The following publication Sumbal, M.S., Tsui, E. and See-to, E.W.K. (2017), "Interrelationship between big data and knowledge management: an exploratory study in the oil and gas sector", Journal of Knowledge Management, Vol. 21 No. 1, pp. 180-196 is published by Emerald and is available at https://dx.doi.org/10.1108/JKM-07-2016-0262

Interrelationship between Big Data and Knowledge Management: An Exploratory study in the Oil and Gas sector

Abstract

Purpose – The purpose of this article is to explore the relationship between big data and knowledge management (KM).

Design/methodology/approach – The study adopts a qualitative research methodology. A case study approach was used by conducting 9 semi-structured interviews with open ended and probing questions.

Findings – Useful predictive knowledge can be generated through big data that can help companies improve their knowledge management capability and make effective decisions. Moreover, a combination of the tacit knowledge of relevant staff with explicit knowledge obtained from big data improves decision making ability.

Research limitations/implications – The focus of the study was on the oil and gas sector and thus the research results may lack generalizability.

Originality/value – This paper fills an identified need for exploring the relationship between big data and knowledge management which is lacking in the literature.

Keywords: Big Data, Knowledge Management, Interrelationship, Oil and Gas

1. Introduction

Organizations have long been aware of knowledge management (KM) and the various ways of managing their knowledge, however, new challenges continually arise and new strategies need to be adopted to cope with these challenges. One major challenge in organizations is the leveraging of big data and, more importantly, its link with the knowledge management capability of the organizations. The huge potential of big data for organizations in valuable knowledge creation and gaining competitive advantage has prompted companies to invest substantially in this area. Big data analytics help in understanding and extracting valuable knowledge from the huge volumes of data, and this knowledge then can be used for enhancing the performance of many

different processes in an organization. This links big data to knowledge management capability as well because of its potential for valuable knowledge generation and the ability to improve the organizational processes. In organizations, big data is especially important as massive amounts of data are generated during the many different processes, for example sales, marketing, transaction records, oil and gas exploration, customer feedback, competitor's data, research and development and other day to day operations. Knowledge generation using big data is a comparatively new field and companies have started experimenting in this field. After a thorough literature review, it has been found that very sparse research has been conducted on the linkage of big data and knowledge management as further explained in section 4. Addressing this shortfall in research, this article studies the interrelationship between big data and knowledge management by conducting an exploratory study in the oil and gas industry to answer the following research question:

What is the interrelationship between big data and knowledge management in a knowledge intensive industry?

The following sections cover the literature review, research methodology, data analysis and discussion of results and finally the conclusion of the study, with future recommendations.

2. Big Data

The three words that are typically used to represent big data are volume, variety and velocity meaning that different types of data are being generated in large volumes at a huge pace; nowadays a fourth word, veracity is also used to represent big data (Schroeck et al., 2012). Due to advancement in technology, huge volumes of data are generated from multiple sources (Waller and Fawcett, 2013). The velocity of data generation these days is far higher than before and a data of 5 exabyte is generated in 2 days which was the total amount of data created by humans up to 2003 (Sagiroglu and Sinanc, 2013). Variety means that data can exist in various forms, structured, semi structured or unstructured and be generated through various means. Veracity is used to describe the quality and usefulness of the captured data for effective decision making. Many industries are increasingly becoming interested in accelerating research on big data (Chen et al., 2014). The huge and rapid growth of data is also supported by the Internet of things and Cloud Computing. Big data has implications on improving services and operations to

meet customer demands as compared to traditional datasets in the past that were mainly used for handling internal business decisions (Davenport, 2014).

Russom (2011) in his survey report mentioned that anything that involves customers can benefit from big data and analytics (for example, customer based segmentation, targeted marketing), secondly, big data analytics help improve business intelligence (for example better planning and forecasting) and thirdly, specific analytical applications can benefit from big data (for example fraud detection, risk quantification). Nowadays, analytics using big data are considered to be at the heart of a business. Organizations can capitalize on big data by i) paying attention to data flows rather than just stocks ii) relying on data scientists as opposed to data analysts and iii) moving analytics away from IT and into core businesses and operational functions (Davenport et al., 2013). However, there are barriers, like lack of expert staff for analytics and insufficient IT infrastructure to manage which some people consider a problem in handling big data.

3. Big Data and Knowledge Management

Big data was initially defined in 2001 by Doug Laney, an analyst of META (now Gartner) with a 3V model i.e. volume, velocity and variety, and this model was used for a decade by IBM, Microsoft and Gartner. Later on, IDC (International Data Corporation) in 2011 defined big data as "big data technologies describe a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling the high-velocity capture, discovery, and/or analysis". According to Chen et al. (2014) this definition is widely recognized and entails a fourth V i.e. value, which highlights the meaning and importance of big data and how to extract the hidden value from large datasets. Knowledge management is a well-known term and it is almost an integral part of every organization in today's world that is based on the knowledge economy. In the classical literature, knowledge management is considered as a process which involves various activities with four major activities: creating, storing, transferring and applying knowledge (Alavi and Leidner, 2001). Similarly, Bassi (1997) explains knowledge management as the process of creating, capturing and using knowledge to enhance organizational performance. It is clear that knowledge management comprises activities or processes of knowledge creation and dissemination to achieve different organizational objectives, such as business success, improved performance of employees and achieving a competitive advantage. Thus if we look in terms of the knowledge

creation or value described above in the big data definition, and compare with KM definitions, a link can be seen as to how knowledge can be created through big data and further used for enhancing organizational performance. Further, if we look at the DIKW model (Ackoff, 1989), most authors understand knowledge by separating it from data and information (Fig.1). Here the challenge is to understand how the raw data is converted into information, and information into knowledge. We can draw an analogy and can link big data with KM using this DIKW model. Data is basically discrete facts, the flow of events and activities, in an organizational system without any context, but when it is given context, it becomes information. This information, based on intuition and personal experience, is then converted into knowledge. The aim of big data and knowledge management is almost same which is to use knowledge for better decision making. However, the difference exists in the way it is performed. Big data is structured, unstructured or semi-structured data available through a variety of sources. When analytics are applied to big data, it becomes information and has some meaning because we try to organize or structure the data to see some underlying trends and patterns. When this information is fed into business intelligence tools to interpret the data and finally through the analysis performed by data analysts, some actionable knowledge is generated which is termed knowledge management. Thus this analogy provides a strong link between big data and knowledge management, where we have data and analysis of this data using analytics, along with the input from the data analysts can result in effective knowledge management, which is very helpful for improving a business.

