Journal of Theoretical and Applied Information Technology
30th June 2023. Vol.101. No 12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A UNIFIED FRAMEWORK FOR BIG DATA-DRIVEN DECISION MAKING SYSTEM IN HIGHER EDUCATION INSTITUTES

MAMOUN ABU HELOU

Faculty of Information Technology, Al Istiqlal University, Jericho, Palestine.

E-mail: mabuhelou@pass.ps (https://orcid.org/0000-0002-8597-3033)

ABSTRACT

Higher education institutions (HEIs) are increasingly using information and communication technologies in their daily operations, including management, financial, and didactic processes. As a result, the amount of produced data has increased dramatically in heterogeneous formats, and the lack of connectivity between the various systems has made it difficult for decision-makers to obtain integrated and relevant information that accurately reflects the institution's actual condition. Business Intelligence (BI) systems and big data technologies can take advantage of the various data sources accessible at HEIs to facilitate the processes of decision-making by enabling managers to make precise decisions that are efficient and effective with respect to time and formats. Several efforts focused on studying and enabling BI capabilities for managing sustainable information in an HEI context. Despite these attempts, integrating big data into HEIs remains an open challenge, and additional studies are required in order to determine the best techniques for employing big data analytics and enhance the results. In this research, qualitative and analytic research methodologies are used to identify and design a theoretical BI framework that incorporates best practices and takes into account big data analysis capabilities. The proposed framework combines the data analysis steps commonly used by HEIs and provides a framework made up of six components: the key stakeholders; the key business operations; the data identification; the data processing; the presentation of data; and the system's evaluation, validation, and monitoring. The framework is designed such that it can be adapted to various data sources and deploy different analytical scenarios, including descriptive and predictive analytical scenarios. The framework can be used as a guideline for current practice for implementing a BI in HEIs. Open research concerns and challenges are also identified.

Keywords: *Higher Education, Decision Making, Business Intelligence, Data Analytics, Big Data.*

1. INTRODUCTION

In the ear of big data and its emerging technologies, higher education institutions (HEIs) need to effectively exploit these opportunities by better monitoring their activities and managing their resources in a sustainable way, aiming to support effective and efficient processes of decision-making [1-6].

Higher Education (HE) (or, so-called Academic) sector, similar to other business sectors, is increasingly depend on the use of information and communication technologies (ICTs) in their processes; including studying, teaching, development, research, and management processes [7-10]. For example, learning management systems, human resources management systems, and financial management systems are ICT-based systems running in HEI, among many others [7].

It is well noted managing sustainability information is considered a crucial task for strategic management in HEIs [9], which includes data collection, data integration, and reporting activities [11]. Goni et al. [12] pointed out that, to support the decision-making process, it is required to develop specific tools that are effectively designed to analyse and manage institutions' information. The role of these tools is to convert the large amounts of data available in HEIs into information and knowledge that can be used by managers to make effective decisions [1, 13].

A key challenge, which increases with time due to the use of ICT, is the massive increase in data that requires processing and storing such a large amount of data to support HEIs in making decisions about their different processes [7]. Business Intelligence (BI) tools are considered a solution to this challenge [14, 15, 16]. BI tools are designed to enable the extraction, storage, and processing of data available in HEIs [17]. According to Sholtez et al. [9], the primary purpose of using the BI tool in HEI is to enable the stakeholders to more effectively monitor and manage sustainability activities by giving them up-to-date, precise strategic sustainability information. BI solutions can give strategic management and institution stakeholders' real-time, accurate access to the information that enables effective decisions [9]. Ultimately, HEIs use BI tools to support strategic management with predictive analytics that provide insights for long-term strategy [18].

In other words, the main issue is data heterogeneity, that is, the size and variety of formats in which educational data are expressed, whether in local or external sources, and the requirement for the development of predictive analytical techniques. To this end, several efforts focused on studying and enabling BI capabilities in managing sustainable information in an HEI context, which identified the impact of BI and its importance in the field of HE [8, 14]. In spite of these efforts, integrating big data into HEIs remains an open challenge [6]. In their investigation of the matter, Khalili et al., [6] concluded that additional research is required in order to determine the optimal methods for utilizing big data analytics and enhance the results. Several initiatives have been conducted to develop BI solutions to address this challenge; however, such initiatives are not mature enough, specifically when managing large volumes of data. To this end, this paper aims at identifying the common processes and operations that are appropriate for HEIs to develop their BI solutions.

This research was motivated by the initiative of the home university to support the decision-making process with an accurate and integrated view of the university's information. To this end, the qualitative and analytic research methodologies were adopted to identify and design a theoretical BI framework in the context of HEIs. Initially, different research databases (including IEEE, Google Scholars, Scopus, Web of Science, ScienceDirect, and PubMed) were used to access published scientific articles. The snowfall approach was adopted to identify related work. Then, inductive and analytical reviews of related literature in the background are conducted. As a result, common factors and open challenges are identified. The proposed framework will outline a roadmap and provide a guide for researchers that plan to design BI solutions to analyse and manage sustainable information in HEI.

The proposed framework unified BI processes, which are common and suitable to manage ad-hoc reporting capabilities as well as to support predictive analysis based on online analytical processing (OLAP) and big data technologies, while considering different types of data sources. Two important factors are considered in the proposed design of a BI solution: the engagement of the stakeholders in the system life cycle and the governance policies (e.g., management standards, quality control, and quality assurance procedures) that are adopted in the institution.

