Identifying Barriers to Big Data Analytics: Design-Reality Gap Analysis in Saudi Higher Education

Bandar Abdullah AlMobark

College of Computer and Information Sciences
Imam Mohammad bin Saud Islamic university (IMSIU), KSA, Riyadh
bamobark@imamu.edu.sa

Summary

The spread of cloud computing, digital computing, and the popular social media platforms have led to increased growth of data. That growth of data results in what is known as big data (BD), which seen as one of the most strategic resources. The analysis of these BD has allowed generating value from massive raw data that helps in making effective decisions and providing quality of service. With Vision 2030, Saudi Arabia seeks to invest in BD technologies, but many challenges and barriers have led to delays in adopting BD. This research paper aims to search in the state of Big Data Analytics (BDA) in Saudi higher education sector, identify the barriers by reviewing the literature, and then to apply the design-reality gap model to assess these barriers that prevent effective use of big data and highlights priority areas for action to accelerate the application of BD to comply with Vision 2030.

Keywords: *Big Data, Big data analytics, Higher Education, Vision 2030, design-reality gap model*

1. Introduction

In 1997, the BD term was used for the first time to explain the visualization of data and the challenges it poses to computer systems [30, 31]. With the technological revolution, vast amounts of data are generated through the increased use of devices, and remote sensors produce from the vast amount of heterogeneous data that is either organized or unorganized, where this data is known as BD. It is one of the most critical enablers that organizations can use to generate value[31]. Governments have begun BD strategies, among them the government of Saudi Arabia, within the vision of 2030, which brings many benefits, including the promotion of research and development, new job opportunities, and better support for decision-making [8]. The fact that we are in the information age is very precious. If the BD is exploited, that will have many benefits for the higher education sector from improving education and curricula, providing a new platform for education, development of the new teaching model, and encouraging the exchange of research and innovation. According to the previous study, BDA in Saudi higher education sector is still in their infancy. Where Saudi higher education needs to identify and address concerns and barriers that related to the implementation and growth of big data to achieve the 2030 Vision [24].

2. Big Data

BD is the data that exceeds the amount of technology in its capacity to store and manage it, and the efficiency in processing it [17]. The term "Big Data" differs from traditional data, as it includes a set of large and growing (structured, unorganized, and semi-structured) data [25]. BD has four diminutions: (Volume, Velocity, Variety, and Visibility) the amount of BD available enormously, the high speed with which the data is available, the great diversity in data forms and types, and finally the increased availability and accessibility of data [12].

2.1 BD Challenges

BD includes many technical challenges that institutions must overcome to benefit from this data [21]. The growth rate of BD generates many challenges, as this data indicates data storage units in a range that exceeds the capacity of current storage and processing systems [17]. BD faces many challenges that limit its exploitation, begin from the processing systems until the means of analysis, these challenges may be caused by the characteristics of the big data itself, the methods of analysis, or because of the restrictions imposed on the processing systems [16].

3. BD in Higher Education

Higher education institutions have access to large data sets and analysis tools, as the amount of student data that needs to be analyzed is increasing [3]. BD is going to revolutionize higher education. Whereas BD will lead to a qualitative leap of technological progress, which will increase academic effectiveness [23]. BD can be applied to multiple aspects in the areas of administrative and educational higher education, such as employment, financial management, admission processing, and student performance monitoring [26].

3.1 Challenges of BD in Saudi higher education

There are various expected complications related to the execution of analytic techniques of BD in Saudi higher education. Previous studies indicated that implementation of BD in the accreditation of higher education in Saudi universities constitutes significant security concerns concerning storage, processing, and management. External data sources such as social media constitute one of the most security concerns as the accumulation of immense and unorganized data resulting from them requires the adoption of more advanced security models and safety solutions conventional is useless in this case [22].

Privacy constitutes one of the challenges of implementing BD in higher education in Saudi universities concerning the protection of user's personal data and intellectual property rights. It is necessary to adopt applications to encrypt private data and hide identity before working on data analysis in order to maintain privacy [28].

Supporting top management plays a crucial and decisive role in approving BD by supporting projects related to BD through financial and administrative support, spreading awareness, and removing obstacles [11].

Analytical skills are among the most critical challenges facing the adoption of BD. The fact is the Saudi higher education is in the cradle stage concerning BD. There is a lack of cadres and expertise in dealing with this modern trend of data [2].

BD collection, storage, analysis, and management needs tremendous capabilities and systems, which poses a challenge to the IT infrastructure in Saudi higher education [3].