Organizations need to remain updated about how they can improve and what is going to happen in future in order to remain competitive in today's rapid and changing business environments. According to Bose (2009, p.157), decision oriented analytic application for big data helps in decision making and thus can be termed a ‘"knowledge management initiative in which the organization's best practices for each decision making process are pushed to the desktops of end users as embedded logic within analytic applications". To integrate the information from different perspectives and analyze it for valuable decision making is one of the goals of knowledge management (Lamont, 2012). We are living in era of big data and thus through proper analysis of the data generated through different sources, valuable knowledge can be generated which organizations can use to make valuable decisions. Through the development of

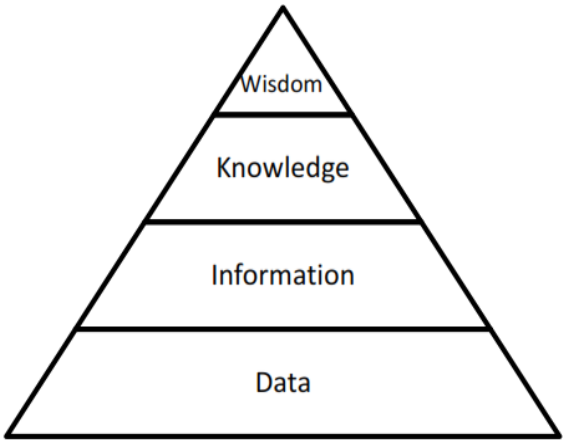

Figure 1: The knowledge pyramid

big data approaches (Vance, 2011), organizations are now focusing on creating new knowledge from big data and in analyzing the patterns for their business growth. The concept of business intelligence is the upcoming trend in which the main emphasis is to utilize and analyze structured and unstructured data using data analytic techniques like data mining, statistical analysis etc. for improved business performance. Managers can readily know about all the details of their businesses through big data analytics and can put the acquired knowledge into practice for improved decision making (McAfee et al., 2012, Waller and Fawcett, 2013). For example, in the tourism industry, a destination management information system was developed by Fuchs et al. (2014) which helps in knowledge creation and application for organizational learning and improved performance. It basically capitalizes on extraction of structured and unstructured data from web based sources, online surveys, social media etc. and analyzes this data for useful knowledge generation which can then be applied in the form of intelligent services for customers and destination stakeholders, for example hotels, guides, recreational service providers etc. Regarding the application of big data in e-businesses for marketing and sales, companies flourish in their businesses by incorporating real time knowledge management through big data; thus significantly gaining competitive advantage over their competitors (McAfee et al., 2012). For example, Amazon.com is a well-known company and has challenged the business of traditional stores using big data analytics, and has capitalized on the knowledge generated from big data. Amazon is continuously optimizing itself and is highly customer focused. It uses machine learning algorithms on the huge amount of web data generated by customers to look for the

trends and personalized preferences and based on that, it reconfigures its processes and services for customers. Similarly, Ali Baba is another e-business giant. The company decided to provide loans to enterprises after analyzing their transaction data and all this is done automatically without any human intervention. About RMB 30 million of loans have been granted through this process, resulting in only 0.3% bad loans. In case of the health sector, at Toronto's Hospital for sick children, effective decision making is enhanced through the usage of machine learning algorithms for discovering patterns that anticipate infectious diseases in babies before they occur (Davenport et al., 2013).

Thus this can be considered as a new form of knowledge management in which the major role is played by the big data and analysis techniques for efficient and effective knowledge generation, on which many industries, as explained above, can capitalize. Effective decision making, which is one of the goals of KM, is also now performed through the use of knowledge generated from big data analytics which provides evidence of the relationship between knowledge management and big data (Murdoch and Detsky, 2013). Chen et al. (2014) also argued that enterprises can use big data analytics to generate knowledge for improving competitiveness and production efficiency, for example in marketing, prediction of consumer behavior, in sales, optimization of commodity prices etc. According to a survey by IBM in 2011, companies, especially top enterprises, are successful because their use of big data and analytics is 5% more than companies lower to them and this performance is because of the competitiveness lying in big data (LaValle et al., 2013). In today's competitive world, executives want the best decisions in real time for their companies in case of disruptions whether it is an earthquake, an epidemic, political disaster or unexpected competitor, and the valuable knowledge gained from big data can provide the answers (LaValle et al., 2013).

Lack of constant knowledge creation leads to poor performance, therefore, proper strategies need to be devised to create new knowledge (Choi and Lee, 2002). Moreover, knowledge fostering enablers such as people, processes, systems and organizations need to work coherently for effective knowledge creation. Studies on intellectual capital and intangible resources (Bassi, 1997, Hall, 1992) explain which type of knowledge is unique and valuable; then how these resources can support an organization's market position and products is explained by resource based theory and knowledge based theory (Choi and Lee, 2002). Finally, how knowledge

creation and learning occurs using these resources is explained by the SECI model. In the traditional SECI model (Nonaka and Takeuchi, 1995), there are two types of knowledge, tacit and explicit (Fig-2). Explicit knowledge is structured knowledge that can be documented, categorized and easily transmitted to others (Duffy, 2000) whereas tacit knowledge resides in the heads of people and is relatively difficult to be codified (Gore and Gore, 1999). Knowledge creation and learning occurs through conversion between tacit and explicit knowledge in 4 modes i.e. socialization, externalization, combination and internalization and such knowledge transitions can occur either in physical space (office), virtual space (email, teleconferencing) or mental models (shared experience, ideas, beliefs) etc. Knowledge management strategies adopted by firms greatly influence the KM processes (Zack, 1999). A big challenge for organizations nowadays is to possess "dynamic capability" in sensing and adapting themselves to new trends as well as learning and creating new knowledge for achieving sustainable competitive advantage (Fuchs et al., 2014, Manyika et al., 2011). According to Teece et al. (1997); this concept of dynamic capability is referred to as the "ability to integrate, build and reconfigure internal and external competencies to address changing environments". Based on this dynamic capability view, big data is also a new trend which has created a lot of opportunities to discover and gain understanding of hidden information in order to improve different business processes and operations, especially in prediction and decision making. In big data based knowledge management, we need to take into account a variety of factors, as the change in data is linked to these factors. For example, oil and gas prices fluctuate based on the political situations in the oil producing countries and the change in oil prices causes changes in almost all sectors. Similarly, knowledge about competitors, trends in the market as per customer feedback, force majeure (epidemics, earthquakes); all these factors affect the generated data. Decisions are made in real time according to the data generated, and timely and accurate decision making is what makes all the difference among companies. Based on the above discussion, Fig-2 and Fig-3 have been developed which give a comparison of this new type of big data based knowledge management with traditional knowledge management.

