Is Big Data for Everyone? The Challenges of Big Data Adoption in SMEs

S. Shah[1], C. Bardon Soriano[1], A.D. Coutroubis[1]
[1]Department of Applied Engineering and Management, Faculty of Engineering and Science, University of Greenwich, Chatham Maritime, Chatham, Kent, United Kingdom
(s.shah@gre.ac.uk; c.b.soriano@gre.ac.uk; a.d.coutroubis@gre.ac.uk)

Abstract – **The aim of this paper is to present an investigative study on the concept of Big Data and its challenges towards implementation in manufacturing SMEs. Big Data aims to facilitate the collaborative approach in SMEs through the creation of real time data visualization to address key challenges to many of the market variations for every sector SMEs. Although, earlier research studies have highlighted the importance of Big Data from technological perspectives, this study focuses towards SMEs due to its feasibility and flexibility within the market space. This research aims to investigate the use of case study approach for the re-use, adoption and understanding of strategic future direction from the findings. The findings and early analysis from this paper could be referred by researchers when addressing the use of big data analytics within manufacturing SMEs. Finally, the paper provides a key strategic point towards the exploration of Big Data within SMEs.**

Keywords - **Big data, big data analytics, manufacturing, small and medium enterprises (SMEs)**

I. INTRODUCTION

In today's modern world, companies need to grow up in a more and more competitive and highly changing environment. It is due to this dynamic world, firms are not only forced to cope with their operational concerns but also to keep an eye in what is happening around the world, what their competitors are doing and how they are doing it, what the costumers feel, what they are expecting to receive or how to improve the suppliers' relationship, making it stronger and more reliable. All these relationships are translated into data, from several different resources and forms, so enterprises must be ready to take in and integrate them so as to find out useful knowledge. Knowledge is fundamental source of competitive advantage and value creation to development of dynamic capabilities and this is reflected in organizational success [1, 2]. Firm's knowledge provides them the capacity to distinguish themselves from competitors getting a competitive edge. Earlier studies identified data as a collection of raw numbers and facts to information which is processed data, and finally knowledge as the validated information [3].

In the last few years, according to a report from International Data Corporation (IDC) the volume of global data has reached 2.8ZB and are estimated to reached 40ZB by 2020, in other words, more than 5,200GB per person [4]. Data volume is predicted to continue growing exponentially because of Internet of Things (IoT) technologies, cloud computing and wearable devices [5]. That gives companies the opportunity to gain access to much more information in a single day that were stored in the entire internet 20 years ago [6]. Analyzing information is not something new. For a long time, businesses have stored, processed and analyzed data sets looking for some business value or monitor transactions that drives business optimizations [7]. It is the form those sets of data are found, how they are analyzed, and the quantity of data, that have changed. While traditional data are structured [8], new data produced consist on massive volume of heterogeneous datasets that traditional technologies are not even capable to process them adequately [5, 9, 10].

Big data is a term using to describe these massive data sets which have more variety and complex structure including semi-structured and unstructured data such as posts on social media as text, photo, audio or video, credit card purchases, science data surveys or customer feedbacks but also include traditional structure data. So an innovative step forward from the traditional IT and both software and hardware tools is required in order to be able to manage not only the variety but also the volume of data produced and the velocity they change with [10, 11]. Almost all data which are produced is under leveraged and only a fraction of them are being explored for analytic value [4]. The overwhelming potential of big data necessitates firms move on not only from traditional analytical and storing technologies but also a change in management and decision making mechanics are required.

Big data analytics (BDA) has been given significant attention by businesses in recent years due to the promise of higher productivity and profitability [12]. Big data can create real-time solutions to challenges in every industrial sector [13]. There are several examples of companies that have been able to become big data drivers successfully. For instance, Amazon uses information such as customer's names, addresses, payments and search histories to improve customer's relations or American Express is using big data to analyze and predict customer's behavior and make an accurate forecast [14]. Although Big Data offers many advantages for business, such as, improved decision making, presents new insights, and make processes more efficient, it has not yet convinced most European organizations of producing worthwhile returns from big data investments [9].

II. LITERATURE REVIEW

The research starts with an approach of big data dimension, features and challenges. Next, the paper illustrates the application of Big Data management within the enterprises, especially SMEs, and their challenges. Finally, the paper shows the current gaps between literature and the state of existing SMEs. The research aims to investigate the use and adoption of Big Data by using a case of study and use the finding from this paper to define a strategy for the implementation of Big Data within SMEs.

