BI refers to the structures, applications, instruments, and methods used to analyse data and provide information useful to business management. BI operates at the data capturing, storage, processing, and goal statement and allows organizations to notice certain performance trends to improve how their operations are being performed. BI in the

past only used data in tabular and relational formats acquired from databases and simple spreadsheets, while current BI targets large data, real-time data, and data with the help of artificial intelligence. Tableau, Power BI, Qlik Sense, and other BI tools have tools inbuilt dashboards, data visualization, and advanced reporting and data handling options to support users in analyzing data. Moreover, BI also carries out descriptive, diagnostic, predictive, and prescriptive analytics, which will help a business learn about past patterns, further recognize operational problems, forecast future results, and select the right approach for taking further action. Further, incorporating Big Data technologies has expanded BI's role and capabilities to process and analyse real-time and unstructured data to move at a faster rate the decision-makings of an organization. Business areas that have applied BI include the finance industry, health care, retail, and manufacturing, but a few get to wake up with a new day of better performance. BI becomes even more important with increased data production and usage as it directly involves planning and performance enhancement in enterprises.

II. LITERATURE SURVEY

A. Evolution of Business Intelligence

- **Traditional BI vs. Modern BI:** Traditional Business Intelligence (BI) is based on structured data in operational systems in relational databases and usually necessary the help of IT to prepare reports. Such information solutions were often hard-copy-based and were comprehensive in how they needed to be used occasionally to extract valuable insights. [7-11] On the other hand, contemporary BI tools utilize new technologies such as self-service analytical tools, cloud, and real-time data dashboards. With the movement towards modern BI, users rather than IT employees give attention to business analytics, and the decision-making process becomes much faster.
- **Emergence of Data Warehouses and ETL Processes:** EDW has proved vital in BI development due to its capability to store and retrieve consistent and large volumes of Structured Data. ETL means to extract, transform, and load through which data is extracted from multiple sources, cleansed, and then loaded into the repository. This structured manner helps to control and organize the data used in various organizational analyses to increase the reliability of conclusions. ETL processes have become far more efficient due to automation and various improvements in real-time data integration.
- **Adoption of AI and ML in BI:** This paper aims to explore the role of AI and ML in the BI process in the contemporary generation of insights about organizational data. BI tools powered by artificial intelligence support real-time data analysis, data outliers identification, and predictive data analysis to give companies a better and more detailed view of the market. They are data-driven and improve with time in making forecasts and enabling prescriptive analytics. Recourse to AI and ML in BI have rendered decision-making from reactive to proactive where organizations can prepare for trends and enhance strategies correspondingly.

B. Role of Big Data in BI

- **Data Lakes vs. Data Warehouses:** Data lakes and data warehouses are two essential concepts used in BI, and they have distinct functions. A data warehouse is more analytic-oriented, which means it is designed to provide the capacity for efficiently querying conventional business data. In contrast, a data lake is a more open and comprehensive solution that can store data. Data lakes are the repositories used for storing raw data that are then analyzed using advanced analytics and machine learning. On the other hand, data warehouses provide value-added data through well-defined schemas. The choice between these depends on the organization's analytical needs and data management strategy.
- **Predictive and Prescriptive Analytics:** Predictive analytics in BI use historical and statistical models to make expected future patterns to allow organizations to have preemptive action plans. In prescriptive analytics, one is also offered the right way to proceed as advised by the data patterns. These analytical techniques are applied to enhance efficiency in business operations, how companies manage their clients and partners, and in gaining methods of managing risks. The combination of AI makes a real change to the effectiveness of prediction and prescription of analytics so that various organisations can gain an edge in competitive markets.
- **Importance of Real-Time Data Processing:** Real-time data processing has gained popularity in today's world, especially for companies that need to receive information and make decisions quickly. Traditional BI systems also used batch processing that slowed up the reporting process. However, today's BI solutions support real-time streaming analytics that can assess performance and identify emerging issues in real-time action. Real-time BI has a broad range of applications in industries like finance, health care, and e-commerce to increase the quality of their customers' experience, increase productivity, and address possible risks.

