DECISION SCIENCES INSTITUTE

Managerial Perceptions of Big Data: A Dynamic Organizational Capability

Canchu Lin
Carroll University
Email: clin@carrollu.edu

Anand Kunnathur
University of Toledo
Email: anand.kunnathur@utoledo.edu

ABSTRACT

This qualitative research project seeks to gain an initial understanding of managerial perceptions of Big Data. Top level managers from large companies were interviewed about their understandings of Big Data and Big Data applications in their companies. Interview results were then thematically analyzed. Four major themes about Big Data emerged and then were further analyzed. The analysis of the four major themes suggests that Big Data constitutes a dynamic capability to companies. Thus, this study develops the concept of Big Data capability.

KEYWORDS: Big Data, Thematic analysis, Big Data capability, Dynamic capabilities, Interview

INTRODUCTION

In recent years, Big Data has represented a high point in the advancement of information technology (IT). Its noted power has attracted the attention of the popular press. *New York Times* published an article by David Carr on February 24, 2013 about how Netflix used Big Data to develop a new TV series named *House of Cards,* which became the most streamed piece of content in the United States and 40 other countries. Similarly, the business press has captured the importance of Big Data. A survey by *Accenture* showed that 89% of the surveyed senior officers from companies in nineteen countries across seven industries indicated that Big Data is very or extremely important to their business digital transformation, and 83% of them agreed that Big Data generates value for their companies (Business Wire, 2014). Likewise, *Harvard Business Review*, starting from 2012, has carried a significant number of articles addressing the role of Big Data in opening up business opportunities for organizations. These articles provided exciting examples about how Big Data helped to boost business performance (see, Barton & Court, 2012; McAfee & Brynjolfsson, 2012). Further, the business research community certainly did not fall behind in reporting the amazing power of Big Data. For example, Waller and Fawcett (2013) showed that Walmart, by using Big Data, outsmarted disaster relief experts in predicting most needed things in areas hit by a hurricane and thereafter improved their demand forecasting and inventory management.

While these examples illustrate some specific business applications of Big Data, even a cursory review of practitioner literature suggests that Big Data can play big roles in supply chain management (SCM): demand forecasting (McAfee & Brynjolfsson, 2012), inventory management (Waller & Fawcett, 2013), production or service scheduling (Barton & Court, 2012), and even new product development (Carr, 2013). Yet, these roles of Big Data are yet to be investigated in business research. So far, the burgeoning academic research on Big Data

has mainly focused on the technical issues, which were explored predominantly by computer science, and information systems researchers (for a review, see Chen, Chiang, & Storey, 2012). Although recent business research has shown strong interest in this topic, Big Data is still treated, largely, as a unidimensional construct. For example, Chen, Preston, and Swink (2015) investigated organizational usage of Big Data analytics, which is just one aspect of Big Data.

However, the above-mentioned discussions by practitioners and academics on Big Data point to some complexity of Big Data. In some conversations, especially among practitioners, Big Data is understood as a huge amount of data. Big Data has also been conceptualized as a technology, an advanced data-processing IT (e.g., see, Esteves & Curto, 2013). Still, it is treated as data analytics (Chen et al., 2015). Moreover, in their discussion of Big Data scientists, Davenport and Patil (2012) described them as a new type of professionals competent in discovering golden opportunities for business in the mines of big raw data. Clearly these discussions reveal some complexity and multi-dimensionality of Big Data. Given the scarcity of research on Big Data, before we move to investigate Big Data's roles in enhancing business functions, it is important for us to gain a deep understanding of Big Data. Thus, an initial step in investigating Big Data is to explore and capture its complexity and multi-dimensionality. An effective way of gaining such an insight into Big Data is to examine how Big Data is perceived, understood, and practiced in organizations.

