Big Data Analytics - Pertaining Technology 'Vocative' to Big Data Enhancing Organisational Capabilities

Deepa Priyanshu[1], Dr. Rubina Liyakat Khan[2]

[1]*Lecturer, Administrative and Financial sciences Department, , Dammam Community College, Imam Abdurrahman Bin Faisal University, P.O. Box 1982, Dammam, Saudi Arabia*

[2]*Lecturer, Computer Science Department Administrative Dammam Community College, Imam Abdurrahman Bin Faisal University, P.O. Box 1982, Dammam, Saudi Arabia*

[1]dpriyanshu@iau.edu.sa, [2]rlkhan@iau.edu.sa

Abstract

Data is a word which is as old as human history. For centuries organisations have been relying on the various types of data for their effectiveness and performance. Big data analytics (BDA) is a technological way to deal with complex situations that persist in a business. These areas have been drawing the interest of various academicians and industrial practitioner from the beginning of these areas. These technologies are used by the organisations in an integrated manner than they can create wonders in the effectiveness and performance of the organisation. This paper is divided in three sections. In the first section of this paper the researcher has tried to throw light on the various aspects of the Big data in terms of characteristics. The researcher has analysed the various aspects of characteristics of big data defined till date discussed the role of data characteristics step by step as they have evolved on the organisational effectiveness and performance. Second section justifies the use of new characteristics in organisations. Third section of this paper proposes a new model using the new characteristics for future use.

Keywords: *Data, Big Data, Big Data Analytics, Characteristics*

Introduction

In the current business scenario, organizations have to take a no of decisions in their daily routine on the basis of facts which must be data driven using BI & BDA or just based on their experience and knowledge (Jaklic, Grubljesic, & Ales, 2018).

The fast advancement of information and communication technology has made it important to give due importance to the study and use of big data and data analytics (BDA). (Jin & Kim, 2018).Various researches have shown that BI&BDA used provides value to these organisations by enhancing their organisations capabilities (Jaklic, Grubljesic, & Ales, 2018). None the less several studies are not correlated to BDA, since corporations do not recognize and employ the concepts in the incorporated way. (Jin & Kim, 2018).

Big Data Analysis (BDA) is the process necessary to understand the conglomerate of data to extract and generate useful information and knowledge which, through interpretation and categorization, lead to more effective management. The main objective of the collection process and big data analysis is developing viable ideas and new insights to establish competitive advantages. Therefore, BDA is becoming an important differentiator between high and low performance, as it allows companies to have a long-term vision, reduce customer acquisition costs. Therefore, managers can use big data to learn more about their business and transform themselves the knowledge generated in efficient decisions, improving performance and the entire decision-making process. However, the management of the knowledge generated by the BDA, as well as its integration and combination with business knowledge, require a structured and integrated approach.(Ferraris, Mazzoleni, Devalle, & Couturier, 2019) Data wrapper is an innovative data visualization software developed by Data wrapper GmbH for journalists, developers and designers who work in frenzied editorial offices. It is an open source solution that allows users to create and view responsive, interactive and embeddable charts and maps on the organization's website. The platform simplifies the conversion of CSV, PDF and Web fonts data into simple, beautiful and accurate graphics and requires no programming or design skills to use.

Data wrapper is designed to offer unprecedented flexibility and you can create custom chart styles for specific newsrooms. Some of its key features include export such as PNG or PDF, maps, live graphics

updates, custom design, fully customized graphic styles and discrete colors.(Andre, Louie , 2020)

"Although big data, big data analytics (BDA) and business intelligence have attracted growing attention of both academics and practitioners, a lack of clarity persists about how BDA has been applied in business and management domains."(Shenga, Amankwah-Amoahb, & Wanga, 2018) (Phillips, 2017)

Efficient decision making on the basis of business intelligence (BI) is important for the development of business. (Jin & Kim, 2018). In this study data wrapper is used as a technical tool in the proposed model for the implication of the new variable named as Vocative in enhancing organisational capabilities.