A. Overview of Big Data Principles

Firms are be forced to leverage their technical and management performance by the continuous increasing amount of data "whose size and complexity is beyond the ability of conventional tools of manage, store and analyze data" [15]. Traditional dataset typically included masses of structured data. After the explosive increase of global data, these datasets have started to be produced in an unstructured form due to the burgeoning used of wearable devices, internet of things technologies, online transactions, and social media websites. These data; which are generated as audios, images, click streams or emails; suppose a challenge for traditional database software tools to capture, store, manage and analyze them.

Big data is an abstract concept. There is not a unique definition for it and different explanations about what it is are given by research scholars, scientific and technological enterprises, data analysts, and technical practitioners [10]. According an IBM report [16], in general terms, we could understand big data as a new approach of how to make decisions using huge amount of data, from structure to unstructured data, which cannot be analyzed using traditional software. Another similar definition has been given by Apache Hadoop which defined big data as "datasets which could not be captured, managed, and processed by general computers within an acceptable scope" or McKinsey & Company that defines "Big Data as the next frontier for innovation, competition, and productivity". It can be extracted more accurate information from a large amount of data than using only a small bunch of data [17].

Gartner analyst Doug Laney defined big data with a 3Vs model in a 2001 META Group research publication [18]. Literature has used the 3V model as a common frame for Big Data. This model identifies Volume, Velocity and Variety as the main features to define Big Data. ***Volume*** refers to the generation or collection of the amount of data by an organization or an individual. Manufacturing companies and retailers gathering massive amount of data from customers, suppliers and operations in order to improve performance and customize actions for suitable service and products. ***Velocity*** refers to both the speed with data arrives, is processed and delivered. The velocity of data generation rises exponentially over time. Velocity is more subjective than volume for many applications [6, 8]. ***Variety*** refers to the different types of data. Data can be produced and stored in multiple formats. As more and more firm activity is digitized, new resources of information are coming up and different approaches and techniques are required for each type of data.

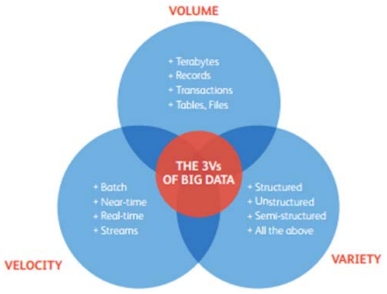

Fig.1. The 3Vs of big data. volume, velocity and variety.

These are considered the main dimensions of big data. In addition to the 3Vs, many authors and companies have been adding features over the time. IBM added *veracity* as a fourth dimension [8], which refers to trustworthiness of the data. As the number of resources grows, unreliability and inaccuracy inherent in some sources of data increase as well. ***Variability*** and ***complexity*** were introduced by SAS as two additional dimensions. While variability refers to the variation in data flow, complexity refers to the fact big data are generated through innumerable sources [19]. From earlier studies to the more recent research, the definition of Big Data has advanced. Research sorts out up to twelve dimensions in three tiers with four dimension each. On the first tier, there are the three main big data characteristics of Volume, Velocity and Variety, and Complexity.

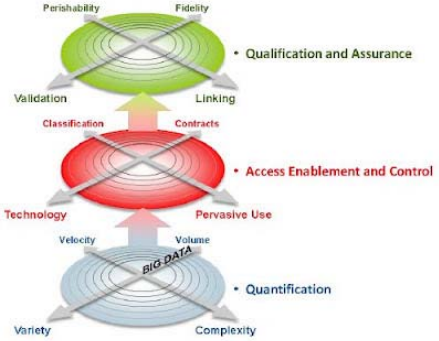

Fig. 2. Extreme information concept [20]

The second tier refers to understand and analyze the information. It is made up of Technology, Pervasive use, Classification and Contrast and the third tier consist of qualification and assurance aspects. Validation, linking, Fidelity and Perishability are the dimensions in this tier. This new classification come up as Big Data is only the beginning of Extreme Information Management (EIM). How to make the best use of this spread information is a hurdle for organizations. This is why some experts break down big data into three subcategories [21]: smart data, identity data and people data. ***Smart data*** is the most

popular of the three terms because of its simple understanding. Smart data is essentially the sub-set of Big Data that can immediately be put into use. ***Identity data*** refers to human behavior it is say that every step an individual takes and everything he makes can be predicted. ***People data*** refers to data sets from social media that helps companies understand their customers as individuals and develop programs that address and anticipate their needs. In recent years many different terms related to information are rising up and we must now mix them up with Big Data concept. ***Business intelligence*** [22] combine large amounts of corporative data with analytical tools to turn it to usable information for planners and decision makers whom are able to generate, aggregate, analyze and visualize data to make better management decisions. ***Data mining*** is the process of going through large set of data to identify relevant patterns which allow enterprises to solve problems, improve themselves or come up with competitive advantages.