C. Challenges in Integration

- **Data Quality and Consistency:** There are several issues with BI, particularly integrating with various data sources One of the biggest problems when working with BI is data quality and data quality. Silos taking in inconsistent data and information results in the production of ineffective and wrong conclusions. Thus, it is crucial to adopt sound policies and methods in data cleansing and standardization in organizations regarding data quality. This article clearly shows that poor-quality data leading to wrong decisions must be avoided at all costs; this can be achieved by proper validation and monitoring.
- **Scalability Issues:** It is especially apparent when searching for information becomes difficult due to the amount of data an organization gathers over time within BI systems. Traditional corporate BI architectures are not designed to meet the increasing data volumes, and this will lead to a slow data processing problem. To address scalability issues, applications and computing technologies, including cloud-based Business Intelligence, Hadoop, and Spark, have been developed. Companies need to have infrastructures that allow the business intelligence systems to increase in capacity in the future.
- **Data Security and Privacy Concerns:** This has given rise to two valuable assets in enterprise environments: business intelligence and data security/privacy. Data organizations must adhere to requirements such as GDPR or HIPAA to safeguard sensitive data. Cybersecurity threats, data breaches, and unauthorized access pose significant risks to BI systems. Encryption, access control, and data masking may improve the security measures being applied to information. Besides, organizations have an important role in enhancing data protection by promoting a culture of data privacy among the workers.

D. Case Studies

- **Retail Industry: Use of Big Data in Customer Analytics:** Big Data is also used in retail to analyze customers' behavior, for advertising purposes, and stock control. In particular, Amazon and Walmart are examples of companies that use BI tools to predict buyers' purchase tendencies, determine demand changes, and improve clients' satisfaction. Implementing RLA and AI helps retailers to provide specific product recommendations to the consumers and, in so doing increases customer satisfaction and sales revenue.
- **Healthcare Industry: Business Intelligence through Artificial Intelligence:** Otherwise, in the healthcare industry, the use of AI in BI to support predictive analysis in the diagnosis and treatment of patients is crucial. The members of a medical facility also apply machine learning to discover symptoms of illnesses and advise treatment. Predictive analytics in health care is vital in preventing hospital release, advancing disease control measures, and quality patient treatment. The efficient integration of BI in healthcare empowers several professionals to make clinical decisions based on the availability of Vital information to enhance the quality and speed of valuable patient care.
- **Finance: Fraud Detection Using Big Data Analytics:** Credit and other financial institutions' executives and boards use Big Data to prevent fraud. The conventional method of performing BI combines various solutions that collect data, such as data mining of transactions done by various banks and the payment service provider company to analyze the pattern of transactions and flag any suspicious activities, all in real-time. The fraud detection system uses machine learning techniques to identify fraudulent purchases from other genuine ones, hence cutting the losses incurred and adding to security measures. Real-time fraud often increases the reliability and credibility of financial services.

III. METHODOLOGY

A. Research Approach

As it concerns BI and Big Data analytics, research methods can be further divided into quantitative and qualitative research. Quantitative research is more of numbers and statistical inferences and touches on the objective elements of comparisons. It is frequently employed in BI studies whereby it is used to analyze and evaluate big data, capture trends, or COMMENT analyse the performance of a given system. It offers generalizable information through experiments, questionnaires, surveys, and statistical modeling techniques. [12-16] Qualitative research, on the other hand, is a collection of information that does not entail numbers like opinions of experts or interviews and case studies. This approach helps grasp the users' experiences, the issues in adopting BI, and the strategic implications of technologies. Qualitative research is used to gain an in-depth understanding of supporting variables, supplementing quantitative analysis. The data-gathering mode depends on the study question or hypothesis under investigation. Questionnaires are popular instruments used in BI research to establish satisfaction rates, efficiency, and utilization rates of BI systems.

Surveys make it easier for the researcher to gain data from a large group of people, given that the responses are organized formally and easier when it comes to trend analysis; as with any methodology, structured or semi-structured interviews allow for acquiring a greater amount of opinion from the experts, published users, and the problems facing organizations using BI. The sections on case studies look at how BI is implemented in practice, which provides detailed information on different industries such as retailing, healthcare, and finance, amongst others, for instance, successes and areas of difficulty that have been encountered. Using these methods makes it possible to have a broad view of BI systems in different aspects. It also employs other forms of analysis like statistical analysis and Machine Learning (ML). Regression analysis, hypothesis testing, and factor analysis provide crucial pointers regarding the correlation, relationship, and characteristic tendencies among BI-adopting companies. It is evident that using decision trees and neural networks, for example, in developing predictive and prescriptive analytics, improves Business Intelligence insights. Such valuable approaches will enable the evaluation of BI effectiveness, identification of Big Data patterns, and presenting findings-oriented recommendations to the companies.