'Big data, Data Science, Big data Analytics (BDA) and business intelligence (BI) – an overview:-

 As per Gartner, "Big data is high-volume, and high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation" (Verma, 2018). Broad-spectrum definitions which were frequently used to define the actual meaning of big data were defined in the Forbes list (Arthur, 2013). Big data in its basic concept is an art of handling a large amount of data processing it and generate insights from it. The concept of big data is generally used in telecommunication, marketing, retail, financial services (Fraud analysis, customer analysis, operational analysis, compliance analysis etc) (Verma, 2018) e-commerce and security services. Major tools used for dealing with the big data are Hadoop, spark and flink. Big data collected by the different organizations may be structured, semi structured and unstructured. It is mined for getting information which is used in machine learning projects, predictive modelling and other analytics application. (Botelho& Bigelow, 2019).

Data Science includes the handling of enormous information (both organized and unstructured) including the planning, investigation, cleansing of the information. It likewise includes programming, mathematical sciences, statistical science, insights for critical thinking, ability to see things in an unexpected way, instinctively catching information and so forth (Verma, 2018). Major techniques used for by data scientists are SAS, SQL and python.

Data analysis is used as a forecasting tool in increasing organizational effectiveness by developing effective methods of data mining. Data analytics involves the technique of drawing inference on the basis of raw data by implementing a mechanical or algorithmic process.

This technology is nowadays very popular among organizations for increasing their performance.

Business intelligence (BI) uses programming and administrations to change information into noteworthy bits of knowledge which may suggest ways to enhance organization's capabilities. BI tools way in and analyse data sets and presents investigative conclusions in reports, outlines, consoles, diagrams, charts and drawings to furnish clients with neck to neck view of the business condition. The term business intelligence generally alludes to a scope of devices that give sharp, simple approach to bits of knowledge about an organization's present conditions, in light of information present. (Pratt & Josh, 2019) There are different types of tools used for Business Analytics such as Dashboards, Visualizations, Reporting, Data mining, ETL (extract transfer load), and OLAP (online analytical processing)   The dashboards and visualization are most popular tools. (ProStrategy, 2019)

Table 1: Use of Big data in various industries

S No	Industry	Use
1.	Education	Reframing course material, profiling of students, Customized and dynamic learning programs Evaluating Systems, grading system, career forecasting (Baumann & Riedel)
2.	Insurance sector	Collecting information, Gaining customer insight, Fraud detection, Threat mapping (Baumann & Riedel)
3.	Government sector	Welfare schemes, Cyber security(Baumann & Riedel)
4.	Banking Sector	The frauds of credit cards and debit cards, Venture credit risk treatment, Business transparency, Customer indicators adjustment, Money legalizing, Risk Alleviation (Baumann & Riedel)
5.	Industry	Logistics, optimization of profits, revenue, working capital, investment (Bumblauskas, Nold, Bumblauskas, & Igou, 2017)
6.	Marketing	Increasing sales, consumers profiling, improved market segmentation (Dhingra & Chaudhry, 2018)
7.	Agriculture	Supply chain (Krishnamoorthi & Mathew, 2018)

Literature Review of Characteristics of Big data

Big data and BDA are commonly used terms which are generally used to describe huge data sets, which require storing, managing, analysis, visualize technology, statistical analysis the advanced and unique data (Chen, Chiang, & Storey, 2012). Previously only Three V's of Big data i.e. volume, variety and velocity were defined (McAfee, Brynjolfsson, Davenport, Patil, & Barton, 2010). After words Four V's of big data and Five V's (5V's) (FossoWamba, Akter, A., & Gnanzou, 2015) were described. Recently (Firican, 2017) redefined the5V's as 10V's"Volume, Velocity, Variety, Variability, Veracity, Validity, Vulnerability, Volatility, Visualization and Value."
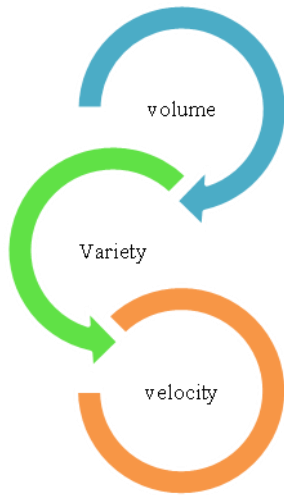
SECTION I