The Saudi Vision 2030 represents a framework for transformation into a knowledge-based society. This transformation is completely different from any transformation that can happen in many countries. It has shifted from primary dependence on oil to other vital resources, and this requires the participation of higher education in this ambitious vision. Saudi higher education is rich in data yet learning analyzes are relatively early in the Saudi Arabia [13].

The challenges facing Saudi higher education can be summarized in security concerns, maintaining privacy, supporting senior management, analytical skills, and IT infrastructure. The problem is that few studies have analyzed the BD barriers in Saudi higher education. In an attempt to address this literature gap, the realistic-design gap model will be applied as a basis for analyzing these barriers that prevent the effective use of BD for the development of Saudi higher education.

4. Conceptual framework

The framework, which was developed by Richard Hick [14], is used for organizational change analysis and risk analysis associated with this change. In the field of applying BD in Saudi higher education, there were no studies that used the design-reality gap model. In this paper, the model used to evaluate the possibility of addressing barriers and profiting from BD to achieve the 2030 Vision by comparing the availability of requirements for applying BDA with the capabilities of the higher education situation in the current reality.
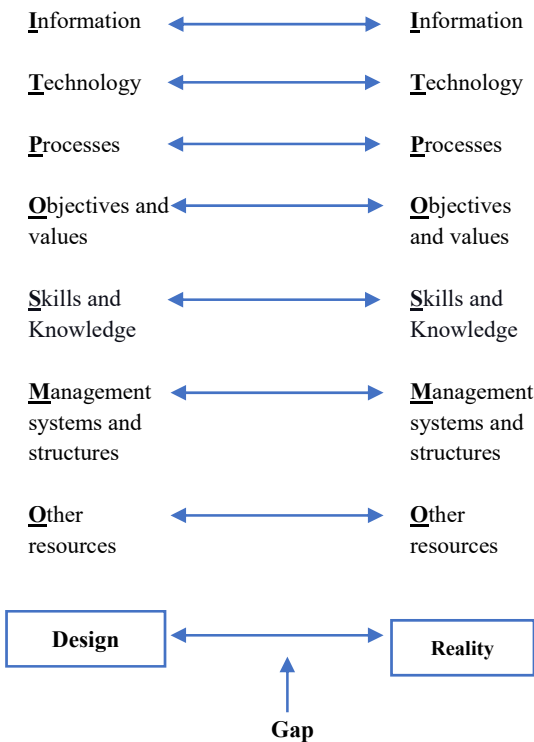

Fig. 1 Design-Reality Gap Model

4.1 Design- Reality Gap Model

Seven dimensions have a design-reality gap around it (Figure 1), which summarized in "**ITPOSMO**":

Information: includes both information and data.

Technology: focuses on information and communication technology (ICT) that handles data.

Processes: activities required to generate, capture, and analyze data.

Objectives and values: both issues of strategies and cultural values .

Skills and knowledge: human capabilities to perform the operations related to data.

Management structures and systems: management and structure of the organization.

Other resources: considerable investments like budget.

1.1. Analysis of Barriers to BD in Saudi higher Education

Due to the limited time available for research, the following analysis is illustrative and not final. The analysis based on sources of evidence from the literature review to assess the gap in each of the ITPOSMO dimensions.

Information: active development of BD requires the 5 V's of BD (Volume, Variety ,Veracity, Velocity, Value) To meet the data quality standards [9]. In reality, Saudi higher education institutions have a massive amount of BD that flows from all directions through internet applications, program-based exercises, online classrooms, social media, blogs, and student surveys. Therefore, the volume of data is enormous. As for speed, the data swells and grows faster and faster, but most data sets are not dynamic; some data are not available in real-time. An example of this blackboard system, the courses not automatically downloaded at the beginning of the term, as is supposed. However, communication channels in higher education site or even on the university's sites, are active and are responded directly through them. Also, there is a variety of mixed data; there are many different categories of forms of data sets that need much effort to control and deal with them. There is also confusion between data quality and other issues such as accuracy, representativeness, and ease of use, which means that there are still some challenges to be addressed. However, it is obligatory developing a plan to reduce design reality gaps in terms of information [6].

