

# Exploring the relevance and correlation between Big Data and Knowledge Management

Muhammad Saleem Sumbal<sup>a,\*</sup>, Eric Tsui<sup>b</sup>, W B Lee<sup>c</sup>

<sup>a</sup>**Research Student**, Dept. of ISE, Hong Kong Polytechnic University, HK

*E-mail address: saleemsumbal@gmail.com*

<sup>b</sup>**Professor**, Dept. of ISE, Hong Kong Polytechnic University, HK

*E-mail address: Eric.Tsui@polyu.edu.hk @polyu.edu.hk*

<sup>c</sup>**Professor**, Dept. of ISE, Hong Kong Polytechnic University, HK

*E-mail address: wb.lee@polyu.edu.hk@polyu.edu.hk*

## Abstract

Big data is buzz word encapsulating the organizations and is represented through volume, velocity and veracity. Huge amount of data is being generated at an astonishing pace which is then analyzed through advanced data analytics. As a result, useful knowledge is gained which then becomes basis for the companies for effective decision making and gaining competitive advantage over the competitors. This article attempts to explore the relevance and correlation between big data and knowledge management and gives a conceptual framework to determine the relationship between big data knowledge management (KM) and competitive advantage. No such work has been found up to author's knowledge and thus this article will provide a basis for the organizations to get to know about the relationship among big data and KM which are mostly considered as two separate entities. Most of the organizations now understand the importance of big data and analytics, thus, by gaining an insight into this relationship between big data and knowledge management, organizations can better maneuver their business activities and engage in better and effective decision making.

**Keyword:** Big Data, Knowledge Management, Competitive advantage, Effective decision making

## 1. Introduction

Big data is a buzz word these days and governments, organizations and media have been encapsulated with this word. With the advancement in IT, huge amount of data is being generated over the web and stored in the cloud. The three words that are typically used to represent big data are volume, variety and velocity meaning, different types of big data is being generated in large volumes at huge pace. This has paved the way for the researchers to focus on topics like data sciences and data analytics and in the upcoming years, how big data will play a revolutionary role in the success of organizations.

Data exists in both structured and unstructured forms. The opportunities lie in analyzing the big data and finding the trends and patterns as huge volumes of data exist but the challenge

is to find the quality information and seek knowledge from it to be used for the success of the organizations. Big data analytics is coming out to be one of the most emerging technology investment by the organizations as the big data has its implications over all areas of an enterprise like marketing, sales, research, finance, supply chain, customer relations etc. The hiring of people for the role of chief data scientist or chief data officers are perhaps the hot issues among organizations these days.

## **2. Aims & Objectives**

There are various industry articles emphasizing the importance of big data but in academics, not much work has been done in this regard. Moreover, there is need to investigate the role of big data and its relationship with knowledge management as the linkage between these two domains is worth exploring. The current article focuses on this topic and gives a detailed view about the relationship of big data with knowledge management. The major commonalities among big data and knowledge management in terms of competitive advantage, effective decision making, IT infrastructure and organizational culture are discussed in detail and finally a conceptual framework is presented for the empirical testing of the relationship between big data and knowledge management.

## **3. Big Data and Knowledge Management**

The knowledge based view of the firm states that it is the amount of knowledge within the organization that determines the value of the organization [1]. Thus the sustainability of the organizations is dependent on the accessible knowledge within the organization and reconfiguration of knowledge based resources for gaining the competitive advantage. Knowledge based resources tend to be valuable, scarce, difficult to imitate and substituted by competitors and thus serve as the basis for competitive advantage [2]. This reconfiguration of knowledge takes place through continuous learning and knowledge acquisition as the business environments change. Thus organizations nowadays need to possess this "dynamic capability" of adopting themselves according to new trends and learning and creating new knowledge for achieving sustainable competitive advantage [3]. This concept of dynamic capability is referred to as "ability to integrate, build and reconfigure internal and external competencies to address changing environments" [4]. If the stocks in the organizations are used for knowledge flow across the organizations in a purposeful way [5], it can bring competitive advantage to the companies [1, 6] resulting in superior place among the competitors in the market. Big data actually falls in this category and with the development of big data approaches [7], there is a similar interest to gain competitive advantage through big data on which organizations are focusing to create new knowledge from big data and analyze the patterns for their business growth. This is especially useful for the e-businesses as they can quickly comprehend the ongoing trends