TABLE I
BUSINESS INTELLIGENCE VS. DATA MINING VS. BIG DATA

Characteristics	Business Intelligence	Data Mining	Big Data
Memory access	Data is stored in a centralized warehouse.	Data is stored in a centralized warehouse	Data is stored on high distributed data sources.
Data Variety	Structured data in spread sheets, relational and dimensional data base.	Structured data in spread sheets, relational and dimensional data base	Structured, semi-structured, unstructured data
Data Volume	Large datasets, but limited to the processing of relational/dimensional databases	Small datasets, work on data sample with high data processing costs	Large datasets based on distributed and highly scalable processing structures
Data Velocity	Bunch of data. Present or past datasets	Bunch of data. Present or past datasets	Often real-time. Requires immediate response
Analysis style	Reflect past business performance in a small or large scale	Prediction and discovery relevant business factors from a small dataset.	Prediction and discovery relevant business factors from a large dataset.
Management	Requires management professionals to interpret the information for decision making.	Requires management professionals working in collaboration with data scientist	Requires management professionals working in collaboration with data scientist

B. Big Data Management & Challenges

The attention of most every industry have been caught by the emerging technological development of big data which is recognized as one of the most important areas of future information technology [8]. Big data promises organizations to create business advantages such as better decision making, yield new insights, cost savings, and optimize business processes getting higher product and service quality [8, 9, 15].

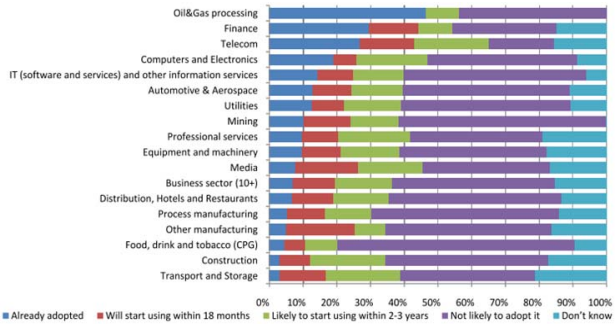

Fig. 3. European companies 'current and planned adoption of big data [23]

Company's investments on big data are growing rapidly due to this promise, but they still have ignored the potential downside [15]. According to a recent survey by Gartner, Inc. in 2016, 48% of companies have invested in big data, a 3% more than in 2015. However, the amount of firms planning to invest within the next two years has fallen from 31% to 25% in 2016 [24]. According to the literature, the challenges firms have to cope with could be sorted out in three main groups as technical, managerial, and economic challenges.

Technical challenges - The challenge of becoming a big data enabled organization can be enormous due to the nature of big data. Data volume, acquisition speed, or data variety limits the capacity of using traditional relational methods to conduct effective analysis of the data.

Managerial challenges - Big data projects usually fail because organization are on technologies and their capabilities instead of the business opportunity [9]. Companies, which are taking innovative approaches and giving a chance to exploration, are creating new values and reaching an advance performance [25]. According to [15], a common mistake is to treat big data like a delimited over time project, rather than a continuous exploration exercise.

Economic challenges - The first obstacle to overcome when implementing big data in a company is not only the lack of budget, but also the lack of time and resources [23].

C. Big Data within Manufacturing SMEs

Research studies suggest that there are vast amounts of data due to the use of multiple storage and recording devices within production and industrial units in manufacturing sectors. Many companies are trying to

implement the advancement of technology and that of big data techniques to leverage the performance optimization of processes, operations and to gauge the benefits through real time application model development. Many studies have presented different case examples and application papers highlighting the importance of use of big Data in manufacturing sectors. For example, pharmaceutical companies have used the technology to optimize manufacturing process; design companies have utilized big data analytics to design and develop next generation of "future smart factories" and rely on managing information through various data sources. Another automotive manufacturing company uses big data and analytics to facilitate the use of smart grids to monitor the energy supply and usage within the production activities [26].

Early research studies and experts have identified that the key strategic decisions for the growth of SMEs is the adoption of technology and innovation. Big data is considered to be the key driver towards this growth by enabling to identify, model, predict and analyse the market data to facilitate the customer service function within the SMEs. It also benefits from the potential of higher proportion of productivity and performance through the capture of right amount of data and information at the right time and place. This allows the key decision makers in the company to make better strategic decisions to facilitate and sustain the future growth [13]. Research studies have long argued that innovation and technology has the potential to increase the future growth of the SMEs based on the company's performance and ability to create new knowledge. However, in order to generate and utilise new knowledge it requires the use of more advanced technology application and big data certainly plays a pivotal part in the development [27].

