2016

Big Data and Strategy: A research Framework

Patrick Mikalef
Norwegian University of Science and Technology, patrick.mikalef@idi.ntnu.no

Pappas O. Ilias
Norwegian University of Science and Technology, pappas@idi.ntnu.no

Michail Giannakos
Norges teknisk-naturvitenskapelige universitet, michailg@idi.ntnu.no

John Krogstie
Norwegian University of Science and Technology, john.krogstie@idi.ntnu.no

George Lekakos
Athens University of Economics and Business (AUEB), glekakos@aueb.gr

BIG DATA AND STRATEGY: A RESEARCH FRAMEWORK

Completed Research

Mikalef, Patrick, Norwegian University of Science and Technology, Trondheim, Norway, patrick.mikalef@idi.ntnu.no

Pappas, O. Ilias, Norwegian University of Science and Technology, Trondheim, Norway, ilpappas@idi.ntnu.no

Giannakos, N. Michail, Norwegian University of Science and Technology, Trondheim, Norway, michailg@idi.ntnu.no

Krogstie, John, Norwegian University of Science and Technology, Trondheim, Norway, john.krogstie@idi.ntnu.no

Lekakos, George, Athens University of Economics and Business, Athens, Greece, glekakos@aueb.gr

Abstract

With big data growing rapidly in importance over the past few years', academics and practitioners have been considering the means through which they can incorporate the shifts these technologies bring into their competitive strategies. To date, there has been an emphasis on the technical aspects of big data with limited attention on the organizational changes they entail and how they should be leveraged strategically. As with any novel technology, it is important to understand the mechanisms and processes through which big data can add business value to companies and have a clear picture of the different elements and their interdependencies. To this end, the present paper aims to provide a theoretical discussion leading up to a research framework that can help explain the mechanisms through which big data lead to competitive performance gains. The research framework is grounded on past empirical work on IT-business, and builds on the resource-based view (RBV) and dynamic capabilities view (DCV) of the firm. By identifying the main areas of focus for big data and explaining the mechanisms through which they should be leveraged, this paper attempts to add to literature on how big data should be examined as a source of a competitive advantage.

Keywords: Big Data, Dynamic Capabilities, Resource-Based View, Competitive Performance, IT Strategy

1 Introduction

The notion of big data and its application in driving organizational decision making has attracted enormous attention over the past couple of years. As the label itself indicates, big data refers to large volumes of data generated and made available online and in digital media ecosystems (Constantiou & Kallinikos, 2015). Associated with the notion of big data are aspects such as the diversity of data, varying data quality, the frequency by which it is updated, and the speed at which it grows (Krogstie & Gao, 2015). Many companies realize that the data they own or can get access to, combining own and open data, and the way they use them can differentiate them from competition, and even provide them with a competitive edge. Thus, todays companies try to collect and process as much data as possible. The need to harness the potential of rapidly expanding data volume, velocity, and variety, has seen a significant evolution of techniques and technologies for data storage, analysis, and visualization. Yet, there is limited understanding of how organizations need to change to embrace these technological innovations, and the business shifts they entail (McAfee et al., 2012). As big data tools and applications spread, they will inevitably change long-standing ideas about decision making, management practices, and most importantly competitive strategy formulation (Kallinikos & Constantiou, 2015). But as with any major change, the challenge of becoming a big data-driven enterprise can be enormous. Nevertheless, it's a transition that executives need to navigate through, with limited empirical knowledge to guide their decisions.

Despite the hype surrounding big data, the aforementioned predicaments still remain largely unexplored (McAfee et al., 2012), severely hampering exploiting the business potential of big data. To date, efforts have been primarily focused on infrastructure, intelligence, and analytics tools; substantially disregarding how these technological developments should be incorporated into strategy and operations thinking. Several research commentaries stress the potential value of big data as a strategic tool (Constantiou & Kallinikos, 2015), but to date no empirical work has delved on this problem, leaving practitioners in unchartered waters regarding their big data deployments. A recent highly-cited framework sketches the emerging trends in big data tools and analytics techniques (Chen et al., 2012). Another highly influential article underscores the importance of the data scientist in the big data-driven enterprise, and emphasizes on the skills and knowledge that a successful data scientist should have (Davenport & Patil, 2012). On a more strategic perspective, Constantiou and Kallinikos (2015) outline how big data can be linked to strategic management theories, and propose a set of directions for future research.