through the analysis for their online sales and purchase and thus reconfigure their resources for product development and the services they provide. All this new knowledge can be generated in real time through real time analytics. Thus reconfiguration capability of organizations is quite important as it determines capacity to absorb external knowledge and deduce generalizable cause effect relationships using existing knowledge for improved performance [4]. Thus organizations are increasingly becoming aware of reaping benefits by investing in big data analytics and getting better return on investment (ROI) and building sustained competitive advantage [8].

There is no doubt that organizations nowadays are immersed in huge expanding sea of data which is very much unstructured and thus can't be managed and analyzed using traditional means. Organizations use this big data by i) paying attention to data flows rather than stocks ii) relying on data scientists and product developers as opposed to data analysts and iii) moving analytics away from IT functions and putting into core businesses and operational functions [9]. In this way, organizations can generate valuable knowledge and thus use it for improved performance and competitive advantage which is same as knowledge management and in other words we can say that big data and analytics contribute towards real time knowledge management using the online unstructured data. Big data can be considered as a knowledge asset and thus field of knowledge management has gained new momentum by the inclusion of big data analytics for knowledge creation.

#### **4. Big Data as basis for knowledge generation for Effective Decision Making and Competitive Advantage like Knowledge Management**

Due to rapidly changing environments, data warehouses need to continuously remain up to date by incorporating new data structures and modifications in the old ones because of the huge amount of data being generated on a very fast pace and also because nowadays employees average time in an organization is far less as compared to before. Thus when employees leave a knowledge gap is created which needs to be filled by continuously by dataset modifications and relationships with each other [10] .

Emphasis has been put on big data and linkage with the knowledge generation as a result of analysis of this data. According to Bose [8], decision oriented analytic application allow informed decisions and can be considered as a "knowledge management initiative in which the organizations' s best practices for each decision making process are pushed to the desktops of end users as embedded logic within analytic applications". A lot of factors are being considered in this regard to have a common understanding of the knowledge generated through use of different datasets and data structures. Sukumar [10] developed a tool called SEEKER which effectively helps in visualizing different datasets together and form a linkage among the data. This eventually helps in better knowledge generation as the graphic visualization of the data linkages is very effective and easily grasped by humans than

the traditional way of databases layouts.

McAfee et al [11] state that one of the goals of knowledge management is to integrate the information from different perspectives and analyze for valuable decision making. It is now much easier to collect information from different sources as we are living in era of big data and thus we can use this information through proper analysis for generating knowledge and utilizing it and as a result, organizations can make valuable decisions and increase their worth. One of the examples can be given of expedia.com which is one of the leading online travel's site and provides information in 26 local languages with over 75 million users each year. Expedia has largely invested on big data analytics to gain useful insights and generate knowledge for understanding the customer behaviors. They generate useful insights on what type of advertisement links attract customers and developed models for determining the causal relationships between marketing effort and customer's response. In this way we can see that big data analytics help in generation of knowledge which is further stored and shared among the employees to see how the business can be improved and what steps to be taken accordingly after knowing the behaviors of the customers. This is very clear example of how knowledge management is related to big data. The main difference between companies in today's highly competitive environment is their ability to make on time, accurate and effective tactical, operational, and strategic decisions.

Another example is of credit card companies which through continuous web monitoring and call centre activities can make personalized offers to customers and then optimize these offers by analyzing the responses from customers and in this way they are capitalizing using effective decision making with the help of real time knowledge generation concept using big data analytics [9]. We are living in the age of knowledge economy and key to survival in this highly dynamic and competitive environment is to capitalize on knowledge and ability of organizations to extract knowledge from pool of big data [12]. According to study done by Hair Jr. [12], predictive analytics and big data are playing a big role in marketing area involving product development, distribution, advertising and retailing.