Many of the early data management applications highlighted earlier within paper are not capable to manage, store and analyse large amounts of data. Hence, newer technology applications are considered as more cost effective through their hardware and software capabilities to analyse both structured and unstructured data sets within the company. It can also be suggested that with better use of technology adoption within the company, firms can focus more towards growth and service aspects of their business. Many of the SMEs collect, store and analyse higher amounts of transactional data in digital formats, and hence with the use of Big Data applications it can enable them to higher strategic and performance related information within their environments. It also allows them to differentiate and segment their customer base more effectively and enable them to provide them more customized solutions, products and services. Some of the key aspects that big data enables SMEs to achieve are towards performance analysis, productivity, efficiency, new market segments, financial stability and control and innovation, hence achieving overall measure in cost, profits, revenues, growth and agility. SMEs are required to see the wider and future perspectives on the use of big data as it allows them to implement data analysis measures for risks, opportunities, predictive analysis measures, demand forecasting, optimization, inventory and resource planning, market segmentation, customer modeling and many more [13].

III. RESEARCH GAPS

After the first approach of Big Data and a study of the state of art, the authors have come up with some gaps. More and more companies are getting caught by the opportunities associated with big data and they are trying to have their own initiatives. One of the first points firms pay attention is the fact they need to acquire new technology resources. It is wrongly believed that the most expensive technology will yield better performance.

- ***Research Gap 1:*** The investment in data scientists, data warehouses or data analytics software needs to be balanced with firm's objectives and its capacity. It requires better understanding that Big Data is not something that you can buy for your company, it is not a certain project in which firms invest money once and they already become data-driven.
- ***Research Gap 2:*** Big Data means a change in the way organizations interact globally; it is a continuous loop from shifts and improvements to learning. And this continuous learning must be accompanied by a change of attitude, a change in the way executives manage this process.
- ***Research Gap 3:*** Companies should express their interest and beliefs towards accepting Big Data as a change tool for future growth and performance.

IV. CONCLUSION

The rise of data generation, the ubiquity of social media and always connected devices are changing the world rapidly and it will never be the same. Information is generated for everyone, everywhere and in a large amount of forms, and companies have to get used to living in this information ecosystem to get more intuitive, anticipatory and personalized performance. It becomes essential to undertake a research study to identify which are the existing challenges firms must face in order to be part of this new era. However, this is not enough. It remains more important to discover and come up with the attributes and aptitudes companies must have to become data-driven. It is not a matter of the most that they have but of how they are able to use the resources that they possess to prepare for how their business behavior will evolve over the next years. Future research will be aiming to describe the current state of SMEs businesses and the progress achieved in meeting new business attributes, identify their strategies and goals. Taking into consideration what literature says and what it is the current situation of SMEs

regarding Big Data, we will be able to find out and develop a feasible and progressive strategy for Big Data implementation. The objective is to provide companies with the tools and methods not only for becoming in data drivers but also to having the capacity to continue changing and adapting to world market needs over the time. Then, a methodology could be developed to provide companies a detailed overview of how they can identify where they are, how they can meet their needs and highlight a key directional shift to do it.

ACKNOWLEDGMENT

This paper is part of an ongoing research project within the research center and includes a preliminary study towards understanding the importance of Big Data applications within SMEs. The framework draws up current literature studies within this context and aims to develop more investigative based case studies within manufacturing SMEs to further develop the framework of these search.

REFERENCES

[1] S. Massa and S. Testa, "A knowledge management approach to organizational competitive advantage: Evidence from the food sector," *Eur. Manag. J.*, vol. 27, no. 2, pp. 129–141, 2009.

[2] A. W. King and C. P. Zeithaml, "Measuring organizational knowledge: A conceptual and methodological framework," *Strateg. Manag. J.*, vol. 24, no. 8, pp. 763–772, 2003.

[3] M. Alavi and D. E. Leidner, "Review: Knowledge management and knowledge management systems:Conce ptual foundations and research issues," *Manag. Inf. Syst. Res. Center, Univ. Minnesota*, vol. 25, no. 1, pp. 107–136, 2001.

[4] J. Gantz and D. Reinsel, "THE DIGITAL UNIVERSE IN 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East," *Idc*, vol. 2007, no. December 2012, pp. 1–16, 2012.