The purpose of this study is to explore this complexity and multi-dimensionality of Big Data and its business implications. To accomplish this research objective, we will take an inductive approach in this study. More specifically, we will interview business executives, for their opinions, with an open question (what is Big Data?), aiming to capture perceptions, understandings, and practices accumulating and exploiting Big Data in business organizations through extended conversations with these executives. This naturalistic inquiry will yield initial knowledge about practitioners' understandings, and more importantly, Big Data practices. Then we will attempt to associate their understandings and practices with regard to business theory. Thus, the goal of this study is to theorize Big Data practices in organizations.

The rest of the paper is structured as follows. First, we detail the methodology for gathering interview data, and present a thematic analysis of our interview data. Next, we discuss the thematic analysis results with regard to business theory. Then we review the theory and relevant literature to develop the construct of Big Data capability.

LITERATURE REVIEW

According to the dynamic capabilities theory (Teece, Pisano, & Shuen, 1997), a firm's competitive advantage is the result of its strategic efficiency, which is achieved because of its capabilities and assets and some isolating mechanisms as key drivers of excellent performance. A firm develops and leverages unique resources, which help them to gain competitive advantage over its rivals (Barney, 1991). Although the resource-based view emphasized the importance of resources and thus identified key resources that a firm relies on in achieving its competitive advantage, it provided inadequate and insufficient explanation for why that happens (Teece et al., 1997). In order to show why a firm gains a competitive advantage, Teece et al. (1997) argued that a firm needs not just resources but have to develop important capabilities. Thus, to extend the resource-based view, they developed the dynamic capabilities theory. Based on this theory, besides acquiring firm-specific resources, firms must develop organizational routines that help to integrate firm-specific assets so as to enable employees to perform distinctive activities, and core competencies that sustain their core businesses. Further,

firms must have the capability to integrate, build, and reconfigure their internal and external competencies.

More importantly, firms' capabilities must be dynamic in nature (Teece et al., 1997). Technology, competitors, regulatory events, significant changes in economic and political conditions, and other factors bring about both evolutionary and revolutionary changes (Tushman & O'Reilly, 1996). Firrms often find themselves in a turbulent business environment, which then poses a great challenge to their competitive advantage. To keep their competitive advantage sustainable, firms have to develop an ability to manage evolutionary changes and make incremental improvements in their products, and to benefit from revolutionary changes to explore new business opportunities such as developing new products (Tushman & O'Reilly, 1996). Product improvement and new product development are processes of intensive knowledge management (Cohen & Levinthal, 1990; Smith, Collins, & Clark, 2005), which deals with acquiring, creating, and sharing knowledge (Nonaka et al., 1996). Knowledge management is enabled and facilitated by IT (Grant, 1996). Product innovation and improvement is a process of knowledge creation and management (Tushman & O'Reilly, 1996). But knowledge creation and management is a process of collecting and analyzing data and refining data into information, which then is then accumulated and integrated into knowledge (Nonaka, 1994).

Further, coping with the turbulent business environment mandates that the firm take quick action to renew and reorganize its resources and competences and develop a capability to adapt, integrate, and reconfigure internal and external skills, resources, and competences to address its unique needs mandated by the changing environment, a process in which the firm develops dynamic capabilities (Teece et al., 1997). The dynamic capabilities theory suggests that firms must continuously develop new capabilities to address changes in the business environment, as any capability has a lifecycle, a progression of growth, maturity, and decline (Helfat & Peteraf, 2003). However, developing and maintaining such dynamic capabilities is inseparable from understanding the turbulent business environment thoroughly. Developing dynamic capabilities depends on knowledge of competition and changes in the environment and their causes. Such knowledge rests on a mastery of skills to collect and analyze information about competitors and market trends with the help of IT. In this sense, firms have to develop a special dynamic capability, an IT-enabled competence and ability to closely monitor and analyze the changing environment with respect to their survival and growth.