While the abovementioned studies attempt to explore the business value and transformations that big data entail, they largely base their arguments on non-theoretical perspectives. The purpose of this paper is to introduce a theoretical research framework for examining how big data lead to competitive performance gains. By building on the theories employed in Management Information Systems (MIS) and novel contributions in the strategic management domains, this paper aims to explain the different levels and mechanisms by which big data can deliver business value.

The rest of the paper is structured as follows. In section 2 the latest literature on big data analytics and management is discussed. Section 3 provides the theoretical foundation upon which the conceptual research model is built, presented in section 4. In closing, section 5 provides a discussion on research directions and areas of future interest.

2 Literature Review

Big data have very quickly demonstrated that they are a great example of the impact of information technology on business and decision making at the enterprise level (Constantiou & Kallinikos, 2015). They transform the way that companies relate to both their customers and employees, and the way

they enact and perform business operations (Wamba et al., 2015). Through appropriate implementation of big data initiatives, companies have the potential of renewed business value creation along with increased productivity and innovativeness (Maglio & Lim). The importance of big data and business analytics is evident throughout the literature, and as they evolve they have various applications creating multiple emerging research areas (Chen et al., 2012). A heated discussion over the past years has been on the opportunities and challenges that big data bring to organizations, communities, and individuals (Constantiou & Kallinikos, 2015).

Past studies have highlighted the importance of big data and business analytics is various areas such as customer relationship management, new business models, and short-term economic predictions (Mcafee et al., 2012; Varian, 2014). The value of big data however has also been noted in terms of combining information from various sources, creating innovative services (Varian, 2014). This suggestion, of combining multiple sources of data in order to derive value, has been advocated by multiple scholars following recent paradigms presented in the business world (Davenport, 2014). Although these articles demonstrate the value of big data in improving several areas within a business, they do not address the issue of how to incorporate big data in the wider change that transcends business strategy. Therefore, it is important to develop a theoretically-driven understanding of how the overarching concept of big data can be incorporated in a firm's business strategy, and under what conditions investments in this direction can lead to a competitive advantage.

3 Theoretical Background

The growing interest in big data requires a focused discussion on how it can be examined empirically and what theories can be used to understand the critical success factors as well as the business value such technologies add. Within the information systems literature, several theoretical perspectives have been used to explore these issues, with recent attempts building on the convergence of theories from strategic management and operations research. When examining the impact of IT investments at the firm level of analysis, the Resource Based View (RBV) of the firm has been one of the most employed theoretical perspectives over the past two decades (Wade & Hulland, 2004). The main argument of the RBV is that resources that are valuable, rare, in-imitable, and non-substitutable are the building blocks of a competitive advantage (Bharadwaj, 2000). In terms of IT resources, they have been distinguished into tangible (IT infrastructure), human (IT human skills & knowledge), and intangible (culture and relationships) (Bharadwaj, 2000). This has enabled researchers and practitioners to identify the different types of IT resources their firms should aim to acquire and strengthen. Nevertheless, despite the RBV providing a basis for identifying the raw materials that are necessary to build a competitive advantage, it fails to explain how these resources are leveraged (Kraaijenbrink et al., 2010). In addition, the RBV provides little explanation as to how companies react in the face of external changes and how the dynamic of resources evolves (Kraaijenbrink et al., 2010). These shortcomings of the RBV have also been documented in the IT literature, and have caused academics to rethink the theoretical perspectives that could complement this gap.

A growing body of literature emphasizes on the role of dynamic capabilities as a source of sustained competitive advantage, especially in turbulent and uncertain environments (Teece, 2007). The dynamic capabilities view (DCV) of the firm posits that the ability to purposefully adapt an organizations resource and capabilities in the face of external pressures is the ultimate source of sustained competitive advantage (Eisenhardt & Martin, 2000). The DCV has only recently begun to attract the interest of IS scholars in terms of helping determine how an IT-based competitive advantage can be achieved in dynamic and rapidly changing environments. Nevertheless, there have been several papers that employ the theory empirically (Pavlou & El Sawy, 2010) with results indicating that thinking of IT through this lens has good explanatory power (Drnevich & Kriauciunas, 2011; Mikalef et al., 2016). The rationale developed in these studies is that IT that is embedded in specific capabilities can provide a

competitive edge. As such, the value of IT does not lie in IT resources per se, although their availability is a prerequisite, but rather, on the process of integrating them into the organizational fabric. The DCV has seen a maturing in terms of its theoretical grounding over the past few years, and has been tested empirically in a range of contexts. In this way, empirical studies have demonstrated how to conceptualize and measure the firms' capacity to do so by virtue of its IT assets. The general consensus in these empirical studies is to identify between the routines or capabilities of which dynamic capabilities comprise (Eisenhardt & Martin, 2000). While there are some differences in terms of the routines used, the underlying philosophy remains the same in most of these studies.

