DEVELOPING AN INTEGRATED FRAMEWORK TO UTILIZE BIG DATA FOR HIGHER EDUCATION INSTITUTIONS IN SAUDI ARABIA

Noura A. Alsheikh

Information Management Department, Al Imam Mohammad Ibn Saud Islamic University (IMSIU), Saudi Arabia.

ABSTRACT

In recent years, there has been widespread use of the Internet, the Internet of things, mobile devices, networks, and applications. All this usage produces daily huge data that cannot be processed using existing database management techniques and tools because of the size, the volume, the heterogeneity, and the unstructured nature of the data. This has led many sectors like healthcare, business, education, and so forth to start using Big Data technologies to analyze, process, decision making and performance. Big Data is “datasets which could not be captured, managed, and processed by general computers within an acceptable scope” [1]. Education sectors are one of the most important sectors that use information and communication technology (ICT). However, the education sector in Saudi Arabia is still behind other developed countries in terms of the adopting and implementation of Big Data techniques. The aim of this study is to develop an integrated framework to utilize Big Data for higher educational institutes in Saudi Arabia and to support decision making and improve performance. While many studies look at data mining and Big Data in the education sector, there are few studies that touch on this issue in Saudi education, especially in universities. The study collected data through self-administered surveys as a principal quantitative method and through semi structured in depth interviews as the follow-up qualitative method. The study used SPSS software to analyze the data from surveys and used manual analysis to analyze the interview data. This study’s major contribution addresses issues related to the development of a research framework that presents factors affecting the adoption and implementation of Big Data.

KEYWORDS

Big data, education, data mining, Saudi Arabia, Riyadh, factors, adoption.

1. INTRODUCTION

In recent years, a big growth in the use of social media, cloud computing, the Internet of things and smart phone devices has produced a big amount of data that is used every day in different industries, sectors, and areas such as business, healthcare and education. According to are port published by IBM, “every day, we create 2.5 quintillion bytes of data—so much that 90% of the data in the world today has been created in the last two years alone”.

Big Data technology blossomed as a promising technology for analyzing and managing the Big Data generated in the digital world. Big Data technology was raised as a solution to handle this tremendous volume of various data produced by users or technology environments. Thus, Big Data could be referred to as the amount of data that eclipsed the processing abilities of current database systems. Big Data is generated through storing large quantities of data in one main storage facility that supports special transactions [2].

DOI: 10.5121/ijcsit.2019.11103
Nowadays, Big Data has become a hot research topic in many sectors such as business, government, and healthcare due to the development of technology which aiming to collect, organize and analyze big amounts of data and convert these data into information that can be used to make decisions, improve performance, increase efficiency, understand customers’ needs, reduce costs, and develop products and services. Higher educational institutions such as universities produce a big amount of data every day, and using such data effectively can improve decision making. However, the education sector still lags behind other industries in the adoption of Big Data technologies, especially in the context of Saudi Arabia in comparison with other countries [3].

Saudi Arabia’s higher education sector has developed significantly in the last few years in the number of students and universities. According to the [4] “General Authority for Statistics”, there were 20 universities in 2009, but by 2015 this increased to 35. Over the same period, the numbers of new students enrolling had increased by almost 40%, from 272,854 to 379,179 [4]. Saudi Arabia’s higher education sector now focusing on finding and developing new methods to improve educational efficiency, solve current problems, and forecast the future to help the university community (students, faculty members, and staff members) make better decisions. These goals necessitate implementing Big Data technology in the higher education sector [5].

Technology makes many changes in the education sector and improves learning methods, student registration, virtual classes, and the LMS (Learning Management System). These systems produce huge amounts of data every day. Analyzing this data will benefit higher education by improving the decision-making process and increasing efficiency.

2. LITERATURE REVIEW

2.1. INFORMATION TECHNOLOGY

In the 21st century, technology is recognized as a basic element of life, and this era is recognized as the technology era. Technology is recognized as a tool to improve individuals’ work and to make work easier and faster, which could lead society's economic growth. Information technology can be defined as a tool that depends on utilizing computers or any related technologies to have the required information to support the organization in easily performing their duties. Such computerized technologies, such as the Internet, could enhance the dissemination of knowledge [6]. These information technologies could aid in collecting, classifying, arranging, keeping, and spreading knowledge in various sounds, images, and texts [7]. All of the above indicate the urgent need to apply new technologies in education according to their ability to quickly and widely spread knowledge in educational fields, especially in distance-learning education. The main purpose of information technology is to make the education process faster and easier by sharing knowledge throughout the world. These can help teachers and students have up-to-date and accurate information and knowledge necessary for effective teaching and learning. The best way to achieve effective teaching is using the Internet to provide information for students. It is a very easy way to learn and quickly find the right information in a short time. It also helps teachers deliver important information to students. Furthermore, information and communication technology (ICT) has a clear influence on enhancing the teaching process's quality and quantity. The dynamic and interactive nature of ICT in teaching could enhance traditional and distance learning outcomes. Technological means used in today’s learning process include:
1. Smart boards, which are white boards that help teachers actively interact with students by displaying information via image-aided tools and digitally drawn images in the classrooms.