As a result, this research proposes a conceptual framework for a higher education business intelligence solution. The proposed framework is composed of six components that combine the data analysis steps commonly used by HEIs: (1) The key stakeholders, which are involved at the various stages of the institutions' business processes, are identified. (2) The key business operations, which depict the preferred outcomes and the potential directions in the future based on the institution's strategic plan. (3) The data identification, in which the relevant sources of data that will be utilized and explored in the BI processes are identified. (4) The data processing, in which data is extracted and stored from the various sources and later the analytical tools are used to recognize patterns and trends in the data. (5) The data presentation, which enables the decision-makers to view the data and analytical results, also provides them with real-time strategic information. (6) The system's evaluation, validation, and monitoring. That is, when the system is operational, these activities should be periodically addressed, especially whenever new or updated requirements are presented. Taken as a whole, the framework is capable of providing the essential tools for evaluating and monitoring the success of realizing the business needs of the institution, including the capability of reporting, analytical and predictive processing, and filtering.

Next, Section 2 provides a background on BI, in particular considering the outstanding state-of-the-art efforts in the HE context. Section 3 details the proposed framework. Section 4 outlines and discusses the analytical capabilities of the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

framework. Section 5 concludes the paper and provides potential future directions.

2. BACKGROUND

Background information on BI systems is given in this section, with special attention paid to the remarkable state-of-the-art initiatives in the HE context.

2.1 Business Intelligence Systems

The business intelligence (BI) system concept has been used by different researchers and vendors when referring to the use of information systems by an organization in various business sectors (e.g., financial [19], banking [20], telecommunications [21], healthcare [22], marketing [23], and small business [24]. Thus, BI systems involve the communication of three core components: 1) people, 2) information, and 3) processes [25].

Three goals are in common for different organizations when using BI systems: "1) allow improved access to data; 2) allow data manipulation; and 3) give to managers and business analysts the ability of performing important analysis" [26]. Ultimately, the use of BI systems aims at improving the decisions and performance of an organization in terms of its effectiveness and efficiency when running its business [27, 28, 15].

An early effort in developing BI systems was initiated in 1958 [29], which were described as "an automatic system is being developed to disseminate information to the various sections of any industrial, scientific, or government organization. Later on, Business Intelligence was broadly defined as a set of "terms and methods to improve business decisions by making use of support system based on business data." [30].

For instance, Reinschmidt and Francoise [31] viewed a Business Intelligence system as “an integrated set of tools, technologies and programmed products that are used to collect, integrate, analyze and make data available”. Negash and Gray [32] declared the Business Intelligence system as a system that “combines data gathering, data storage and knowledge management with analytical tools to present complex and competitive information to planners and decision makers”. That is, a Business Intelligence system enables users to access reliable information and utilize useful tools for data analysis [33]. Further details on the historical perspective in developing BI systems are reported in [34].

On the other hand, from a technical perspective Olszak and Ziemba defined BI systems as a system that “offer an integrated set of tools, technologies and software products that are used to collect heterogenic data from dispersed sources in order to integrate and analyse data to make it commonly available” [35]. In this regards, In [36], Olszak and Ziemba defined a technological BI structure that consists of six components; namely, 1) extract, transform and load (ETL) tools, which responsible for transforming data in different data sources into data repositories (warehouses); 2) data warehouses, in which data are aggregated and stored; 3) analytic tools, (e.g., OLAP) which enable users accessing, analysing and modelling business problems and share information that is aggregated in data warehouses; 4) data mining tools, enable users to discover various trends and patterns in data resources; 5) tools for reporting, they allow users for inquiring, creating and utilizing various kind of reports; and 6) presentation layer, a user friendly graphical interfaces that provide users with a integrated view and flirting capabilities.

2.2 BI in Higher Education Sector

In the last two decades, Business Intelligence systems have been increasingly adopted to utilize data sources available at HEIs (e.g., [7-10, 14, 33, 36], which were developed to produce significant information to be used in their strategic decision making.

According to Muntean et al. [33], HEIs can rely on the BI systems to provide the institutions’ top management levels, as well as the administrative and organizational structures, with centralized access to the institutions’ information. Additionally, it supports the analytical operations of the primary components of the HEIs, including didactics, research, and management, and provides the data required for reporting, analysis, and formulating the institutions' strategic plan.

Zulkefli et al. [36], emphasized that, as the amount of data obtained and disseminated from external sources grows, HEIs are continually searching for innovative management strategies and techniques to improve the utilization of available data and information. That is, HEIs frequently employ BI tools to access and manage massive amounts of data. Agustiono et al. [34], also stated that the education business has recently experienced a great rise in data, demanding an advanced analysis process of such data to support the management and operational activities of HEIs [37].

Identifying and assessing crucial metrics and key performance indicators (KPIs) of HEIs is greatly facilitated by the use of business intelligence tools [33, 38, 39]. As they enable the institutions to manage and maintain consistent levels of institutions' efficiency and effectiveness, these metrics and KPIs play a crucial role in their competitive advantage [40, 41].

The work of Muntean et al. [33] is regarded as one of the primary references for BI system research for HEIs. They provided a framework to direct the application of BI solutions for institution governance in HEIs [33]. The framework consists of eight components;“ (1) the mission and vision, (2) the strategic plan, (3) the performance indicators, (4) the key business processes, (5) the personnel, (6) the processes, (7) the technologies BI, and (8) the information infrastructure” [33]. Also, Muntean et al. [42] studied the performance of BI's dashboards for HEI using six layers framework: “(1) the extract, transform and load (ETL) process layer; (2) the data layer; (3) the reporting layer; (4) the analytical layer; (5) the monitoring layer; and (6) the presentation layer”.

Pinheiro [43] analyzed the EUNIS (European University Information Systems) survey 2013 [57], which examined the level of BI system maturity at a number of EU HEIs, and proposed a kit for BI in HEIs, based on best practices in order to start their BI initiatives.

