Obitade *J Big Data* (2019) 6:71
https://doi.org/10.1186/s40537-019-0229-9

Journal of Big Data

RESEARCH Open Access

Big data analytics: a link between knowledge management capabilities and superior cyber protection

Peter Oluseyi Obitade*

*Correspondence:
Petetola@hotmail.com
College of Business,
University of North Texas,
Dallas, 7300 University Hills
Blvd, Dallas, TX 75241, USA

Abstract

As cybersecurity threats increase in frequency and sophistication, organizations are realizing that one of their strongest resources to combat cyberattacks lies in the growing volume of data at their disposal. However, traditional knowledge management technologies have limited capabilities to effectively process and analyze these larger data volumes that can provide managers with pertinent information to make better-informed security decisions. Using survey data from 479 business and information security executives and drawing from the resource-based view, I find that 69% of organizations that deployed big data analytics reported significant improvements in their cyber knowledge management capabilities. Second, the findings also confirm a significantly positive association between big data analytics and cyber agility. In particular, 72% of organizations that deployed big data analytics solutions reported significant improvements in their ability to detect and respond faster to cyber threats. Lastly, compared to 56% of organizations without big data analytics, 78% of companies that have deployed big data analytics technologies considered their cyber security capabilities adequate and robust enough to protect critical information assets from cyberattacks. Overall, the results demonstrate that organizations need advanced data analytics capabilities to effectively leverage IT resources, allowing them to respond to cyber incidents with speed and agility which ultimately cumulates in the superior cyber protection of valuable information assets. The results also provide some useful implications for research and managerial practices.

Keywords: Big data analytics, Business intelligence, Knowledge management processes and structures, Cyber agility, Information technology capability, Organizational capabilities

Introduction

The rapid rise of the Internet and the digital economy has fueled exponential growth in data amassed by organizations. This growth has also been accompanied by an increase in the frequency and sophistication of cyberattacks. As a result, many organizations have invested heavily in knowledge management capability to acquire, store, and secure their high-value information assets. However, the complex and changing nature of cyber threats and attacks has rendered traditional knowledge management capabilities obsolete. These constantly evolving cybersecurity threats facing organizations are forcing

managers to rethink and reevaluate the tools and tactics they deploy to address the cybersecurity threat [1]. To improve cybersecurity and reduce risk, therefore, organizations are pivoting from a reactive to a proactive approach toward securing information assets, wherein they identify and respond to threats before an attacker can cause damage. Switching to this approach requires an organization to leverage advanced threat detection functionalities, access real-time identification of risks, and implement the rapid deployment of countermeasures to contain cyberattacks before their negative effects can materialize. To execute this approach, the organization requires an updated toolset consisting of an Information Technology (IT) resource that is able to analyze, describe, and combine a large volume of data in a diverse format and from multiple sources. This paper proposes big data analytics as an IT capability that can be leveraged for this function.

Big data analytics is considered one of today's newest technologies designed to transform the way in which organizations manage and utilize information assets. The term "big data" has been used to describe data that is "massive, complex, and real-time… that requires sophisticated management, [as well as] analytical and processing techniques to extract insights" ([2], p. 34). In the current interconnected global economy, data sources have extended beyond structured database records to unstructured data lacking a standardized format [3]. Big data analytics, with its embedded predictive analytical capabilities, is designed for organizations to process and integrate large, diverse volumes of highly detailed data and present that data in a common, familiar format so that businesses and IT managers can make informed strategic decisions, including how to protect valuable information assets.

The purpose of this study is to better understand the ever-changing cybersecurity threat landscape and the role that big data and predictive analytics play in mitigating the threats and risks faced by organizations. To investigate the relationship between IT capabilities and the ability of organizations to effectively protect high-value information assets, I draw on the resource-based view (RBV). In particular, I develop a conceptual model of big data analytics-enabled (BDET) methodology and use this framework to examine how big data analytics enhances cybersecurity capabilities.

This paper identifies and fills two gaps in the existing literature. First, prior studies on big data analytics have focused either on its conceptual and technological aspects, or its general business benefits on adopting organizations [4, 5]. Therefore, little is known about the way in which organizations are leveraging big data analytics functionalities to achieve greater protection from cyber intrusions. Second, while academic studies have documented the importance of knowledge management capabilities and organizational agility to efficient business operations, none has explored these capabilities from a cybersecurity point of view. To fill these gaps, therefore, this paper provides some empirical evidence that demonstrates that big data analytics: (1) enhances organizational capabilities to rapidly respond to cyberattacks and (2) has essential functionalities that allow an organization to effectively harness knowledge management capabilities to provide better cyber protection of valuable information assets. Rather than examining the general business values of big data analytics, this study focuses on how big data analytics has been

deployed to effectively leverage cyber knowledge[1] to improve cyber agility[2] and overall cyber protection. Hence, this is the first paper to empirically test the hypothesis that big data analytics augment organizational cybersecurity capabilities. With organizations' growing interest in extracting business benefits from large accumulated data, the value of big data analytics and the factors that drive its value need to be examined.