2. Projectors, which are tools that connect the teachers’ laptop and the class white board. This device enables students to see the explained information on teacher laptops.

3. A learning management system (LMS) is technological software used for documenting, managing, reporting, and delivering educational and training programs in an interactive way [8].

2.2. BIG DATA

Consistently, 2.5 quintillion bytes of data were made. About 90% of the data were created in the previous two years. Managing and utilizing such a tremendous amount of rapidly created data will be a difficult challenge. It would be difficult to acquire, preserve, distribute, investigate, and envision information with the current data processing tools[9].

The advanced data management age that was beginning is called the Big Data era. In this era, Big Data involves vast volumes of complicated and developing data sets with various independent sources. Big Data gains a consistent and burgeoning interest from researchers and experts in the scholarly community, industries, and governments where Big Data facilitates novel development, reproduction, and progression in the majority of societies’ fields [10]. [9] illustrated the drastic change of data size and characteristics as a —data revolution‖ or a —data deluge‘. Previously, the data created over the world were small, analog, and provided within determined carriers. In today’s digital era, a tremendous quantity of data is created and produces on a regular basis from miscellaneous resources through various carriers within minutes. The speed and frequency of data diffusion and the ascending number and assortment of sources from which it radiates comprise the data deluge. The digital data available in the world in 2005 was 150 exabytes, and this total increased to 1200 exabytes in 2010. The estimated annual growth percentage of data is 40% for the incoming few years, which corresponds to 40 times of the much-argued world’s population growth. This indicates that the increasing rate of digital data is estimated to rise 44 times from 2007 to the 2020, duplicated every 20 months.[11]. Big Data is a sequence of huge and complex forms of digital information that are unable to be analyzed or processed using conventional techniques. Authorities on Big Data advocate that Big Data is the revolutions that will changes the ways people live, work, and think. Big Data has three distinguished characteristics:
Volume, velocity and variety, which are denoted by the three Vs of Big Data shown in figure 2 below [12].

![Fig.2: The Three Vs of Big Data](image)

[10] Described the three Vs as follows:

1. Volume represents the data units that quickly advance through forms of GB, TB, PB, and more. 2. Variety is the aggregation of high numbers of various files, such as kilobytes file, megabytes audio file, and gigabytes video file.

2. Velocity is described as the processing speed of the data stream, including the data capture and transformation toward the data’s center that indicates launched velocity.

### 2.3. THEORETICAL MODELS OF IT INNOVATION ADOPTION

For this study’s aim, the authors used the framework of Technology, Organization, and Environment (TOE). The TOE model can be clarified as follows. First, calculating information systems’ success is complicated because the effects of information technology are affected by human, organizational, and environmental factors [13], and these factors sit in the framework of TOE. Second, as stated by [14], various researchers used the TOE framework to grasp various IT adoptions like business to business (B2B), ecommerce, enterprise resource planning (ERP), and knowledge management systems (KMS). Third, this model can be used at an organizational level [14], which is appropriate for this research. The framework is applied to build this study’s theoretical framework. It is applied for the aim of the theoretical framework of this research to identify factors that affect the adoption of Big Data technologies in Saudi higher education institutions.

The technological context contains technologies pertinent to the organization—technologies that are used at the organization and technologies that are available in the marketplace but are not actually in use.

The organizational context shows the organizational environment and the organization’s resources and features that ease or constrain the adoption of technological innovations [15]. The mutual method for defining the organizational context is through descriptive measures that IT literature has specified as organizational determinants. Lastly, the environmental context includes
the inter organizational and the external environments in which the firm manages its business [15]. The researcher observed the usage and adoption of the competitor and the partner in addition to the governmental support and policies in the environmental context.