Zulkefli et al. [36] proposed a theoretical framework by adopted Muntean et al. framework [33]. Scholtz et al. [9] extended the BI architectural framework in [42] to enable strategic sustainability information management in HEIs. Two layers were added; the users and data sources layers. Six stages of the BI framework were adopted by Villegas et al. [8], which are; data sources selection, data storing, data pre-processing, data transformation, data mining, and finally discovering knowledge [8].

Hamed et al. [10] build decision support systems using data mining algorithm (Support Vector Machine) to predict students' annual grade point average. Ali Sorour et al. [4] concluded such BI frameworks typically consist of three layers: (1) the data source, (2) the extract, transform, and load (ETL), and (3) the data presentation.

An important aspect to consider in the design of BI solutions is the linkage (relations) among the BI system components [4], in particular the role of the stakeholder in validating and monitoring the system quality.

Recent works have given a particular attention to the involvement of the educational big data analytics processes. For instance, Berges et al. [15] show how a framework for managing strategic intelligence has been implemented in the Madrid community's system for managing education. Munshi et al. [1] implemented educational big data platforms that consisted of five stages, namely: (1) data production; (2) data storage; (3) data processing; (4) data querying; and (5) data analytics.

However, in spite of these efforts, the integrating big data into HEIs remains inadequate and considered among the technical challenges [6, 44]. Khalili et al. investigated the issue and found that more study is needed to identify best practices for employing big data analytics and maximize the outcomes in terms of cost-effectiveness [6].

To this extent, the paper analyses the aforementioned state of the arts literature in the field of Business Intelligent in higher education context with a particular attention of providing big data analytics capabilities. The next section presents a Business Intelligent framework adequate for different analytical scenarios considering different sizes of data in HEIs' (i.e., small, medium, and large data) [15, 34].

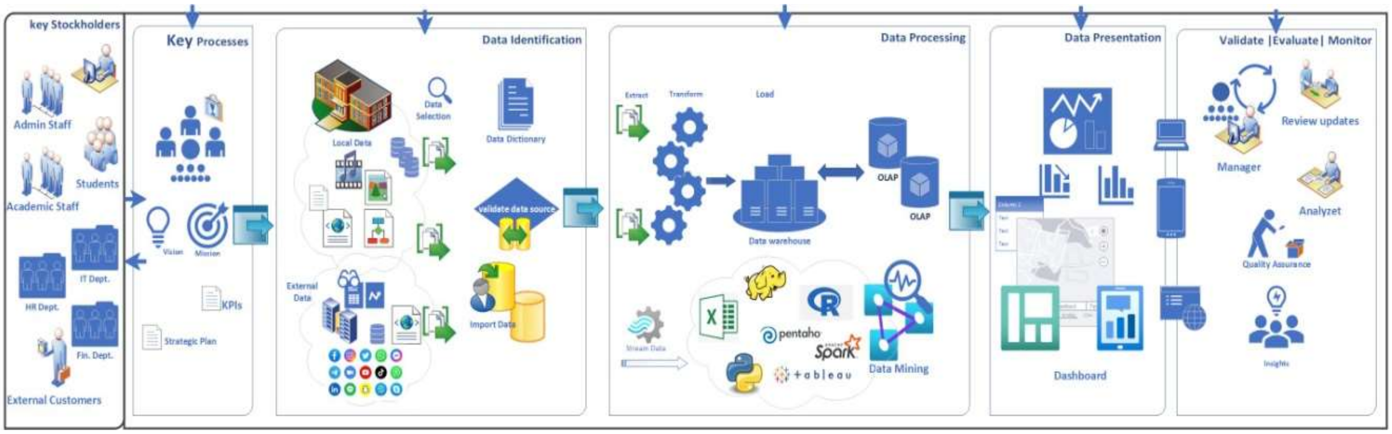

Figure 1: An overview of the BI framework in HEIs

3. BIG DATA-DRIVEN DECISION MAKING SYSTEM IN HEI FRAMEWORK

The BI framework illustrated in Figure 1 is composed of six components, namely: HEIs' key stakeholder; HEIs' key processes; data identification; data processing; data presentation; evaluation, validation, and monitoring of the system. The next sections describe the function of each component.

3.1 HEIs' Key Stockholder

The stakeholders of HEI are made up of a wide range of users with various business, analytical, and decision-making roles and responsibilities [33, 36]. These users need various kinds of analysis, information distribution, and data [36]. They should be able to access real-time information that is relevant for their activities and organized in an integrated view [9].

To realize the capabilities of a Business Intelligent system the various stakeholders should support and be engaged at the different stages of the organization businesses processes [35]. Scholtz et al. [9] have underlined that the success of a HEI is tied to the participation of diverse stakeholders in the institution. Figure 1, shows the relation and involvement of the users with the different components' operations.

In this framework, users who participate in management processes in both academic and non-academic contexts are among the stakeholders; including, HEI's top management, decision makers, administrative and academic staffs, students, operational workers, IT staff, and different departments involved the knowledge management and strategic operations [8, 9, 36]. From a technical perspective, a special attention should be given to the IT staff [8,35]. In [35], Olszak and Ziemba stated that IT professionals should have comprehensive knowledge of how users and analysts work, as well as their responsibilities in processes. Additionally, they should have a thorough understanding of pedagogical and other relevant processes. Users who participate in the formation of the organization's vision, mission, and strategy are also included among the stakeholders [36].

Cates, et al. [25] in the LOBI framework summarize the roles of users as: "(1) Agree on business vision, mission, core questions. (2) Agree on enterprise roadmap: vision, objectives, metrics, processes, initiative alignment. (3) Form enterprise process management team and program management office. (4) Executive business role intelligence analysis".