Critical Evaluation of effect of characteristics of Big Data on organizational effectiveness

A. 3V's Model

3Vs concept was introduced by Gartner and Doug Laney in 2001 in their Meta Group research paper, *"3D data management: Controlling data volume, variety and velocity"*. (Wigmore, 2013). Each V has its own significance in Big data.

Figure 1: 3V's Model

Volume: It provides the amount and form of the data in terabytes (Baumann & Riedel)or data sets minimum in petabyte; (Bumblauskas, Nold, Bumblauskas, & Igou, 2017)

Velocity: Velocity deals with the speed at which data flows in from the sources like business activities, networks, social media sites, sensors, mobile devices etc. (Bumblauskas Nold, Bumblauskas, & Igou, 2017).

The flow of data is vast and unbroken (Baumann & Riedel)

Variety: Variety refers to the heterogeneous sources and the nature of data both from structured and unstructured. Previous form of data, like spreadsheets has been replaced by emails, photos, videos, monitoring devices, PDFs, audio etc. which is to be considered in the analysis applications (Baumann & Riedel).

All three provide answers to different questions as shown below;

Figure 2: Showing the type of questions asked by Volume, velocity and variety
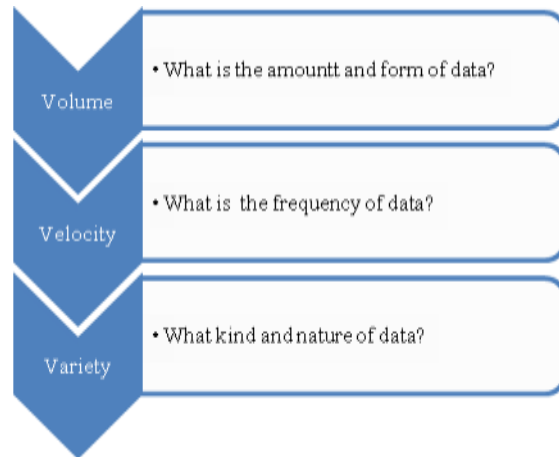
Since big data's biggest use is in increasing the organizational effectiveness therefore for improving performance the biggest requirement is of the outcome. As the questions which are answered by the big data characteristics 3V's viz. volume, velocity and variety could only answer the questions about the amount, form, frequency, kind and nature of data but they are unable to answer about the outcome. This led to the requirement for adding on the characteristics of big data.

According to Grouses Michael Whitehead, "Big data is a lot more interesting when you bring in 'V' for value.""Does new data enable an organization to get more value, and are we doing enough to get to that value quickly?" (Swoyer, 2012)

B. 4V's Model

Now next question arises how trust worthy and clean the data is. What is the accuracy and applicability of data? **Veracity helps selecting the important data to** generate a deeper understanding of data and how to contextualize it in order to take action. Data veracity of big data defines not just the quality of the data but how

trustworthy the data source, type, and processing of it is (Veracity: the most importan V of Big Data, 2019).

Figure 3: '4V's of Big data

C. 5V's Model

Even After defining the data's trust worthiness the question which still stands what will be the benefit or Out come from the data? Value defines the economic benefits from the data (FossoWamba, Akter, A., & Gnanzou, 2015). "Given all the complexity involved at every large enterprise today, perhaps the third 'V' should be the virtualization required to simplify and accelerate these efforts." (Swoyer, 2012)