2.4. The Application Of The Big Data Mining Technique In Higher Education

Big Data literature includes numerous studies and relevant research on Big Data in education fields. [16]’s study on Big Data technology's adoption in education sectors found that education sectors utilized only 3% of analytics data. This percentage is not low, but it is not sufficient compared to other sectors. In terms of the empirical findings and outcomes about Big Data in education sectors, this percentage has to be enhanced and advanced. The authors stated that Big Data has powerful capabilities to introduce a new experience for children, to make the learning content more subjective (which will improve the instruction), to supply parents and children with adequate information and data for each program, and to boost the program selection decision. Big Data can also upgrade the recruitment rates by fulfilling occupations with required aptitudes and recognizing the jobs’ requirements, clarify the financial aspects of education, promote the educational management platforms’ competence through viable allotment of educational resources and create proficient arrangement for future education interests. [17] Characterized Big Data as a promising revolution in higher education based on advanced technologies in big data that can enhance academic employees’ competence. Further advantages for higher education include: enhanced leadership and administration, distinguishing understudies’ successes and hazards at an early stage, using —consider the possibility that” tests for decisions, and improving foundational productivity through viable reactions based on constant access to information. [16] Emphasized the ability of data mining and data analytic to quickly provide feedback for pupils and instructors to comprehend their academic performance. This feedback will give an in-depth perspective on specific educational patterns and capture worthy information. Big Data can be used to estimate the assistance the education system needs to offer students, which can and avert failures and drop out risks. Instructors can take advantage of big data’s scale by exploring, revealing, and learning novel approaches and educational methods. This possibility advances the education system by enhancing teachers’ ability to align the appropriate teaching technologies with each student’s knowledge and demands. [18] Highlighted two main benefits of Big Data in higher education. First, Big Data can be used for prediction tools to enhance individual students’ academic performance. Second, Big Data guarantees high-quality standards in academic programs by constructing academic programs based on student performance data gathered at each level and addressing students’ requirements through customized modules, assignments, feedback, and learning trees that drive better and richer learning. Furthermore, for the purpose of creating new learning experiences, enhancing pupils’ achievements in all classes, diminishing dropout rates, and boosting graduation rates, higher education can utilize Big Data technologies to assess and interpret the abilities and achievements of each pupil and to establish an individual learning experience aligned with their educational needs.

3. Research Methodology

The methodology represents the plan to answer the research questions. The research questions, by nature, dictate the design. The researcher usually selects explanatory, descriptive, or exploratory, and the research designs are often selected based on the study’s objective and the extent of the currently available research on the subject[19]. This research was conducted to identify the factors impacting the implementation and adoption of big data techniques in higher education institutions in Saudi Arabia and to identify the use of Big Data in higher education institutions to support the decision-making process for staff members, student and faculty member. The study’s methodology is based on a mixed methodology that includes qualitative and quantitative methods. To first collect quantitative data from the participants (staff members and faculty
members who work in the college of computer and information science and in the deanship of information technology), the researcher made a self-administered survey. Then, the researcher made in-depth and semi-structured interviews. In this research, the researcher conducted self-administered surveys as a key quantitative method and in-depth interviews as a qualitative follow-up method. This research begins by determining the factors that impact the adoption of Big Data using three frameworks: Information Technology Innovation Adoption, Rogers’s Diffusion of Innovations (DOI), and Technology–Organization–Environment (TOE). The researcher determines these factors using the TOE framework, and Figure 3 below shows the research procedure.

![Research procedure](image)

The survey was completed by 69 participants who work in the deanship of computer technology, the data center, and the college of computer and information science in three Government Riyadh University (Imam Muhammad Ibn Saud Islamic university, King Saud University and Prince Noura Bint Abdulrahman University) over two months from October 2018 to December 2018. The interview was conducted with six interviewees from three Saudi universities (two interviewees from each university).

4. Results and Discussion

This section reports the results of the study and discusses them in light of the literature and related works.

A. Factors Influencing the Adoption of Big Data Technology

This section describes the most important factors that influence the implement of big data, the factors is determine base on the mean and standard deviation value, if the means and standard deviation is high that indicated this factor has high influence.

1) Environmental Factors:

Governmental support and policies refer to the degree to which governmental support and policies for education influence a university on Big Data adoption. Awareness refers to what degree the university believes in the importance of Big Data adoption. Table 1 shows that the most important issue in the environmental factors was a market-related lack of awareness about Big Data technology. The second most important factor, based on responses, is related to university policy; this strongly influences the adoption of big data technology, but, based on the standard deviation value, the respondents have various opinions. This may indicate there are various sources that universities can collect data from to store in a data warehouse. The results revealed that the university does not have a ready data warehouse that can store big data.
Moreover, all participants agree that Big Data is very important to support decision making in the university, however, top management needs to recognize this benefit.