The above examples clearly provide an insight of how big data and knowledge management are interlinked and e-businesses are paving the way to success through knowledge generation by the help of big data. In fact this can be termed as "Real time knowledge management" or "Big data based knowledge management" as a comprehensive insight is provided by big data analytics to track opportunities for revenue generations, enhancing sales, mitigating cost risk and put focus on customers and resources same like knowledge management. Figure 1 gives a comparison of this new type of knowledge management and traditional knowledge management. Then, table 1 shows the major differences in these two types of knowledge.

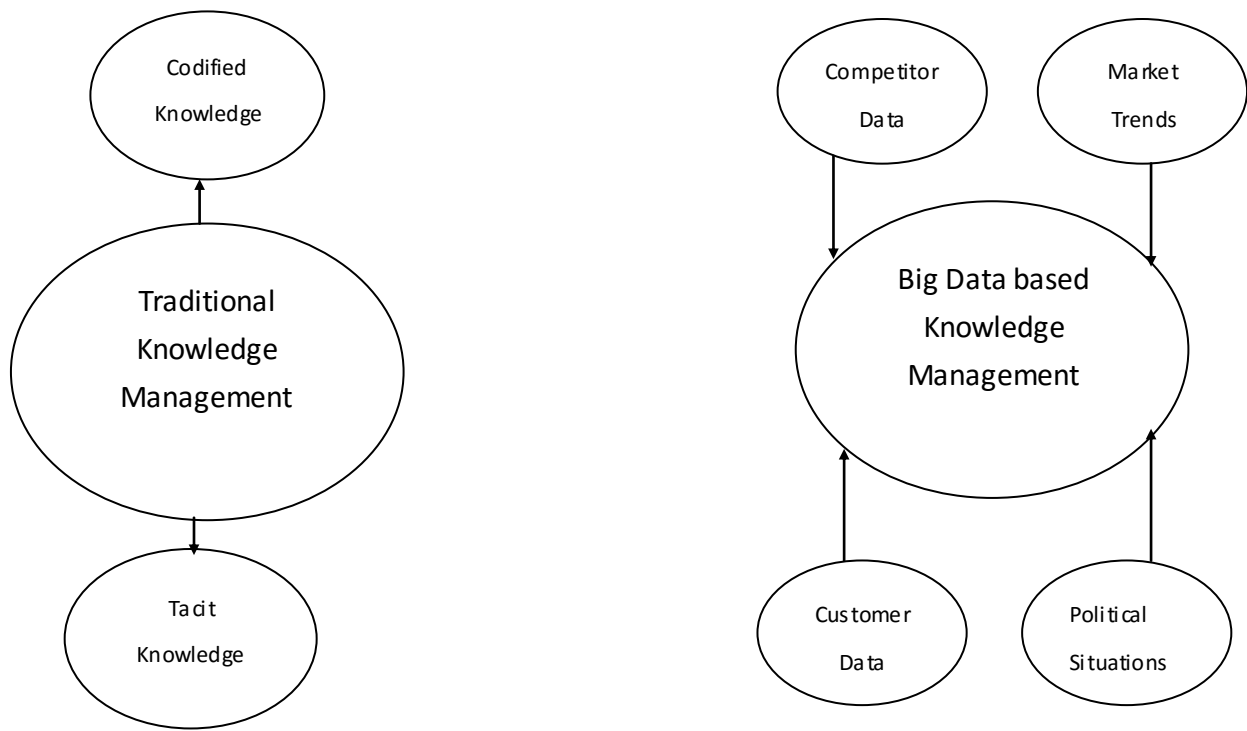


Figure 1: Comparison of traditional knowledge management and big data based knowledge management

### **5. Big Data based knowledge management linked to IT, organizational culture and people same like traditional knowledge management**