Technology: BD requires that the ICT infrastructure be robust and support all stages of the data chain from its generation, capture, analysis, storage, presentation to visualization, and dissemination. It should be widely distributed and interoperable [15]. Saudi higher education institutions have a massive amount of BD flowing from all directions, and there is inflation in public data. All of this data is putting pressure on the current IT infrastructure, the software solutions in place that have been around since the 1970s cannot meet current data demands. Because of that, many institutions of higher education did not reap the real value of BD. The Ministry of Higher Education also

indicated that one of its plans to achieve the 2030 Vision is to improve services and infrastructure .However, cloud computing, for example, provides an ideal solution to avoid the challenges of IT infrastructure and to reduce infrastructure costs [24]. Cloud computing is the perfect substitute for educational institutions that do not have a sufficient budget for investments [18, 31].

Fig. 2 Information Value Chain

Processes: BD requires enacting all of the information value chain processes (Figure 2). The higher education in Saudi Arabia is wealthy in data. However, learning analytics is in its relative earliest stages in Saudi Arabia [7]. There is significant schizophrenia within the chain, as the existence of data sets is much more potent than their use [4].

Objectives and Values: BD needs a set of laws and regulations covering data security and privacy issues [20]. BD culture should be spread and confidence in the role and appreciation of data enhanced. Saudi higher education system promotes dependence on safe and reliable sources as it has obtained the certificate (ISO 27001), which is the international standard for information security. The IS team is also working on many initiatives in the field of security. Among these initiatives are periodic evaluation of information security by an independent authority specialized in information security, digital authentication, confidentiality and integrity of information, and electronic signature. The Ministry also built and operated the Information Security Operations Center. One of its objectives is to monitor and implement information technology policies and information security measures. As for the promotion of the data culture, for a very long time, higher education institutions did not have the resources to control the process of storing and analyzing data effectively, but with the emergence of next-generation technologies such as Hadoop, this is no longer an issue. Because ignoring data is more expensive than the cost of maintaining. Accordingly, it has become necessary to create a culture that respects data and makes data analysis an integral part of every decision taken by the academic institution. Therefore, the higher education sector seeks to achieve this to keep pace with Vision 2030.

Skills and Knowledge: BD requires that there are full awareness and knowledge of BD and the availability of the necessary capabilities that must work properly across the

value chain of information. Although some BD workshops have started as a system ("Itqan") developed by King Saud University to analyze BD, however, there is still a lack of awareness and knowledge of BD in Saudi higher education and in the experiences to deal with this data. Where an evaluation study conducted for the "Itqan" system, showed that the university's use of the system services in analyzing the critical evidence for the university is less than expected compared to the capabilities of the system that was built, which weakened its usefulness in supporting the decision-maker [1]. Higher education institution must extend their effort to bridge the skills gap and take the necessary measures to control the significant changes that are likely to happen as a result[19].

Management Systems and Structures: It involves a BD strategy requirement and governance across the higher education sector and issues related to data standards, sharing, security, and privacy [27]. The higher education sector has adopted a policy of acceptable use of resources and technical services, which is concerned with the governance of safe use of the services of the information technology sector. This policy aims to provide secure electronic services in line with government regulations and directives issued in this field, and the Ministry hopes that this policy will help in protecting technical and conservative resources and services on confidentiality, integrity, and privacy of data, and to ensure that information sources are available when needed while clarifying the importance of users understanding of their responsibilities to protect information.

Other Resources: BD requires a considerable investment of money, time, and funding has always been a problem in higher education institutions in developing countries [10]. About 192 billion riyals has been allocated to the public education, higher education, and workforce training sector, which included the amounts allocated to the initiative to achieve the Vision 2030, estimated at 4.89 billion riyals.

Although BD analyzes can become part of the solutions for the evolution of higher education sector development to achieve vision 2030, however, the utilization of data for the advancement and development of Saudi higher education is still in its very early stages. Much remains to be done to obtain the promised value to benefit from the data in organizational development, whether by analyzing BD, building knowledge, or using open data [5].

5. Conclusion

The data is increasing and spreading in various formats within the higher education institutions, which makes it challenging to retrieve or consolidate it. Consolidated data from different sources provide a basis for making better decisions regarding main business and technical needs, reducing redundancy, and wasting valuable time to recover data from multiple sources, which reveals the importance and value of BD.

This paper presented the current state of Saudi higher education regarding BD and barriers that prevent its adoption and gain it worth, which affects keeping up with the Saudi vision. BD has a clear role in the future of development, but this role will not be achieved unless it is possible to overcome the barriers that prevent BDA. By observing reality, tallying the past, looking ahead, and by reviewing the design-reality gap, it can see that there is still a wide range of severe barriers to BD in the Saudi higher education sector. From comparing the actual (reality) performance with the desired performance, the presence of gaps has been revealed (Table 1).