While the RBV and DCV build on different ideas to support how a company can achieve competitive advantage, there is a growing body in literature which identifies their complementarities (Helfat & Peteraf, 2003). Despite the DCV being more appropriate in explaining competitive advantage in turbulent and unpredictable business environments, it is noted in literature that the types of resources that a firm possesses will ultimately have an effect on the responses they can initiate. The types of resources and their influence on responsive actions have also been noted in IT literature, thus reinforcing the theoretical linkages (Wang et al., 2012).

4 Conceptual Research Framework

A vast amount of data is available through the internet and as part of daily operations, and most companies are already trying to process it in order to extract specific information that will increase their value. In order to make sense out of this vast amount of data, various tools, methods and analytical concepts are applied such as text mining and sentiment analysis. However, the amount of data that can potentially be processed is vast, leading to information overload for company managers, decision makers and executives. Decision makers also have limited understanding on how to adopt and implement big data initiatives that can help drive their business strategies. To this end, a theoretical framework is put forth that aims to delineate the core areas that should be considered when adopting big data initiatives with a strategic perspective.

The theoretical discussion of the previous section provides a basis for understanding how to approach big data projects for business purposes and develop a research framework for future reference. The proposed research framework can therefore be used to avoid common pitfalls that have been documented in past IT research. This issue is particularly prevalent when novel solutions are promoted, in which case it is important to understand the boundaries of their business value and under what circumstances they can result in competitive performance gains. To do so, we demonstrate how the theoretical perspectives are associated, what core notions and aspects are relevant, as well as what contingency factors shape these relationships in the context of big data. Already there have been some studies that attempt to describe the theoretical basis on which big data can be examined (Constantiou & Kallinikos, 2015), yet, there is no coherent theoretical framework, or an underlying unifying framework to provide a clear view of the overall business potential and strategic value.

We start this discussion by isolating the different levels and dimensions pertinent to each theoretical perspective, identify how they are relevant to the context of big data, and finally attempt to describe how they are inter-related. The starting point is to analyse the RBV in the IT context. Over the past decade there have been numerous studies that define the levels at which the RBV can be decomposed as aforementioned in the previous section. Building on the distinction between IT infrastructure, IT human skills and knowledge, and relational IT resources, we overview some of the work that could be used to guide researchers in the big data and business analytics area. In terms of IT infrastructure and the different types that exist several papers have proposed aspects that should be considered (Bharadwaj, 2000). In the context of big data there have been some attempts to describe the necessary infrastructure which span hardware, software, and mathematical and analytical tools (Chen et al., 2012).

Employees working in this area should have analytical thinking, ability to handle large amounts of data, knowledge of analytics techniques and statistical modelling, as well as the capacity to tackle problems with a data-driven approach (Davenport & Patil, 2012). In terms of relational IT resources, in the context of big data and business analytics, it is important that companies establish a virtual presence on mediums where they can interact with involved parties, and clearly articulate communications and interactions strategies with their existing or potential customers (Grégoire et al., 2015). An additional aspect that has been noted as being a core resource when dealing with big data, pertains to the 4 V`s, volume, velocity, variety, and veracity (Erl, 2016). Volume concerns not being restricted to samples but observing all available information. Velocity has to do with getting data on real-time. Variety in big data is achieved when having access to various informational format resources (text, images, audio, log-file data etc). Finally, veracity has to do with the quality of the data and its accuracy.

Although the four afore-mentioned aspects of big data are critical in order to gain a competitive edge, they are of limited value if not leveraged appropriately. This means that the infrastructure, human skills and knowledge, relational resources, and data must be put into action and into specifically directed initiatives. To do so, a firm must have the IT competencies, i.e. the collective capacity to coordinate activities and transform and bring together individual IT resources into IT-enabled dynamic capabilities. While IT competencies have been described in previous studies (Cragg et al., 2011), they are not the main area of focus of the present paper. This does not mean that they are of lesser importance; in fact, they are particularly relevant and important in the process of transforming big data IT resources into a potentially competitive asset. Yet, there is considerable qualitative research that needs to be conducted to define them and they are highly likely to be context-specific.