[5] R. Y. Zhong, S. T. Newman, G. Q. Huang, and S. Lan, "Big Data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives," *Comput. Ind. Eng.*, vol. 101, pp. 572–591, 2016.

[6] A. McAfee and E. Brynjolfsson, "Big Data: The Management Revolution," *Hardvard Bus. Rev.*, no. October 2012, pp. 60–68, 2012.

[7] J. Cawthorne, "Knowledge Management and Bid Data: Strange Bedfellows?," [Online]. Available: http://www.cmswire.com/social-business/knowledge-management-and-big-data-strange-bedfellows/ [Accessed 28 April 2017] 2015.

[8] I. Lee, "Big data: Dimensions, evolution, impacts, and challenges," *Bus. Horiz.*, vol. 60, no. 3, pp. 293–303, 2017.

[9] F. Portela, L. Lima, and M. F. Santos, "Why Big Data? Towards a Project Assessment Framework," *Procedia Comput. Sci.*, vol. 98, pp. 604–609, 2016.

[10] M. Chen, S. Mao, and Y. Liu, "Big data: A survey," *Mob. Networks Appl.*, vol. 19, no. 2, pp. 171–209, 2014.

[11] S. Sagiroglu and D. Sinanc, "Big Data : A Review," in *Collaboration Technologies and Systems (CTS), 2013 International Conference on. IEEE, 2013*, pp. 42–47. 2013.

[12] S. Akter, S. Fosso Wamba, A. Gunasekaran, R. Dubey, and S. J. Childee, "How to improve firm performance using big data analytics capability and business strategy alignment?," *Int. J. Prod. Econ.*, vol. 182, pp. 113–131, 2016.

[13] D. Sen, M. Ozturk, and O. Vayvay, "An Overview of Big Data for Growth in SMEs," *Procedia - Soc. Behav. Sci.*, vol. 235, pp. 159–167, 2016.

[14] E. O'Neill, "10 companies that are using big data,". [Online]. Available: https://www.icas.com/ca-today-news/10-companies-using-big-data [Accessed 4 May 2017], 2016.

[15] L. Analytics, "Why Most Big Data Projects Fail - Big Data into Insights," pp. 1–7, 2014.

[16] "What is Big Data?" [Online]. Available: https://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/. [Accessed 5 May 2017], 2012.

[17] C. Kenneth Neil and V. Mayer-Schoenberger, "The Rise of Big Data," [Online]. Available: https://www.foreignaffairs.com/articles/2013-04-03/rise-big-data. [Accessed 5 May 2017]. 2013.

[18] D. Laney, "3D data management: Controlling data volume, variety and velocity," *META Gr.*, no. February 2001, p. 4, 2001.

[19] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," *Int. J. Inf. Manage.*, vol. 35, no. 2, pp. 137–144, 2015.

[20] M. J, Michael., "Beyond Big Data: Gartner on 'Extreme' Information Management," [Online]. Available: http://forwardthinking.pcmag.com/show-reports/289401-beyond-big-data-gartner-on-extreme-information-management. [Accessed 29 May 2017], 2011.

[21] "the evolution of big data, and where we're ahead." [Online]. Available: https://www.wired.com/insights/2014/03/evolution-big-data-headed/. [Accessed 15 May 2017], 2014.

[22] S. Debortoli, O. Müller, and J. vom Brocke, "Comparing Business Intelligence and Big Data Skills," *Bus. {&} Inf. Syst. Eng.*, vol. 6, no. 5, pp. 289–300, 2014.

[23] IDC, "Business opportunities : Big Data," [Online]. Available: http://ec.europa.eu/enterprice/dem/sites/default/files/pages-files/big_data_v1.1.pdf [Accessed 26 April 2017]. 2013.

[24] R. van der Meulen, "Gartner Survey Reveals Investment in Big Data Is Up but Fewer Organizations Plan to Invest," [Online]. Available: http://www.gartner.com/newsroom/id/3466117. [Accessed 29 May 2017]. 2016.

[25] M. J. Mazzei and D. Noble, "Big data dreams: A framework for corporate strategy," *Bus. Horiz.*, vol. 60, no. 3, pp. 405–414, 2017.

[26] R. Y. Zhong, S. T. Newman, G. Q. Huang, and S. Lan, "Big Data for Supply chain Management in the Service and Manufacturing Sectors," *Comput. Ind. Eng.*, vol. 101, pp. 572–591, 2016.

[27] P. Thompson, R. Williams, and B. Thomas, "Are UK SMEs with active web sites more likely to achieve both innovation and growth?," *J. Small Bus. Enterp. Dev.*, vol. 20, no. 4, pp. 934–965, 2013.