Although IT has empowered firms to collect, store, retrieve, and analyze information, transforming such practical information into enlightening knowledge requires a capability that is beyond implementing and using IT. Indeed, the use of IT has generated big challenges in managing data, the raw element of information (Alavi & Leidner, 2001; Devlin, 1999; Nonaka, 1994). Because of its wide use and advancement, IT now generates large volumes of various types of data with a quick speed for firms, now commonly known as Big Data (McAfee & Brynjolfsson, 2012). Big Data are usually scattered in different organizational information systems, unconnected in structure (Daveport & Patil, 2012). Transforming data into information, and then turning information into knowledge requires an integration of these scattered Big Data in the first place. Additionally, knowledge creation entails the ability to further identify rich data sources, integrate them with current data sources, analyze them effectively, and communicate the analysis results with their business implications to decision makers (Davenport & Patil, 2012). To address all these challenges and others emerging in the process of dealing with Big Data, firms have to develop a dynamic organizational capability.

METHODS

To gain some insightful understanding of how Big Data is perceived and practiced in organizations, we conducted qualitative interviews. We selected seven large companies for this study. They represent industries of insurance, telecommunications, petroleum, new energy, information technology, packaging (Fortune 500), and finance (Fortune 100). We first contacted top management (CEO, President, or Vice President) of these companies with an intent of investigating their Big Data operations. Then they sent us contact information of top level managers responsible for the oversight of these operations. We then made appointments to interview these top level managers. The interview questions are open-ended, with a tunnel approach (from general to specific). The questions ask about general definitions and perceptions of Big Data, purposes of using Big Data for, issues, concerns, and challenges associated with Big Data, and resources for Big Data, as well as corporate strategies. The interviews were recorded, each lasting more than an hour. Notes were also taken to capture nonverbal cues accompanying interviewees' talks. An undergraduate research assistant who is a native speaker of American English transcribed the interviews. The interviews were then thematically analyzed using Owen's (1984) method.

Thematic analysis has been extensively used in business research (for recent examples, e.g., see, Hu et al., 2015; Rishi, Jauhair, & Joshi, 2015; and Tate, Ellram, & Kirchoff, 2010). As a qualitative research method, thematic analysis has been elaborated in great detail in Boyatzis (1998). But for this study, we mainly employed Owen's approach, a method of naturalistic inquiry, which was originally developed in his relationship research. Although his research was in a different field, Owen's method of exploring, developing, and categorizing family themes is transferrable to our current research on Big Data. Owen (1984) argued that thematic analysis is particularly useful in identifying major themes that emerge in the discourses of the studied that. reflect their understandings, perceptions, definitions, attitudes, and even feelings. This method particularly suits well our purpose of capturing firms' understandings, definitions, and practices of Big Data. Owen's approach identifies a theme based on three criteria: recurrence, repetition, and forcefulness. Recurrence means that a same meaning which may be expressed in different words occurs multiple times in the discourses. In our context, a certain meaning will count as a theme of Big Data if it is identified in responses of multiple interviewees. Repetition is an extension of the first criterion in that words indicating a theme are repeated multiple times throughout the responses of the interviewees. In our context, this criterion is observed by identifying same words used multiple times to refer to or comment on Big Data from the interview responses. The third criterion, forcefulness, refers to the emphasis or stress on certain words that constitute a major meaning through vocal inflection, volume, or pauses, or use of color or bold words. This criterion is observed by identifying places in their responses where interviewees appeared to emphasize certain words about Big Data. Our study reports an interrater reliability of .86 for the Big Data related themes.

RESULTS

Following Owen's method and procedure, we thematically analyzed all the interview data. In our analysis, four major themes have emerged: unprecedentedness, hi-techness, transformativenss, and strategicness. These four themes are discussed in the following paragraphs.

Unprecedentedness refers to the changed nature of raw data. It is shown in several ways. First, unprecedentedness is reflected in the scale that huge amount or volume of data is created and stored. All the interviewees commented on the huge volumes of data that are now readily

available to their companies regardless of where they are generated. One company representative commented that Big Data is “about the proliferation of recording data, saving data, … there are more and more devices that are generating data.” These instantly generated data are accumulated to a level that is highest in history. Second, unprecedentedness is revealed in the variety of data and number of ways that data is generated. Data is now multi-dimensional; it is visual, audio, as well as textual. Further, it is both structured and unstructured. This is because an increased number of automation devices are now available for generating data, such as “phones, websites, and log files,” according to one interviewee, and even “process equipment and transportation equipment,” based on another interviewee. Third, unprecedentedness is also shown in the speed that data is generated. For example, according to a high-level manager of an energy company, they now “bring in 700,000 different process variables in less than a second frequency with which it’s coming in.”