While the ability to effectively orchestrate IT resources may also result in operational capabilities, the focus of our research framework will be on dynamic capabilities due to their importance in contemporary businesses. Therefore, we seek to explain the routines of which IT-enabled dynamic capabilities comprise, and how big data initiatives can be infused into them with the purpose of augmenting them. Researchers have sought to quantify the notion of dynamic capabilities by identifying distinct and measurable dimensions, or else, capabilities (Teece, 2007; Pavlou & El Sawy, 2010; Mikalef et al., 2016; Mikalef & Pateli, 2016). These capabilities include sensing, learning, coordinating, integrating, and reconfiguring (Mikalef et al., 2016). A sensing capability concerns the capacity of a firm to spot, interpret, and make sense of opportunities and threats in the business environment (Teece, 2007). Big data and business analytics can be leveraged to enhance a firms sensing capability by helping identify customer requirements, gaining feedback on existing products or services, or even monitoring competitor moves and their customers' responses (He et al., 2013; Risius & Beck, 2015). A learning capability is defined as the capacity to acquire, assimilate, transform, and exploit new knowledge that enables informed decision making (Zahra & George, 2002). While this capacity closely resembles a sensing capability, it differs in that it doesn't solely rely on spotting trends, but creates distilled information that can be used in competitive actions. A coordinating capability is defined as the ability to orchestrate and deploy tasks and resources, and synchronize activities with involved stakeholders (Pavlou & El Sawy, 2011). Through feedback iterations with customers over social media and developing meaningful analytics, firms can coordinate their efforts with various departments and at different stages of development, and come up with products or services that are tailored to their likings. In addition, by analysing in real-time log files of business processes, coordination can be enhanced and improved (Vera-Basquero et al., 2013). An integrating capability includes the capacity to evaluate external resources and competences, and embed and exploit them in new or revamped ways (Woldesenbet et al., 2012). Big data and business analytics can be employed towards strengthening this capability by gathering information from multiple sources and utilize them in combination, or else, through the process of bisociation. Finally, a reconfiguring capability is defined as the ability of a firm to effectuate strategic moves and demonstrate agility when there is a need to change existing modes of operation (Lin &

Wu, 2014). For this particular capability, big data and business analytics are particularly relevant since by generating information at a constant flow, or else nowcasting, decision makers are equipped with knowledge that allows them to respond instantaneously, thus increasing operational agility.