4. Research Context and Motivation

An extensive literature review was conducted using various databases such as Emerald, ScienceDirect, Scopus and Google Scholar and it was found that very little work has been done

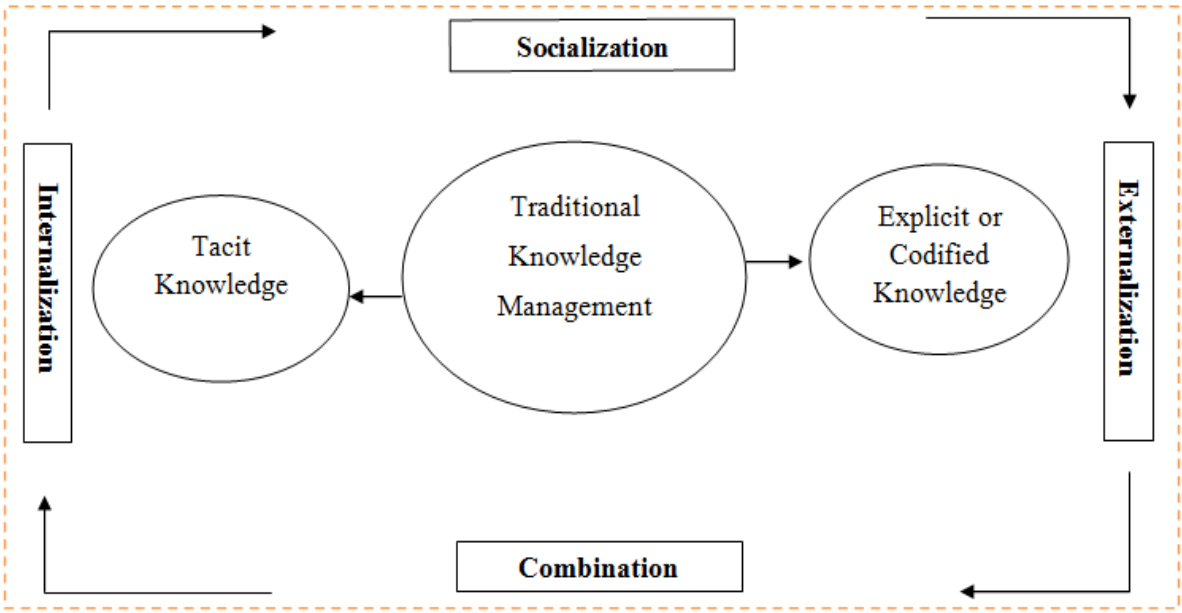

Figure 2: Traditional Knowledge Management

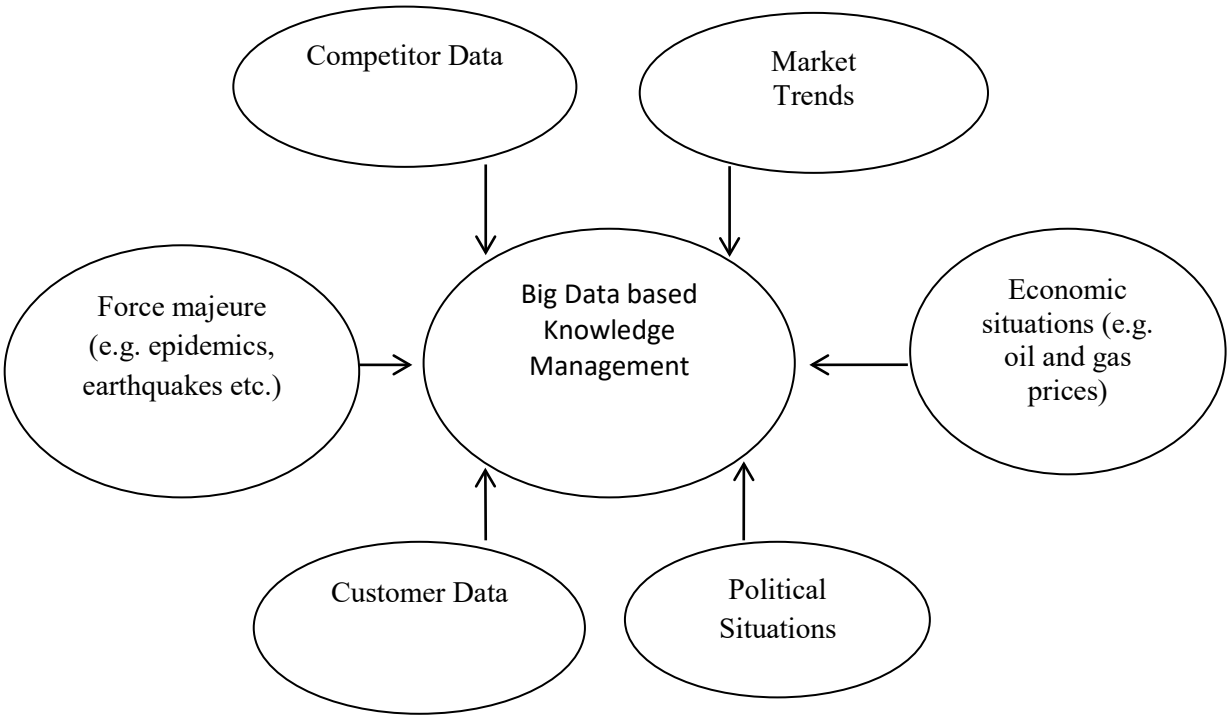

Figure 3: Big Data based Knowledge Management

on linkage of big data to knowledge management. Special attention was paid to journals on knowledge management such as Journal of Knowledge Management, Knowledge Management Research and Practice, Journal of Information and knowledge management, VINE etc. Various combinations of key words, such as big data, analytics, knowledge management, knowledge generation, knowledge creation were used to search the articles. Articles on big data from the field of computer sciences were not included. Being a relatively new field, a total of 60 articles were found. Further review and refinement of these articles finally yielded 2 articles, which tried to explain the relationship between big data and knowledge management. One study was conducted by Erickson and Rothberg (2015), in which they took information about various industries from the McKinsey Global Institute report on big data, and using a strategic protection factor framework, developed a link between big data, intellectual capital and knowledge management. In the other study by Chan (2014), a framework was developed for the integration of customer knowledge management (CKM), big data channels and technology platforms such as social media and machine to machine platforms (M2M). The main idea was to capture structured and unstructured data from big data platforms and its transformation into actionable business insights through the CKM process of knowledge creation, dissemination, storage and utilization. These studies provided a hint but still they did not focus specifically on how big data is linked to knowledge management. Also, most of the work discussed in the literature review just provides an overview and some theoretical evidence of the linkage of big data to knowledge management, revealing the gaps in terms of empirical research on exploring the relationship between big data and knowledge management. Thus, in order to fill this gap on how organizations and employees link big data to knowledge management and to gain some practical evidence of value creation and leveraging of big data to create useful knowledge is explored in this article. The oil and gas industry is one of the most complex, most important and one of the world's largest industries globally (Inkpen and Moffett, 2011). It is a highly knowledge intensive industry and thus efficient knowledge management is key to success of the organizations (Nonaka and Takeuchi, 1995). Effective knowledge management helps organizations in enhancing their overall performance and gaining competitive advantage (Grant, 1996). The oil and gas industry has been long working on knowledge management and has effective ways of managing knowledge, however, big data is a new challenge for these companies and how effectively it can be utilized for effective knowledge management is vital. Processes and

decisions related to oil and gas exploration, development and production generate large amounts of data and with new technologies, this data is being generated at a far higher rate (Farris, 2012). Studies have shown that oil and gas is a highly data intensive sector and companies are well aware of the benefits associated with big data, thus, there is great potential in generating predictive knowledge for improved performance through big data (Farris, 2012, Feblowitz, 2013). All the above factors provided a strong context and motivation for research in this industry.