To bridge that, must develop a plan with main dimensions lie in setting a time frame for achieving the goals, adjusting the goals in a way that suits the requirements and resources, undertaking the redistribution of resources to ensure the achievement of the goals, and getting additional resources.

Vision 2030 is one of the incentives for data attention in the Kingdom of Saudi Arabia, which represents a long-term strategic plan for the Kingdom. Where significant goals defined in the vision, its progress is measure through a set of performance indicators that rely heavily on collecting data from the target indicator and then analyzing it. Therefore, BDA is an essential tool for government agencies to draw their plans and indicators, and to measure and follow up these indicators periodically. The process of applying the future vision is one of the essential stages to start translating the vision, mission, and goals into practice.

The study limited in the case of higher education sector because the approach followed in this study is a theoretical review, and qualitatively, it is possible to conduct an in-depth study in the future study by surveying and interviewing officials to obtain more accurate results and to measure the development of Saudi Arabia BDA in different sectors.

References

[1] Alaklubi A., " دراسة: البيانات الضخمة واتخاذ القرار في جامعة الملك سعود: تقييمية لنظام اتقان," Journal of Information Studies & Technology (JIS&T), vol.2018, no.2, 2018.

[2] Ahmed A. I., "BIG DATA FOR ACCREDITATION: A CASE STUDY OF SAUDI UNIVERSITIES," Theoretical & Applied Information Technology, vol.91, no.1, pp.130-136., 2016.

[3] Aldiab A., et al., "Prospect of eLearning in Higher Education Sectors of Saudi Arabia: A Review," Energy Procedia, vol.110, pp.574- 580., 2017.

[4] Alharthi A., et al., "An exploratory study for investigating the critical success factors for cloud migration in the Saudi Arabian higher education context," Telematics and Informatics, vol.24, no.2, pp.664-678., 2017.

[5] Alsayd M. E.," الاستفادة من البيانات الضخمة والمفتوحة في التنمية الإدارية ودورها في تحقيق رؤية المملكة 2030," Electronic Interdisciplinary Miscellaneous Journal (EIMJ), 2019.

[6] Alsheikh N. A.," Developing an Integrated Framework to Utilize Big Data for Higher Education Institutions in Saudi Arabia," International Journal of Computer Science & Information Technology (IJCSIT), vol.11, no.1, pp.31-42, 2019.

[7] Alsurehi H. A., and Youbi A. A., "Towards applying social networking in higher education: Case study of Saudi universities," International Journal of Academic Research, vol.6, no.5, pp.221-229, 2014.

[8] Amatriain X.," Beyond data: from user information to business value through personalized recommendations and consumer science," In Proceedings of the 22nd ACM international conference on Information & Knowledge Management, Los Gatos, CA 95032, USA, pp. 2201-2208, 2013.

[9] Beyer M. A., and Laney, D. "The importance of 'big data': a definition," Gartner ,2012.

[10] Dasuki S., et al., "Evaluating the implementation of international computing curricular in African universities: A design-reality gap approach," International Journal of Education and Development using Information and Communication Technology (IJEDICT), vol.11, no.1, pp.17-31, 2015.

[11] Dong L., Neufeld, D., & Higgins, C.," Top management support of enterprise systems implementations, "Journal of Information technology, vol.24, no.1, pp.55-80, 2009.

[12] Dumbill E., "Volume, velocity, variety: What you need to know about big data," Forbes, 2012.

[13] Habibi N. "Implementing Saudi Arabia's Vision 2030: An Interim Balance Sheet," Middle East Brief, no.127, pp.1-9, 2019.

Table 1: Design-Reality Gap Summarization

Design-reality gap diminutions	BD Design needs	The reality of BD in Saudi higher education
Information	5 V's of BD	-Most data sets are not dynamic. -There are many different categories of forms of data sets that need much effort to control. -There is also confusion between data quality and other issues.
Technology	Robust ICT infrastructure	The Ministry of Higher Education indicate that one of its plans to achieve the Vision is to improve infrastructure.
Processes	Enact in all of the information value chain processes	There is significant schizophrenia within the chain, as the existence of data sets is much more powerful than their use.
Objectives and Values	Laws, regulations and culture	-Saudi higher education system has obtained the certificate (ISO 27001). -Ministry IS team working on many initiatives in the field of security. -The Ministry built and operated the Information Security Operations Center. -The higher education sector seeks to promote the data culture to keep pace with Vision 2030.
Skills and Knowledge	Awareness and capabilities	There is still a lack of awareness and knowledge of BD in Saudi higher education and in the experiences to deal with this data.
Management Systems and Structures	Strategy requirement and governance	The higher education sector has adopted a policy of acceptable use of resources and technical services, which is concerned with the governance of safe use of the services of the information technology sector.
Other Resources	Budget	192 billion riyals has been allocated to the public education, higher education, and workforce-training sector.