3.2 HEI Key Processes

A set of key processes in the HEI are identified on the basis of the institution's vision and mission, which describe the desired outcomes and the future direction of the institution, and relay on a well-defined and detailed strategy plan. The key processes are recognized to achieve the institution strategy, and ultimately the HEI sustainability [36]. According to Cates et al. [25] HEI's possesses can be viewed as a "set of activities that carried out in the HE context. Such activities are designed to utilize the HEI resources aiming to convert inputs into desired outputs; including financial and operational activities".

Although, in educational institute environments, didactic processes are considered the core activities [36], other related processes also exist. Pires and Lourenço [45] defined the support and integrated processes to achieve the HEI macro-operational processes that include didactics, service delivery, research, and development processes. The support processes involve the administrative and financial management processes, whereas the integrated processes involve the strategic and global management processes, among others [45, 46].
Mihaela et al. [26] proposed a framework for HEI processes that composed of four levels, including the management level (3P's: People, Process, and Performance Management), the human resources management level, the teaching and learning process development level, and the measurement analysis, improvement level.

Another comprehensive process framework was developed by the HEI-UP project [47]. The framework is composed of five main processes, namely: teaching, studying, research, development, management, and supporting processes.

A crucial aspect for HEI, in order to achieve its strategy and sustainability, is its ability to measure and evaluate the institution's performance level. To this end, key performance indicators (KPIs) should be defined. Accordingly, this performance data (i.e., KPIs) must be presented to the institution's top management-level in a clear, understandable format to support the institution's management processes. [33].

Ultimately, the key business processes employed in the proposed framework can assist HEI in carrying out its operational tasks and achieving the top management-level strategy. And the KPIs enable strategic decision-making by monitoring performance levels and ensuring the sustainability of HEI.

3.3 Data Identification

To meet the HEI objectives as well as manage the institution's resources in a sustainable way, identifying the data sources that could be utilized to support the institution's business needs is crucial. Specifying the relevant sources of data, which will be used and explored in the BI processes, is a primary step to avoid ambiguity and guarantee quality results [8].

Data is available from different sources, including internally stored data in the databases, external sources, social media, and data stored in the cloud [4]. According to [9], these data sources fall into four categories: educational, environmental, economic, and social. Another important aspect to consider is the use of big data in the best possible way, i.e., the management of large volumes of raw data [49]. That is, to convert the huge amounts of data available at the institution into information and knowledge for top management-level to make good decisions. [13].

To summarize, in this context the type of the data sources can be distinguished based on the following characteristics of data:

1) Format: Structured, unstructured (repetitive, unrepetitive), semi-structured [49].
2) Size: Big data, traditional (small, or medium) data [34].
3) Storage: internal source, external source.
4) Category: educational, environmental, economic, social [9].

Villages et al. [8] recommend the available data sources should be evaluated to identify the degree of relevant, efficient, impactful or sustainable that each data source impacts the BI process. Also, this include verifying reliability of the sources; define relations between data; describe the logical structure of data; data consistency [35].

Once the data sources are specified, the selection of the data is made. Olszak and Ziemba [35] stated the realization of this stage needs a remarkable amount of support and engagement provided by the different stakeholders. To manage this step efficiently, Villages et al. [8] advocate using a data dictionary, which is simply a technical description of each data resource. The data dictionary includes information about the data fields, their source, their accessibility, and the person in charge. Also, different variables concerning the educational data, including academic, psychosocial, and financial information, need to be specified in the dictionary [8].

3.4 Data Processing

After the selection of the relative data sources that match the HEI's business needs, it is time to collect the data from the different types of sources, integrate it, and make it ready for analytical activities. This stage comprises the full processes of discovering knowledge in databases (so-called "KDD") [4], which start with data collection and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

involve integrating and analysing all the data from the many sources [50]. The extract, transfer, and load (ETL) process essentially gathers the data required for analysis and transforms it into the data repositories (DW, Data Warehouse), which are in charge of storing data for additional analysis. Considering both traditional OLTP data and big data technologies, which have in common the following processes: data ingestion (ETL processes), data warehousing (storing), and data analytics [8], they are introduced next.

Data Ingestion: At first, the extract, transfer, and load (ETL) processes take place, in which data is extracted from the different sources. Then the data is passed to the second stage, i.e., the transformation stage, which ensures the data is complete, clean, and consistent by performing different data cleansing techniques. The data transformation stage may also include the creation of new features generated from the original attributes for further analysis. According to Venkatadri et al. [51], these operations will maintain the quality and integrity of data, increase data efficiency, reduce data storage [52], and improve information correctness [36]. To this extent, the output provides clean, high-quality data for use in analytical tasks.

Data Storing: in this phase, the processed and verified data is sent to a repository, which enables convenient exploitation and reliability. To this end, the data is loaded into the Data Warehouse (DW), which is responsible for saving data for later analysis [53]. The DW "is a data repository that is populated from the integration of different operational data sources maintained in different business units of organizations" [54]. When designing a DW, scholars propose many approaches [49,51,55]. Yet, two approaches stand out: Kimball proposes a bottom-up approach, and Inmon proposes a top-down approach. The top-down method is adaptable to support the management of change since it views the organization as a whole [54]. With the emerging of big data technologies, two major strategies can be used to implement big data warehouses [55]. The first is called "lift and shift," where traditional data warehouses are augmented with big data analytical technologies (e.g., [56]). The second strategy is called "rip and replace," where traditional data warehouses are replaced with big data technologies (e.g., [57]). Based on the requirements of the instructions' business needs, practitioners and researchers can decide which approach fits a particular scenario [55].