5. Methodology

Qualitative methods, such as interviews, are considered more appropriate for obtaining a deeper understanding of a subject under study as compared to quantitative methods such as questionnaires. Therefore, interviews are the most appropriate method when detailed insights are required from individual participants (Bruce and Berg, 2001). For the current research, semi-structured interviews were conducted with managers, directors and senior staff of the oil and gas sector where open ended and probing questions were asked to gain a deeper insight on the research topic (Gill et al., 2008). Due to nature of the research questions, scattered knowledge and the applicability of big data in various operations of oil and gas sector, multiple companies were selected to gain a practical insight of the patterns of relationship between big data and knowledge management. The interviewees were selected based on the available contact points of the research team and their vast amount of experience in the area under research. The interviewees also represented “organizational elites” or “key informants” working in responsible and key positions, as elite interviewing is a common qualitative method with unique benefits of yielding insightful information (Marshall and Rossman, 2011). The target people were contacted through emails, LinkedIn profiles and phone calls. Based on availability; a total of 9 interviews were conducted (Table 1). The duration of each interview was between 50-70 minutes. The interview questions were sent in advance to the respondents, so they could prepare themselves accordingly. The interviews were recorded and transcribed afterwards. Moreover, field notes were also taken during the interviews and later on matched with the transcribed data for analysis purposes. Unclear and ambiguous answers were also determined from interviewees and amended according to the responses received. In order to validate the study and avoid bias, the transcribed interviews were verified from the interviewees and in some cases, interviewees made

amendments in the transcribed data. In addition, the validation of the results was also performed through grounding in theory and the literature (Daghfous et al., 2013). Thus a critical analysis of the results was performed to obtain the answers for our research question. Through careful analysis of the arguments made by respondents and noting down the similarities and differences, the underlying concepts were understood. These concepts were further compared with existing literature to see what has already been discussed and what new ideas have emerged from these results in the light of practical examples provided by participants. The responses indicated not only the significance of the research but also helped in elaborating the linkage through rich practical examples provided by respondents, and the underlying concepts were further complemented by the existing literature which helped in establishing the validity and reliability of the data.

Interviewee	**Years of Experience**	**Position**	**Company**	**Location**	**No of employees in company**
1	8	Managerial	A	Russia	50k-100k
2	30	Managerial	B	USA	5k-10k
3	10	Managerial/ Consultant	C	USA	20k-50k
4	35	Director/ KM Lead	D	Netherlands	50k-100k
5	6	Junior Manager	E	Nigeria	5k-10k
6	16	Managerial	F	Middle East	20k-50k
7	10	Managerial	G	Middle East	20k-50k
8	7	Managerial/KM Coordinator	H	Australia	50k-100k
9	26	Director/ KM Lead	I	USA	50k-100k

Table 1: Interviewees and participating companies in the research project

6. Results

This section describes the results obtained from the interviews. Six of the interviewees stated that they have dedicated units working on big data and analytics, and the rest mentioned that big data operations are normally handled by the IT department. All interviewees agreed that big data is of major concern in the oil and gas industry as huge volumes of data are generated during exploration and drilling, coupled with the fact that more advanced technology and tools are being used these days which can generate higher volumes of data. Addressing the main research question on the linkage of knowledge management and big data, the interviewee from Company A stated that data management isn't a new activity for them. They have made good progress in both technology and management of big data, and big data is seen as a complementing rather than disruptive technology. There are 15 big data projects going on in the upstream sector. Big data can be linked to knowledge management as it is all about generating useful predictive knowledge and utilizing it effectively, for example, in performing predictive maintenance of rotating machines such as turbines and pumps. The company has a scheduled maintenance every 6 months during which the machines are stopped and inspected. Based on huge amounts of sensors data, there are softwares to analyze the parameters of machines and provide recommendations on parts maintenance and when to do maintenance. The crew then performs maintenance accordingly. This project started 3 years ago is now in the application phase. Another project in the trial phase uses a huge collection of satellite images for decision taking related to the pollution of assets. The company has a lot of installations under water and the idea is to monitor the assets such as oil traces in the water by monitoring the leakage of oil pipelines. So using the images from satellites, the company is trying to automatically detect the leakages. Another example is digital scouting in which company uses text mining and semantic analysis of publicly available information to analyze its competitors. Completed trials clearly showed a big role in open-source technologies for generating valuable knowledge to enhance productivity and competitiveness. The interviewee added that company achieved promising results on big data experiments for value creation and improving business operations.

According to interviewee from company B**,** big data and knowledge management are very much related. Once you discover the hidden secrets or knowledge in the data, how you propagate it is all about management of this knowledge. In his company, for example, they wanted to know,

what constitutes a good fracturing job. Fracturing is a technique in which fractures in deep rock formations are opened and widened by injecting liquids and chemicals at high pressure for easier extraction of oil or natural gas. Fracturing jobs are very expensive costing a lot of money. There are about 32 critical parameters in the operation of a fracturing job. The team had to look at these parameters, for example pressure, flow rates, proppant size etc. and decide which ones were the most important in terms of the design of the fracturing job configuration down in the well. There is also potential for environmental damage during fracturing. So getting it right is extremely important and “having the right knowledge from experienced people and using it for analysis of data is extremely valuable and people don’t realize it that much”. The team found that results were consistent in a particular field but when they moved to a different field, they could not use the same process. The operating parameters and completion designs used in one field for best results were different in other fields. They were highly dependent on the geology and structure of the actual formations and you couldn’t see that from the data originally**.** Thus along with the volume of data, it is also the variety of data and how much knowledge can be gained and effectively utilized through this data also depends on the analysis of this data for which we also need to rely on the tacit knowledge of experienced employees. So in such processes, we can see how big data and KM fit with each other.

The interviewee from company C revealed that the way big data gets sold to companies through software vendors is like a wonderful black box in which the data goes in, conclusions come out and everything is ok. The general perception among average employees is that whatever is coming out is the right thing. According to him, "Organizations don't typically brand the big data initiatives with knowledge management, however big data can be linked to knowledge management very clearly around decision making. It provides the basis for decision making. You analyze the data, you give a prediction and you execute what the prediction means, then you analyze in the after action review if it went well and then you feed it back to the big data people to make sure prediction is better next time. Thus, at the end, big data is purely the specific information provided for decision making and can be vetted by the community of experts, and you can put knowledge management personal element into there to give feedback"**.** For example, an experiment was performed in the company. From the analysis of huge data collected through different sensors, a 70 % chance of multiple failures during drilling was predicted for next hour. So, a decision could be made by an individual to shut down the well or change the motor or take

a risk and continue the work. The company instead put it on a social media group with all the experts. They reviewed and, based on their expertise, decided that the chance was correct and suggested a way to execute it. Then when it happened or didn't happen, the group would discuss it and feed the conclusion back to the staff generating the reports that come out of big data. So it was an experiment using the knowledge of experts through social media platform to converge fast on the predictive events that are happening in almost real time. Similarly, the geographical environment such as the sub surface environment is important; for example, if you drill wells in Oman and in West Texas US, the sub surface is similar and thus you can do analytics and learn lessons by pulling the data from both environments. Thus knowing the results at one location, you can use that knowledge to better predict a similar operation at another location. These examples of company C provide a practical insight on the relationship of knowledge management and big data. The interviewee further stated that in the case of drilling, they can use big data to execute a well faster. If they execute a well faster, instead of drilling a well in 10 days, they drill a well in 5 days, that means saving the cost of rig for 5 days period. The cost of a rig can be up to hundreds of thousands of dollars each day. Thus if well is drilled faster, you go into production faster, you can get profits and returns on investment much faster. So, everything is about optimization, how fast you drill the well, and in the right location, and this knowledge is gained through the use of analytics combined with the experience of employees who can judge and make better decisions out of the analysis. He added “A major concern is that people get excited about analytics and they think that analytics are the solution to everything but analytics cannot do all the work”. Thus experts' insights are also needed along with the predicted outcomes to avoid any failures. Analytics can fail on prediction. Understanding why or why not predictions are correct, is linked with the organizational cultural and how people understand this value generation from big data.