B. Data Processing Techniques

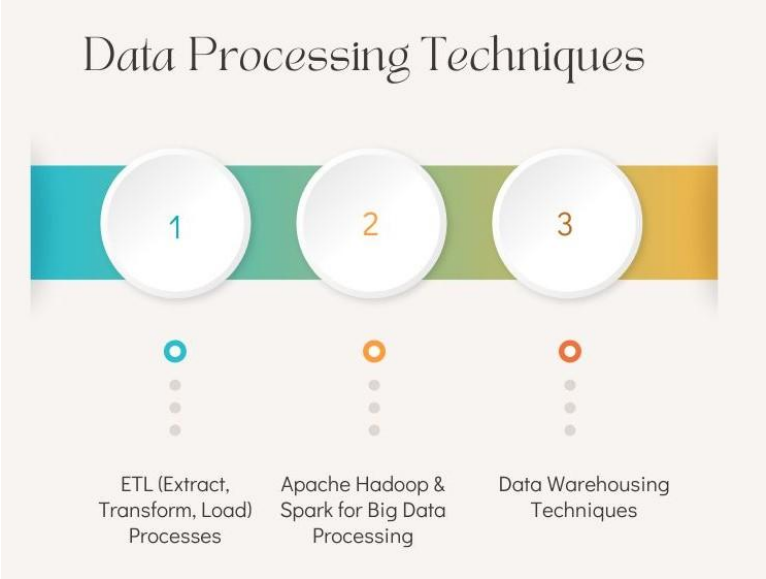

Figure 2. Data Processing Techniques

- **ETL (Extract, Transform, Load) Processes:** ETL (Extract, Transform, Load) is a data integration methodology used in business intelligence and analytics practices. Extracting is the process of obtaining data from various sources like databases, APIs, and peripheral or cloud storage. The Transform step aligns the information acquired by transforming it from one format to another as needed for analysis. This may comprise data deduplication, normalization as well as aggregation. Finally, the Load phase transfers/loads the transformed data into a consolidated/distributed location, the data warehouse, or data lake for reporting and analysis. ETL processes ensure data consistency to be used across different systems and improve how data is used to make decisions.
- **Apache Hadoop & Spark for Big Data Processing:** Apache Hadoop and Apache Spark are two popular plat forms in the advanced programming framework for Big data analysis is thus. Hadoop can be described as a data processing framework that implements HDFS and MapReduce to process and store structured and unstructured data. It also supports forming distributed and robust computations over a number of message-passing nodes. However, in some cases, one disadvantage of Hadoop is that it is a batch-processing system. Apache Spark is a much faster in-memory computing engine for real-time and big data processing and machine learning. The traditional model used in Spark is known as the Resilient Distributed Dataset (RDD), which enables complex data to be processed quickly and, hence, is suitable for applications where many mappings, such as fraud detection, predictive analytics, and artificial intelligence big data intelligence are needed.
- **Data Warehousing Techniques:** Data warehousing can be described as the process through which data is collected and processed for storage and easy access. Star schema is a kind of schema design where data is arranged to measure performance facts and dimensions or descriptors to enhance the efficiency of queries. OLAP (Online Analytical Processing), unlike OLTP, allows a business to make multidimensional analysis thereby permitting

analysis at more inherent levels. Some examples of currently available cloud data warehousing solutions include – Amazon Redshift and Google BigQuery; they are both effective and affordable data storage solutions with fast query response. These methodologies allow businesses to deal with large data and gain profitable information.