Figure 4: 5 V's of Big Data
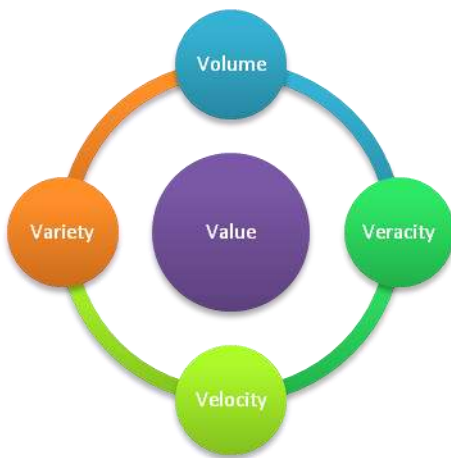
The main objective of big data evaluation is to get an economically benefit value for the organisation carrying out the analysis. (Williamson)

"Shouldn't the first 'V' be the business value derived from analyzing big data? Might we be better served if the second 'V' were the vision required to successfully synthesize business needs, analytic capabilities and source data?" (Swoyer, 2012)

D. 10V's model

Even after getting the 'Value' there still remain some question like format and layout of data analysis result (Huang, McIntosh, Patrick, & Hung, 2017) ,security of data, accuracy and correctness of data and variable dimensions of data (Huang, McIntosh, Patrick, & Hung, 2017). These answers were provided by Firican after defining 10V model. (Firican, 2017)

The 7V's viz Volume, Velocity, Variety, Variability, Veracity, Visualization and Value are able to define the data in a systematic manner (Impact, 2016) but to be more précised 10V's were suggested as "Volume, Velocity, Variety, Variability, Veracity, Validity, Vulnerability Volatility, Visualization and Value" (Firican, 2017)

All of these 10V's answer different questions required in any data analysis undergone by any organisation. These questions are summed up in the table given below;

Table 2: 10 V Characteristics of Big Data

S. NO	Characteristics of Big Data	Questions cleared	Explanation
1.	Volume	Amount and form of data	Huge amount of data per second is created in business.
2.	Velocity	Frequency of data	Pace of the data establishment has added significance than volume, due to increasing competition in the economic world which makes decision making process fast (Ferraris, Mazzoleni, Devalle, & Couturier, 2019)
3.	Variety	Kind and nature of data	Different digital platforms help in generating data through surveys conducted through websites to know about the requirements of the consumers. (Ferraris, Mazzoleni, Devalle, & Couturier, 2019)

4.	Variability	Data Diversity	Variability i.e. variation if the meaning is constantly changing it can have a huge impact on your data homogenization. (Impact, 2016)Follow-on numerous dissimilar sources and data types. (Huang, McIntosh, Patrick, & Hung, 2017)
5.	Veracity	Quality of data	Accurate analysis of collected data is worthless if it's not from accurate and trustworthy source. (Ferraris, Mazzoleni, Devalle, & Couturier, 2019)
6.	Validity	Data Genuineness	It refers to the precision and correctness of the data. (Huang, McIntosh, Patrick, & Hung, 2017)
7.	Vulnerability	Data Security	It defines to the safety in the data source.(Huang, McIntosh, Patrick, & Hung, 2017)
8.	Volatility	Duration of Usefulness	It defines the duration to which the data is useful to the organisation.
9.	Visualization	Data Act/ Data Route	It is the process of defining data path or abstract (Ferraris, Mazzoleni, Devalle, & Couturier, 2019)
10.	Value	Outcome of the data	Economic benefits which can be reaped from the big data is used to enhance the organisational capabilities.(Ferraris, Mazzoleni, Devalle, & Couturier, 2019)

Figure 5: Showing 10 V Model
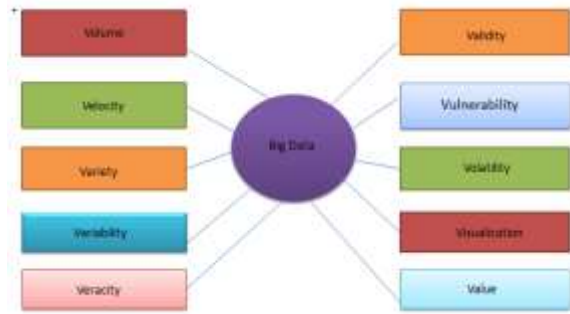
Although some researchers have defined 14V's and 1C, in big data characteristics viz. Virality (spreading speed), Viscosity (Lag of Event), Venue (Different Platform), Vagueness (Data Terminology), Complexity (Correlation of Data). (Ferraris, Mazzoleni, Devalle, & Couturier, 2019)

10Vs of Big Data characteristics has tried to full-fill the requirements of organisations dealing with data to increase their effectiveness and performance.