IT trends have helped shaping the present big data era and advancement in big data analytical techniques. According to Thomas Friedman [13], high speed network connections, international travel and outsourcing have created the tremendous opportunity for IT to boom. Moreover, development of business related standards, electronic data interchange formats, business databases and information systems have played a huge role in data creation and utilization [14]. Due to advancement in IT, highly rich and contextual content relevant to a specific business or organization can be obtained by analyzing the tremendous amount of web-based, sensor and mobile generated data. Thus information and communication technology (ICT) is an enabler for big data generation and big data analytics same like knowledge management. In order to handle such massive amount of data, both structured and unstructured, highly optimized databases are required like EMC Greenplum, HP Vertica etc and furthermore, after the processing of data, it is required to present this data in an easy to understand way for the end users and this can be accomplished using open, popular, light weight and web oriented architectures [15]. As the users at the end normally want rich interactive web interfaces, thus popular and light weight architectures need to be used which will facilitate the knowledge discovery process so that the end users

<b><i>Bid Data based knowledge management</i></b>	<b><i>Traditional knowledge management</i></b>
Online processing of data to extract useful knowledge. Real time and may be codified later.	Tacit knowledge and explicit knowledge
Less involvement of people / Machine focused	Relies more on people/ People focused
High analytical skills for knowledge extraction	No need of high skills
Continuous flow and processing/ Cloud storage	Huge Repositories plus tacit knowledge in heads of people
Less face to face interaction with customers	More face to face interaction with customers

Table 1: Comparison of Big Data based knowledge management and traditional knowledge management

can understand the final output let's say in the forms of graphs, figures, tables etc . Thus ICT has a pivot role in this regard and thus in order to take full advantage of big data, smart optimized processes need to be considered for data processing and results display.

Robert W Gehl [16] thrives on the concept of sharing and states that one of the products of knowledge economy is sharing and it produces huge amount of data. Then there comes the concept of knowledge worker who basically finds out the value from this data and information called knowledge and applies this knowledge into action. This strengthens the idea of knowledge economy which centralizes on the concept of sharing and communication to expand and thus gave birth to the concepts of "story telling" and "best practices sharing ". Sharing means growing knowledge and thus the knowledge economy [16]. To produce value from this, knowledge workers are the people as knowledge can be evaluated by the valuable actions which lead to new products, services and potential vectors for value realization.

Organizational culture also effects the knowledge extraction using big data analytics same like traditional knowledge management. Big Data analytics are powerful enough but a number of steps must be ensured as explained by Bose [8] to cultivate a proper organizational culture necessary for efficient knowledge generations, utilization and dissemination. These include i) a cross functional team should be made for proper analytics project including technical staff along with executives to monitor the whole process ii) Project success to be measured through properly defined metrics and iii) there should be incentives on performing well on the project. Along with that specialists need to be hired or employees need to be trained to handle such systems. The output as discussed earlier should be presented in usable and understandable formats to the users in the form of graphs or visuals.

## 6. Conceptual Framework

Based in the light of above discussion, a conceptual framework is presented (Figure 2) to have an empirical investigation of the linkage of big data and knowledge management and other relating factors. The main hypothesis to be investigated are as follows :

H1 : Big data is directly correlated with knowledge management and helps in knowledge creation sharing and utilization for effective decision making.

H2 : Big Data is the basis for competitive advantage among organizations nowadays.

H3 : Big Data is highly related to organizational culture and IT infrastructure same like knowledge management.

