We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

5,900
Open access books available

144,000
International authors and editors

180M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter

Modern Business Intelligence: Big Data Analytics and Artificial Intelligence for Creating the Data-Driven Value

Ahmed A.A. Gad-Elrab

Abstract

Currently, business intelligence (BI) systems are used extensively in many business areas that are based on making decisions to create a value. BI is the process on available data to extract, analyze and predict business-critical insights. Traditional BI focuses on collecting, extracting, and organizing data for enabling efficient and professional query processing to get insights from historical data. Due to the existing of big data, Internet of Things (IoT), artificial intelligence (AI), and cloud computing (CC), BI became more critical and important process and received more great interest in both industry and academia fields. The main problem is how to use these new technologies for creating data-driven value for modern BI. In this chapter, to meet this problem, the importance of big data analytics, data mining, AI for building and enhancing modern BI will be introduced and discussed. In addition, challenges and opportunities for creating value of data by establishing modern BI processes.

Keywords: Business Intelligence, Big Data Analytics, Artificial Intelligence, IoT, Data mining, Data governance

1. Introduction

Recently, in the fourth industry revaluation, there is a very huge amount of created and generated data by computer machine such as GPS, sensors, website or application systems or by people through social media (twitter, Facebook, Instagram, or LinkedIn) [1]. Every moment, the data servers store huge amount of data which are produced by organizations. This is a huge amount of data comes from website, social media, tracking, IoT applications, sensors, and online news articles. Also, the advancement in computing and communication technologies have facilitated collecting a large volume of heterogeneous data from multiple sources. This data consists of structured and unstructured, complex and simple information.

Currently, business gets a revenue from the analysis of such data with unstructured form up to 80% [2]. So, the organization can improve the business productive process due to this analysis of unstructured data that contains valuable information. In addition, it is significant for education, security, healthcare, and manufacturing.

This can be achieved through big data analytics, artificial intelligence, and data management in order to achieve the business intelligence.

BI is the technologies, tools, systems, and applications for the compilation, analysis, combination, and exhibition of the business report with active business decision executing way. This way will give unlimited help to gain, learn and control their data to further decision making for developing business processes and procedures [3]. Also, BI can be described as the ability of a firm to make meaningful data using which is collected every day from business processes and operations [4].

Business intelligence (BI) plays an importance role for helping the decision marker to get the insights for improving productive or better and fast decision. In addition, BI can enhance and assist the effectiveness of operational rules and its impression on corporate-level decision-making, superintendence system, administration, budgeting, and financial recording which gives better strategic alternatives in dynamic business environments [5]. Also, BI can improve the organizational performance by identifying new opportunities, revealing new business insights, highlighting potential threats, and enhancing decision making processes among many other benefits [6, 7].

The first issue in business is big data management with various data formats which is the serious management problem due to that the current tools are not adequate for managing such massive big data volumes [8]. The new challenges in terms of data integration complexity, storage capacity, lack of governance, and analytical tools gives an importance for solving the big data management problem related to pre-processing, processing, security and storage. The big data management in abundant data, generated by heterogeneous sources for using in BI and decision-making, is a complex process. Therefore, some form of big data may be managed by 75% of organizations. The goal of managing big data is ensuring the effectiveness of security, storage, and analytics applications of big data [9].

Unfortunately, the practical implications using big data analytics in enhancing business intelligence remains comparatively immature and under-researched because the existing research models are mainly focused on the benefits and challenges of business intelligence and big data. So, the most important issues are studying the implications of big data analytics on business intelligence in for data collected from various sources and exploring the future directions to find further developments in use of big data analytics for business intelligence.

The second issue in BI is the determination of the most appropriate data mining technique, which is one of the most critical responsibilities. Based on business nature and difficulty suffered or object kind in the business there is a need for determining the data mining optimal technique [10]. In the data mining process, most core techniques identify the character of the reclamation option of data and its mining process. Based on the results, the data mining technique will be highly productive [11]. There are many data mining techniques as association rule, clustering, classification, decisions tree, and neural networks are profoundly successful and practical.

Data Mining is regarding to interpret the huge volume data and extract knowledge from its various objects. For some businesses, the purposes of data mining are recognized to identify different trims, develop marketing abilities, and predict the prospect based on earlier observations and modern inclinations. There is a requirement for examining the data for sustaining divestments and additional purposes of an entrepreneur. Furthermore, data mining could continue practiced for recognizing unusual performance and a strange behavior of representatives practicing on some technologies could be identified [12].

The third issue in BI is artificial intelligence (AI). AI is the main step in the technology evolution that has been actively pursued since British mathematician

and code breaker Alan Turing envisioned a clear way forward in his groundbreaking 1950 paper, "Computing Machinery and Intelligence." At the time, computer technology could not keep up with Turing ideas. But, due to the advancement in computing, AI was established. At Oxford University, the Future of Humanity Institute introduced a 2018 report for surveying a panel of AI researchers on timelines for Strong AI. This report found that in 45 years, 50% chance of AI will outperform humans in all tasks and in 120 years it will automat all human jobs. As well as, AI will bring many opportunities for creating new jobs. Also, removing the need to do tedious and repetitive tasks is one of the great values of AI, as many experts said. Instead, users can focus on their main skills and values. For reducing human error, shrinking labor costs, and subsequently increasing profit, the application of technology in many industries and business has been aimed. This was true for the advancements made during the fourth Industrial Revolution (FIR) on through to the birth of the computer, and still true for the era of AI.

In this chapter, the importance of big data analytics, data mining, AI for building modern BI and enhancing will be introduced and discussed. In addition, challenges and opportunities for creating value of data by establishing modern BI processes.

2. Business intelligence (BI)

Business Intelligence (BI) can be described as an automated process for deriving models and insights form raw data that are collected from heterogeneous data sources and are organized in a systematic way for improving business operations and processes. In enterprise BI architectures, the best practice is splitting the data collection and data organization processes that are associated with back-end architecture from data analysis and display to a user through the frontend. In BI, the processed transactions generate data, which are stored in Operational Data Sources called Online Transaction Processing servers (OLTP). With OLTP, the data is stored in a structured data repository called data warehouse after extraction and transformation processes. With data warehouse, there are different query optimization techniques can be applied for speeding-up of data analysis and running the analytics query. To achieve this speed-up, data warehouse creates subsets of the data warehouse called data marts. Also, reporting mechanisms for accessing transaction data stored in data warehouse are used in traditional BI systems. Therefore, analyzing these transaction data can help us for detecting patterns and predicting business trends.

Recently, the data sources of BI are not only traditional data sources as transaction data, but they include modern data sources as mobile devices and sensor data, and web messages which were sent by company intranets and profiles of employees and customers. Most of modern data sources are unstructured, for example, posted messages in online social networks (OSN) and data from various sensors. Therefore, the main challenge is how to maintain these modern data sources as traditional relational database and achieve query efficiency. From the data analysis perspective, additional data means additional opportunities for discovering more insights. However, the big data challenges remain the big problem from the analytic perspective.