SECTION II

In this section of the paper, the researcher has tried to analyse the characteristics and use of Big data and tried to establish their relationship with organisational effectiveness. Researcher has studied different question arise from big data characteristics and how they answer the problems arises from business situations. Author has defined a new characteristic named Vocative.

Vocative

In any organisation for making decisions, it is required to have data in a large amount of different kind and nature, with good frequency according to our needs from the homogeneous source. Now the quality of data should be good and from a trustworthy & secure source. Data is useful in certain duration. After that duration, it becomes obsolete. In the same time span, it should be used to yield an economically beneficial outcome. Only in this condition, it is worthwhile for a company to spend a huge amount of money on it. Here for an organisation, it is also important to collect data homogenous source. In the characteristics defined till date there is no characteristic which is used for homogeneous data. Main reason behind it is Big data is in such a big amount that there is no way to collect homogeneous data. The only way left to solve the problem of homogeneous data is using technical tools to make it homogeneous. This gives rise to our next question that is how technical the data is?A

heterogeneous data should be collected in a technical way to enhance the organisational capabilities. Researcher has named this component as **Vocative** which mean technicality of data.

SECTION III

A model was designed defining the variables between data, information, Knowledge, suggesting that actionable Knowledge could be helpful in enhancing the process of decision making. According to data is converted to get information which in turn is converted into actionable knowledge and provides feedback about data (Davenport & Prusak, 1998).

Figure 6: Davenport Model
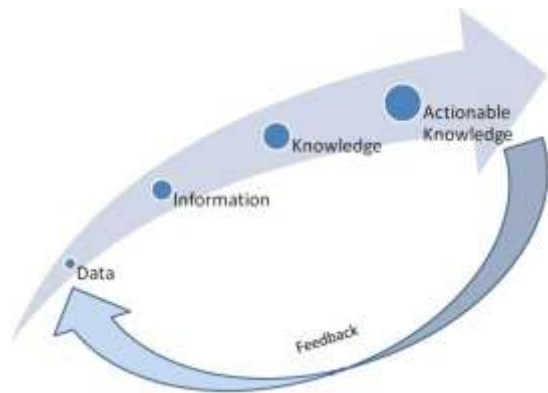
Researcher has proposed the model given below.

Figure 7: Organisational Capabilities Enhancer
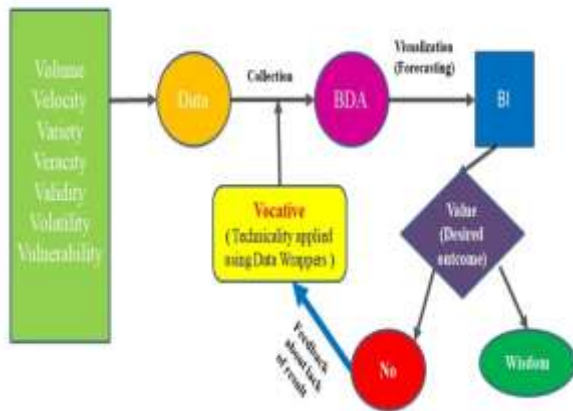
In this model the 7 V's viz. Volume, Variety, Variability, Veracity, Vulnerability and volatility are used to collect and interpret data to apply BDA tools. Organisations used 'Visualization' characteristic to forecast the organisation's requirement from the data. Now the role of BI starts as Business intelligence (BI)

that uses programming and administrations to change information into significant 'bits of knowledge' that yields the economically required outcome as 'value' enhancing the organisations capabilities to take fruitful decisions. Getting the value according to the requirement of the organisation is 'Wisdom'. If desired results are not achieved, than feedback about the lack of result will help at the collection stage to apply the 'Vocative' characteristics to improve the homogeneity of data. Here it is worthwhile to mention that at this stage vocative component will be technically applied using data wrappers.

