Journal of Engineering and Applied Sciences 14 (4): 1203-1210, 2019
ISSN: 1816-949X

Creation of Value Through Big Data in Higher Education Institutions: A Survey

[1]Camilo Guerrero, [2]Alejandra Hernandez Acosta and [3]Javier E. Sierra
[1]Faculty of Administrative and Financial Sciences,
Corporacion Universitaria Del Caribe-CECAR, Sucre, Colombia
[2]Faculty of Administrative and Financial Sciences,
Universidad Abierta Para Adultos-UAPA, Dominican Republic
[3]Faculty of Engineering, Universidad De Sucre, Sucre, Colombia

Abstract: Big data proposes the analysis of large volumes of information in companies or organizations. An analysis is made to different researches related to the application of big data for the creation of value and in this way, to contribute to decision making, focused on how this contributes to higher education institutions. The objective of the article is to provide information, through a bibliographic review, on how it is possible to create value through big data. The creation of value is the most important feature of big data. In this study we show how the theories and concepts, applied to different organizations and companies can be applied to higher education institutions. The application of the most relevant theories for the generation and creation of value, will allow the institutions of higher education to improve the decision making and to be more efficient in the investment of the resources.

Key words: Big data institutions of higher education, creation of value, decision making, information, application, ibliographic review

INTRODUCTION

Big data involves the collection and analysis of large quantities of data to reveal patterns and trends. The accumulation of large amounts of information in the data processing centers of companies institutions and government organizations which have been carried out in recent years is one of the most relevant events that have occurred in Information and Communication Technologies (ICT) recently (Agarwal and Dhar, 2014).

For the specific case of higher education institutions (Universities), academic information systems for student enrollment processes, compilation of grades, financial processes for educational credits or university welfare software (for tracking students with physical, cognitive or economic deficiencies), generate a large amount of information. The analysis of the generated information can be used to propose strategies that benefit the academic development of the students, better manage the human talent of the universities and above all for the making of strategic decisions of higher education Institutions.

The scientific literature on information and communication technologies and the management of universities, review the relevance of big data (Big Data Analytics BDA) as a great tool for:

- Detect preferences of study programs chosen by the students (Tan *et al*., 2017)
- Improve campus efficiency and reduce costs (Yan and Hu, 2016)
- Detect preferences of research lines for the realization of the student's degree thesis (Xi, 2017)
- Identify expectations and needs of the people in front of the web pages of the universities (Barcellos *et al*., 2017)
- Conduct analysis of the relationship between student learning and reading data on the internet or in the university databases and the attendance at classes (Yan and Hu, 2016), to validate the quality of education by performing evaluations of teachers based on FAHP (Fuzzy Analytical Hierarchy Process) among others

The application of big data in the universities allows:

- Investigate and analyze the generated information, coming from the large amounts of heterogeneous data, to detect study preferences, lines of research, expectations and needs of students, administrators, parents and the general public

Corresponding Author: Camilo Guerrero, Faculty of Administrative and Financial Sciences,
Corporacion Universitaria Del Caribe-CECAR, Sucre, Colombia

- Analyze, effectively and quickly, the information, to improve the efficiency in the administrative and financial universities
- The study about the relationship between the learning of the students and the existing information in the academic database of the universities
- Determine the impact on the quality of education by assessing the different systems of evaluation of the same in universities

This study seeks to show through a bibliographic review how the most relevant theories of generating value through big data can apply to higher education institutions

Big data: Big data is a trend that has opened the doors to a new approach of understanding and decision making in companies and is used to describe huge amounts of structured, unstructured and semi-structured data which would take a lot of time and resources to load them in a relational database for its analysis (Shafakhatullah Khan (Member IAENG), Khan *et al*., 2018). Big data is a term used to refer to the increase in the volume of data which presents storage, processing and analysis difficulties, through the traditional databases technologies (Hashem *et al*., 2015).