Due to the increase in data, there are expanded opportunities within the scope of BI, which is not only a mechanism to analyze historical data trends, but it can combine data from sensors and other real time personal information for inferring insights that are not commonly available that is called situational BI [13]. For business operations, BI is called operational BI, which provides insights in real time to
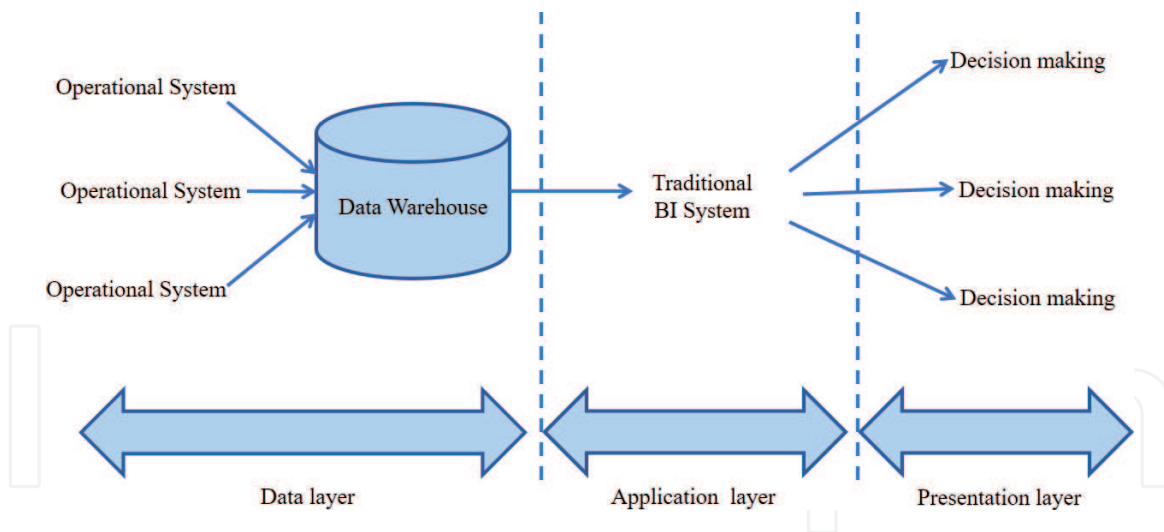
Figure 1.
A traditional BI system.

these operations as getting instant feedback for a call center operation as benefits from their work. In addition, the analytics rules may be composed depend on meta-information of the exposed data to his/her which can be considered as a self-service BI. Therefore, these new BI approaches must be managed carefully such that the compliance models and governance of enterprise are not violated.

The three-tier architecture of traditional BI system is shown in **Figure 1**. This architecture consists of three layers: 1) Presentation layer, 2) Application layer, and 3) Database layer. The main challenge with this three-tier architecture, is how to fulfill service level objectives such as minimal throughput rates and maximal response time. This is because, the data storage management at the low-lever layers is hidden from the application layer which makes some difficulties to predict execution times.

However, traditional BI systems are efficient in extracting and analyzing data, but they are rigid, slow, time-consuming, and requiring knowledge experts for maintenance. Therefore, many research works have been done for adding modern features to improve the three-tier architecture, which will establish the next generation BI.

3. Modern business intelligence (MBI)

In the traditional BI platforms, the main goal is giving answer "What happened?" Question by providing the efficient analyses. While, the BI modern platforms are giving the answer for "What is happening, what will happen, and why?" which offers the ability to monitor and obtain a continuous development of organization within fast analytics, while for accomplishing objectives of mission using predictive analytics.

Traditional business intelligence platforms over the past two decades have mainly succeeded to provide users with historical comprehensive reports and easy-to-use custom analysis tools. Due to the underlying data architecture, which consists of a central data storage solution such as an enterprise data warehouse (EDW), the availability of BI functionality is largely. EDWs form the backbone of traditional data management platforms and usually connect vast network systems of data source into a central data warehouse. The data is then consolidated, refined, and pulled into different reports and dashboards after converting data in EDW to display old business information, such as weekly revenue metrics or quarterly sales.

Although, this traditionally BI provides a basis for these types of dashboards and interim reports.

While users have gained immense value from traditional platforms for historical reports capabilities, there are more users now require data analysis technologies that need direct access to data without depending on IT professionals. Federal agencies highlighted the following challenges associated with traditional BI solutions in analytics [13]:

1. *On-Demand Analysis Capabilities Lacking*: advanced users of BI today do not need to wait for answers to more business complex problems. Additional users need capabilities of self-service for linking and analyzing specific datasets depend on their own understanding, for any purpose, and at any time.

2. *Needed Predictive Analyses*: Historical reports capabilities provide just one puzzle piece: an insight about what happened in the past. Companies look to predictive analytics or insight about the future to forward thinking and truly be driven by data. With predictive models, the companies can use patterns and forecasts to get next actionable steps using their data.

3. *Mixed Data Types Analysis:* Traditional BI platforms have largely focused on structured data, but today users require the ability for viewing and analyzing semi-structured, unstructured data and third-party data. In recent years, the massive number of produced information has increased, partly due to new data mining technologies, the Internet of Things (IoT), the proliferation of data sensors and automated data collection tools. Now, advanced BI users and data scientists need access to unutilized data in different formats to mix data types and create their own algorithms, where on demand insights are available to make accurate and quick decisions. A lot of organizations that lack the processes, technology, and people needed to extend data-analysis capabilities to the next level become frustrating. These challenges need a strategy and platform for analytics that goes faraway the traditional BI platforms scope, as shown in **Figure 2**.
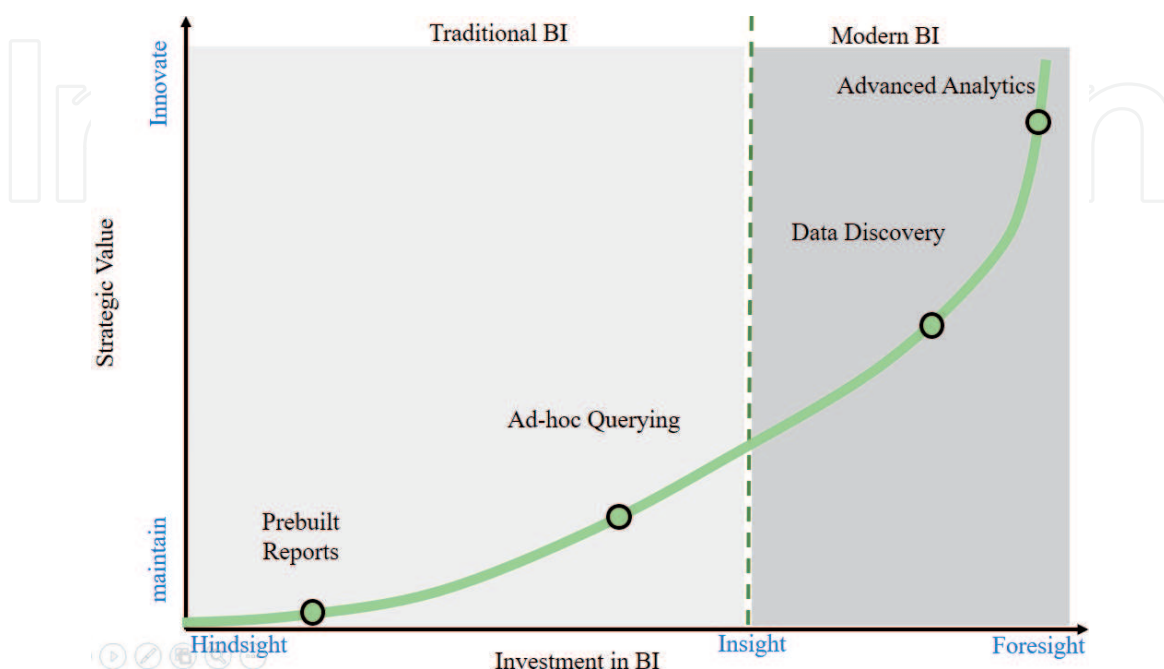
Figure 2.
Grows of BI platforms based on insights: Hindsight to insights to foresight [14].