Conclusion

In this paper different concepts of the Big data, use of Big Data, Big data characteristics and the concept of business intelligence have been discussed in detail. A new characteristic has been introduced. Vocative the new proposed characteristic will be helpful by segregating data according to the organisations requirement. This proposed model named 'Organisational Capabilities Enhancer' will be helpful in enhancing Organisational Capabilities.

Future scope of the study

In the first section of the paper we have talked about the scope of Big data analytics in our study in various fields. Today marketing is the most important area for any business. Organisations are spending ample of money on research and development. If we consider the case of marketing there is footfall of various types of customers in a mall. Marketing firms are interested in analysing customer's behaviour for different products so that they can enhance their products accordingly. Let's us consider the case of research under going for garments. They use to collect data from the consumer behaviour which is mixed for different variable like size of garments, taste & preferences, colour choice, size, new innovative products, discount & pricing seasons & festivity, occasion, fashion trend, celebrity influence etc. The concern firm can segregate this Big data with the help of data wrappers using the new variable vocative. This research will pave a path for all organisations which deal with big data and want to segregate it on the basis of their requirement. Likewise it can be applied for other management areas like insurance and banking etc. Talking about the future scope this model could be applied effectively to industry.

REFERENCES

[1] Andre, Louie . (2020, 2 29). *Datawrapper Review*. Retrieved 5 15, 2020, from Finances on Line Reviews for Business: https://reviews.financesonline.com/p/datawrapper/

[2] Arthur, L. (2013, August 15). *What Is Big Data*. Retrieved March 21, 2020, from Forbes:https://www.forbes.com/sites/lisaarthur/2013/08/15/what-is-big-data/#658bc26d5c85

[3] Baumann, P., & Riedel, M. (n.d.). *Big Data - Definition, Importance, Examples & Tools*. Retrieved March 21, 2020, from RDA: https://www.rd-alliance.org/group/big-data-ig-data-development-ig/wiki/big-data-definition-importance-examples-tools

[4] Botelho, B., & Bigelow, J. (2019, october). *Guide to big data analytics tools, trends and best practices*. Retrieved March 18, 2020, from searchdatamanagement: https://searchdatamanagement.techtarget.com/definition/big-data

[5] Buhl, H., Röglinger, M., Moser, D., & Heidemann, J. (2013). Big data. *Business & Information Systems Engineering , 5*, 65-69.

[6] Bumblauskas, D., Nold, H., Bumblauskas, P., & Igou, A. (2017). Big data analytics: transforming data to action. *Business Process Management Journal , 23* (3), 703-720 .

[7] Chen, H., Chiang, R., & Storey, V. (2012). Businessintelligenceandanalytics:frombigdatatobig impact. *MIS Quarterly , 36* (4), 1165-1188.

[8] Davenport, T., & Prusak, L. (1998). *Working Knowledge: How Organizations Manage What they Know* (Vol. 1). Boston: Harvard Business School Press.

[9] Dhingra, S., & Chaudhry, K. (2018). A Study of the Impact of Data Warehousing and Data Mining Implementation on Marketing Effort. *INTERNATIONAL JOURNAL OF ADVANCED STUDIES IN COMPUTER SCIENCE AND ENGINEERING , 7* (1), 13-20.

[10] Dong, X., & Srivastava, D. (2013). Big data integration. *IEEE 29th International Conference on Data Engineering* (pp. 1245-1248). Brisbane: IEEE.Ferraris, A., Mazzoleni, A., Devalle, A., & Couturier, J. (2019). Big data analytics capabilities and knowledge management: impact on firm performance. *Emerald Insight , 57* (8), 1923-1936.