The big data conceives the information characterized by (Berman, n.d.), like this: volume (large amounts of data), variety (numbers, images, documents, others) and speed (content constantly changes for information from many sources)

Big data is not only characterized by the three previous ones but it can also be extended to a fourth one the value (Gantz and Reinsel, 2011) understanding the latter as the most important aspect of big data, since, it refers to process of discovering large hidden values, from large amounts of data with various forms of rapid generation. In this sense, the perspective of Barnes (2013) in which big data is the digital trace left by all the activities that are generated, since, we connect to an intelligent device is what makes it useful to create mechanisms that are capable of process all that information using structured data. For example, demographic data such as name, age, gender, date of birth, preferences or economic transactions. And also, unstructured data for example, clicks on "like", links, Tweets, voice messages, text messages, video, audio, online games, blogs and in general all information generated by users of social networks or of the Internet of Things (IoT).

The contribution to the massive accumulation of data is found in various industries in which companies maintain large amounts of transactional data, gathering information about their customers, suppliers, operations and others (Puri and Haritha, 2016). On the other hand in the public sector in many countries databases are administered that contain data on population census, medical records, taxes and others. The above in addition to financial transactions made online or by mobile devices, analysis of social networks (on Twitter are about 12 Terabytes of Tweets created daily and Facebook stores around 100 petabytes of photos and videos), geographical location using GPS coordinates in other words, all those activities that are carried out several times a day with smartphones, generate around 2.5 quintillion bytes daily in the world.

With the large volumes of information handled, it becomes an obligation to know what data the company must manage which requires us to classify them and integrate various components and projects that, together, will form the necessary system to analyze large amounts of data which is not an easy task. In any case, companies and governments are grappling with the problem of analyzing data for two important purposes: to be able to detect and respond to current events in a timely manner and to be able to use predictions of historical learning.

Data available in a big data environment

Types of data: As explained above, there is a large amount of data that is prone to analysis, so, it is necessary to classify them to better understand their sources and what they represent. In a big data environment, the large amounts of data that can be handled can come from different sources. Akter and Wamba (2016) classify them according to their origin while (Hashem *et al*., 2015) classify them according to 5 aspects in order to better understand the characteristics of them. The data available in a big data environment are classified, according to (Fig. 1) in:

Data sources: Data generated by activity in social networks, monitors and intelligent machines that generate information, sensors, transactions and information from the internet of things, generated by smartphones, cameras, tablets and others.

Content format: Structured data coming from databases such as numbers, words, dates, managed with programming language, semi-structured data, coming from systems that are not conventional data bases and unstructured data such as text messages, geolocation, videos and data on social networks.

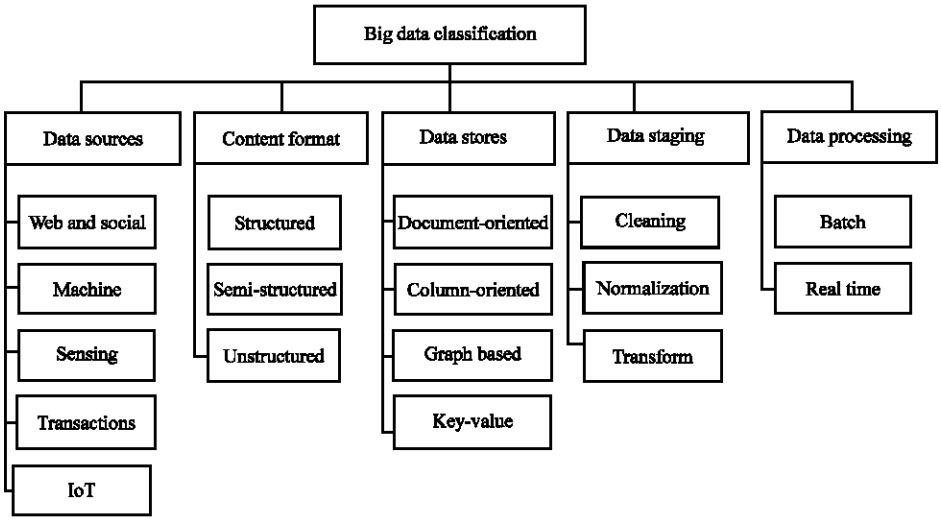

Fig. 1: Classification available data in a big data environment

Data stores: Data from data warehouses designed to store and retrieve collections of documents or form information, graphic databases and scalable pool databases, to perform large tasks with data sets.