Integration of traditional and modern BI Platforms is essential to laying the groundwork for enterprise-wide data transformation and organizations are truly concerned for getting rid of IT infrastructure and starting over. Data warehouses play a major role in existing data platforms, which provide the data that fully cleaned, organized, and managed for most businesses and companies. The data warehouse gives business managers, executives and others ability to obtain insights from historical data with relative ease without deep technical knowledge. The obtained data from data warehouse is very accurate due to careful testing, IT cleaning, and accurate knowledge of data layers. However, traditional BI challenges create a demand to increase EDW with different form of optimized architecture for fast access to ever-changing data: Lake Hadoop Data.

Organizations look to upgrade their platforms of analytics are beginning to adopt the data lakes concept. Data lakes store information in its raw and unfiltered form, whether structured, semi-structured, or unstructured. Unlike the stand-alone EDW, the data lakes themselves perform little of the automated data cleaning and transfer operations, allowing data to be swallowed more efficiently, but they transfer the greatest responsibility for preparing and analyzing data to business users.

Data Lakes can offer a low-cost solution by using Hadoop's Distributed File System (HDFS) for efficiently storing various types of data and analyzing them in their original structure. As shown in **Figure 3**, a data lake coupled with the data warehouse to identify the next generation of BI and provide the optimal basis for data analysis.

In the system shown in **Figure 3**, EDW receives system data from different sources through the ETL process (Extract, Transform, and Load). After the data is cleaned, transformed, and standardized, it will be ready for analysis by a diverse group of users using dashboards and reports.

In the interim, a data lake collects raw data from single or multiple source systems or all systems, and the data is absorbed and ready for discovering or analyzing processes. The result: a broader user base for exploring and creating relationships between vast amounts of various data for individual analyzes, upon request.

3.1 Features of modern BI

1. *Operational BI (real-time):* Today, the competitive pressure of businesses has increased the requirement for almost real-time BI, which called operational BI. The operational BI goal is reducing latency between data analysis time and data acquisition time. Reducing response time enables the system to take suitable action when an event exists. With operational BI realization, companies can discover patterns or time trends across flow of operational data.

2. *Situational BI:* it enables situational awareness. In companies, BI positioning is important where a rapid turnaround in positions, commonly external business trends, has affected business [15]. However, this external data, which mostly comes from the intranet of company, external vendor, or the Internet, is unstructured. Moreover, this unstructured data must be combined with other structured data from the local data warehouse of the company for supporting real-time decision marking. For example, the company may want to know if its users and customers are posting negative or positive comments about their new products. Through the analysis of these comments, companies can provide immediate comments to the development team for making the product more competitive and qualified. Another example is important for a company
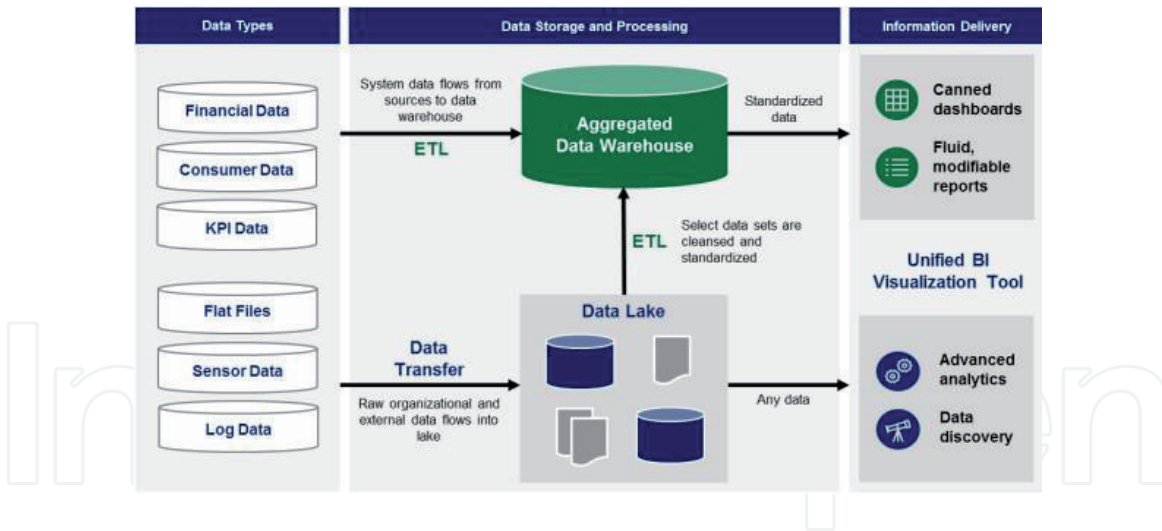
Figure 3.
Data sources, data warehouse, and data analytics in modern BI [14].

to know whether natural disasters have affected their contract suppliers. Recognizing natural disasters and enable businessmen to take appropriate measures to reduce losses [16].

3. *Self-Service BI* (SSBI)*:* it enables end users for generating analyzes and analytical queries without involving of the IT department. In SSBI, the user interface of applications must be easy to use and intuitive, therefore technical knowledge of the data repository is not needed. In addition, the user should be allowed for accessing or expanding data sources organized by IT, but also non-traditional sources.

3.2 Data architecture

1. *Background:* Traditional business application architecture has three layers: data, application, and presentation. In the three-tier architecture, execution time is very difficult for predicting, due to the relationship between processes of low-level data management and operations of high-level. Usually, workload management solutions are built on top of general-purpose DBMS, which need time delays for executing parallel requests. With modern business applications, this will create challenges for functions as operational information in real time. Therefore, technologies that enable simultaneous business transaction and analytical queries to be performed on the same data are important. Organizations today use the ETL to extract data, make transfers, and upload data that is converted into a data warehouse. This model is based on two types of business process critical processes: Online Analytical Processing (OLAP) and Online Transaction Processing (OLTP). OLTP is used for managing business operations, such as processing of an order. OLAP is used for supporting strategic decision making as sales analytics.

2. *Challenges:* Traditionally, OLAP and OLTP workloads are performed on the same database system. However, workloads of OLAP mostly consist of bulk reads on data only that is updated by OLTP, constantly. Therefore, transaction-processing performance may be unexpected due to competition for a resource when both workloads are performed in a single database. Thus, it is necessary to separate workloads from OLAP and OLTP. **Figure 4-a** describes ETL-based background information where OLAP and OLTP are separated.
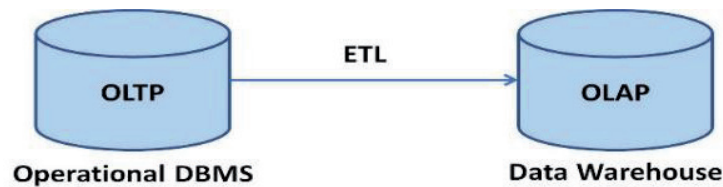
Figure 4.
Classification of database systems.

In this architecture, each workload of OLAP must wait until the data in the date pool is completely refreshed and visible which will cause delays. Today, for reducing the delay, BI operating systems execute OLTP and short-term analytical queries together on the DBMS, as shown in **Figure 4-b**. These workloads are called short OLAP workloads. However, long-term OLAP workloads may be conflicted with many short OLTP transactions that make changes to the database. So, high synchronization is needed to deal with resource competition, which produces lower utilization of all resources.

Also, the commercial database management system (DBMS) uses special techniques as shadow copy [17], for handling mixed workloads with lower over-heads. That is, on different logical versions of the data, different workloads will be separated and performed. Therefore, the additional space may be increased, which increases the infrastructure costs and requirements. Therefore, in current disk-based DBMSs a major challenge is managing these mixed workloads (OLAP and OLTP) [18].