C. Integration Architecture

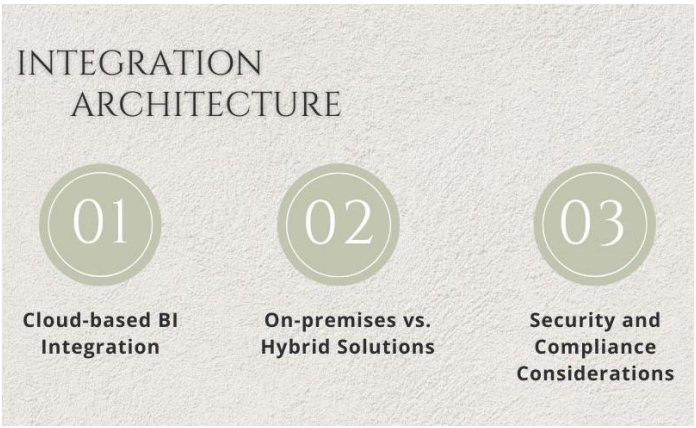

Figure 3. Integration Architecture

- **Cloud-based BI Integration:** Real-time web-based BI integration is widely preferred because it is more scalable, affordable, and easily accessible. Some popular cloud BI software are Microsoft Power BI, Google Data Studio, and Tableau Cloud, offering a company an environment where it can store and process data in a near real-time fashion without requiring a high level of local infrastructural support. These platforms are highly compatible with different cloud-based data feeds, APIs, and other applications, which allow organizations to use AI analysis and report automation.
- **On-premises vs. Hybrid Solutions:** It should be noted that depending on the needs of organizations and the security levels required, one can implement the BI integration either in on premise, cloud, or both. On-premises BI solutions refer to the hosting of data and analytics within an organization's infrastructure which provides organizations with full control together with data security but comes with high IT costs. However, cluttered BI solutions ensure that the applications exist both on-premises and in the cloud to enable organizations to compile and store sensitive data within the company servers while using the cloud for analytical purposes and real-time access. They offer companies the best of both worlds since they satisfy the security issues that arise from using cloud computing solutions while still offering maximum benefits and advantages to the organization. So they are highly suitable in industries such as health and finance since they are highly sensitive.
- **Security and Compliance Considerations:** Security and compliance play a key role in BI integration, especially when the business deals with sensitive information. Encrypting data, access control solutions, and multi-factor identification are necessary to control risks posed by cyber threats in the organization. Legal requirements concerning data protection include GDPR, HIPAA, and SOC 2 to ensure the maximum legal and ethical approach to working with data. Finally, on data protection, it is recommended that data governance policies incorporate roles and responsibilities of accessing or monitoring the data, thus minimising cases of leakage. For the BI architecture to be secure, there should be data masking where one can only get a bounded amount of data, log auditing where all the activities should be recorded, and real-time threat detection should also be incorporated.

D. Data Visualization and Reporting

The essentials of BI involve data visualization and reporting, where raw information is analyzed properly to be informative. [17-20] Business intelligence ensures organizations gather data that portrays a particular trend, monitors how an organization performs or even determines certain patterns from large datasets. Another popular BI method is the use of cross-functional, multi-tab, and level interfaces for KPI, metrics, and reports, collectively called interactive dashboards. Dashboards are a real-time performance measurement system that presents current business status, potential trends, and variances from past performance to help decision-makers identify problems and opportunities and make timely decisions. It supports different types of charts, such as bar charts, heat charts, line charts, map charts, and so on, allowing the user to view different aspects of the data set. Flexible and easy to use, dashboards enable the presentation of the required data without the help of a technical person; organizations can customize the way data is presented in the form of a dashboard according to their requirement.

These advanced analytics tools include Tableau, Power Business Intelligence, and Google Looker to improve BI reporting and predictive attributes. These tools use AI and ML algorithms to analyze data and create alert systems and insights far from human intervention. Functionality such as natural language query enables the users to ask questions in the form of data interrogatives, making it easier for the BI solution to be adopted by non-IT personnel. Moreover, AI capabilities embrace predictive and prescriptive analytics that enable the enterprise to forecast the market and plan its activities more effectively. BI solutions also increase accessibility, enabling more people to easily access data, collaborate in real time, and share across the organisation or department. In this way, using AI tools for BI, organizations can go beyond reporting and ceding to ever-changing and dynamic real-time insights for better decision-making and befitting the competitiveness of the business.