Hi-techness, the second major theme, means that Big Data presupposes state of the art technology that includes software, hardware, and infrastructure. For software, for example, according to the interviewee representing a telecommunications company, the formerly relational database, is no longer able to store Big Data, as part of it is unstructured. Thus, new types of databases have to be created. Or new types of software have to be developed so as to handle data of diverse formats. Most of the interviewees used Hardoop as an example of this. Besides, Big Data requires updated technology equipment such as large technology repositories or storage repositories which make it possible for companies to store big amounts of data at one location and to run analytics. Further, advanced infrastructure is also part of this high-techness theme as well. For example, when companies do not have certain types of data, which they really want, they have to have networking infrastructure such as the internet of things to help to connect them to the sources of their wanted data, according to a representative of a technology company.

The major theme of transformativeness refers to the process of effectively and efficiently making sense of Big Data. Such a process is deemed as being more important than having merely huge volume of data. As a new energy company representative put it, there should be steps, devices, mechanisms, tools, and algorithms developed to manipulate Big Data, more specifically, “to query, aggregate, analyze, and display that data.” Transformativeness also means that these so-called steps, devices, mechanisms, tools, and algorithms must be new and revolutionary enough for companies to process Big Data and make them meaningful in a business sense. For example, the traditional data warehousing concept no longer applies in the context of Big Data. What is needed instead is a new system that enables real time analysis of Big Data. Relating to the theme of unprecedentedness, the changed nature of data requires skills, knowledge, or more specifically algorithms for analyzing Big Data. In the new energy company, for example, a new trend has emerged that IT professionals “spend less time in creating reports but more time in positing data and end users create data analysis reports.” More importantly, the transformative power of Big Data is made possible mostly by highly advanced soft skills, which, some company representatives called analytics. For example, as a petroleum company representative put it, “when I hear Big Data, we usually say Big Data analytics. We call it advanced analytics.”

Lastly, the theme of strategicness refers to the decision making component of Big Data. In this sense, Big Data serves to accomplish strategic goals for companies. Big Data undergirds strategic decision making regarding both external and internal issues. External oriented strategic decision making focuses on monitoring and thereafter responding to the market. This includes using Big Data to capture market trend and develop new products/services. According

to the telecommunications company representative, they analyze Big Data mainly to see where the market is going. “Let’s say that a big thing right now is that people are dropping cable television and viewing more online. We see that in our usage statistics.” Thus, they started “making new products with higher download speeds, higher usage allowances, because people are streaming more and more. They need high bandwidth products.” As far as internal strategic decision making is concerned, Big Data is used to augment and strengthen operations of production as well as human resources management. For example, for the telecommunications company, Big Data is used to safeguard their operational network, “keeping it to be as reliable as possible.” More specifically, Big Data “allows us to monitor files, log files, and to do predictive modeling about potential failures, as well as scheduling preventive maintenance.” For the petroleum company, for example, they use Big Data to predict the likelihood of employees leaving the company.

To sum up, the thematic analysis of company executives’ discourses about Big Data shows that Big Data is a meaning-laden business concept. The four major themes suggest that Big Data encompasses resources, knowledge and skills, as well as strategic thinking in business. The results of the thematic analysis represent new discoveries of Big Data practice in organizations, but they are not theoretically affirmed yet. Thus, it triggers us to connect the four major themes of Big Data to theory. Given their meanings, these four major themes are implicitly tied to the dynamic capabilities theory (Teece et al., 1997).