Data staging: Data that goes through processes of cleaning, transformation and normalization to be analyzed.

Data processing: Data that are processed by groups of machines that comprise thousands of nodes. It is important to highlight that all these types of data, from the technical point of view are very heterogeneous which complicates their treatment, management and the technological infrastructure required for their correct exploitation and treatment.

Characteristics of big data: The characteristics with which a big data environment is defined are:

Data volume: Refers to the quantity of all the types of data generated from different sources. Some researchers collected a collection of data from intelligent mobile devices and made them available to the research community (Laurila *et al*., 2012). The results of the manipulation of these data produced among other things, a predictability of human behavior patterns or means to share data based on human mobility.

Variety of data: Refers to the types of data collected through sensors, smartphones or social networks. These include video, image, text, audio and data records in structured format and unstructured format. For example for, (O'Leary, 2013) internet users also generate a large number of structured data and unstructured data.

Speed of data generation: Speed refers to the speed at which the data is transferred with which for (Berman and Morgan, 2013) in that transfer changes are also presented in the contents when the recently transferred data is mixed with data previously archived.

Value: It is the value that is generated when discovering hidden value information in large amounts of data. This information, then can serve to improve business management. For example, the personalized Communication and marketing strategies which cover the basic needs of potential clients can allow, at the same time, to obtain better information, value and knowledge of said target with the objective of specifying sales that allow greater profitability in the user company of big data (Akter and Wamba, 2016).

The number of participants: Since, it is known that technology companies were created that offered big data in 2009 until 2016, there has been an increase in the number of companies dedicated to the sale of this type of technology services with the result that these companies, at least would give to generate a graph like the one shown in Fig. 2 ("Is Big Data Still a Thing? (The 2016 Big Data Landscape) (Turck, 2016). The trend in the opening of companies dedicated to offering big data technology has been moving from left to right, from the infrastructure layer (essentially the world of developers/engineers) to the analytical layer (the world of data scientists and