2) Technological Factors:

In a technological context, the researcher studied perceived benefits from big data, the simplicity of using the system, the compatibility with current systems, security and privacy, and integration as the adoption factors for big data. The perceived benefits from big data are related to how much universities perceive receiving benefits such as, improved student outcomes, decreased dropout rates, increased teaching quality, and improved performance. There is no reason to implement big data in universities without recognizing the benefits, usefulness or relative advantages. The simplicity of system usage refers to the Big Data system’s ease of implementation and operation. The simplicity of system usage is related to the usability and availability of the system. Therefore, the system’s simplicity in implementation and utilization would promote the adoption of Big Data in universities. Compatibility with the existing system is the degree to which Big Data technologies are compatible with the existing system. Security and privacy are associated with big data’s obstacles in

<table>
<thead>
<tr>
<th>Technological Factors</th>
<th>Descriptive analysis</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applying big data technology will be difficult to understand and use.</td>
<td>3.07</td>
<td>1.12</td>
<td></td>
</tr>
<tr>
<td>IT infrastructure does not support big data technology.</td>
<td>3.27</td>
<td>1.17</td>
<td></td>
</tr>
<tr>
<td>The complexity of dealing with unstructured data.</td>
<td>3.86</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>Security concerns while using and analyzing data using big data techniques</td>
<td>3.78</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>The current system is not compatible with big data technology.</td>
<td>3.18</td>
<td>1.03</td>
<td></td>
</tr>
<tr>
<td>The university does not realize the benefit of utilizing big data</td>
<td>3.30</td>
<td>1.25</td>
<td></td>
</tr>
</tbody>
</table>

To summarize, the interview results showed that universities have several technological challenges when considering the adoption of big data, including inappropriate IT infrastructure with a shortage of storage space and processor speed, a lack of ready data to apply to Big Data technologies, and a problem in integrating data from other databases.

3) Organizational Factors:

Management support, technological competence, and financial investment competence are considered parts of the organizational context. Management support refers to how important

Table 1: The influence of environmental factors.

<table>
<thead>
<tr>
<th>Environmental Factors</th>
<th>Descriptive analysis</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of awareness about big data technology.</td>
<td>3.68</td>
<td>1.14</td>
<td></td>
</tr>
<tr>
<td>There is no resource for data collection and processing</td>
<td>2.84</td>
<td>1.39</td>
<td></td>
</tr>
<tr>
<td>The university policy does not support big data technology.</td>
<td>3.23</td>
<td>1.12</td>
<td></td>
</tr>
</tbody>
</table>
management perceives the tasks of Big Data to be. Financial investment competence refers to how much a university is able to invest in introducing and operating Big Data at the university. This includes equipment, software packages, and consulting. Table 3 shows the most common organizational challenges reported by the respondents. These are based on the mean score and include cost, experience, strategic support, and the availability of support. The lack of experience issues have a higher value based on the mean, and the size of the organization has the lowest value based on the mean. Also, the availability of support has a higher value based on the standard deviation. The researcher believes the university’s strategies for adopting and implementing this technology differ from one organization to another.

Table 3: The influence of Organizational Factors

<table>
<thead>
<tr>
<th>Organizational Factors</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of experience in dealing with Big Data technology</td>
<td>4.01</td>
<td>0.83</td>
</tr>
<tr>
<td>Big Data technology is costly to implement.</td>
<td>3.23</td>
<td>1.17</td>
</tr>
<tr>
<td>The size of your organization is too small for Big Data adoption.</td>
<td>2.01</td>
<td>1.13</td>
</tr>
<tr>
<td>The Big Data strategy is not supported by top management.</td>
<td>3.11</td>
<td>1.10</td>
</tr>
<tr>
<td>Availability of support for implementing and using Big Data technology.</td>
<td>3.37</td>
<td>1.15</td>
</tr>
</tbody>
</table>

In conclusion, the interviews revealed challenges related to organizational factors, including management support, finances, and awareness of Big Data technology.

B. Perceived Benefits of Applying Big Data Technology

Education is often connected with new technologies like ubiquitous computing devices and flexible class room design. But the most important and overlooked aspect that would revolutionize the field is data. Higher education is a field where tremendous amounts of data are available. Analyzing this data would lead to a great benefits.