3.3 Current BI systems

1. *Extended systems of traditional BI*: Current traditional BI technologies can perform OLAP queries and OLTP transactions on the same database without interfering with each other. Combining these mixed workloads with the same system needs extreme performance improvements due to the huge explosion in dynamic data size.

- "In-memory database (IMDB)": Today, in most BI systems, OLAP and OLTP mixed workload on a one system can be handled using an In-Memory database (IMDB) (also, called Master-Memory). This technique needs that the system stores all data in the main memory, because it is faster than the optimized databases on disk and the internal optimization algorithms use less CPU instructions and are simpler. In case of querying data, this technique provides more predictable and faster disk performance by reducing the time of search. However, the IMDB systems can lack durability because of stored information losing when the device is reset or loses power. Many IMDB systems have proposed various mechanisms for supporting durability such as snapshots, non-volatile DIMM, non-volatile RAM, transaction logging, and high availability.

System	Type	Methods	Achievements
H-store [20]	IMDB	Distributed, row-store technique	High OLTP throughput rate
Radu Stoica [21]	Hybrid	Data reorganization	High performance, reduce paging I/O, and improve memory hit rate
Siberia [22]	Hybrid	Cold data access and migration mechanisms	Acceptable access rates with 7–14% throughput loss

Table 1.
Systems of modern BI that use various methods to hold most or all the data in the main memory.

Table 1 shows systems of modern BI that use various methods to hold most or all the data in the main memory for obtaining high OLTP productivity. For example, a distributed set of shared devices is used to run H-Store system, where the data is completely located in the main memory. The H-Store can execute transaction processing at high productivity rates, by removing traditional DBMS features as buffer management, lock and close. Recently, the H-Store prototype was marketed by a startup called VoltDB [19].

- "Hybrids with on-disk database": The main-memory has become big enough for handling most OLTP databases, nevertheless this may not constantly be the best choice. For OLTP workloads by using access patterns, where some records are "cool" (rarely or not accessed at all), others, "hot" (accessed frequently). So, the coldest records are stored on fast secondary storage devices in the modern systems to ensure good performance. For example, Stoica and Ailamaki [19] suggested a way to migrate primary memory DB data to cheaper and larger secondary storage. In [20], for improving major memory heart rates and reducing I/O operating system migration, relational data structures are reorganized using access statistics for workloads of OLTP. Recently, Siberia was introduced as a cold data management framework in Microsoft Hekaton IMDB [21]. Like [19], it does not require storing an entire database in the main memory.

Hekaton focuses on how records are migrated to and from a cold store and how records are accessed and updated in a cold store in a consistent manner for transactions. So, only some tables can be declared and managed in the main memory by Hekaton. Experience evaluation shows that when cold storage is located on commodity ash, Siberia can lead to an appropriate productivity loss of 7–14%, given that cold data access rates are for an improved main memory DB.

2. *Modern features* of *BI Systems:* There are three modern information survey indicators: operational biological investigation, situational temporary survey, and self-service self-examination. Whereas, the H-Store system is only for OLTP transaction processing, a modern system called HyPer can handle it mixed workloads of both OLTP and OLAP are extremely high throughput rates using a low-overhead mechanism to create differential shots [22]. This system is used an unlocked approach which allows all OLTP transactions to be carried out in sequence or on special sections. In parallel with OLTP processing, HyPer system performs OLAP queries on the same shot and consistent.

Castellanos et al. [23] proposed a new platform called to notify business managers to situations that could affect their business. SIE-OBI integrates the functions

required for exploiting relevant rapid flow information from the web. They proposed new schemes for extracting and linking information that obtained from the web with the stored historical data in the data warehouse to reveal position patterns. The relevant information is extracted only from two or more different unstructured data sources, usually one stream of internal slow text and stream of external fast text. This time and effort minimization platform were built to build slow and fast data streams that integrate structured and disorganized flows, and to analyze them in almost real time.

3.4 Data governance

1. *Background*: in DAMA I [24] data governance is defined as "the exercise of authority and control over the management of data assets, planning, supervision and control over the management and use of data". Data governance describes the responsibilities and roles of the organization in promoting desired behavior in the use of data [25]. Data management differs from data management, which involves setting data quality standards, making decisions and implementing them [26]. It is also different from BI Governance, which aims to provide a dedicated decision-making framework through the governance of all activities within the BI environment [27]. DAMA I [28] identifies 10 data management functions as shown in **Figure 5**. The data management function is high-level supervision, planning, and control of all other functions. There are four data management functions related to the next generation of biological information that requires fast access to data, external data utilization, and analyzing data by users, generally. Data architecture management includes setting of data standards, maintaining, and developing structures of enterprise data and linking application projects and architecture. The department of data quality focuses on planning, implementing, and controlling activities that apply techniques of quality management to measure, evaluate, improve and ensure the use of data. Data storage and business intelligence management focus on providing decision support data for reporting, query and analysis. Metadata management focuses on activities to enable easy access to high-quality metadata, such as architecture, integration, control, and delivery.

2. *Deploying Next Generation BI in Data Governance*: Data management has become vital for the organization as the data becomes inherent. The business derives its business value and decides based on the information derived from the data. Consequently, data control is required to ensure the quality of the data that directly affects the quality of the decisions made by the organization [29]. More effective data governance (DG) can lead to a higher scale of decision-making. To achieve effective data governance, maturity models of enterprise data governance help to understand DG and to determine what the next expected plan is [30]. Many data management maturity models [31] have been proposed for directing an organization to understand what data management level is. In [29], Oracle anticipated that the maturity model of data governance would help the organization in locating it in its data governance system evolution, identifying steps of short-term needed to reach the next level, and enhancing capabilities of data management. In the Oracle model, the highest level of maturity is the integration of data management with BI.

The next generation of BI supports almost insights of real-time with using of external information that generates a large data amount and its manipulations. So, this requires very mature DG for providing data quality, reliability, and integrity.
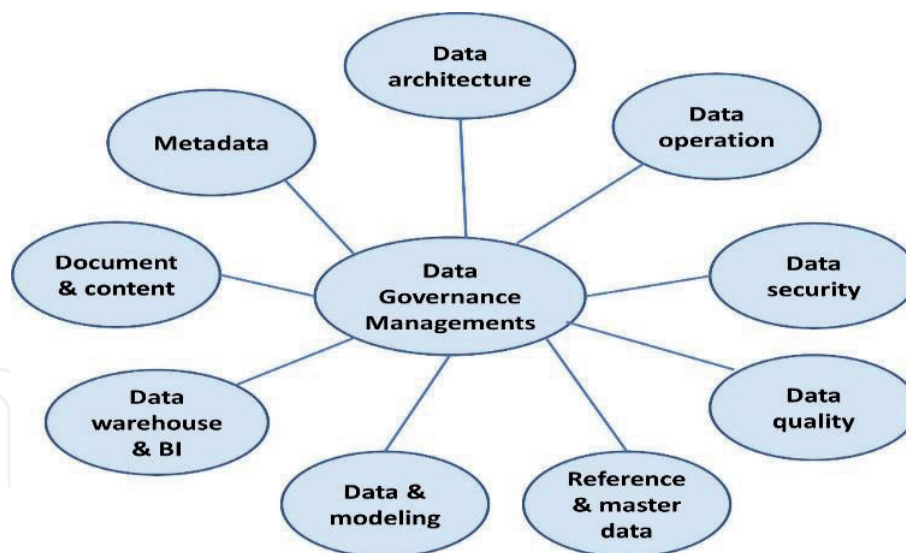
Figure 5.
Framework of data governance as defined in Dama I [27].