[14]   Heeks R., "Information Systems and Developing Countries: Failure, Success and Local Improvisations," The Information Society, vol.18, no.2, pp. 101-112, 2002.

[15]   Helbing D., and Balietti, S., "From social data mining to forecasting socio-economic crises," The European Physical Journal Special Topics, vol.195, no.1, pp.3-68, 2011.

[16]   Jin X., et al., "Significance and challenges of big data research," Elsevier, vol.2, no.2, pp.59-64, 2015.

[17]   Kaisler S., et al., "Big data: Issues and challenges moving forward," Hawaii International Conference on System Sciences, pp. 995-1004, 2013.

[18]   Mohamed F AlAjmi, Arun Sharma and Shakir Khan, "Growing Cloud Computing Efficiency" International Journal of Advanced Computer Science and Applications(IJACSA), 3(5), 2012. http://dx.doi.org/10.14569/IJACSA.2012.030526

[19]   Khan S., "Modern Internet of Things as a Challenge for Higher Education," International Journal of Computer Science and Network Security, vol.18, no.12, pp.34-41, 2018.

[20]   Kitchin R. "The real-time city? Big data and smart urbanism" GeoJournal, vol.79, no.1, pp.1-14, 2014.

[21]   Ma M., Wang P., and Chu C. H. "Data management for internet of things: Challenges, approaches and opportunities," IEEE International conference on green computing and communications and IEEE Internet of Things and IEEE cyber, physical and social computing, pp. 1144-1151, 2013.

[22]   Manyika J., et al., "Big data: The next frontier for innovation, competition, and productivity," McKinsey Global Institute, 2011.

[23]   Moreno J., Serrano M. A., & Fernández-Medina E., "Main issues in big data security," Future Internet, vol.8, no.3, 2016.

[24]   Mukthar M. A., and Sultan M. M., "Big Data Analytics for Higher Education in Saudi Arabia," International Journal of Computer Science and Information Security (IJCSIS), vol.15, no.6, 2017.

[25]   Oussous A., et al., "Big Data technologies: A survey," Journal of King Saud University-Computer and Information Sciences, vol.30, no.4, pp.431-448, 2018.

[26]   Picciano A. G., "The evolution of big data and learning analytics in American higher education," Journal of asynchronous learning networks, vol.16, no.3, pp.9-20, 2012.

[27]   Shlomo N., and Goldstein H., "Big data in social research," Royal Statistical Society, pp.787-790, 2015.

[28]   Tulasi B., "Significance of Big Data and analytics in higher education," International Journal of Computer Applications, vol.68, no.14, 2013.

[29]   Wang Y., Kung L., and Byrd T. A., "Big data analytics: Understanding its capabilities and potential benefits for health care organizations," Elsevier, vol.126, pp.3-13, 2018.

[30]   Ylijoki O., and Porras J., "Perspectives to definition of big data: a mapping study and discussion," Journal of Innovation Management, vol.4, no.1, pp.69-91, 2016.

[31]   S. Khan, A. S. Al-Mogren and M. F. AlAjmi, "Using cloud computing to improve network operations and management," 2015 5th National Symposium on Information Technology: Towards New Smart World (NSITNSW), Riyadh, 2015, pp. 1-6

Dr. Bandar Almobark received his Bachelor, Master and PhD in 1997, 2002 and 2007 respectively. He is member of many scientific committees in Saudi Arabia. He is currently working as Assistant Professor at College of Computer and Information Sciences in Imam Mohammad Ibn Saud Islamic University, Riyadh (Saudi Arabia). His research interest are Information Retrieval, Digitization, Electronic Record Management, Software and Application Development, Content Analysis and Data Analysis, Emerging Technology, Open Source Software, Library Automation and Mobile / Web Application. He published many research papers in international journals and conferences in his research domain. He has around 15+ years of experience in managerial, teaching and research. Dr. Mobark is teaching bachelor degree courses in the college of computer at Imam University.