DISCUSSION AND CONCLUSIONS

Prior research, building on the argument of the resource-based view, already developed a concept of IS capability. According to Peppard and Ward (2004), development of IS capability is an emergent process: initially, in their interactions with IT, individual employees acquire skills, knowledge, experience, and display behaviors relating to IT, which become individual level resources; resources are further integrated, organized, and utilized through organizational processes governed by structural rules; these integrated and synergized resources are then turned into IS competencies; finally, these IS competencies, aligned with firm business strategy, and supported and enhanced by allocated IS investment, grow into IS capability at the firm level. Although it reveals the emergent process, the concept of IS capability, however, does not hold promise in capturing a firm’s capability of handling Big Data. IS capability is a generic, and broad concept, as it relates to IT in general (Peppard & Ward, 2004). Building on the concept of IS capability, we treat Big Data capability as a specific type of IS capability. Viewed from that lens, Big Data capability consists of skill, knowledge, and behavioral resources and competences, development of which depends on material resources such as IT hardware and software. The four major themes identified and analyzed in the interview data give credence to the conceptualization of Big Data as a capability.

First, the two themes of unprecedentedness and hi-techness point to the basis of the material resources of Big Data capability. While the unprecedentedness theme characterizes the nature of the material resources of raw data, the hi-techness theme indicates the technological tools used to collect, store, and analyze data, and present analysis results to a right audience. As shown in the thematic analysis, Big Data as raw data exists in multiple formats such as numerical, text, visual, and audio. Big Data is also carried in structured as well as unstructured formats (Chen, Chiang, & Storey, 2012). Further, Big Data is produced rapidly in multiple sources. Moreover, Big Data flows out in big volumes. These constitute the three basic characteristics of Big Data: volume, velocity, and variety (McAfee & Brynjolfsson, 2012), which are all unprecedented in scale. Technological tools, the essence of the hi-techness theme,

cover both hardware and software, including but not limited to IT infrastructure, computing equipment, and all software used for collection, storage, and analysis of Big Data such as Hardoop, NoDQL and SAS (Chang, Kauffman, & Kwon, 2014). Some of these material resources are possessions of a firm, while others can only be acquired externally.

Second, the transformativeness and strategicness themes refer to the non-material resources, which include individual employees' experiences, skills, knowledge of using and handling the material resources, and their behaviors such as managing information security, and policy compliance. These non-material resources can also be acquired both internally and externally. Some typical Big Data non-material resources include data analytics, text analytics, web analytics, and mobile analytics (Chen et al., 2012). In the spirit of the dynamic capabilities theory (Teece et al., 1997), and consistent with the premise inherent in IS capability (Peppard & Ward, 2004), development of Big Data capability is an emergent process. Non-material resources are then turned into Big Data competencies when they are used in organizational processes where the technical skills, experiences, and knowledge of Big Data interact with firm-based skills, knowledge, and behaviors such as communication, relationship building, and rule compliance, which are firm specific. Thus, Big Data competencies are more firm specific. Finally, when Big Data competencies are aligned with the firm's business strategy, and shaped by the firm's Big Data related investments, they grow into its Big Data capability. Thus, *Big Data capability* is defined as a firm's capability of identifying sources where large volumes of various kinds of data flow out in high speed, collecting, storing, and analyzing such Big Data for the purpose of accomplishing the firm's strategic as well as operational goals.

The development of this concept of Big Data capability has important implications for both theory and practice. From a theoretical perspective, this study helped to integrate what used to be fragmented views and definitions of Big Data emerged in the popular and applied business literatures into a well-refined and systematically developed concept. As this concept is developed based on a thematic analysis of practitioners' narratives, its theoretical rigor has a solid practice foundation. In this sense, the concept best captures practitioners' understandings and practices of Big Data in organizations. More importantly, this concept will further contribute to Big Data research in that it enables us to investigate Big Data's roles in various aspects of business performance. The development of this concept helps to link Big Data to other functions of business. From a practice point of view, this study suggests that companies have to treat Big Data at a strategic level. The concept implies that by viewing and developing Big Data as a dynamic organizational capability, firms can expect to cope with external changes in a positive way and capture business opportunities, as they will occur.