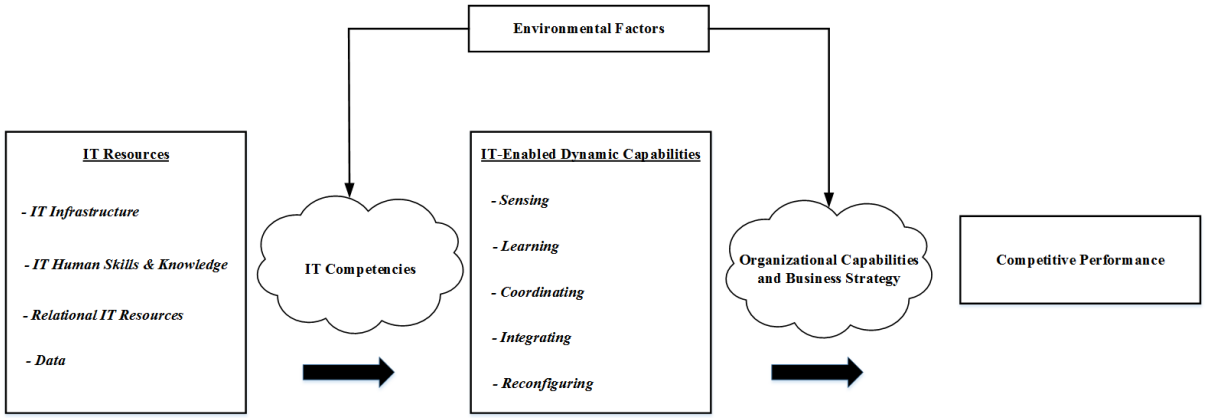

Figure 1 Conceptual Research Framework

Building on the main theoretical arguments presented and their underlying concepts and notions, the research framework presented above and the main associations can help guide future studies in determining the business value of big data and business analytics. By separating IT resources from IT-enabled dynamic capabilities, it is possible to understand the nexus of relationships through which competitive performance gains can be realized. While IT resources related to social media analytics are necessary, they are of very limited value if they are not transformed into IT-enabled dynamic capabilities through effective orchestration and management (i.e. to develop the necessary IT competencies at the group or business unit level). IT-enabled dynamic capabilities have been shown to be an important part of gaining a competitive edge, especially in turbulent and highly dynamic markets. Nevertheless, their effect on competitive performance has empirically been proven to be and indirect one, mediated by other organizational capabilities and contingent upon business strategy (Barreto, 2010). It is therefore important to examine how IT-enabled dynamic capabilities relevant to big data and business nalytics operate, in terms of changing the existing modes of operation and decision making.

5 Discussion

An ever increasing numberof companies are attempting to use big data and business analytics in order to analyse available data and aid decision making. For these companies, it is important to leverage the full potential that big data and business analytics can offer in the aim of gaining a competitive advantage. Nevertheless, since big data and business analytics are a relatively new technological and business paradigm, there is little research support on how to effectively manage them and use them in the most effective manner. Early studies have shown the benefits of using big data in different contexts, yet, there is a lack of theoretically driven research on how to utilize these solutions in order to gain a competitive advantage. This work identifies the need for a paradigm shift on firms' view of big data, and focuses at the same time on the yet underserved but highly needed and requested area, that of big data and business strategy.

To this end, this study proposes a conceptual framework that is based on concepts from strategic management and management information systems literature, building on the core areas of big data as identified by early studies. The framework therefore provides a reference for the broader implementa-

tion of big data in the business context. While the elements present on the research framework are on a high-level and can be interpreted as quite abstract, they are purposefully described in such a manner so that they can be adapted depending on the company at hand. This poses a novel perspective on big data literature, since the vast majority focuses on tools, technical methods (e.g., data mining, text analysis, and sentiment analysis), network analytics, and infrastructure. Thus, the proposed framework contributes to big data and business strategy literature by covering the aforementioned gap. It is more important for managers and decision makers to learn how to implement big data and business analytics in their competitive strategies, than to simply perform raw data analysis on large data sets without a clear direction of where it contributed to the overall business strategy.

Furthermore, this study argues that the main source of a competitive edge, especially in highly dynamic and turbulent environments will stem from companies being able to reinforce their dynamic capabilities through targeted use of big data and business analytics. This of course does not lessen the importance of IT resources, since their availability and VRIN (Valuable, rare, in-imitable, and non-substitutable) characteristics can determine the strength of the associated IT-enabled dynamic capabilities developed (Bowman & Ambrosini, 2003). The concepts used in the proposed framework may help managers better understand, plan and organize the process of implementing big data analytics within a business strategy.

This paper offers a theoretical framework on how to increase business value and competitive performance through targeted application of big data. Future studies should empirically test and evaluate this framework by using surveys, interviews, observation, focus groups with experts (e.g., managers, decision makers) and with customers', as well as case studies from the industry. Also, both qualitative and quantitative methods of data collection should be employed. For each different type of data, more than one ways of analysis should be used (e.g., structural equation modelling, qualitative comparative analysis). The main argument made in this paper is that the value of big data does not solely rely on the technologies used to enable them, but are apparent through a large nexus of associations that are eventually infused with organizational capabilities. Strengthening these core dynamic capabilities by virtue of big data is what will lead to competitive performance gains.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 704110

This work was carried out during the tenure of an ERCIM 'Alain Bensoussan' Fellowship Programme.

References

Barreto, I. (2010). Dynamic capabilities: A review of past research and an agenda for the future. Journal of management, 36(1), 256-280.

Bharadwaj, A. S. (2000). A resource-based perspective on information technology capability and firm performance: an empirical investigation. MIS quarterly, 24 (1), 169-196.

Bowman, C., & Ambrosini, V. (2003). How the resource-based and the dynamic capability views of the firm inform corporate-level strategy. British journal of management, 14(4), 289-303.

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. MIS quarterly, 36(4), 1165-1188.

Constantiou, I. D., & Kallinikos, J. (2015). New games, new rules: big data and the changing context of strategy. Journal of Information Technology, 30(1), 44-57.