In company D, the representative revealed that they have high volumes of structured data in their databases. These are not really one big data system but a cluster of databases. The company conducted big data experiments with IBM Watson and results revealed a link between big data and the knowledge of employees. The interviewee explained the linkage through example of monitoring machines; for example, in the case a compressor is not working, they can collect the data from all other 50-60 compressors around the world through monitoring systems and

compare all the parameters and performance indicators for good performance. Through this, the people having trouble can sort out the problem by using machines exactly in the same way to get maximum operating efficiency. About a year ago, they didn't have that much data about compressors to analyze and make decisions. Now, they can shut compressors down for less time and they learned how quickly to start them again after maintenance and how to get maximum performance from them. People who work within specific areas, for example an engineer who works with compressors, through his experience and expertise, can build on how data is linked with compressors. Another area of improvement was the development of a catalyst for gas liquefaction. A catalyst is a substance that speeds up a chemical reaction and you need to look at various parameters for developing the catalyst. The company uses a combination of parameters to develop various catalysts and knowledge gained from this mass experimentation through big data, company can find out which catalysts are more efficient. Thus this has speed up the process and now a catalyst is developed in 13 months, which took years before the usage of big data. The company uses this catalyst in a physiotropic chemical process to speed up the process for gas liquefaction. The interviewee stated that “point here is that data scientists can't work alone; they don't understand the nature of work and thus the relevant persons with subject matter expertise knowledge help in getting the true value out of these experiments”. However, there is still a lot of scope and company needs to work more in this area and also needs to hire more data scientists.

Interviewee from company E mentioned that big data is a key thing for every industry and especially in the upstream oil and gas sector, although it is not being used to its full potential. He added that people in his company are well aware of big data. A clear example which he provided regarding the linkage of big data and knowledge management was reservoir facilities management. By analyzing the data acquired through the larger number of instruments installed in the reservoirs, the company produces valuable knowledge about the capacity and weakness of reservoir, how to get optimal production from reservoir by continually adjusting the operating settings and what are the reasons for variance between production plan and actual production going on and as a result of managing this knowledge, the company devises more efficient working patterns for the reservoirs. All this became possible through the valuable knowledge generation from data which is being produced continuously in large amounts due to advanced technology that has helped in permanently monitoring the reservoirs, which is crucial for

efficient production. However, there is issue of integration, the data for all the assets should be defined, organized and stored in proper way as it is vital for understanding the asset. In oil and gas, a lot of historical data is generated related to sub-surface properties, and the structure of the surface which is key source for drilling wells and thus if this data is properly stored and classified, it can be tracked easily when someone wants to use it.

The interviewee from company F stated “this is the right problem to tackle”. All the rich oil fields in Middle East have deep access to big data but organizations don’t know how to use it exactly. The other reason is cyber security as they are scared to open up access because of cyber attacks. Organizations have always been using data in highly digitized oil fields on reservoir modeling and sub surface analysis but through rise of big data, they are trying to further enhance it; for example, big data analytics are applied on different CAD diagrams of the fields to have meta data of drawings and see what parts of the drawings are affected by health, safety and environmental (HSE) concerns. The interviewee stated that companies don’t really understand the linkage between big data and KM in Middle East. Companies are struggling with big data in a way for it to be meaningful and are trying to figure out how to use big data and thus it will take some time to understand its relationship with KM. He further mentioned that because of large number of lays offs of experienced employees due to economic downturn and low oil prices, companies need to find some alternatives such as big data to maintain the competitiveness through effective knowledge utilization.

The interviewee from company G revealed that there is an indirect relationship between big data and knowledge management as it is an enabler to produce structured knowledge material that will be helpful for the organization. There is a lot of data; however, it is the component or valuable knowledge extracted from it that can be used for business improvement. The idea of utilizing company’s big data in a structured way is still not clearly adopted by the oil and gas companies in the Middle East; however, companies do realize that there is great value in aggregating their big data, including for example, reservoir data from new/old fields and to analyze this data for future enhancement of their core operations.

The representative from company H stated that “it is all about decision making”. Companies won't make a decision until they are satisfied and this doesn't come out from unstructured information but from structured information. "Company won't make decisions based on poor

data. They want authoritative, confirmative data that's probably trended over long period of time so that they are comfortable in decision making or process improvement, thus the tacit knowledge of decision making comes from explicit knowledge. From big data, we acquire tacit knowledge and make decisions based on this analysis. We do not use the term knowledge management much in company anymore." An example in this case is a pump struggling to maintain its performance and it is observed that pump is not working properly even after multiple replacements and thus will affect production. From the data, we then analyze whether to redesign the pump or fit new equipment in the plant. The participant further added that "main issue is integration, we have got different applications, vendors, and thus it is hard to integrate data; for example, in our case, our metadata not aligned with document management systems and share point collaboration solution, we don't have security aligned nor the information architecture aligned, so for complex systems in oil and gas, it will be even harder". Thus integration of data from multiple sources is important to properly harness the value from big data.

Representative from Company I said that they have a large team specifically working on big data and have been successful in some projects. She is the KM lead in the company and argued that a clear connection can be seen between big data and knowledge management. She said that "One strong use of big data is to create knowledge buried or hidden in data and data analytics can certainly enhance our knowledge for example in case of hardware maintenance, we have knowledge manuals or books which tell us when to perform scheduled maintenance and how to maintain a piece of equipment, so we know this already but through big data, we can have advanced predictive knowledge if some equipment needs maintenance apart from regular schedule and if there is any issue with the equipment while it is running, and thus we can take measures accordingly to prevent failure". Thus from this predictive knowledge, big data enhances the company's knowledge management capability and it can timely work on a piece of equipment to save time, money and avoid temporary stoppage of operations.

7. Analysis

The results reveal that oil and gas sector has been working with data right from the beginning, however the big data term is now more widely used due to advancement in technology, as fields are equipped with a far greater number of tools and sensors that continuously generate large

volumes of data (Feblowitz, 2013). Now, the challenge for oil and gas companies is to effectively utilize this data for knowledge discovery and value creation. There was a strong consensus on the linkage of big data and knowledge management on the whole, however, for companies in Middle East, this linkage is unclear primarily because these companies are still struggling with knowledge management (Geisler and Wickramasinghe, 2015, Pfeffer and Sutton, 2013) and big data is one step further. Results further reveal that super major companies involved in the study clearly understand the linkage between big data and knowledge management but most of the projects are in experimental phase with few projects in the development phase. Based on the responses, the linkage between big data and KM can be categorized into following different but related aspects. Here an attempt has been made to explain the relationship in the light of existing literature and further discussion on new discoveries which is still unexplained by current theory.