The three characteristics are crucial to extract accurate insight through techniques of data mining. For example, in a "self-service" BI (for example Tableau and QlikTech), allows users to discover insight from many data sources without modeling the data environment and implementing complex ETL operations, which is one of the most time-consuming and difficult tasks in BI. So, these new features allow users to easily access data, get quick results and visual data visualization. To enable the evolution of the next generation of biological information, data management is critical to the reliability of data from the discovered vision. For example, in the case of BI self-service, the fact that end users can access and process their data reduces the reliability of BI results [32]. In data management, useful functions to ensure reliability can be considered such as tracking data ratios to source and creating records of how data is processed or transferred. However, integrating data governance into the next generation of biological information has faced some challenges due to the requirements of flexible and reliable responses while there is an enormous amount of external data and public user engagement.

3. *Data governance challenges*: There are two main advantages to the next generation of biological information that affects the data management model. Decision making in the next generation BI, should be more effective and faster between a huge amount of data that comes from many data formats and sources. However, data from many sources makes data management more difficult to manage and sophisticated to control properly. This can also lead to ineffective decisions being made. In case of data comes from different conflicting sources, more research and analysis of the data and the different sources of that data to determine what is true and accurate or its approximation must be done by the decision-maker, which will be costly operations. Therefore, management of data across heterogeneous sources in the next generation BI system is very important. In the next generation of personal information, especially "self-service" business users participate in procedures of decision-making.

In general, the central IT organization and many data supervisors have been involved in data management initiatives and have a metadata repository for the data management platform and a set of data management tools to deal with varied data. In advance, they standardize common data definitions of master data and reference data that are widely shared across many enterprise applications. When they receive

disparate data, they match it to define predefined shared data, determine its quality, determine which rules, convert, and merge them. However, in the next generation of BI, users also select, manipulate or merge their data names themselves using various "self-service" tools of BI. They may want to upload to the DB and share their vision with others. Participation of business user in the data process can lead to data in a mess where the same data can be converted and combined in various ways through data managers and a central organization using tools of data management and by business users who have tools of BI for "self-service". Consequently, meta-data sharing criteria are crucial through this sharing to transfer shared data, shared data names, and shared integration rules [33].

4. *Data governance model for next generation BI*: The data management model design is designed to centralize versus. Decentralization and hierarchy versus cooperative. Central design assigns all decision-making authority in the central IT department while decentralized design assigns authority to individual business units [25].

The term big data is a group of huge and complex data sets from various sources where data the management and traditional application processing techniques face difficulties to process it. Big data is a collection of a large amount of structured or unstructured data that is processed and analyzed for informed decision-making or evaluation. These data can be taken from various sources including browsing history, geographic location, social media, medical records, and purchasing record. Big data is made up of complicated data that will smash the processing power of traditional simple database systems [34]. In [35], the authors mentioned that, there are three main characteristics associated with big data: (1) Volume is a feature used for describing the vast data amounts that big data uses. Usually, the range of data amounts starts from GB to YouTube. Big data should be able to handle any data amount even with its highly anticipated growth. (2) Variety is a feature used for describing various types of data sources that are used as portion of a large data analytics system. Currently, there are many data storage formats used by computers all over the world. One format is the structured data such as databases and. Csv, video, short message service (SMS) and excel papers. Unorganized data can be in the handwritten notes form. All data from these sources will be ideally used for Big Data Analytics. (3) Velocity is a feature used to describe the speed at which data is generated. It is also used to describe the speed at which generated data is processed. With the click of a button, an online retailer can quickly view big data about a specific customer. Speed is also important to ensure that data is updated and updated in real time, allowing the system to perform at its best. This speed is necessary as real-time data generation helps organizations accelerate operations. Which can save institutions a large amount of money.

Today, many companies are increasingly interested in using technologies of big data to support their BI, so that it becomes very important to understand the different practical issues from previous experiences in BI systems. Today's BI systems sense the world and harness these data points for recommending the best possible options and forecast results, accurately. As BI systems continue to be built in real time, the demand for data collection, integration, processing, and visualization increases almost in real time. BI systems are characterized by high sensitivity opportunities as seen in sensors with the rich diversity of sensors ranging from mobile phones, personal computers and health tracking devices to technologies of Internet of Things (IoT) designed to give contextual and semantic sound to entities that could not previously contribute Intelligent in key decisions. So, many companies are analyzing big data today.

Big data analytics is needed and is machine learning techniques because of often distributed data sets, and its privacy and size considerations are evidence of distribution techniques, where data is on platforms with different computing capabilities and networks. The benefits of application diversity and big data analytics pose challenges. As an example, every hour the servers of Walmart handle more than million transactions for a customer, and this information is stored into databases that contain larger than 2.5 petabytes of data, which is 167 times the number of books in the Library of Congress. Herein, CERN's Collider Hadron Collider produces around 15 petabytes of data annually, and that is enough to fill over 1.7 million double layer DVD discs annually [36]. Big data analytics are used for education, health care, media, insurance, manufacturing and government. Big data analyzes of business intelligence and decision support systems that enable healthcare organizations to analyze data size, diversity and tremendous speed have been developed across a wide range of healthcare networks to support evidence-based decision-making and action [37]. Therefore, it is clear from the discussion that data management and big data analytics [38] are important in BI for 4 reasons:

1. *Better-decision-making (BDM)*: Big data analytics can analyze current and old data for making predictions about the future. So, companies can make not only better current decisions, but also preparing for the future.

2. *Cost reduction (CR)*: Big data technologies like cloud-based analytics and Hadoop offer great cost advantages when storing large data amount. In addition, he provided insights on the effect of various variables.

3. *New products and services (NPS)*: With the ability to measure needs and satisfaction of customers through analytics, the strength comes for giving customers what they want. So, more companies are creating new products and services to meet customer needs.

4. *Understand the market conditions (UMC)*: By analyzing big data, we can get a better understanding of current market conditions for retrieving important information. In addition, there are a few features and challenges that must be considered in the tools and techniques of big data analytics, and they include scalability and fault tolerance as well [39–41]. The following **Table 1** represents a few of the widely used tools with the advantages of Big Data Analytics.

The rapid development of business intelligence and analysis attracted the attention of researchers. The reason is that organizations no longer rely on traditional technologies as data grows exponentially. This huge amount of data requires advanced analytical techniques in order to convert it into valuable information that helps organizational growth. BI&A is the contemporary methodology for extracting value from this vast amount of data, driving strategic decision-making, and forecasting and benefiting from future opportunities.

BI&A is necessary in most organizations. BI&A has proven effective support in decision making. In addition to that data and IT infrastructure is clearly influenced by the good use of BI&A practices. Nowadays, business intelligence and analysis have played a vital role in most institutions and sectors due to their value and benefits. BI&A helps organizations gain a better view of their private data and thus improves fact-based decision-making. These methodologies and data analysis also help to maintain competitive advantage in addition to resolving technical and quality problems that will enhance the performance and productivity of enterprises [42, 43].

According to Abai et al. [44] BI&A helps to build an integrated framework that supports speeding up organizational performance. Many factors and technological developments have shaped the past and present trends of BI&A. With the rapid development of technology, it is not enough to use traditional analytical techniques. The future direction of business intelligence and analysis will expand to include areas of diversity. According to Chen et al. [45]. The success opportunities associated with data analysis technologies have generated future interest in business intelligence and analytics. Additionally, BI&A contains different practices and methodologies that can be applied to different sectors; Health care, security, market intelligence, e-government, and others. According to Mohammed and Westbury [46] BI&A is contributing to future development systems. By mapping all the facts, BI&A has become soon biotechnology in developing cities by supporting real-time information that will turn countries into smart cities.