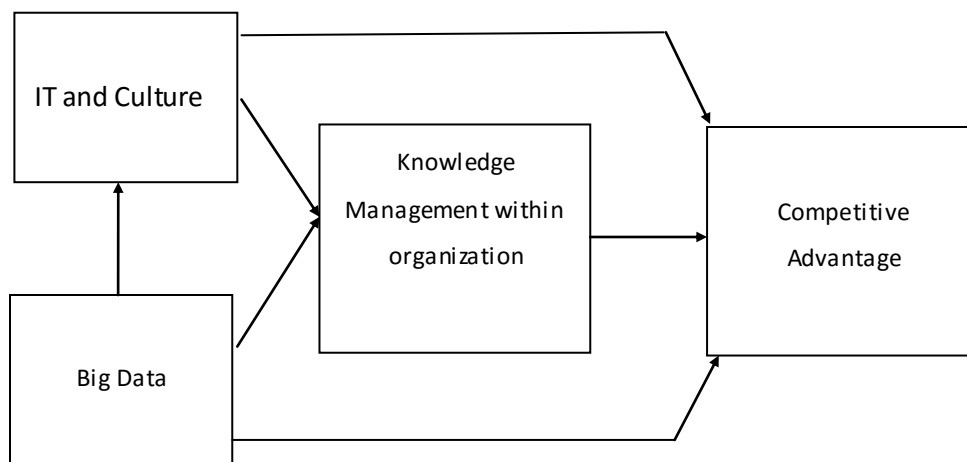


Figure 2: Conceptual framework

## 7. Conclusion

The current article focuses on the relationship between big data and knowledge management. An attempt has been made to interlink the two concepts. Big data and knowledge extraction from big data mainly serve the purpose of providing competitive advantage and business performance to the organizations. Evidences show that valuable knowledge can be extracted from big data and it has become necessary to keep track of the big data due to rapidly changing environments and markets. In this regard, there are challenges like lack of data scientist, learning new technologies and privacy of data. Data driven decisions tend to be better and help companies shape their businesses in a better way along with the experiential knowledge of the employees. Market dynamics involving globalizations, political uncertainty, commoditization and massive generation of data, all are the reasons which compel successful companies to incorporate business intelligence using big data analytics for growth. Especially for online businesses, efficient knowledge management is only possible through big data analytics.

## References

- [1] Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic management journal*, 17(S2), 109-122.
- [2] Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of management*, 17(1), 99-120.
- [3] Fuchs, M., Höpken, W., & Lexhagen, M. (2014). Big data analytics for knowledge generation in tourism destinations—A case from Sweden. *Journal of Destination Marketing & Management*, 3(4), 198-209.
- [4] Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic management journal*, 18(7), 509-533.
- [5] Teece, D. J. (1998). Capturing value from knowledge assets: The new economy, markets for know-how, and intangible assets. *California management review*, 40, 55-79.
- [6] Zack, M. H. (1999). Managing codified knowledge. *Sloan management review*, 40(4), 45-58.
- [7] Vance, A. (2011). The data knows. *Bloomberg Businessweek*, 70-74.
- [8] Bose, R. (2009). Advanced analytics: opportunities and challenges. *Industrial Management & Data Systems*, 109(2), 155-172.
- [9] Davenport, T. H., Barth, P., & Bean, R. (2013). How 'big data' is different. *MIT Sloan Management Review*, 54(1).
- [10] Sukumar, S. R., & Ferrell, R. K. (2013). 'Big Data' collaboration: Exploring, recording and sharing enterprise knowledge. *Information Services and Use*, 33(3-4), 257-270.
- [11] McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D., & Barton, D. (2012). Big data. *The management revolution. Harvard Bus Rev*, 90(10), 61-67.
- [12] Hair Jr, J. F. (2007). Knowledge creation in marketing: the role of predictive analytics. *European Business Review*, 19(4), 303-315.
- [13] Friedman, T. L. (2006). *The world is flat [updated and expanded]: A brief history of the twenty-first century*. Macmillan.
- [14] Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS quarterly*, 36(4), 1165-1188.
- [15] Begoli, E., & Horey, J. (2012). *Design principles for effective knowledge discovery from big data*. Paper presented at the Software Architecture (WICSA) and European Conference on Software Architecture (ECSA), 2012 Joint Working IEEE/IFIP.
- [16] Gehl, R. W. (2014). Sharing, Knowledge Management, and Big Data: A Partial Genealogy of the Data Scientist. *Forthcoming in the European Journal of Cultural Studies*.