[11] Firican, G. (2017, Ferburary 8). *The 10 Vs of Big Data*. Retrieved March 23, 2020, from tdwi -upside where data means business: https://tdwi.org/articles/2017/02/08/10-vs-of-big-data.aspx

[12] FossoWamba, S., Akter, S., A., C. G., & Gnanzou, D. (2015). How 'bigdata' canmake big impact: findings from a systematic review and a longitudinal case study. *International Journal of Production Economics , 165*, 234-246.

[13] Hitzler, P., & Janowicz, K. (2013). Linked data,bigdata,andthe4thparadigm. *SemanticWeb , 4* (3), 233-235.

[14] Huang, S.-C., McIntosh, S., Patrick, S. S., & Hung, C. K. (2017, October 17). Big Data Analytics and Business Intelligence in Industry. *Springer Science+Business Media, ,* 1229–1232.

[15] Impact. (2016, April 7). *The 7 V's of Big Data*. Retrieved March 26, 2020, from Impact: https://impact.com/marketing-intelligence/7-vs-big-data/

[16] Jaklic, J., Grubljesic, T., & Ales, P. (2018). The role of compatibility in predicting business intelligence and analytics use intentions. *International Journal of Information Management ,* 305-318.

[17] Jin, D.-H., & Kim, H.-J. (2018, October 19). Integrated Understanding of Big Data, Big Data Analysis, and Business Intelligence: A Case Study of Logistics. *MDPI* .

[18] Krishnamoorthi, S., & Mathew, S. K. (2018, Janurary 31). Business analytics and business value: A comparative case study. *Information & Management ,* 643–666.

[19] McAfee, A., Brynjolfsson, E., Davenport, T., Patil, D., & Barton, D. (2010). Big data: the management revolution. *Harvard Business Review , 90* (10), 61-67.

[20] Phillips, F. (2017). A perspective on 'Big Data'. *Science and Public Policy , 44*, 730-737.

[21] Pratt, M. K., & Josh, F. (2019, October 16). *What is business intelligence? Transforming data into business insights*. Retrieved March 23, 2020, from CIO India: https://www.cio.com/article/2439504/business-intelligence-definition-and-solutions.html

[22] ProStrategy. (2019). *Business Intelligence*. Retrieved April 10, 2020, from Pro Strategy Business solutions: https://www.prostrategy.ie/business-intelligence/

[23] Shenga, J., Amankwah-Amoahb, J., & Wanga, X. (2018, June 12). Technology in the 21st century: New challenges and opportunities. *Technological Forecasting & Social Change ,* 321-335.

[24] Swoyer, S. (2012, July 24). *Big Data -- Why the 3Vs Just Don't Make Sense*. Retrieved March 26, 2020, from TDWI: https://tdwi.org/articles/2012/07/24/big-data-4th-v.aspx

[25] *Veracity: the most importan V of Big Data*. (2019, August 19). Retrieved March 23, 2020, from Gut Check: https://www.gutcheckit.com/blog/veracity-big-data-v/

[26] *Veracity: The Most Important "V" Of Big Data*. (2019, August 19). Retrieved March 23, 2020, from Gut Check: https://www.gutcheckit.com/blog/veracity-big-data-v/

[27] Verma, A. (2018, October 9). *Data Science vs Big Data vs Data Analytics*. Retrieved March 20, 2020, from Whizlabs: https://www.whizlabs.com/blog/data-science-vs-big-data-vs-data-analytics/

[28] Wigmore, I. (2013, feburary). *3Vs (volume, variety and velocity)*. Retrieved March 26, 2020, from Techtarget: https://whatis.techtarget.com/definition/3Vs

[29] Williamson, J. (n.d.). *https://www.dummies.com/careers/find-a-job/the-4-vs-of-big-data/*. Retrieved March 26, 2020, from Dummies-A willey brand: https://www.dummies.com/careers/find-a-job/the-4-vs-of-big-data/