Data Analytics: at this stage, in order to help an institution improve its decision-making abilities, the analytical tools are utilized to identify patterns and trends in the data. Enabling predictive analytics would help create an institution's strategic intelligence. The process of analysing data to find trends, group people into categories, or make projections is accomplished by subjecting it to data analytics algorithms. This stage may involve specialized software, tools, and libraries to analyse data from a range of educational sources using both traditional and contemporary approaches. For example, using programming languages like R and Python and using open-source or commercial platforms, including Hadoop, Weka, Spark, OLAP, HDFS, SPSS, Tableau, MS Excel, and Pentaho [1, 4].

In [58], they classified the methods of analysis into three categories: (1) Machine learning, which includes clustering, prediction, association, and text analysis algorithms. (2) Multi-criteria decision making is used to determine ranking and gives a selection of the top candidates from a set of options that are evaluated using a variety of criteria (e.g., Near Field Communication). (3) Big data (streaming) analytics; which enable real-time data processing (e.g., Stream Clustering , Stream k-means, and Very Fast Decision Tree). It can be challenging to choose an algorithm for a data analysis task because each method generates a distinct output [8, 59]. This is not to say that an analysis cannot utilize different algorithms at the same time; instead, the key is to choose the relevant features and then apply a certain method to draw a given judgment utilizing the existing information [60].

4.5 Data presentation (reporting and visualization)

In this stage, the data and analytical results are presented to users (i.e., decision makers) usually through reports and interactive graphical interfaces, so-called dashboards [9, 42]. Dashboards present a summary of these analytics, enabling users to drill down to finer granularities and filter the resulting data [4]. This provides decision makers with real-time strategic information and gives them the capability to identify trends and patterns that support management in assessing institution objectives and monitoring the performance of KPIs [33].

Users should be able to access these details and generate reports through different devices, including mobile, desktop, or Web applications [19] and [42]. Different software tools were used in the literature to support the querying, filtering, and reporting capabilities. Typically, such reports are generated in CSV and PDF formats. Scholtz et al. [9] used Tableau software to create and visualize reports and took advantage of several filter features, which provide users with the ability to edit the information they display and avoid unwanted data. For instance, in [9], they suggested employing a toolbar with checkboxes and radio buttons where users could select the necessary details for filtering alternatives. In this way, by choosing one of the interactive, coordinated dashboard elements utilized for the filtering features, the entire dashboard is filtered.

Scholtz et al. [9] in their study defined three groups of guiding principles for the design of an efficient dashboard: (1) interaction (involving users through a variety of interactions and visualization strategies), (2) media (to present information by choosing the suitable media), (3) visualization (to reduce the amount of time needed to comprehend the message), and (4) feedback (there should be rules governing interaction). (2) Aesthetics (the presented information ought to be attractive to users). (3) information detail, relevance, and purpose (avoid providing unnecessary detail).

3.5 Evaluate, Validate and Monitor the system

An essential role in this framework is to ensure that the developed system meets the BI system requirements. In [9], they define several vital requirements that such systems should realize. This includes visual and interactive dashboards that allow reporting and filtering capabilities, analytical processing capabilities, and predictive and forecasting capabilities.

Different approaches could be applied to evaluate if the BI system meets its requirements. For instance, [9] conducted two phases. At first, a set of heuristic evaluations was performed to gather feedback from expert users using several heuristic characteristics adopted from [61], namely: "information coding; minimal actions; flexibility; orientation and help; spatial organization; consistency; recognition rather than recall; prompting; removing the extraneous; data set reduction." As a result, the experts identify problems and recommend fixing them. Accordingly, an improved version of the system is produced. In the second phase, the updated system goes through a usability evaluation that covers the main capabilities of the system, such as the ability to generate reports, apply filters, view information at different levels of detail, and forecast. Five evaluation criteria were adopted from [62], which are relevant for evaluating BI solutions: 1) the visibility of the information, 2) the flexibility to operate and maintain the system, 3) the user's ability to learn, 4) operability and error control, and 5) providing appropriate support.

Accordingly, the system should be updated and improve its capabilities to meet the indented requirements. In fact, these activities should be carried out continuously while the system is live, and in particular when new requirements are introduced or updated.

4. RESULTS AND DISCUSSION

Based on the review and analysis of related research papers, each of the many categories of data is characterized in terms of the format, size, storage, and category. Besides that, the key stakeholder (i.e., the user who is involved and intends to use the system) is identified. Based on the available data sources and considering the business needs, the IT staff should decide whether to select the data processing approaches (i.e., OLAP or big data technologies) or even integrate the power of each technology. The system should ultimately assist in data exploration to find patterns, aid in decision-making, and produce statistics and performance metrics.

Based on the evaluation and analysis of related research publications the institutions' data is described in terms of format, size, storage, and category. In addition, the key stakeholder is identified, that is, the user who will utilize the system and is involved. The IT staff must choose whether to use data processing methodologies (such as OLAP or big data technologies) or even combine the strengths of each technology based on the data sources that are available and the business demands. Ultimately, the system should support pattern-finding data exploration, decision support, and the generation of statistics and performance indicators.

4.1 Analytical Scenarios

The framework can support different data analytics tasks, including descriptive and predictive analytics. For instance, descriptive analysis involves providing statistical measures for the student: their GPAs, demographical data, drop rate, financial status, etc.; also for administrative staff: working hours, leaves, vacations, demographical data, etc.; and also presenting technical reports and statistics on the usage of different resources: internet, electricity, gas, and water. This also includes the categories of web sites students and staff access, locations they visit, and workshops they hold or attend, among many other KPIs in HEIs using different statistical measures.

And besides, the framework also has predictive capabilities, which it achieves by carefully planning and picking the right mining algorithm. For example, recommend the courses a student can take in order to possibly achieve a better grade [1]. In order to forecast flu epidemics, analyse social media (such as tweets). Look into the interactions between the departments. Predict students' admission, performance, and dropout. Determine the moral standing and skill of teachers [6]. Find out which students are most likely to give up on their education. Find out if the pupils are ready for distant learning [53]. Learn what the students think about online education.