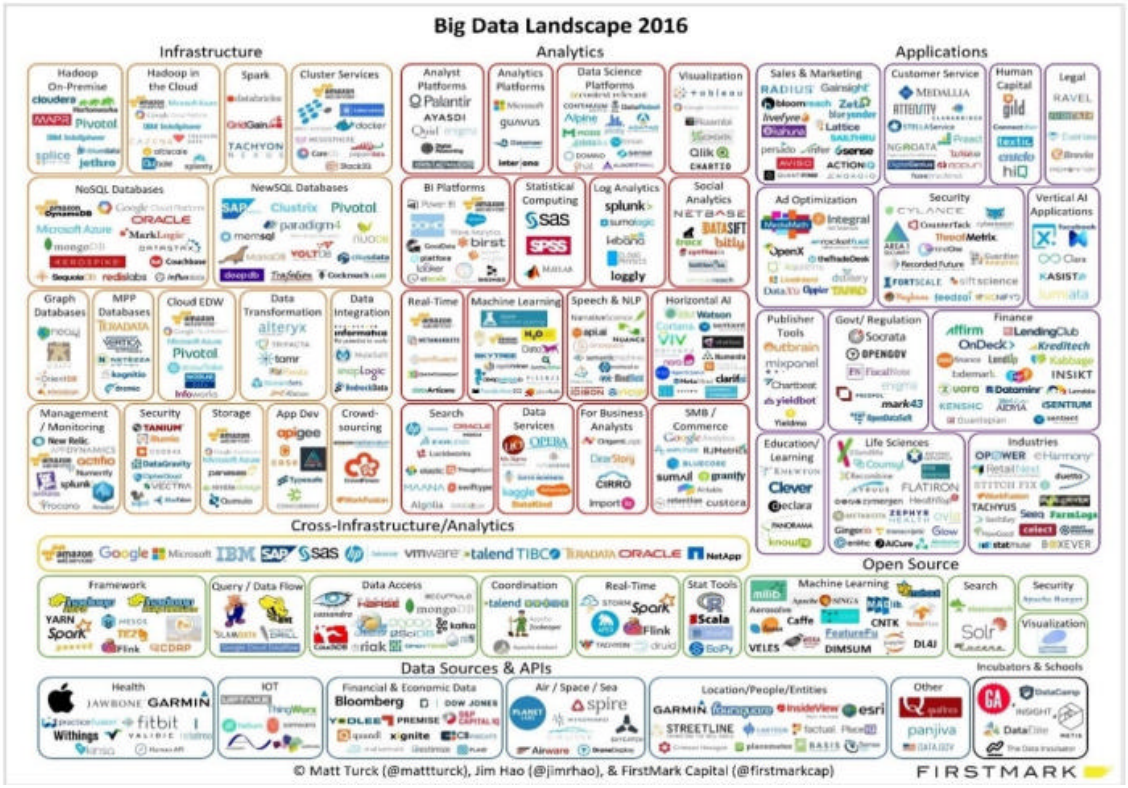

Fig. 2: Big data landscape ("Is Big Data Still a Thing? (The 2016 big data landscape) (Turck, 2016)

analysts), to the application layer (the world of business users and consumers) where "native big data applications" have been emerging rapidly, following more or less the pattern shown.

The alliances: Few organizations have been able to obtain value from their investments in businesses which according to several researchers could be due to an inadequate understanding of the necessary investments to obtain value of the analysis of the information that the data generate in the companies (Anand *et al*., 2016). It is determined that implementation in different businesses requires an initial investment of technological infrastructure capital to build a good platform, followed by investments in innovation capital to create value through competitive actions with statistical information enabled for analysis. Value is generated by the agility of the process of allocating resources for IT innovation and by the efforts made by managers to find and select the ideas that can together with the data analyzed generate value.

The literature on the Commercial Value of Information Technology (BVIT) has attempted to theorize the relationship between IT capabilities, IT assets and organizational performance (Anand *et al*., 2013a, b). An important argument in this line of research is that competitive actions mediate the relationship between IT investments and performance. Similarly, Piccoli and Ives (2005) argue that competitive movements, dependent on IT are key to the creation and appropriation of economic value and delivery of high levels of performance of the company.

To complement the above, one can cite the literature of dynamic capabilities which takes a process-oriented approach to highlight the strategic role of managers in the creation of value through "orchestrate complementary and co-specialized assets invent and implement new business models and make astute investment decisions in situations of uncertainty and ambiguity" (Helfat *et al*., 2007). Specifically, this literature identifies management's key roles in creating value from IT resources (Anand, *et al*., 2013a, b; Helfat *et al*., 2007; Sirmon *et al*., 2011).

As an example of the above, the case study in (Anand *et al*., 2016) about United Parcel Service (UPS) which highlights the role of investments in information technology to identify and exploit value creation opportunities can be taken. In this case, IT provided UPS with integrated data flows that allowed managers to estimate the costs and profitability of package delivery

routes which became the basis for subsequent investments in competitive actions to improve performance. To achieve the above, it is required that the participants (understood as a participant to all companies that are part of the big data project) in the development of big data projects provide the following elements:

- Infrastructure (hardware, fiber optic networks, servers)
- Operating systems
- Databases (relational and non-relational)
- Business management software
- Free software for data migration
- Security
- Web services provided by third parties
- Languages for charging qualitative or non-numeric data
- Analysis tools for texts that are generated around a brand or a name to generate consumable information
- Social networks as sources of data, feelings and trends

Each of the above components of the alliances, outlined in this list are important to respond to the needs of big data in the companies but they are not enough to provoke complete solutions that respond to all market needs. Therefore, it is foreseeable that we will meet different suppliers that will contribute new solutions to the management of big data.