Big Data and predictive analytics are increasingly used in higher education to classify students and to predict student behavior. Institutions of higher education, sometimes working with commercial providers, have begun using these methods to recommend courses, monitor student progress, individualize curriculum, and even build personal networks among students. Data mining, as a major part of business intelligence, is held to be part of a radically different future for higher education in which the ability to predict individual outcomes revolutionizes management and allows institutions to understand better about their students and their needs by taking advantage of the vast trove of data that institutions generate in their operations. Saudi universities have started realizing the importance of Big Data and its utilization in providing a whole vision for decisions, saving time, making decisions based on accurate data that will produce good student results, reducing the dropout rate, enhancing the productivity of students and staff members, reducing costs, organizing the university’s jobs, and increasing the quality of the final semester report.

This finding corresponds to the findings of other studies including [20] and [14]. The results indicate that the adoption and implementation of Big Data technology helped find the right decisions for the right person in a timely manner. It also helps predict the future, like the number
of students who will pass the course, the dropout rate, and the number of students who will register for this course the next semester. This is in agreement with [21] and [22].

Applying Big Data technology will also help reduce student queries regarding registration, exams, and more. It's indicated that adopting and implementing data mining techniques helps to accelerate responses to queries, to avoid errors, and to improve the reliability and availability of information. These issues mean that Big Data techniques have a good reputation.

Finally, the researcher believes that when a university wants to enhance quality, make the right decisions, increase student performance, improve staff members’ efficiency, and reduce costs, it can be achieved by adopting big data technology. This framework presented all issues in three stages: preparation and extraction, construction, benefits, and outcomes.

C. Framework Development

The final result from this study is an integrated framework that describes the factors affecting the adoption and implementation of Big Data and the utilization of Big Data in higher education institutions. It is arranged in the form of a project cycle where the adoption of Big Data technology can be considered a project that contains a number of stages, with each stage affected by a number of factors.

5. LIMITATIONS AND FUTURE RESEARCH

The researcher worked hard to make this study as comprehensive as possible but certain limitations we represent. The study has several limitations that should be considered. These issues are related to time, context, methods used, and other issues.

The study was constrained to IT departments, datacenters, and faculty members in the college of computer and information sciences at Riyadh Government University. Thus, the results may not be generalizable for other countries where the adoption and implementation of technology varies. The number of participants in the survey was lower than was hoped for; this reflects the
policies and regulations of the university. A strategy was implemented to distribute the questionnaire by submitting an approval letter to the department heads asking for their assistance in distributing questionnaires; this was not sufficient to encourage a higher response rate. The other limitation is the fact that Saudi universities had little understanding of Big Data concepts and how they can be applied in education. They tend to focus on the extracted information from structured data to enhance educational quality, whereas Big Data can also use unstructured data.

Based on the study’s findings, the researcher provides some suggestions and recommendations for future research.

- All universities thinking of applying Big Data in the future should consider the technical plans for the coming years. How they can build a data warehouse? What types of data will be used? Mining for new sources of data and analysis in real-time will have a very important role.

- Some universities are still looking for the best mining software and the proper mining technologies to find new insights according to their goals and strategies, so the need for a deeper study about the technologies of data mining.

- As mentioned in the study limitations, the study participants were staff working in the IT departments and data center and faculty members in the college of computer and information sciences in a university. Further research is necessary in other sectors to know more about issues influencing the adoption of big data technology.

- The framework could be used as a starting point for further research, and it could be developed through the participation of other cities in Saudi Arabia to measure the results of the adoption and implementation of Big Data technology.

6. CONCLUSION

This study explored the factors that affect the adoption and implementation of Big Data in higher education institutions in Saudi Arabia. This chapter outlines the research's aim and objectives by answering the research questions through analysis of the data collected during the survey and interviews. It was found that some factors identified in this study have already been mentioned in the literature, and this confirms previous research and findings. In addition, new factors have emerged from this research. Factors were integrated from the previous literature and current research in the research framework presented in Figure 4 in the previous section.

The research was conducted with great care, but some limitations remain. In addition, this research has provided several new contributions to the context of Saudi education, and some recommendations have been made.

ACKNOWLEDGMENTS

First and foremost, I praise and thank Allah Almighty for giving me the strength, knowledge, ability and opportunity to complete this thesis. During my research, many essential individuals made an impact in my life to reach this result. I would like to send them my genuine appreciation: Firstly, I would like to express my sincere gratitude to my advisor Prof. Dr. Adulrauf Baig for his continuous support on my work and related research, for his patience, motivation, and immense knowledge. His guidance helped me in preparing and writing my research. I would also like to offer my warm thanks to Dr. Mohammed Altyar for his advice and guidance, who assist me to accomplishment this work successfully.
REFERENCES