There are numerous considerations to make when creating a BI solution, but it is crucial to take into account the unique business requirements of each institution. Such a project requires the adoption of management techniques for its creation and maintenance, as well as governance rules to define agreements and lines of communication among the numerous stakeholders. In [49], they make the addition of the management and governance components, which cover the various stakeholder functions covered by the framework, highly recommended.

5. CONCLUSION

This research proposes a comprehensive theoretical framework which offers integrated operation of transformation and data quality, with an emphasis on the predictive analytical operation. Using qualitative and analytical research methodologies, the proposed framework presents a unified analytical approach considering educational big data.

The framework was derived from the analysis and observation of best practices. The framework unifies the steps of data analysis available in HEIs and provides a business intelligence framework composed of six components: HEIs' Key Stakeholder; HEIs' Key Processes; Data Identification; Data Processing; Data Presentation; Evaluate, Validate and Monitor the System. The framework is designed such that it can be adapted to different data sources (i.e., in terms of the data formats, size, storage, and category) and deploy different analytical scenarios, including descriptive and predictive analytical scenarios.

To this extent, the proposed Business Intelligent Framework offers the HEIs a road map and guidelines to efficiently manage their sustainable information and measure their performance. As a future directions, the author plan to evaluate the proposed system through the implementation of several analytical scenarios, in the context of real-life use cases, in particular, considering home university Al Istiqlal University (PASS), as well as higher education institutions in Palestine. Further, is the development of a methodology to assess the current BI maturity and set higher maturity goals in the context of under development countries (e.g., HEIs in Palestine) [63].

ACKNOWLEDGMENTS

The author thanks Al-Istiqlal University for supporting this research. This research was conducted as part of PASS-DDDM research project "Towards Quality Improvement of Education and Managerial Processes at PASS - Based on Data Driven Decision Making System".

REFERENCES

[1] Munshi and A. Alhindi, "Big Data Platform for Educational Analytics," in IEEE Access, vol. 9, pp. 52883-52890, 2021, doi: 10.1109/ACCESS.2021.3070737.

[2] Y. Nieto, V. Gacia-Diaz, C. Montenegro, C. C. Gonzalez and R. G. Crespo, "Usage of machine learning for strategic decision making at higher educational institutions", *IEEE Access*, vol. 7, pp. 75007-75017, 2019.

[3] M. Nauman, N. Akhtar, A. Alhudhaif and A. Alothaim, "Guaranteeing Correctness of Machine Learning Based Decision Making at

Higher Educational Institutions," in *IEEE Access*, vol. 9, pp. 92864-92880, 2021, doi: 10.1109/ACCESS.2021.3088901.

[4] Sorour, A. S. Atkins, C. F. Stanier and F. D. Alharbi, "The Role of Business Intelligence and Analytics in Higher Education Quality: A Proposed Architecture," 2019 International Conference on Advances in the Emerging Computing Technologies (AECT), 2020, pp. 1-6, doi: 10.1109/AECT47998.2020.9194157.

[5] Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly: Management Information Systems*, *36*(4), 1165-1188. https://doi.org/10.2307/41703503

[6] Alkhalil, Adel and Magdy Abd Elrahman Abdallah, Azizah Alogali, and Abdulaziz Aljaloud. "Applying Big Data Analytics in Higher Education: A Systematic Mapping Study," *International Journal of Information and Communication Technology Education (IJICTE)* 17, no.3: 29-51. http://doi.org/10.4018/IJICTE.20210701.oa3

[7] Villegas-Ch., W.; Jácome-Vásconez, S.; García-Ortiz, J.; Calvache-Sánchez, J.; Sánchez-Viteri, S. Proposal of a Model for the Analysis of the State of the Use of ICT in Education Applied to Technological Institutes of Higher Education. *Computers* 2022, *11*, 112. https://doi.org/10.3390/computers11070112

[8] Villegas-Ch, William, Xavier Palacios-Pacheco, and Sergio Luján-Mora. 2020. "A Business Intelligence Framework for Analyzing Educational Data" *Sustainability* 12, no. 14: 5745. https://doi.org/10.3390/su12145745

[9] Scholtz, Brenda, Andre Calitz, and Ross Haupt. "A business intelligence framework for sustainability information management in higher education." *International Journal of Sustainability in Higher Education* (2018).

[10] Hamed, M., Mahmoud, T., Gómez, J.M., Kfouri, G. (2017). Using Data Mining and Business Intelligence to Develop Decision Support Systems in Arabic Higher Education Institutions. (eds) Modernizing Academic Teaching and Research in Business and Economics. Springer Proceedings in Business and Economics. Springer, Cham. https://doi.org/10.1007/978-3-319-54419-9_4

[11] Frost, G., Jones, S. and Lee, P. (2012) The Collection, Integration and Reporting of Sustainability Information within an Organisation, CPA, Sydney.

[12] Goni, F.A., Sahran, S., Mukhtar, M., Shukor, S.A. and Chofreh, A.G. (2013), "Aligning an information system strategy with sustainability strategy towards sustainable campus", International Conference on Research and Innovation in Information Systems (ICRIIS), Kuala Lumpur, pp. 245-250.

[13] D. Kabakchieva, "Business Intelligence Systems for Analyzing University Students Data," Cybernetics and Information Technologies, vol. 15, pp. 104-115, 2015.

[14] L. Chen, P. Chen and Z. Lin, "Artificial Intelligence in Education: A Review," in IEEE Access, vol. 8, pp. 75264-75278, 2020, doi: 10.1109/ACCESS.2020.2988510.