One of the most important responsibilities in the data mining process is choosing the appropriate data extraction technology. The nature of work and the type of object or difficulty experienced by the work provides appropriate guidance for identifying the best techniques [47]. Application of data mining techniques There are some generalized approaches that can indicate enhanced efficiency and cost-effectiveness. Many of the basic techniques that are performed in the data mining process, determine the nature of the mining process and the option of data recovery.

Artificial intelligence (AI) represents a step in the evolution of technology that has been actively pursued since the British mathematician and code-breaking Alan Turing was conceived as a clear way forward in his pioneering research of 1950, "Computing and Intelligence." At the time, computer technology could not keep up with Turing's ideas. But as computing advanced, Amnesty International advanced. Most of the artificial intelligence that we see today is narrow artificial intelligence (ANI), which means it can perform a well-defined task. A 2018 report by the Future of Humanity Institute at Oxford University has surveyed a group of AI researchers in the schedules of strong AI. She found "50% chance of artificial intelligence outperforming humans in all tasks in 45 years and automating all human functions in 120 years." However, AI will bring with it many opportunities to create new business opportunities as well. As many experts have pointed out, one of the great values of artificial intelligence is its ability to eliminate the need for strenuous and repetitive tasks. Alternatively, users can focus on their core values and skills. Technology was applied in many industries mostly aimed at reducing human error, reducing labor costs, and thus increasing profit. This was true of the progress made during the Industrial Revolution until the birth of the computer, and still true of the emergence of artificial intelligence.

Artificial intelligence has advanced significantly in the past few years due to a number of factors, starting with a massive increase in the computing power available. The once-trained AI model now takes days or even hours with machine learning (more on this soon). Another factor is wider data access. You may have heard that the data is "new oil" or something similar. However, the data must be processed using advanced tools such as analyzes and machine learning algorithms to reveal useful information. This processing is where the AI in BI becomes an invaluable tool.

Machine learning is the engine of artificial intelligence systems. It strengthens artificial intelligence models by analyzing complex data sets. Machine learning enhances models by analyzing complex data sets through a set of self-acquired rules and knowledge as shown in **Figure 6**. The machine learning model learns from big data and from frequent human interactions so that it can provide information and answers related to the user's interests or goals. Big data refer to very large data sets
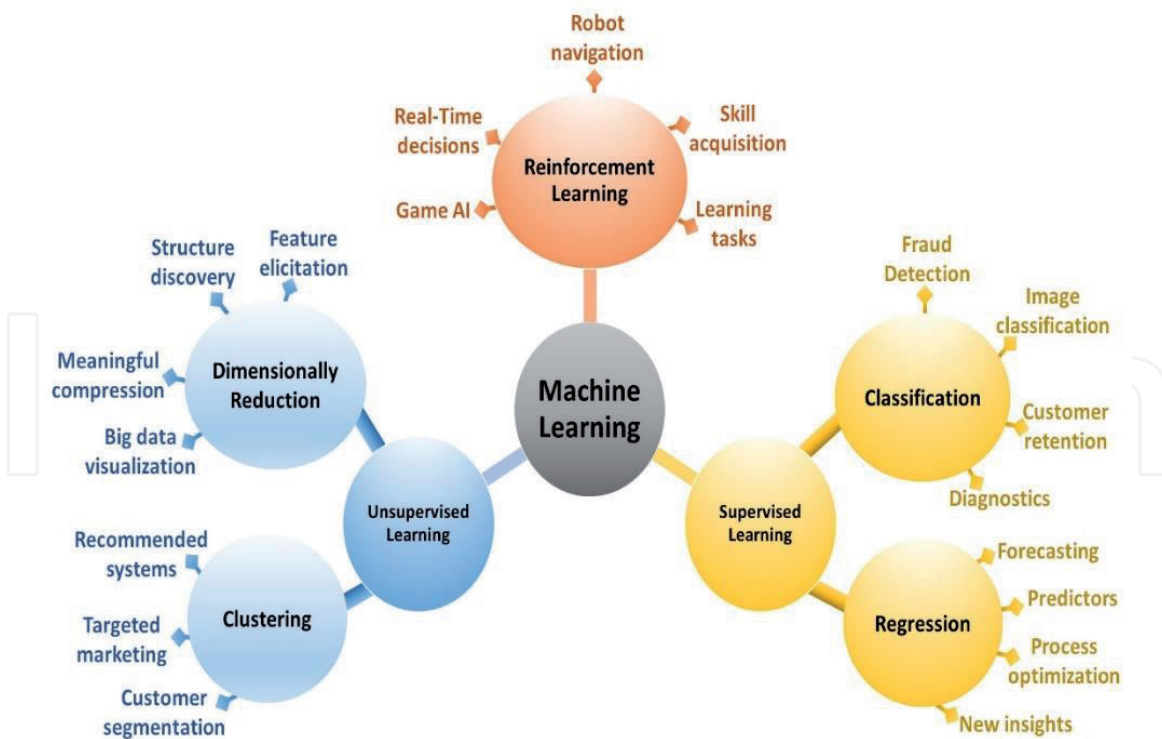
Figure 6.
Machine learning model.

that can be mathematically analyzed to reveal patterns, trends, and correlations, especially about human behavior and interactions. In the space of artificial intelligence, deep learning represents a major leap forward in technology. As we just touched on, programmers write a code that directs the device how to interpret a series of words, pictures, or commands to reach a decision and execute an order. The end user then introduces the entry (data), while internal engineers may define more specific rules for interpreting and analyzing that data. Finally, the system provides outputs (analysis) based on the specific inputs and defined rules. In [48], the authors proposed a demand-forecasting model by BI with machine learning.

3.5 Why BI needs AI?

Does it matter if the constant awareness of the original, or is the copy going to be alive anyway 19? For better or worse, the future comes faster than we realize. There will be no before or after artificial intelligence but a slow transition for a decade or more. As we have seen with Google Glass, it's currently impossible to guess what acceptable results would look like. But how much can we trust in our future assistants? Will they work with us or unknown entities? If we do not ask the right questions now, we'll get the default app. It will be free, but what will include small prints? Good morning John. Here's today's program. Any questions? Perhaps it does not matter after all: Using a good learning algorithm, the program will know what we need and what we need to do, better than we can ever guess. The power of statistics will win the war against the gods and we will lose our soul. It is known that job candidates can lose their chances in a decisive way when they think that no one is watching by bad behavior or rejection of reception staff and waiting staff. Once NLPs and other AIs are widespread, it will not be long before the same literature test is introduced. Looking at 2050, the future of humanity lies in the transition to a civilization of the first kind. We are type 0, extinct. We are about to become half-gods. Most likely, we will merge with our own processing technology and each of us will have our own virtual world to dominate it with absolute control in every aspect

of it, and the countless millions of planets of "life" that we may control or merge with as well. Just as video game programmers have absolute control over the worlds they create. Immortals, omniscient and omnipresent, are all capable of our universes. Of course, he can explore this universe as well, maybe contact directly with his creative being and know that we are characters in his game. Our last question will be morality and maturity. Will we only have one universe? Or does force drive us into madness and transform us into "invaders of the universe" and penetrate the universes of others, based on greed, against the desire for more force? Will we be good? Or evil? Or both? Will we be able to achieve wisdom, and secure peaceful and harmonious coexistence with all other demigods, or will we go to war? Or will we merge into one excessive force? Or are we tired one day from the divine and start the final game again, and transform ourselves into a universe that we will have to evolve for billions of years for us to be re-created one day? Maybe this is exactly what is happening.

3.6 Improving BI with AI

In this section, we explore how BI's AI raises and improves the way of an organization that are used for analysis and interpretation the lifeline of its business.