IV. RESULTS AND DISCUSSION

A. Cloud BI vs. On-Premise BI vs. Hybrid BI

- **Cost:** Cloud BI operates mainly on a subscription basis, making it easy to budget and avoiding expenses associated with the direct purchase of hardware or software. Since the service provider has to bear the establishment costs of infrastructure and maintenance, the organizational costs are considerably reduced regarding equipment and software. Indeed, compared to cloud BI, which is cheaper and provides easy access to the software and data, on-premise BI is quite expensive, which entails high costs to be incurred on servers, software, and the personnel involved in managing the BI systems of an organization. Still, it is very secure and more suitable for large-scale organizations. A moderate cost characterizes hybrid BI since the organizations can keep the crucial information in-house and use cloud services to analyze large portions of data.
- **Scalability:** Another benefit of Cloud BI is that this solution has high flexibility and easy scaling up and down. Resources in a business entity can be rapidly escalated or de-escalated depending on the need for performance and productivity, thus avoiding the wastage of resources. Concerning on-premise BI, it is bound by infrastructure, which necessitates the acquisition of more gears if a firm increases its demands for data processing capabilities. Hybrid BI provides the best of both worlds and allows the cloud to increase its scale while keeping its critical workloads in-house.
- **Security:** The last concern that has been on the list of BI integration for quite some time now is security. Some aspects of cloud security that form the framework of cloud BI are the encryption process, multiple-factor authentication, and compliance standards. However, offsite data storage poses some risks that businesses that opt to store their data offsite consider as threats; data breaches and unauthorized access. On-premise BI solutions can be considered compliant-oriented and highly secure since the data is physically located within the business's internal network. Hybrid BI comes in the middle of the security level based on the BI architecture; storing the most critical data on the company's infrastructure is possible, while less critical data can be processed on the cloud.
- **Maintenance:** Also, the vendor is in charge of maintenance, so internal IT departments are free and do not need to grind software updates or security patches. On the other hand, on-premise BI needs in-house IT personnel to manage upgrades, security, and resolution of any problems. This leads to increased operation costs and congestion of the health workforce. Hybrid BI uses both cloud and on-premise structures, but while the dealings with the workings of the application involve the cloud, some IT management is still needed on the hardware.
- **Accessibility:** Sometimes, it becomes challenging to ease upper management into the concept of Cloud BI, but its principal advantage is the ability to be accessed from anywhere in the world with an internet connection. On-premise BI is only accessible to individuals on the company's internal networks, and therefore, it only supports users within the company's physical premises. Regarding the hybrid model of BI, there is the ability to choose which data has to be stored and processed locally and which data can be accessed without being within the company's network.

B. Traditional BI vs. Big Data BI: Performance Comparison

Table 1: Traditional BI vs. Big Data BI: Performance Comparison

Parameter	Traditional BI	Big Data BI
Processing Speed	3	9
Scalability	4	9
Data Sources	5	10
Real-time Analytics	3	9

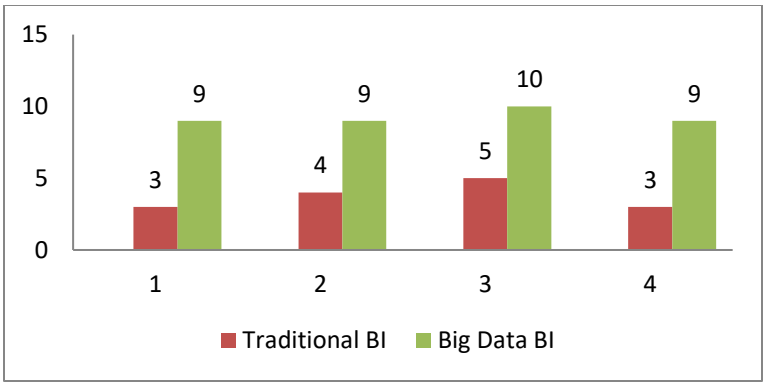

Figure 4. Graph representing Traditional BI vs. Big Data BI: Performance Comparison

- **Processing Speed:** Traditional BI system has inadequate processing capability since it requires structured data and operates on a batch basis. These systems tend to employ a number of databases based on relational databases and SQL queries, which are always slow when handling larger data sets. Thus, traditional BI gets 3 out of 10 in speed of processing. On the other hand, Big Data BI uses distributed systems, such as Apache Hadoop or Spark, where data crunching occurs concurrently and faster. This has the added benefit of increasing speed, so Big Data BI gains the maximum score of 9 on this criterion.
- **Scalability:** There are also limitations related to scalability in the case of traditional BI solutions. In fact, traditional BI solutions can only be scaled with great difficulty since they can only analyze data structures selected in advance, and the systems that support them are on-premise. Extending these systems can be very costly by making hardware additions by which the application achieves a scalability score of four out of ten. On the other hand, Big Data BI incorporates scalability, which enables it to handle structured and unstructured data across multiple nodes in a distributed platform. Another important criterion of Big Data solutions shows high scalability by allowing businesses to add or remove resources depending on current usage and offering the flexibility of cloud solutions. So, therefore this criterion receives a score of of 9/10.
- **Data Sources:** While traditional BI can analyze data from structured sources, such as relational databases, spreadsheets, Enterprise Resource Planning (ERP) systems, etc., it lacks versatility. It is rigid for general usage and has been rated 5 out of 10 in terms of handling different kinds of data. However, BI applies well to structured, semi-structured, and unstructured data sources like social media feeds, IoT, video streams, and Web interactions in real-time, etc. This makes it possible to have a more complete study, giving it a full score of 10/10.
- **Real-time Analytics:** In their current form, BI systems are mainly based on the analysis of historical data, which leads to the fact that reports contain information now and no longer reflected in the current processes. Most of these systems cannot handle and analyze the flow streams in real-time, hence being accorded a rating of 3/10. On the other hand, Big Data BI includes engines such as Apache Kafka, Apache Flink, and Google BigQuery, through which businesses can monitor data changes in real-time. We believe that Big Data BI deserves a high score of 9/10 because it is crucial for, for instance, financial (fraud detection), health (patient supervision), and commercial (flexible preceding of pricing) services.