RESULTS AND DISCUSSION

Creation of value in institutions of higher education: Different researchers have published research that explores the extent to which the development of big data techniques is part of the creation of value of different companies and institutions. Some concentrate on describing new business models derived from the application of big data, the impact in the form of profitability in the results of the organizations and the social transformation that can be derived from it by the substitution of the human resource by computers in tasks of marketing and sales (Loebbecke and Picot, 2015). To further support the above statement, we can refer to the work of (Akter and Wamba, 2016) in which 6 mechanisms are presented to increase the value of e-Commerce companies in this era of data economy:

Personalization: The researchers classify different works that have been done explaining the application of big data in the offer of personalized services to customers of different products or services, using the traceability that the payments of said customers leave with their credit cards.

Promotion of prices: In this area of e-Commerce, the researchers classify different works in which companies like amazon do big data processes in which they take into account information on the price of competitors, sale of products, actions of consumers and all kinds of regional and geographical preferences with which the prices of the products are changed, according to the results of the processing of the previous information in real time.

Customer service: In this other key area of e-Commerce, companies can use all communication channels they have with their customers to process the information that customers generate in terms of opinions, complaints or claims they have of the products. Or services of an organization. Processing and addressing all these requests can offer innovative sales and after sales service solutions.

Visibility of the supply chain: In this case, the researchers validate with different jobs the advantages generated by purchases of goods or services online, since, this allows, from the beginning of the purchase, to make traceability of the state of the same and even of the prediction of the day and time of delivery of said purchase as long as they are interconnected to the companies with their systems of storage and transport.

Detection of security and fraud: For e-Commerce companies it is important to be able to detect fraud in real time by means of the combination of the data of the transactions of the clients with their purchase history, their social networks and their geolocation, using your smartphone.

Predictive analysis: In this case, big data is used in e-Commerce companies to identify events even before they occur, based on the analysis of customer purchase patterns with the historical information stored in the bases of data.

Given that the objective of this stduy is the analysis of the creation of value of big data in higher education institutions as well as the definition of business models for them, a first step would be to analyze the most relevant works that can help demonstrate this value creation. To do this, it is necessary to read all the documents that speak of the benefits of the software used by universities to provide a better service to their students and their professors and the administrative staff that works in them. In Table 1, the relevant studies in scientific literature on the subject are presented.

Table 1: Relevant studies in value contribution

Study	Business area	Category	Case (s) study
Adamko (2017)	Efficiency in campus Reduction of operating expenses Software systems for service architecture	Smart University Smartness characteristic Smart campus Smart community Cloud archuitecture XMPP Software system internet of things Smart education Service-oriented adaptation Crowdsourcing	University of Debrecen in Hungary
Barcellos *et al.* (2017)	Websites as a communication channel with clients Satisfaction of expectations and needs of customers with the web	ICT Universities Universities Visual communication Web sites	Sao Paulo State University en Brazil
Paul and Ghose (2018)	Data administration and management Generation of information Data analysis in engineering faculties	Big data, cloud computing Emerging educational programs Human computer interaction	Raiganj University, Raiganj, West Bengal, India Sikkim Manipal University, Gangtok, Sikkim, India
Wang *et al.* (2018)	Information security Security in the databases of the libraries in universities Library as support to teaching and research	Big data, storage of information and access to knowledge Reliable technologies, advanced equipment and effective administrations Prevention of the use of false information sources	Harbin University of Commerce, Harbin, China
Tan *et al.* (2017)	Admissions of students in universities using big data Selection of postgraduate courses using big data	Analysis of real admission data by algorithm A priori, k-means and KNN	School of Administration, Capital Normal University, Beijing, China
Xi (2017)	Evaluation of the direction of research in a university using big data Decision making in education based on data mining Connection of relevant areas of education and finding key educational variables for decision making with big data Master's thesis title analysis	Educational data mining Word segmentation algorithm, association rule and RStudio tool	School of Science, University of Posts and Telecommunications of Xian, Xian, China