1. ***Big data, An Inimitable Heterogeneous Resource and Catalyst for Knowledge Management Capability***

Fahey and Prusak (1998) are of the view that knowledge is not data or information and it doesn't need to be considered as a stock but as a flow which is embedded in different day to day organizational processes, the same as argued by Nonaka and Takeuchi (1995) and Wiig (1999). Moreover, knowledge is a justified belief that helps an entity to take effective decisions (Nonaka and Takeuchi, 1995). The distinction of knowledge being different from data and information has been argued a lot in the literature (Alavi and Leidner, 2001, Davenport and Prusak, 1998) in which knowledge is considered at the highest level as derived from information and information in turn is derived from data. If we evaluate this on the basis of these knowledge management concepts, this connection is supported from the literature in which knowledge is everywhere in organizations in various forms (Ball, 2011). Big data is one form in which large pools of data from multiple sources can be evaluated to generate predictive knowledge (Bose, 2009). Once this useful knowledge is extracted from the pool of unstructured data through the application of analytics, it can be used for the improvement of different business activities (McAfee et al., 2012) as depicted in the results, for example, predictive maintenance of machines for case A and I are wonderful examples in which data from sensors helps in predicting the required maintenance of machines apart from regular maintenance, thus avoiding a loss of equipment

through timely decisions. Same is the situation using a large number of satellite images for tracking oil traces in water for company A. Through analytics these high resolution images can be processed faster to determine leakages and thus faster decisions can be made if there are any potential leakages, saving money and time as well as protecting the environment. Similar is the situation for cases E, B and D. Thus generation of predictive knowledge from big data can be clearly seen as supporting the goal of knowledge management in improved decision making (Bassi, 1997) and acting as a catalyst for enhancing the knowledge management capability for organizations thus supporting the knowledge based theory of the firm (Grant, 1996) which states that knowledge based resources are most difficult to imitate and complex, heterogeneous knowledge bases are determinants of superior performance which is evident from results for example, in case of company D, the development of catalyst through mass experimentation of data helped in improved organizational performance by reducing the catalyst development time from years to months. Same is the case for fracturing job in company B. The data of these companies are unique and inimitable, thus becoming a resource for valuable knowledge generation with which these companies achieved superior performance, falling in line with knowledge based theory of the firm. On the other hand, the uniqueness of big data also lies in increasing knowledge management capability of the firm through the application of this generated predictive knowledge. Thus according to Teece et al. (1997), usage of big data can be termed as reconfiguration capability of organizations to determine the capacity to absorb untapped knowledge in big data and deduce generalizable cause effect relationships with existing knowledge for improved performance. Another goal of knowledge management is to integrate and analyze the information from different perspectives for valuable decision making (Lamont, 2012). Similarly, organizations also want data to be consistent and in an integrated form because then it is easier to extract knowledge from it (LaValle et al., 2013). Analysis of big data from multiple sources for decision making is clearly in line with this goal as per example in Case C about comparing the data from different geographical locations for the drilling operation to find out the similarities and differences and based on that making an informed decision. Same was the case with company D in which the compressor data across different geographical locations were integrated and compared to determine the most appropriate parameters for obtaining maximum efficiency.

An effective KM system discards unneeded information, makes the right information available at right time to those who need it and focuses on people who can give value and act upon the valuable information (Cowley-Durst, 1999). Thus predictive knowledge is generated through big data but people support, especially the knowledge of experienced people is crucial to properly apply this knowledge which is discussed further.

2. *Combination of Tacit Knowledge (Existing Concept) and Predicative Knowledge (New Concept)*

Another aspect of big data and knowledge management linkage as revealed from results is the combination of tacit knowledge of experienced employees (Ball, 2011) with the knowledge obtained from big data (Feblowitz, 2013). In big data analytics, one step is automated learning through "machine learning" algorithms to identify non-obvious, hidden patterns of information in data that have potential of creating knowledge; then the second step is to test and confirm these relationships and this involves "human learning" and human insights to understand, revise and confirm or reject these relationships (Hair Jr, 2007). Thus, there is a difference here from the traditional concept that knowledge is created and held by individuals (Grant,1996) as here, in the first step, knowledge is created from heterogeneous and voluminous data using analytics. Then later on, the insights of experienced employees is most of the time necessary to better understand and decide on the results obtained through big data, as was in the case of maintenance of compressors in company D and fracturing job in company B. Case C is also a great example in which it was revealed that the data were obtained from sensors and then predictive knowledge was obtained from it which advises the company to shut the machine down but then a further step is taken to get this information vetted by a community of experts. The final decision is made using the knowledge of these experts. This knowledge of experienced people is tacit knowledge (Nonaka and Takeuchi, 1995) which is one of the most important concepts in knowledge management. Thus this existing concept of tacit knowledge is combined with new concept of predictive knowledge using analytics. The results also reveal that predictive knowledge cannot always be correct and it should be further examined by experts to reach a final decision. Also like traditional knowledge management, a proper organizational culture (Davenport et al., 1998) needs to be cultivated to understand the underlying concepts of big data. The machine learning algorithms are becoming more and more robust and thus the main aim is to make decisions

without human intervention, but in case of companies C, I and H, the emphasizes have been put on the decisions made through subject matter experts and not blindly following the predictive knowledge. On the whole, learning still is largely a human function but this learning can continuously be embedded in the big data algorithms to make the predictive knowledge more and more accurate. This can be explained further through traditional SECI model for knowledge creation and learning through big data analytics using case C. First, predictive (explicit) knowledge was generated through data in which there was 70% chance of failure in the drilling operation, then this predictive knowledge was analyzed and transformed into tacit knowledge gained by employees during decision making process to decide on what to do. Then, after the decision was implemented, the obtained results were then further discussed by the experts thus going through the socialization cycle. In the after action review process, i.e. what went right and how to perform the task next time, codification of the obtained knowledge (externalization) took place to be used in future. This codified knowledge was integrated with the existing predictive knowledge obtained through analysis (combination). This stored knowledge or the explicit knowledge sources, are then further used and learned by data scientists/experts to refine or modify their existing tacit knowledge (internalization). Thus a whole cycle of socialization, externalization, combination and internalization is completed illustrating the interrelationship revealed through the results. Further, in the light of this discussion, table 2 shows the steps involved in big data based knowledge management and Fig. 4 explains the relationship of big

Big Data Based Knowledge Management
1. Processing of data to extract useful knowledge. High analytical skills required for knowledge extraction. Knowledge extraction is performed in real time but may be codified later.
2. Principally knowledge creation, knowledge discovery and reasoning with knowledge which later can transform to Nonaka's SECI model.
3. Machine focused initially and people focused on later stages. In some cases, might be purely machine focused not requiring any human intervention depending on the maturity of the system.