REFERENCES

Alavi, M., & Leidner, D. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly, 25 (1)*, 107-136.

Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17,* 99-120.

Barton, D., & Court, D. (2012 October). Making advanced analytics work for you: A practical guide to capitalizing on big data. *Harvard Business Review*, 79-83.

Boyatzis, R. E. (1998). *Transforming qualitative information***.** Thousand Oaks, CA: Sage.

Business Wire (Sept. 10, 2014). Companies are satisfied with business outcomes from Big Data and recognize Big Data as very important to their digital transformation, Accenture study shows; security, budget and a shortage of skilled professionals to implement and run big data are the top challenges for organizations.

Carr, D. (Feb. 24, 2013). Giving viewers what they want. *New York Times*.

Chang, R. M., Kauffman, R. J., & Kwon, Y. O. (2014). Understanding the paradigm shift to computational social science in the presence of big data. *Decision Support Systems, 63*, 67-80.

Chen, D. D., Preston, D. S., & Swink, M. (2015). How the use of Big Data analytics affects value creation in supply chain management. *Journal of Management Information Systems*, 32 (4), 4-39.

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly, 36 (4),* 1165-1188.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective in learning and innovation. *Administrative Science Quarterly*, *17*, 178-184.

Davenport, T. H., & Patil, D. J. (2012 October). Data scientist: The sexiest job of the 21st century. *Harvard Business Review,* 70-76.

Devlin, K. (1999). *Inforsense – turning information into knowledge.* New York, NY: Freeman and Co.

Esteves, J., & Curto, J. (2013). A risk and benefits behavioral model to assess intentions to adopt big data. *Journal of Intelligence Studies in Business, 3*, 37-46.
Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal, 17 (winter special issue)*, 109-122.

Helfat, C. E., & Peteraf, M. A. (2003). The dynamic resource-based view: Capability lifecycles. *Strategic Management Journal, 24,* 997-1010.

Hu, Q., Mason, R., Williams, S. J., & Found, P. (2015). Lean implementation within SMEs: A literature review. *Journal of Manufacturing Technology Management*, *26 (7)*, 980-1012.

McAfee, A., & Brynjolfsson, E. (2012 October). Big Data: The management revolution. *Harvard Business Review,* 60-68.

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science, 5*, 14–39.

Nonaka, I., Umemoto, K., & Senoo, D. (1996). From information processing to knowledge creation: a paradigm shift in business management. *Technology in Society, 18 (2)*, 203–218.

Owen, W. F. (1984). Interpretive themes in relational communication. *Quarterly Journal of Speech*, *70*, 274-287.

Peppard, J., & Ward, J. (2004). Beyond strategic information systems: Toward an IS capability. *Journal of Strategic Information Systems, 13*, 167-194.

Rishi, M., Jauhari, V., & Joshi, G. (2015). Marketing sustainability in the luxury lodging industry: A thematic analysis of preferences amongst the Indian transition generation. *Journal of Consumer Marketing*, *32 (5)*, 376-391.

Smith, K. G., Collins, C. J., & Clark, K. D. (2005). Existing knowledge, knowledge creation capability, and the rate of new product introduction in high-technology firms*. Academy of Management Journal, 48 (2),* 346-357.

Tate, W., Ellram, L. M., & Kirchoff, J. F.. (2010). Corporate social responsibility reports: A thematic analysis related to supply chain management. *Journal of Supply Chain Management*, *46 (1)*, 19-44.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal, 18 (7)*, 509-533.

Tushman, M. L., & O'Reilly III, C. A. (1996). Ambidextrous organizations: Managing evolutionary and revolutionary change. *California Management Review, 38 (4),* 8-30.

Waller, M. A., & Fawcett, S. E. (2013). Click here for a data scientist: Big data, predictive analytics, and theory development in the era of a maker movement supply chain. *Journal of Business Logistics, 34 (4)*, 249-252.