1. *Turning Business Users into Data Experts (TBUDE):* Typically, business analysts (BA) and IT officials control the access to data and its interpretation. Although these occupations are crucial until now. With the AI tools in today's BI tools, including LOB, NLIs users no longer need to depend on data science experts for analyzing their data. AI allows users to obtain actionable answers easily and directly for helping "democratize" data. In other words, it gives users a two-way conversation ability with their data and feel empowered for acting with answers in a reliable way. Here is an example of how AI works in practice: a certain organization is deploying a BI solution that uses an advanced NLI and instead of waiting for system administrators or data scientists for analyzing data, the manager of business unit arrives at the BI solution directly. The manager makes the data available by calling or downloading and asks questions in simple language. Then, the user receives insight into these questions along with a dashboard and visuals ready for presentation to help communicate these answers. A pre-trained model of AI can target even specific tasks of BI such as visualization recommendations, scenarios of "what-if", and prediction for helping managers to make important decisions for their business.

2. *Helping You Explore Your Data (HYEYD):* There is something inherently satisfying to explore your data with the right tool of AI that supports artificial intelligence. In minutes, you can move from loading data sets to revealing hidden facts in the data and introducing these results in beautiful visualizations. At a starting moment the data is available, the AI in the system of BI heavy lifting by sorting automatically, marking columns and joining matching data across groups. Accessing the NLI is the first step in data exploration for the user. The AI tool will suggest questions that might be helpful if you get stuck. You can also start with the basics, like "How did the retail store department perform during the X period?" The AI will provide answers and suggest ways to explore data to get additional insights into performance. Exploration is exciting because you can continue to delve deeper into visions that only AI can achieve. What embodies the imagination of users is imagination. Visuals are an essential feature of all modern BI solutions, but with AI enabled AI solutions, users receive suggested, automated visualizations that best fit the answers to their questions.

3. *Learning from the End User (LFEU):* The leading AI systems in BI systems are customized and improved all times through machine learning that indexes and learns traditional questions and behaviors of a user. The more user interacts with the tool of BI, the better the AI will know what this user wants in the presentation and analysis. If the user usually uses forecast data, the system will begin to prepare and present the data in the prediction model via dashboards.

4. *Automatically Cleansing and Prepping Data (ACPD):* To successfully interpret it, your data must be organized in a unified and searchable manner. As any business knows well, multiple datasets cause multiple headaches. What if names are formatted as first name/last name in one spreadsheet, and last name/first name in another? What if there are duplicate records? What if there are records in one dataset and not the other? What if the data in one set is daily, and the other is monthly?
AI in BI reduces data cleansing and contact preparation and provides massive aspirin for headaches. By setting up data automatically (one of the biggest artificial intelligence in saving time), you can move from making data available to working with it in minutes, instead of hours or days. The future AI function will allow users to enter structured and unstructured data without skipping any win; A big change since most of the data being created today - such as photos, videos and audio - is disorganized. Removing barriers to effective analysis is one of the ways in which the advanced AI in BI tool helps users who are not data scientists to access and interpret their data.

5. *Gaining Competitive Advantage (GCA):* AI now makes a critical difference between the companies that enable it to succeed and those that will be left behind soon. Gartner predicts that by 2021, 75% of pre-prepared reports - such as those used to extract data - will be either replaced or strengthened using automated insights. The robust AI in BI tools also provides improved accuracy for critical operational use reporting. If they do not, the data and analytics leaders should plan to adopt Enhanced Analyzes (AI) immediately in their business as the capabilities of the platform mature. Rita Sallam, vice president of Gartner Research, warned at a recent conference that "data and analytics leaders should examine the potential impact of business" from increasing reliance on predictions using enhanced and automated insights "and adjusting business and business models accordingly, or risking losing the competitive advantage of Those who do. " AI are already offered in BI solutions today, and those companies that adopt technology are poised to succeed more safely than those that do not. By uncovering trends and correlations in data and proposing ways to interpret results in natural language along with providing the best coordination for presenting these results, AI saves time and provides actionable insights to increase profitability and avoid potential problems before they arise.

4. Conclusion

In this chapter, the traditional and Modern BI were reviewed in detail which became a critical and important process and received a great interest in both industry and academia fields. So, the data management, data mining and machine learning techniques are needed for extracting the insights from big data. By using such techniques, business intelligence gets better decision making, cost reduction, new products and services and understand the market conditions. In addition, the importance of big data analytics, data mining, AI for building modern BI and

enhancing were introduced and discussed. Also, challenges and opportunities for creating value of data by establishing modern BI processes were described and how AI raises and improves the way an organization analyzes and interprets the lifeline of its business are explored. In the future work, we will study more AI tools to enhance the processes of BI and solving the cybersecurity problems in modern BI.

Acknowledgements

Author details

Ahmed A.A. Gad-Elrab[1,2]*

1 King Abdul-Aziz University, Jeddah, Saudi Arabia

2 Faculty of Science, Al-Azhar University, Cairo, Egypt

*Address all correspondence to: asaadgad@azhar.edu.eg

IntechOpen

References

[1] Wael M.S. Yafooz Abidin, S. Z., & Omar, N. (2011, November). Challenges and issues on online news management. In Control System, Computing and Engineering (ICCSCE), 2011 IEEE International Conference on (pp. 482-487). IEEE.

[2] https://technative.io/unstructured-data-the-hidden-threat-in-digital-business/

[3] Balachandran, B. M., & Prasad, S. (2017). Challenges and Benefits of Deploying Big Data Analytics in the Cloud for Business Intelligence. Procedia Computer Science, 112, 1112-1122.

[4] Kimble, C. and Milolidakis, G. Big Data and Business Intelligence: Debunking the Myths. Global Business and Organizational Excellence. 35, (2015), 23 – 34.

[5] Richards, G., Yeoh, W., Chong, A. Y. L., & Popovic, A. (2017). Business intelligence effectiveness and corporate performance management an empirical analysis. Journal of Computer Information Systems, 1-9.

[6] Xia B.S. and Gong P. Review of business intelligence through data analysis. Benchmarking: An International Journal. 21, (2014), 300-311.

[7] Kowalczyk M. and Buxmann P. (2014). Big Data and Information Processing in Organizational Decision Processes: A Multiple Case Study. Business & Information Systems Engineering. 5, (2014), 267-278.

[8] Wael M.S. Yafooz Abidin, S. Z., & Omar, N. (2011, November). Challenges and issues on online news management. In Control System, Computing and Engineering (ICCSCE), 2011 IEEE International Conference on (pp. 482-487). IEEE.

[9] Siddiqa, A., Hashem, I. A. T., Yaqoob, I., Marjani, M., Shamshirband, S., Gani, A., & Nasaruddin, F. (2016). A survey of big data management: Taxonomy and state-of-the-art. Journal of Network and Computer Applications, 71, 151-166.

[10] Fahad, S. A., & Alam, M. M. (2016). A modified K-means algorithm for big data clustering. International Journal of Computer Science Engineering and Technology, 6(4), 129-132.

[11] Fahad, S. A., & Yafooz, W. M. (2017). Design and Develop Semantic Textual Document Clustering Model. Journal of Computer Science and Information Technology, 5(2), 26-39. doi:10.15640/jcsit.v5n2a4 .

[12] Thuraisingham, B. (2014). Data mining technologies, techniques, tools, and trends. CRC press.

[13] A. L̈oser, F. Hueske, and V. Markl, "Situational business intelligence," in Business Intelligence for the Real-Time Enterprise. Springer, 2009, pp. 1-11.

[14] Deloitte report, "Modern Business Intelligence: The Path to Big Data Analytics",April 2018.