Table 2: Big Data Based Knowledge Management in the light of obtained results

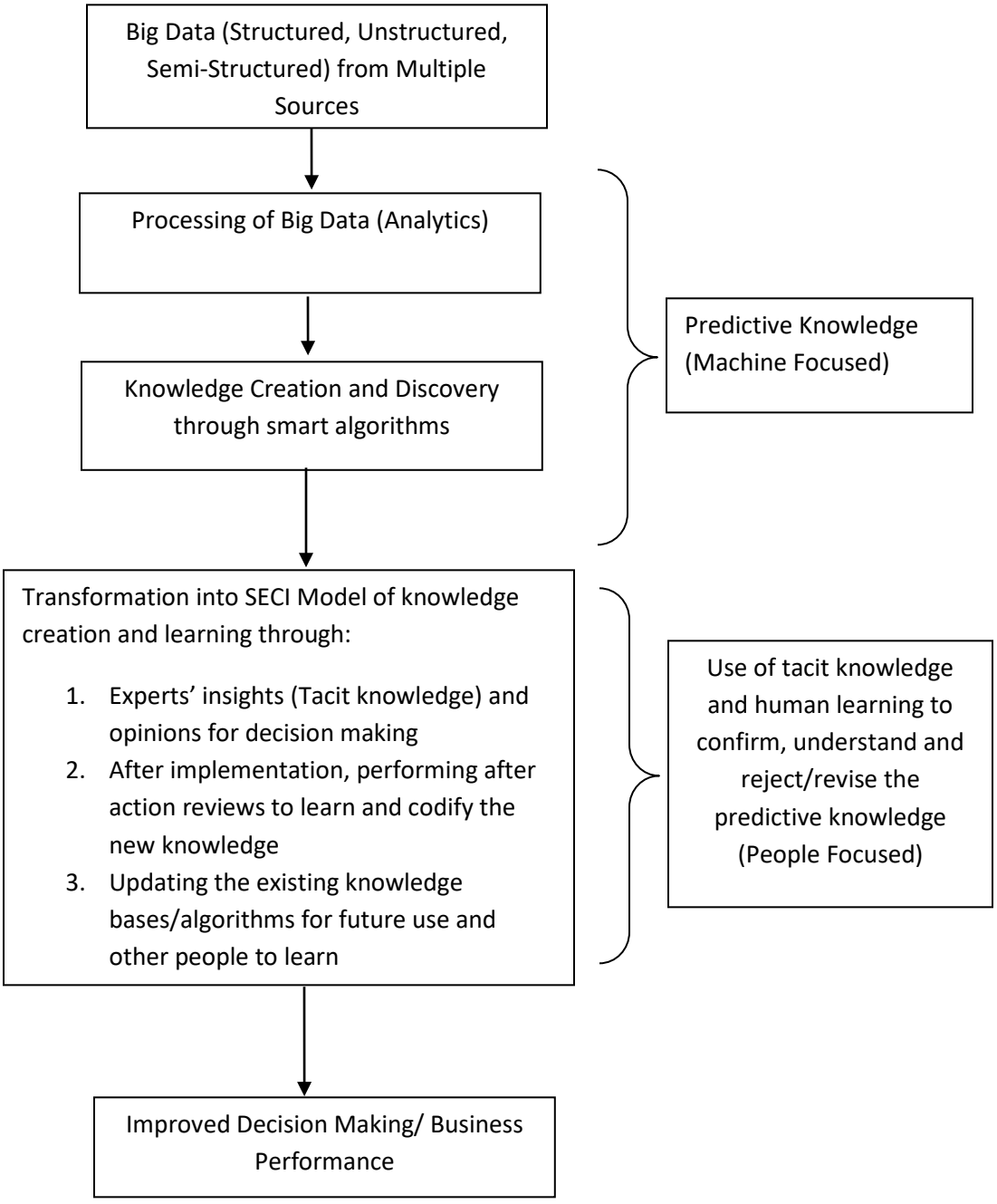

Figure 4: Framework on linkage of Big Data to Knowledge Management

data and knowledge management. Another important point worth discussing, as mentioned by the interviewee from company F, is that if using big data, knowledge can be discovered and applied automatically without human intervention, the knowledge gaps created by layoffs and

retirements can be filled. This challenges the classical concept of tacit knowledge and human capital which needs to be preserved for organizations to flourish. With the current results, it is not supported significantly as human insights and experience seem to play an important role in final decision making as discussed above. However, it could be a possibility in future as organizations become more advanced in big data activities. Due to large number of layoffs as a result of low oil prices along with the aging workforce issue, big data and analytics can provide access to searchable institutional knowledge that can compensate for limited expert staffing (Feblowitz, 2013).

Some major challenges that seem to be worth mentioning for organizations to consider include the applicability of big data as it is revealed from the results that organizations are not very clear about what can be achieved from big data (Russom, 2013) as there seemed no coherency in the responses from the employees. Moreover, employees are not very familiar with the technologies and tools for big data. Another major challenge is integration. Organizations have different databases but main issue is integration of these databases to extract the knowledge from them (LaValle et al., 2013). Technology infrastructure doesn't seem to be a problem, however companies need to hire more data specialists (Davenport and Patil, 2012). Another aspect that has been observed is that companies from developed countries have much stronger focus on big data and can relate it to knowledge management in a more convincing approach as compared to companies from developing countries, apart from the exceptional case of Nigeria where the company is sub section of a super major. However, this argument can be further strengthened by conducting more interviews in future, and including more companies from developing countries.

8. Conclusions

This study attempts to bridge the gap on the interrelationship of big data and knowledge management by conducting an exploratory study in oil and gas sector. Big data is well regarded and well supported in oil and gas sector and is more about the ability to link information from different systems and be able to analyze it in a way to enhance organizational performance. Organizations, especially super majors, are at an experimental phase and have achieved some success in this regard whereas on the whole, little work is being done in field of big data. More dedicated efforts are required to achieve a high level performance from big data and it will definitely take some time to reach this goal. There is a strong linkage between big data and

knowledge management, as apparent from both the literature and evidence provided by the interviewees. Knowledge discovery with reasoning is an intriguing factor for improved business performance. Moreover, the tacit knowledge of experienced employees holds importance when making decisions using big data. Actionable knowledge is generated through combination of tacit knowledge of relevant staff with explicit knowledge obtained from big data. However, there is a hint that in future, big data might eliminate the need of tacit knowledge when the algorithms become more advanced. This tension between the two domains i.e. tacit knowledge replacement with predictive knowledge from big data is worth investigating as not much clear evidence is obtained from results and thus this needs to be further explored as companies become more mature in usage of big data. Integration is main challenge for big data and thus it needs a lot of efforts in terms of which tools to use and how to integrate data from various clusters. Data scientists are crucial for properly analyzing the data and there is a need for well-defined planning and recruitment policies in this regard**.** Culture is in general supportive but organizations need to understand and inquire why something happened or why something didn't happen through predictive knowledge from big data. It will eventually help companies and employees to establish a well-defined and clear connection between big data and knowledge management.