C. Challenges and Mitigation Strategies

Despite the establishment of BI, there are differentials that companies face in their data management, processing, and quality. The best approach to combating these problems Laud with a strategic framework, high-performance infrastructure, and artificial Intelligence.

- **Data Governance Frameworks:** Data governance is very important in ensuring that the data used in BI systems is accurate and meets the intended use. Handling data from various sources, inconsistencies in format, data duplicity, and unauthorized access may cause inconsistencies and security concerns. In addition, extra weight is given to regulation standards, including GDPR, HIPAA, and USA's CCPA. To manage these risks effectively, companies should aim to develop consistent governance procedures, adopt the use of RBAC, and use metadata management tools for tracking data lineage and quality for BI systems.
- **Real-time Data Processing Techniques:** Some of the important features that every BI system must include are its ability to work in real-time, which is critical in several industries such as finance to detect such practices as fraud, health care to monitor patients, and retail to change the price of goods and services. However, real-time processing

is a demanding task that needs tremendous supporting structures like distributed computing and in-memory processing. The batch-processing BI systems are not suited to work in real-time requirements and thus have a problem providing timely insights. However, to overcome this, businesses are shifting to streaming platforms such as Apache Kafka, Apache Flink, and Google BigQuery for continuous data intake and analysis. Adding to the concept of edge computing can take real-time processing to another level, and the decision-making process becomes faster since data analysis is performed closer to its source.

- **AI-Enhanced Data Cleaning:** This article examines the leading impediments to BI and identifies that poor data quality is one of the primary issues organizations face. Lack of data integration and data quality problems such as incomplete, inconsistent, or duplicate data often resulting these cases that can lead to wrong analysis and wrong Manual data cleaning is very cumbersome and also possesses the disadvantage of being very time-consuming and inefficient. Tools like Trifacta, Talend, and IBM InfoSphere are some examples used for data cleansing, where algorithms are used to correct such data on their own. These AI solutions continuously gain experience out of such patterns and from past mistakes; the quality of data thus tends to rise with time and so does the reliability of business intelligence generated.

D. Future Trends in Business Intelligence (BI)

The generation of data is increasing day by day, and due to this, new technologies are coming up to give BI a face-lift by speeding up the processing part of it, ensuring better security on the data front and thereby making BI analytics scalable. Technologies, including quantum computing, blockchain technology, and edge computing, have been predicted to drastically change data management.