Table 2: Most relevant theories in value creation

Area of creation of value in big data	Relevant theories
Strategy, culture, leadership, organization	Resources and capabilities (Barney, 1991) Competitive strategy (Porter, 1980)
Marketing and sales	Market orientation (Kohli and Grover, 2008) e-Commerce relations Value of the client
Operations and production management	Transaction costs
Human resources management/talent	Resources and capabilities

In Table 1 it can be seen that there are many areas in universities where value contribution can occur. And they stand out as the most benefited, the areas of decision making to determine academic offer and of information management for effective and efficient decisions. Information is an intangible asset that makes the difference between companies with this type of capabilities and those that do not. For this, the research of (Windsor, 2017) is very relevant which generated a cross between the areas of creation of value in a company and the relevant theories of business management as shown in Table 2.

CONCLUSION

Information is the most important intangible asset of the organizations. That is why the application of big data in higher education institutions will allow the generation of value in decision making. Institutions must apply the most relevant big data theories (market orientation, marketing, customer value, costs, resources and capabilities) in the different academic-administrative areas, to achieve competitive advantage and also, improve the quality of services offered.

REFERENCES

Adamko, A., 2017. Building Smart University Using Innovative Technology and Architecture. In: Smart Education and Smart e-Learning, Uskov, V., J. Bakken, R. Howlett and L. Jain (Eds.). Springer, Berlin, Germany, ISBN:978-3-319-59453-8, pp: 161-188.

Agarwal, R. and V. Dhar, 2014. Editorial big data, data science and analytics: The opportunity and challenge for IS research. Inf. Syst. Res., 25: 443-666.

Akter, S. and S.F. Wamba, 2016. Big data analytics in e-Commerce: A systematic review and agenda for future research. Electron. Markets, 26: 173-194.

Anand, A., F.S.Wamba and R. Sharma, 2013a. The effects of firm IT capabilities on firm performance: The mediating effects of process improvement. Proceedings of the 24th Australasian Conference on Information Systems, December 4-6, 2013, RMIT University, Melbourne, Australia, pp: 1-10.

Anand, A., R. Sharma and R. Kohli, 2013b. Routines, reconfiguration and the contribution of business analytics to organisational performance. Proceedings of the 24thAustralasian Conference on Information Systems, December 4-6, 2013, RMIT University, Melbourne, Australia, pp: 1-10.

Anand, A., R. Sharma and T. Coltman, 2016. Realizing value from business analytics platforms: The effects of managerial search and agility of resource allocation processes. Proceedings of the 37th International Conference on Information Systems (ICIS), December 11-14, 2016, Association for Information Systems, Atlanta, Georgia, USA., pp: 1-12.

Barcellos, L.I., E.E.I. Barcellos and G. Botura, 2017. Analysis of communicational design and information on university websites. Proceedings of the International Conference on Applied Human Factors and Ergonomics, July 17-21, 2017, Springer, Berling, Germany, ISBN:978-3-319-60476-3, pp: 64-72.

Barnes, T.J., 2013. Big data, little history. Dialogues Hum. Geogr., 3: 297-302.

Barney, J.B., 1991. Firms resources and sustained competitive advantage. J. Manage., 17: 99-120.

Berman, J.J. and K. Morgan, 2013. Principles of Big Data: Preparing, Sharing and Analyzing Complex Information. Morgan Kaufmann, Burlington, Massachusetts, USA., ISBN:9780124045767, Pages: 261.

Gantz, J. and D. Reinsel, 2011. Extracting value from chaos. IDC. Iview, 1142: 1-12.