There are also limitations to this study; first the sample size is small, a bigger sample size with more diverse range of organizations might reveal more insights. This study does not compare the upstream, midstream and downstream sectors of the oil and gas to explore the big data and the KM connections in terms of operations at each of these sectors. Further, the respondents in the study might not have the same expertise in both big data and knowledge management and thus results could be slightly biased. Finally, the current article only focuses on the oil and gas sector which has been targeted because the operations are highly data dependent. Similarly, other data intensive sectors such as finance and retail can also be explored.

References

ACKOFF, R. L. 1989. From data to wisdom. *Journal of applied systems analysis,* 16**,** 3-9.

ALAVI, M. & LEIDNER, D. E. 2001. Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS quarterly***,** 107-136.

BALL, K. 2011. *Surviving the Baby Boomer Exodus: Capturing Knowledge for Gen X & Y Employees*, Cengage Learning.

BASSI, L. J. 1997. Harnessing the Power of Intellectual Capital. *Training and development,* 51**,** 25-30.
BOSE, R. 2009. Advanced analytics: opportunities and challenges. *Industrial Management & Data Systems,* 109**,** 155-172.
BRUCE, B. & BERG, M. 2001. Qualitative research methods for the social sciences. Needham Heights, MA: Allyn and Bacon.
CHAN, J. O. 2014. Big Data Customer Knowledge Management,. *Communications of the IIMA: Vol. 14: Iss. 3, Article 5.*
CHEN, M., MAO, S. & LIU, Y. 2014. Big data: A survey. *Mobile Networks and Applications,* 19**,** 171-209.
CHOI, B. & LEE, H. 2002. Knowledge management strategy and its link to knowledge creation process. *Expert Systems with applications,* 23**,** 173-187.
COWLEY-DURST, B. 1999. Gathering knowledge for your knowledge management system. *Performance Improvement,* 38**,** 23-27.
DAGHFOUS, A., BELKHODJA, O. & C. ANGELL, L. 2013. Understanding and managing knowledge loss. *Journal of Knowledge Management,* 17**,** 639-660.
DAVENPORT, T. H. 2014. How strategists use "big data" to support internal business decisions, discovery and production. *Strategy & Leadership,* 42**,** 45-50.
DAVENPORT, T. H., BARTH, P. & BEAN, R. 2013. How 'big data'is different. *MIT Sloan Management Review,* 54.
DAVENPORT, T. H., DE LONG, D. W. & BEERS, M. C. 1998. Successful knowledge management projects. *Sloan management review,* 39**,** 43-57.
DAVENPORT, T. H. & PATIL, D. 2012. Data scientist. *Harvard business review,* 90**,** 70-76.
DAVENPORT, T. H. & PRUSAK, L. 1998. *Working knowledge: How organizations manage what they know*, Harvard Business Press.
DUFFY, J. 2000. Something funny is happening on the way to knowledge management. *Information Management,* 34**,** 64.
ERICKSON, S. & ROTHBERG, H. 2015. Big Data and Knowledge Management: Establishing a Conceptual Foundation. *Leading Issues in Knowledge Management, Volume Two,* 2**,** 204.
FAHEY, L. & PRUSAK, L. 1998. The eleven deadliest sins of knowledge management. *California management review,* 40**,** 265.
FARRIS, A. 2012. How big data is changing the oil & gas industry. *Analytics Mag., Dec*.
FEBLOWITZ, J. Analytics in oil and gas: The big deal about big data. SPE Digital Energy Conference, 2013. Society of Petroleum Engineers.
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**,** 198-209.
GEISLER, E. & WICKRAMASINGHE, N. 2015. *Principles of knowledge management: Theory, practice, and cases*, Routledge.
GILL, P., STEWART, K., TREASURE, E. & CHADWICK, B. 2008. Methods of data collection in qualitative research: interviews and focus groups. *British dental journal,* 204**,** 291-295.
GORE, C. & GORE, E. 1999. Knowledge management: the way forward. *Total quality management,* 10**,** 554-560.
GRANT, R. M. 1996. Toward a knowledge-based theory of the firm. *Strategic management journal,* 17**,** 109-122.

HAIR JR, J. F. 2007. Knowledge creation in marketing: the role of predictive analytics. *European Business Review,* 19**,** 303-315.

HALL, R. 1992. The strategic analysis of intangible resources. *Strategic management journal,* 13**,** 135-144.

INKPEN, A. C. & MOFFETT, M. H. 2011. *The Global Oil & Gas Industry: Management, Strategy & Finance*, PennWell Books.

LAMONT, J. 2012. Big data has big implications for knowledge management. *KM World,* 21**,** 8-11.

LAVALLE, S., LESSER, E., SHOCKLEY, R., HOPKINS, M. S. & KRUSCHWITZ, N. 2013. Big data, analytics and the path from insights to value. *MIT sloan management review,* 21.

MANYIKA, J., CHUI, M., BROWN, B., BUGHIN, J., DOBBS, R., ROXBURGH, C. & BYERS, A. H. 2011. Big data: The next frontier for innovation, competition, and productivity.

MARSHALL, C. & ROSSMAN, G. B. 2011. *Designing qualitative research*, Sage.

MCAFEE, A., BRYNJOLFSSON, E., DAVENPORT, T. H., PATIL, D. & BARTON, D. 2012. Big data. *The management revolution. Harvard Bus Rev,* 90**,** 61-67.

MURDOCH, T. B. & DETSKY, A. S. 2013. The inevitable application of big data to health care. *Jama,* 309**,** 1351-1352.

NONAKA, I. & TAKEUCHI, H. 1995. *The knowledge-creating company: How Japanese companies create the dynamics of innovation*, Oxford university press.

PFEFFER, J. & SUTTON, R. I. 2013. *The knowing-doing gap: How smart companies turn knowledge into action*, Harvard Business Press.

RUSSOM, P. 2011. Big data analytics. *TDWI Best Practices Report, Fourth Quarter*.

RUSSOM, P. 2013. Managing big data. *TDWI Research. TDWI Best Practices Report*.

SAGIROGLU, S. & SINANC, D. Big data: A review. Collaboration Technologies and Systems (CTS), 2013 International Conference on, 2013. IEEE, 42-47.

SCHROECK, M., SHOCKLEY, R., SMART, J., ROMERO-MORALES, D. & TUFANO, P. 2012. Analytics: The real-world use of big data: How innovative enterprises extract value from uncertain data. *IBM Institute for Business Value*.

TEECE, D. J., PISANO, G. & SHUEN, A. 1997. Dynamic capabilities and strategic management. *Strategic management journal,* 18**,** 509-533.

VANCE, A. 2011. The data knows. *Bloomberg Businessweek***,** 70-74.

WALLER, M. A. & FAWCETT, S. E. 2013. Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. *Journal of Business Logistics,* 34**,** 77-84.

WIIG, K. M. 1999. What future knowledge management users may expect. *Journal of knowledge management,* 3**,** 155-166.

ZACK, M. H. 1999. Managing codified knowledge. *Sloan management review,* 40**,** 45-58.