- **Quantum Computing in BI:** BI systems can also greatly benefit from quantum computing since it can solve some problems faster than traditional computing. Unlike classical computers that employ two values, namely '0' and '1', quantum computers use qubits to solve problems simultaneously. This can improve BI systems' data mining, machine learning, and other analytical orientations by processing large datasets in real-time. Appreciable scholarly research efforts from Google's Sycamore, IBM Q, and D-Wave Quantum AI aim to converge quantum computing in data analytics. Despite the current novelty of its implementation, quantum-powered BI is expected to revolutionize industries that would need to analyze large data sets quickly, including, for instance, the banking sector and finance, genetics, and cybersecurity.
- **Blockchain for Data Integrity:** Data accuracy and security are crucial in BI systems since companies mean more and more reliable information in real time. Blockchain offers a decentralized platform that provides a record of the events that happen and the information that is shared in the Blockchain, hence making it hard for anyone to alter the content of the information that has been recorded in the Blockchain. With the implementation of cryptographic hashing and distributed consensus, it becomes impossible to make alterations to business data, improving BI reports' reliability. The finance and healthcare sectors are examples of industries keenly adopting blockchain-based BI to fight fraud, adhere to the laws, and securely share data. For instance, blockchain-based BI can be used by the banking sector for checking fraud and the healthcare sector for managing the patient's health records across the chain. Blockchain technology will enhance the quality of data gathered, analyzed, and processed in BI platforms since occurrences recorded in the blockchain will be accurate and definite.
- **Edge Computing for Faster Processing:** Analytics is no longer an exclusive privilege of desktop computers or highly developed supercomputers; edge computing is vital in permitting BI ahead of IoT (Internet of Things). Legacy BI solutions require cloud installation or are hosted in centralized servers, which can lead to responsiveness problems when working with extensive datasets. Edge computing minimizes the time needed to obtain data since it can be analyzed at the edge level through the devices or servers of edge computing before forwarding only the results to the BI system. This improves operation excellence through real-time decisions and reduces the communication data load. Some of the benefits of edge computing are most prone to some areas such as Retail (Smart shelves to analyse user behaviour), Health care (Patient monitoring), and autonomous vehicles (on-the-go navigation adjustments). Thus, edge computing will become one of the key enablers in BI performance enhancement and reduction of reliance on central data processing centers by organizations.

V. CONCLUSION

This paper presented the concept of Big Data with BI, its strengths, drawbacks, and future inclusion. The result discusses that, with the help of the Big Data BI solution, the organization receives real-time information and business

insights, improved organizational decision-making, and a way for organizational scalability. An extended from the traditional process of business intelligence, new AI and big data integrated BI platforms are used to analyse structured, unstructured, and semi-structured data. However, businesses face issues, challenges, accompanying solutions, and impacts on data management and processing, such as data governance, data security, integration, and data processing speed. It is still important for organizations to pay attention to data quality and ensure information correlation where the institutions get the information needed for planned objectives. Moreover, the available data on a large scale can be efficiently processed with the help of scalable cloud-based infrastructure. Therefore, addressing these factors should be the goal of organizations to use all the opportunities of modern BI solutions in activities.

A. Recommendations

For a business to build a stronger BI system and overcome the challenges currently in BI solutions, it is time for businesses to opt for AI-based BI solutions. Nowadays, advanced AI and Machine Learning (ML) algorithms provide numerous opportunities for better data accuracy, predictive analytics, and automation, resulting in better and more proactive decision-making systems. Moreover, there are some ideas that organizations need to pursue more secure business data, such as blockchain, encryption, and RBAC. Since scalability has become a core component that every modern BI system must possess, businesses need to embrace cloud-based BI to improve the flexibility, cost, and computation of the BI systems. Flexible cloud solutions may also enhance operations, security, and compliance in large-scale big-data processing for organizations through hybrid and multiple-cloud integration.

B. Future Research Directions

There is, therefore, a need to conduct more research on what is emerging as one of the most popular technologies out there. Some are well established, such as, the development and use of quantum computing, which has a tremendous anticipation to revolutionize BI by optimizing data handling and machine learning algorithms in BI systems. Now, understanding the practical usage of the tool for high-speed analytics, financial and numerical modeling, and cybersecurity will be inevitable. Additionally, ethical concerns in AI-driven decision-making need deeper investigation, particularly regarding bias in algorithms, transparency in data processing, and regulatory challenges. This is especially important because it separately an important foundation on which organizations apply artificial intelligence in the business intelligence context to deliver fair and ethical environments through decision-making processes. Therefore, future research should consider the further application of AI and ML in managing data by automating data cleansing and metadata generation, as well as automating compliance checks. Settling these areas will help businesses build effective, secure, and smart BI environments that will correlate to the progressive or dynamic evolution of data environments.