[15] A. Lser, F. Hueske, and V. Markl, "Situational business intelligence," in Business Intelligence for the Real-Time Enterprise, ser. Lecture Notes in Business Information Processing, M. Castellanos, U. Dayal, and T. Sellis, Eds. Springer Berlin Heidelberg, 2009, vol. 27, pp. 1-11. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-03422-0 1

[16] M. Castellanos, C. Gupta, S. Wang, and U. Dayal, "Leveraging web streams for contractual situational awareness in operational bi," in Proceedings of the

2010 EDBT/ICDT Workshops, ser. EDBT '10. New York, NY, USA: ACM, 2010, pp. 7:1-7:8. [Online]. Available: http://doi.acm.org/10.1145/1754239.1754248

[17] R. Elmasri and S. B. Navathe, Fundamentals of database systems. Pearson, 2014.

[18] H. Kuno, U. Dayal, J. Wiener, K. Wilkinson, A. Ganapathi, and S. Krompass, "Managing dynamic mixed workloads for operational business intelligence," in Databases in Networked Information Systems, ser. Lecture Notes in Computer Science, S. Kikuchi, S. Sachdeva, and S. Bhalla, Eds. Springer Berlin Heidelberg, 2010, vol. 5999, pp. 11-26. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-12038-1 2

[19] "Voltdb," https://voltdb.com/.

[20] R. Stoica and A. Ailamaki, "Enabling efficient os paging for mainmemory oltp databases," in Proceedings of the Ninth International Workshop on Data Management on New Hardware, ser. DaMoN '13. New York, NY, USA: ACM, 2013, pp. 7:1-7:7. [Online]. Available: http://doi.acm.org/10.1145/2485278.2485285

[21] A. Eldawy, J. Levandoski, and P.-A. Larson, "Trekking through siberia: Managing cold data in a memory-optimized database," Proc. VLDB Endow., vol. 7, no. 11, pp. 931-942, Jul. 2014. [Online]. Available: http://dx.doi.org/10.14778/2732967.2732968

[22] A. Kemper and T. Neumann, "Hyper: A hybrid OLTP & OLAP main memory database system based on virtual memory snapshots," in Proceedings of the 2011 IEEE 27th International Conference on Data Engineering, ser. ICDE '11. Washington, DC, USA: IEEE Computer Society, 2011, pp. 195-206. [Online]. Available: http://dx.doi.org/10.1109/ICDE.2011.5767867

[23] M. Castellanos, C. Gupta, S. Wang, U. Dayal, and M. Durazo, "A platform for situational awareness in operational {BI}," Decision Support Systems, vol. 52, no. 4, pp. 869 – 883, 2012, 1)Decision Support Systems for Logistics and Supply Chain Management 2)Business Intelligence and the Web. [Online]. Available: http: //www.sciencedirect.com/science/article/pii/S016792361100217X

[24] D. M. International, The DAMA guide to the data management body of knowledge. Technics Publications, Bradley Beach, 2009.

[25] K. Weber, B. Otto, and H. ¨Osterle, "One size does not fit all— a contingency approach to data governance," Journal of Data and Information Quality (JDIQ), vol. 1, no. 1, p. 4, 2009.

[26] V. Khatri and C. V. Brown, "Designing data governance," Communications of the ACM, vol. 53, no. 1, pp. 148-152, 2010.

[27] R. S. Seiner, "Real-world data governance bi governance and the governance of bi data," http://www.slideshare.net/Dataversity/realworlddata-governance-bi-governance-and-the-governance-of-bi-data14889552 Accessed:2015-11.

[28] D. M. Association et al., "Dama dmbok functional framework (version 3.02)," DAMA International, 2008.

[29] NASCIO, "Data governance - managing information as an enterprise asset part 1 - an introduction," NASCIO Governance Series, 2009.

[30] ——, "Data governance part iii: Frameworks - structure for organizing complexity," NASCIO Governance Series, 2009.

[31] P. Aiken, M. D. Allen, B. Parker, and A. Mattia, "Measuring data

management practice maturity: a community's self-assessment," Computer, vol. 40, no. 4, pp. 42-50, 2007.

[32] B. Potter and R. Software, "Self-service bi vs. data governance," https://tdwi.org/articles/2015/03/17/self-service-bi-vs-datagovernance.aspx, Mar. 17. 2015.

[33] M. Ferguson, "Is self-service bi going to drive a truck though enterprise data governance?" http://intelligent business.biz/wordpress/?p=489 Accessed:2015-10.

[34] Hung, P. C. K. Big data applications and use cases, the springer international series on applications and trends in computer science. Switzerland: Springer International Publishing AG, 2016

[35] Dave, P. What is big data - 3 vs of big data. Retrieved from SQL Authority, 2013 Blog: http://blog.sqlauthority.com/2013/10/02/big-datawhat-is-big-data-3-vs-of-big-data-volume-velocity-and-varietyday-2-of-21/.

[36] Sparks, B. H., & McCann, J. T. (2015). Factors influencing business intelligence system use in decision making and organisational performance. International Journal of Sustainable Strategic Management, 5(1), 31-54.

[37] Wang, Y., Kung, L., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. Technological Forecasting and Social Change, 126, 3-13

[38] Wael M.S. Yafooz., Abidin, S. Z., Omar, N., & Idrus, Z. (2013, December). Managing unstructured data in relational databases. In Systems, Process & Control (ICSPC), 2013 IEEE Conference on (pp. 198-203). IEEE.

[39] Kaisler, S., Armour, F., Espinosa, J. A., & Money, W. (2013, January). Big data: Issues and challenges moving forward. In System sciences (HICSS), 2013 46th Hawaii international conference on (pp. 995-1004). IEEE.

[40] Zhou, Z. H., Chawla, N. V., Jin, Y., & Williams, G. J. (2014). Big data opportunities and challenges: Discussions from data analytics perspectives [discussion forum]. IEEE Computational Intelligence Magazine, 9(4), 62-74.

[41] Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. Journal of Business Research, 70, 263-286.

[42] Lautenbach, P., Johnston, K. and Adeniran-Ogundipe, T.. Factors influencing bussiness intelligence and analytics usage extent in south african organaisations. S.Afr.J.Bus.Manage, 48(3): 23-33, 2017.

[43] Wang, Te-Wei,et al. "Depicting Data Quality Issues in Business Intelligence Environment Through a Metadata Framework." Applying Business Intelligence Initiatives in Healthcare and Organizational Settings, edited by Shah J. Miah and William Yeoh, IGI Global, 2019, pp. 291-304. http://doi:10.4018/978-1-5225-5718-0.ch016.

[44] Abai, N. H., Yahaya, J. and Deraman, A. "An integrated framework of business intelligence and analytic with performance management system. A conceptual framework." In Proceedings of the 2015 Science and Information Conference . London. pp. 452-56, 2015.

[45] Chen, H., Chiang, R. H. and Storey, V. C. Business intelligence and analytics. From Big Data To Big Impact, 36(4): 1165 – 1188, 2012.

[46] Mohammed, J. and Westbury, O. Business intelligence and analytics evolution, applications, and emerging research areas. International Journal of Engineering Science and Innovative Technology (IJESIT), 4(2): 193-200, 2015.

[47] M. A. Khan *et al.*, "Effective Demand Forecasting Model Using Business Intelligence Empowered With Machine Learning," in IEEE Access, vol. 8, pp. 116013-116023, 2020, doi: 10.1109/ACCESS.2020.3003790.

[48] Fahad, S. A., & Alam, M. M. A modified K-means algorithm for big data clustering. International Journal of Computer Science Engineering and Technology, 6(4), 129-132, 2016.