Hashem, I.A.T., I. Yaqoob, N.B. Anuar, S. Mokhtar and A. Gani *et al*., 2015. The rise of big data on cloud computing: Review and open research issues. Inf. Syst., 47: 98-115.

Helfat, C.E., S. Finkelstein, W. Mitchell, M.A. Peteraf and H. Singh *et al*., 2007. Dynamic Capabilities: Understanding Strategic Change in Organizations. Blackwell, London, England, UK.,.

Khan, M.S., M.A. Kausar and S.S. Nawaz, 2018. Bigdata analytics techniques to obtain valuable knowledge. Indian J. Sci. Technol., 11: 1-14.

Kohli, R. and V. Grover, 2008. Business value of IT: An essay on expanding research directions to keep up with the times. J. Assoc. Inf. Syst., 9: 23-39.

Laurila, J.K., D. Gatica-Perez, I. Aad, O. Bornet and T.M.T. Do *et al*., 2012. The mobile data challenge: Big data for mobile computing research. Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland. https://infoscience.epfl.ch/record/192489

Loebbecke, C. and A. Picot, 2015. Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. J. Strategic Inf. Syst., 24: 149-157.

O'Leary, D.E., 2013. Artificial intelligence and big data. IEEE. Intell. Syst., 2: 96-99.

Paul, P.K. and M.K. Ghose, 2018. A Novel Educational Proposal and Strategies toward Promoting Cloud Computing, Big Data and Human-Computer Interaction in Engineering Colleges and Universities. In: Advances in Smart Grid and Renewable Energy Proceedings of ETAEERE-2016, SenGupta, S., A. Zobaa, K. Sherpa and A. Bhoi (Eds.). Springer, Berlin, Germany, ISBN:978-981-10-4285-0, pp: 93-102.

Piccoli, G. and B. Ives, 2005. Review IT-dependent strategic initiatives and sustained competitive advantage: A review and synthesis of the literature. MIS. Q., 29: 747-776.

Porter, M.E., 1980. Competitive Strategy: Techniques for Analysing Industries and Competitors. The Free Press, New York.

Puri, G.D. and D. Haritha, 2016. Survey big data analytics, applications and privacy concerns. Indian J. Sci. Technol., 9: 1-8.

Sirmon, D.G., M.A. Hitt, R.D. Ireland and B.A. Gilbert, 2011. Resource orchestration to create competitive advantage breadth, depth and life cycle effects. J. Manage., 37: 1390-1412.

Tan, Z., J. Wang, Y. Peng and F. Ma, 2017. The admissions big data mining research based on real data from a normal university. Proceedings of the International Conference on Chinese Intelligent Systems, October 14-15, 2017, Springer, Berlin, Germany, ISBN: 978-981-10-6495-1, pp: 531-542.

Turck, M., 2016. Is big data still a thing?(The 2016 Big Data landscape). FirstMark Capital, New York, USA. http://mattturck.com/big-data-landscape/

Wang, X., L.T. Yang, H. Liu and M.J. Deen, 2018. A big data-as-a-service framework: State-of-the-art and perspectives. IEEE. Trans. Big Data, 4: 325-340.

Windsor, D., 2017. Value Creation Theory: Literature Review and Theory Assessment. In: Stakeholder Management, David, M.W. and J. Weber (Eds.). Emerald Group Publishing, Bingley, UK., pp: 75-100.

Xi, X.Q., 2017. A method to evaluate the research direction of university. Proceedings of the Euro-China Conference on Intelligent Data Analysis and Applications, October 9-11, 2017, Springer, Berlin, Germany, ISBN:978-3-319-68526-7, pp: 157-166.

Yan, H. and H. Hu, 2016. A study on association algorithm of smart campus mining platform based on big data. Proceedings of the 2016 International Conference on Intelligent Transportation, Big Data and Smart City (ICITBS), December 17-18, 2016, IEEE, Changsha, China, ISBN:978-1-5090-6061-0, pp: 172-175.