VI. REFERENCES

1. Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. MIS Quarterly, 1165-1188.
2. Kimball, R., & Ross, M. (2013). The data warehouse toolkit: The definitive guide to dimensional modeling. John Wiley & Sons.
3. Inmon, W. H. (2005). Building the data warehouse. John Wiley & sons.
4. Sun, Z., Zou, H., & Strang, K. (2015). Big data analytics as a service for business intelligence. In Open and Big Data Management and Innovation: 14th IFIP WG 6.11 Conference on e-Business, e-Services, and e-Society, I3E 2015, Delft, The Netherlands, October 13-15, 2015, Proceedings 14 (pp. 200-211). Springer International Publishing.
5. Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., & Shahabi, C. (2014). Big data and its technical challenges. Communications of the ACM, 57(7), 86-94.
6. Russom, P. (2011). Big data analytics. TDWI best practices report, fourth quarter, 19(4), 1-34.
7. Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). Strength in numbers: How does data-driven decision-making affect firm performance? Available at SSRN 1819486.
8. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). Big data: The next frontier for innovation, competition, and productivity.
9. Luhn, H. P. (1958). A business intelligence system. IBM Journal of research and development, 2(4), 314-319.
10. Chen, C. P., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques, and technologies: A survey on Big Data. Information sciences, 275, 314-347.
11. Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. Health information science and systems, 2, 1-10.
12. Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. International journal of information management, 35(2), 137-144.

13. Agrawal, D., Das, S., & El Abbadi, A. (2011, March). Big data and cloud computing: current state and future opportunities. In Proceedings of the 14th International Conference on Extending Database Technology (pp. 530-533).
14. Larson, D., & Chang, V. (2016). A review and future direction of agile, business intelligence, analytics and data science. International Journal of Information Management, 36(5), 700-710.
15. Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact**.** MIS Quarterly, 36(4), 1165–1188. https://doi.org/10.2307/41703503
16. Jovanovič, U., Štimec, A., Vladušič, D., Papa, G., & Šilc, J. (2015). Big-data analytics: a critical review and some future directions. International Journal of Business Intelligence and Data Mining, 10(4), 337-355.
17. Sun, Z., Sun, L., & Strang, K. (2018). Big data analytics services for enhancing business intelligence. Journal of Computer Information Systems, 58(2), 162-169.
18. Liang, T. P., & Liu, Y. H. (2018). Research landscape of business intelligence and big data analytics: A bibliometrics study. Expert Systems with Applications, 111, 2-10.
19. Chen, Y., Li, C., & Wang, H. (2022). Big data and predictive analytics for business intelligence: A bibliographic study (2000–2021). Forecasting, 4(4), 767-786.
20. Wixom, B. H., Yen, B., & Relich, M. (2013). Maximizing Value from Business Analytics**.** MIS Quarterly Executive, *12*(2), 111–123.
21. Cherekar, R. (2020). DataOps and Agile Data Engineering: Accelerating Data-Driven Decision-Making. International Journal of Emerging Research in Engineering and Technology, 1(1), 31-39. https://doi.org/10.63282/3050-922X.IJERET-V1I1P104
22. Cherekar, R. (2020). The Future of Data Governance: Ethical and Legal Considerations in AI-Driven Analytics. International Journal of Artificial Intelligence, Data Science, and Machine Learning, 3(2), 53-60. https://doi.org/10.63282/3050-9262.IJAIDSML-V3I2P107
23. R. Daruvuri, "An improved AI framework for automating data analysis," World Journal of Advanced Research and Reviews, vol. 13, no. 1, pp. 863–866, Jan. 2022, doi: 10.30574/wjarr.2022.13.1.0749.
24. Cherekar, R. (2022). Cloud Data Governance: Policies, Compliance, and Ethical Considerations. International Journal of AI, BigData, Computational and Management Studies, 3(2), 24-31. https://doi.org/10.63282/3050-9416.IJAIBDCMS-V3I2P103
25. Cherekar, R. (2021). The Future of AI Quality Assurance: Emerging Trends, Challenges, and the Need for Automated Testing Frameworks. International Journal of Emerging Trends in Computer Science and Information Technology, 2(1), 19-27. https://doi.org/10.63282/3050-9246.IJETCSIT-V1I2P104
26. Cherekar, R. (2020). Cloud-Based Big Data Analytics: Frameworks, Challenges, and Future Trends. International Journal of AI, Big Data, Computational and Management Studies, 1(1), 31-39. https://doi.org/10.63282/3050-9416.IJAIBDCMS-V4I1P107
27. Cherekar, R. (2023). A Comprehensive Framework for Quality Assurance in Artificial Intelligence: Methodologies, Standards, and Best Practices. International Journal of Emerging Research in Engineering and Technology, 4(2), 43-51. https://doi.org/10.63282/3050-922X.IJERET-V4I2P105
28. Cherekar, R. (2023). Automated Data Cleaning: AI Methods for Enhancing Data Quality and Consistency. International Journal of Emerging Trends in Computer Science and Information Technology, 5(1), 31-40. https://doi.org/10.63282/3050-9246.IJETCSIT -V5I1P105