[15] Berges, P. Ramírez, I. Pau, A. Tejero and Á. G. Crespo, "A Framework for Strategic Intelligence Systems Applied to Education Management: A Pilot Study in the Community of Madrid," in IEEE Access, vol. 9, pp. 75313-75323, 2021, doi: 10.1109/ACCESS.2021.3081734.

[16] O. Moscoso-Zea, J. Castro, J. Paredes-Gualtor and S. Luján-Mora, "A Hybrid Infrastructure of Enterprise Architecture and Business Intelligence & Analytics for Knowledge Management in Education," in IEEE Access, vol. 7, pp. 38778-38788, 2019, doi: 10.1109/ACCESS.2019.2906343.

[17] Bosire, S. (2014), A Sustainability Reporting Framework for a Higher Education Institution, Nelson Mandela Metropolitan University, Summerstrand

[18] Stewart, Christopher Lee, and M. Ali Akber Dewan. "A Systemic Mapping Study of Business Intelligence Maturity Models for Higher Education Institutions." Computers 11, no. 11 (2022): 153.

[19] Shao, Cuili, Yonggang Yang, Sapna Juneja, and Tamizharasi GSeetharam. "IoT data visualization for business intelligence in corporate finance." Information Processing & Management 59, no. 1 (2022): 102736.

[20] Vaidya, Tejas, Sagar Ahirrao, Kiran Kadam, and Kundan Lugade. "Business intelligence system for banking and finance." Int J Comput Sci Inf Technol 5 (2014): 4336-4349.

[21] Tutunea, Mihaela Filofteia. "Business Intelligence Solutions for Mobile Devices–An

Overview." Procedia Economics and Finance 27 (2015): 160-169.

[22] Gaardboe, Rikke, Tom Nyvang, and Niels Sandalgaard. "Business intelligence success applied to healthcare information systems." Procedia computer science 121 (2017): 483-490.

[23] Kim, Ki Youn. "Business intelligence and marketing insights in an era of big data: The q-sorting approach." KSII Transactions on Internet and Information Systems (TIIS) 8, no. 2 (2014): 567-582.

[24] Guarda, Teresa, Manuel Santos, Filipe Pinto, Maria Augusto, and Carlos Silva. "Business intelligence as a competitive advantage for SMEs." International Journal of Trade, Economics and Finance 4, no. 4 (2013): 187.

[25] J. E. Cates, S. S. Gill, and N. Zeituny, "The Ladder of Business Intelligence (LOBI): A Framework for Enterprise IT Planning and Architecture," International Journal of Business Information Systems, vol. 1, p. 220, 2005.

[26] Mihaela Drăgan, Diana Ivana, Raluca Arba, Business Process Modeling in Higher Education Institutions. Developing a Framework for Total Quality Management at Institutional Level, Procedia Economics and Finance, Volume 16, 2014, Pages 95-103, ISSN 2212-5671, https://doi.org/10.1016/S2212-5671(14)00779-5.

[27] Herschel, Richard T., and Nory E. Jones. "Knowledge management and business intelligence: the importance of integration." *Journal of knowledge management* (2005).

[28] Hostmann, Bill, Nigel Rayner, and Gareth Herschel. "Gartner's business intelligence, analytics and performance management framework." *Gartner Research Note* (2009).

[29] Luhn HP. A business intelligence system. IBM Journal of research and development. 1958;2(4):314-9.

[30] Phuong, H.M., Hoan, P.M., Tuan, N.T., Tung, D.T. (2022). A Proposed Business Intelligence Framework for Autonomous and Non-public Higher Education Institutions in Vietnam. In (eds) Intelligent Systems and Networks. Lecture Notes in Networks and Systems, vol 471. Springer, Singapore. https://doi.org/10.1007/978-981-19-3394-3_20

[31] Reinschmidt, J., & Francoise, A. (2000). Business Intelligence Certification Guide. San Jose, CA: IBM, International Technical Support Organization.

[32] Negash, Solomon and Paul Gray. "Business Intelligence." Handbook on Decision Support Systems 2: Variations. Eds. Frada Burstein and Clyde W. Holsapple. Berlin: Springer-Verlag, 2008. 175-193.

[33] M. Muntean, A.-R. Bologa, R. Bologa, and A. Florea, "Business Intelligence Systems in Support of University Strategy," Recent Researches in Educational Technologies., pp. 118-123, 2011.

[34] Agustiono, Wahyudi. "Academic Business Intelligence: Can a Small and Medium-sized University Afford to Build and Deploy it within Limited Resources?" *Journal of Information Systems Engineering and Business Intelligence,* Vol.5, No.1 (2019)

[35] Olszak, C., & Ziemba, E. (2007). Approach to Building and Implementing Business Intelligence Systems. Interdisciplinary Journal of Information, Knowledge, and Management, vol. 2, 135-148.

[36] *Zulkefli, Nur Ain, Suraya Miskon, Haslina Hashim, Rose Alinda Alias, Norris Syed Abdullah, Norasnita Ahmad, Nazmona Mat Ali and Mohd. Aizaini Maarof. "A business intelligence framework for Higher Education Institutions." ARPN journal of engineering and applied sciences 10 (2015): 18070-18077.*

[37] Guster, Dennis, and Christopher G. Brown. "The application of business intelligence to higher education: Technical and managerial perspectives." *Journal of Information Technology Management* 23, no. 2 (2012): 42-62.

[38] Wixom, Barbara, Thilini Ariyachandra, David Douglas, Michael Goul, Babita Gupta, Lakshmi Iyer, Uday Kulkarni, John G. Mooney, Gloria Phillips-Wren, and Ozgur Turetken. "The current state of business intelligence in academia: The arrival of big data." *Communications of the Association for information Systems* 34, no. 1 (2014): 1.