Cragg, P., Caldeira, M., & Ward, J. (2011). Organizational information systems competences in small and medium-sized enterprises. Information & Management, 48(8), 353-363.

Davenport, T. H., & Patil, D. J. (2012). Data scientist. Harvard business review, 90, 70-76.

Davenport, T. (2014). Big data at work: dispelling the myths, uncovering the opportunities. Harvard Business Review Press.

Drnevich, P. L., & Kriauciunas, A. P. (2011). Clarifying the conditions and limits of the contributions of ordinary and dynamic capabilities to relative firm performance. Strategic Management Journal, 32(3), 254-279.

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they?. Strategic management journal, 21(10-11), 1105-1121.

Erl, T., Khattak, W., & Buhler, P. (2016). Big Data Fundamentals: Concepts, Drivers & Techniques. Prentice Hall Press.

Grégoire, Y., Salle, A., & Tripp, T. M. (2015). Managing social media crises with your customers: The good, the bad, and the ugly. Business Horizons, 58(2), 173-182.

He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: A case study in the pizza industry. International Journal of Information Management, 33(3), 464-472.

Helfat, C. E., & Peteraf, M. A. (2003). The dynamic resource-based view: Capability lifecycles. Strategic management journal, 24(10), 997-1010.

Kallinikos, J., & Constantiou, I. D. (2015). Big data revisited: a rejoinder. Journal of Information Technology, 30(1), 70-74.

Kraaijenbrink, J., Spender, J. C., & Groen, A. J. (2010). The resource-based view: a review and assessment of its critiques. Journal of management, 36(1), 349-372.

Krogstie, J., & Gao, S. (2015). A Semiotic Approach to Investigate Quality Issues of Open Big Data Ecosystems. In Information and Knowledge Management in Complex Systems (pp. 41-50). Springer International Publishing.

Lin, Y., & Wu, L. Y. (2014). Exploring the role of dynamic capabilities in firm performance under the resource-based view framework. Journal of Business Research, 67(3), 407-413.

Maglio, P. P., & Lim, C. H. (2016). Innovation and Big Data in Smart Service Systems. Journal of Innovation Management, 4(1), 11-21.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data. The management revolution. Harvard Business Review, 90(10), 61-67.

Mikalef, P., Pateli, A., & van de Wetering, R. (2016). IT flexibility and competitive performance: The mediating role of IT-enabled dynamic capabilities. In Proceedings of the 24th European Conference on Information Systems (ECIS).

Mikalef, P., & Pateli, A. (2016). Developing and validating a measurement instrument of IT-enabled dynamic capabilities. In Proceedings of the 24th European Conference on Information Systems (ECIS).

Pavlou, P. A., & El Sawy, O. A. (2010). The "third hand": IT-enabled competitive advantage in turbulence through improvisational capabilities. Information Systems Research, 21(3), 443-471.

Pavlou, P. A., & El Sawy, O. A. (2011). Understanding the elusive black box of dynamic capabilities. Decision Sciences, 42(1), 239-273.

Risius, M., & Beck, R. (2015). Effectiveness of corporate social media activities in increasing relational outcomes. Information & Management, 52(7), 824-839.

Teece, D. J. (2007). Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. Strategic management journal, 28(13), 1319-1350.

Varian, H. R. (2014). Big data: New tricks for econometrics. The Journal of Economic Perspectives, 28(2), 3-27.

Vera-Baquero, A., Colomo-Palacios, R., & Molloy, O. (2013). Business process analytics using a big data approach. IT Professional, 15(6), 29-35.

Wade, M., & Hulland, J. (2004). Review: The resource-based view and information systems research: Review, extension, and suggestions for future research. MIS quarterly, 28(1), 107-142.

Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data'can make big impact: Findings from a systematic review and a longitudinal case study. International Journal of Production Economics, 165, 234-246.

Wang, N., Liang, H., Zhong, W., Xue, Y., & Xiao, J. (2012). Resource structuring or capability building? An empirical study of the business value of information technology. Journal of Management Information Systems, 29(2), 325-367.

Woldesenbet, K., Ram, M., & Jones, T. (2012). Supplying large firms: The role of entrepreneurial and dynamic capabilities in small businesses. International Small Business Journal, 30(5), 493-512.

Zahra, S. A., & George, G. (2002). Absorptive capacity: A review, reconceptualization, and extension. Academy of management review, 27(2), 185-203.