[39] Wang, Yichuan. "Business intelligence and analytics education: Hermeneutic literature review and future directions in is education."

In *Proceeding of Twenty-First Americas Conference on Information Systems (AMCIS), Puerto Rico*. 2015.

[40] Bogza, Rodica M., and Dorin Zaharie. "Business intelligence as a competitive differentiator." In *2008 IEEE International Conference on Automation, Quality and Testing, Robotics*, vol. 1, pp. 146-151. IEEE, 2008.

[41] Njenga, James K., Ildeberto A. Rodello, Karin Hartl, and Olaf Jacob. "Identifying opportunities and challenges for adding value to decision-making in higher education through academic analytics." In World Conference on Information Systems and Technologies, pp. 474-480. Springer, Cham, 2017.

[42] Muntean, Mihaela, Gheorghe Sabau, A. Bologa, Traian Surcel, and Alexandra Florea. "Performance dashboards for universities." In Proceedings of the 2nd international conference on manufacturing engineering, quality and production systems, pp. 206-211. 2010.

[43] Marta R. F. Pinheiro, "Developing a Business Intelligence Initiative in Higher Education," PhD thesis, 2014.

[44] M. I. Baig, L. Shuib, and E. Yadegaridehkordi, ``Big data in education: A state of the art, limitations, and future research directions," Int. J. Educ. Technol. Higher Educ., vol. 17, no. 1, pp. 1_23, Dec. 2020.

[45] Pires, A. & Lourenço, R., 2010. Aplicabilidade do modelo de gestão por processos às Instituições de Ensino Superior.

[46] Gonçalves, Nuno Pina, Sérgio Paulino and Tiago Santos da Silva. "Business Processes Underpinning Information Systems - A Higher Education Institution's Framework." In Proceedings of the Twelfth International Conference on Informatics and Semiotics in Organisations, ICISO (2010).

[47] EUNIS, 2013, https://op.europa.eu/en/publication-detail/-/publication/ceb8a8b5-f342-4890-8323-4000e99deb3d (online accessed; Nov, 2022).

[48] HEI-UP : https://www.bicero.com/training-services/ecbpm-hei (online accessed; Nov, 2022)

[49] Sonia Ordoñez Salinas and Alba Consuelo Nieto Lemus, Data Warehouse And Big Data Integration, International Journal of Computer Science & Information Technology (IJCSIT) Vol 9, No 2, April 2017.

[50] Fayyad, U.M.; Piatetsky-Shapiro, G.; Smyth, P. Knowledge Discovery and Data Mining: Towards a Unifying Framework. KDD 1996, 96, 82–88.

[51] M. Venkatadri, G. S. Hanumat, and G. Manjunath, "A Novel Business Intelligence System Framework.," Universal Journal of Computer Science and Engineering Technology, vol. 1, pp. 112-116, 2010.

[52] A. Mohamadina, M. R. B. Ghazali, M. R. B. Ibrahim, and M. A. Harbawi, "Business Intelligence: Concepts, Issues and Current Systems," International Conference on Advance Computer Science Applications and Technologies, pp. 234-237, 2012

[53] O. Moscoso-Zea, J. Paredes-Gualtor, and S. Lujan-Mora, A holistic view of data warehousing in education, IEEE Access, vol. 6, pp. 64659_64673, 2018.

[54] O. Moscoso-Zea, Andres-Sampedro and S. Luján-Mora, "Datawarehouse design for educational data mining," 2016 15th International Conference on Information Technology Based Higher Education and Training (ITHET), 2016, pp. 1-6, doi: 10.1109/ITHET.2016.7760754.

[55] Nogueira, Marta, João Galvão, and Maribel Y. Santos. (2020). A Data Modelling Method for Big Data Warehouses. In: Themistocleous, M., Papadaki, M. (eds) Information Systems. In European, Mediterranean, and Middle Eastern Conference on Information Systems 2019. Lecture Notes in Business Information Processing, vol 381. Springer, Cham. https://doi.org/10.1007/978-3-030-44322-1_7

[56] Clegg, D.: Evolving data warehouse and BI architectures: the big data challenge. TDWI Bus. Intell. J. 20, 19–24 (2015)

[57] Russom, P.: Data Warehouse Modernization in the Age of Big Data Analytics, The Data Warehouse Institute (2016)

[58] Ilieva, G.; Yankova, T.; Klisarova-Belcheva, S.; Ivanova, S. Effects of COVID-19 Pandemic on University Students' Learning. Information 2021, 12, 163. https://doi.org/10.3390/info12040163

[59] Karr, Alan F., Ashish P. Sanil, and David L. Banks. "Data quality: A statistical

perspective." Statistical Methodology 3, no. 2 (2006): 137-173.

[60] Buenaño Fernandez, Diego, William Villegas-CH, and Sergio Luján Mora. "The use of tools of data mining to decision making in engineering education—A systematic mapping study." Computer applications in engineering education 27, no. 3 (2019): 744-758.

[61] Forsell, Camilla, and Jimmy Johansson. "An heuristic set for evaluation in information visualization." In Proceedings of the International Conference on Advanced Visual Interfaces, pp. 199-206. 2010.

[62] Jooste, Chrisna, Judy Van Biljon, and Jan Mentz. "Usability evaluation for Business Intelligence applications: A user support perspective." South African Computer Journal 53, no. si-1 (2014): 32-44.

[63] Stewart, Christopher Lee, and M. Ali Akber Dewan. 2022. "A Systemic Mapping Study of Business Intelligence Maturity Models for Higher Education Institutions" Computers 11, no. 11: 153. https://doi.org/10.3390/computers11110153.