To increase our understanding of big data analytics capabilities, this paper examines the association between big data analytics functionalities and two organizational resources: cyber knowledge management capability and cyber agility. Specifically, I hypothesize that big data analytics, with its advanced, predictive, and discovery capabilities, can help bridge significant information gaps regarding optimum cyber protection. I refine the conceptualization and measurement of knowledge management capability practices and application as a latent construct reflected in three dimensions—acquisition, conversion, and application. In addition, drawing on the U.S. Department of Commerce's National Institute of Standards and Technology (NIST) Cybersecurity Framework,[3] I conceptualize two dimensions of cyber agility: pre- and post-cyber incidence agility.

The findings provide some empirical evidence that the predictive and analytic tools of big data analytics are instrumental to firms' abilities to improve cybersecurity capabilities.

Specifically, I find four interesting results. First, the results show a significant positive relationship between big data analytics and cyber knowledge management capabilities.

Specifically, 69% of organizations that deployed big data analytics reported improvements in their ability to effectively manage critical knowledge acquired from cyber activities and threats. Second, the findings also confirm a significantly positive association between big data analytics and cyber agility. In particular, 72% of organizations that deployed big data analytics solutions reported significant improvements in their ability to detect and respond faster to cyber threats and attacks.

Third, compared to 56% of organizations without big data analytics, 78% of companies that have deployed big data analytics technologies considered their cyber capabilities to be adequate and robust enough to protect information assets from cyber threats. Lastly, the findings show the positive, joint effects of big data analytics, knowledge management capabilities, and cyber agility on organizations' capabilities to proactively detect and respond to cyber threats, which ultimately reduces cyberattacks' impact on business operations.

Overall, these findings demonstrate that big data analytics functionalities enhance the cyber knowledge management system architecture of acquisition, conversion, and application, which in turn accelerates organizational cyber agility. In essence, big data analytics technologies are instrumental for organizations to improve their capabilities in discovering potential threats, detecting actual threats, gathering intelligence about

[1] *Cyber knowledge* is a knowledge acquired from both internal and external sources about cyber activities, threats, and attacks that, in most cases, are detrimental to efficient business operations.

[2] *Cyber agility* is the organizational capability to respond quickly and rapidly to cyber threats and attacks. Njilla et al. [6] defined cyber agility as an attack avoidance technique designed to render cyberattacks ineffective.

[3] The *cyber agility* is based on the *National Institute of Standard and Technology (NIST) cybersecurity framework*. The framework has been accepted and adopted by both governmental entities and private companies as the de facto cybersecurity standard.

attacks, and deploying a comprehensive response to minimize the business impacts of cyberattacks. Together, these enhanced capabilities ultimately improve the effectiveness of cybersecurity capabilities.

By utilizing these theoretical foundations, this paper provides an empirical context regarding organizational capabilities, while demonstrating the direct positive impact of IT functions—in particular, the benefits of implementing big data analytics and knowledge management to better protect valuable information resources.

The remainder of the paper is organized as follows: "Theoretical background" section contains the related literature. "Research method" section contains a description of the research methodology. "Results" section contains empirical tests and results, and "Implication for practice" section contains discussion, contributions, and the conclusion.

Theoretical background

Big data analytics

This paper explores the issue of effective knowledge management from the perspective of big data analytics capabilities. Big data analytics is one of the recent advances in technologies that support high-velocity data capture, storage, and analysis. The definition of big data analytics has been an important focal point for the limited research on the subject. Scholars have proposed varying but related definitions for big data analytics. Cox and Ellsworth [7] have been credited as the first to use the term "Big Data." They defined big data analytics as a "challenge for computer systems: data sets are generally quite large, taxing the capacities of main memory, local disk, and even remote disk" (p. 4). In their own contribution to the research, Gandomi and Haider [8] described big data analytics in terms of three characteristics—velocity, variety, and volume (3 V's). Boyd and Crawford [4] considered big data as an interplay between technology, analysis, and mythology.

In its simple form, big data analytics is a "focus on very large, unstructured and fast-moving data" ([9], p. 10). Big data analytics is a product of previous IT capabilities and concepts such as "decision support," "online analytical processing" and "business intelligence" [9, 10]. Big data analytics evolved as a response to the recent ability by organizations to collect, mine, and exploit data that are increasingly available from an enormous variety of internal and external sources. For instance, data collected from non-traditional sources such as smart phones, "apps," and other social media devices contain valuable information that could be utilized to develop innovative products or discover new business opportunities (e.g., [11]). In particular, data from non-traditional sources could be instrumental for organizations to not only to 'better understand the present' but also, contingent on certain assumptions, "predict the future" [12] and, hence, positively influence the ability of organizations to compete.

Cyber agility

Today's ever-changing business environment makes it imperative for firms to be agile and develop the capability to effectively adjust to extreme changes, survive unprecedented threats, and capitalize on emerging business opportunities [13]. Organizational

agility has been defined as the ability of organizations to cope and thrive in a rapidly changing, relentless, and highly uncertain and competitive business environment [14].

To better understand how organizations are leveraging big data analytic capabilities, I examine a sub-set of organizational agility that is related to cybersecurity–cyber agility. A unique feature of today's dynamic business environment is the rapid changes in the cyber threat landscape. Specifically, organizations are faced with an exponential increase in the number and frequencies of cyberattacks and the types of threat actors. The attack vectors have also grown from simple cyber incidents to more sophisticated advanced persistent threats (APTs) [15]. Organizations are finding out that the traditional security solutions and processes are not equipped to sufficiently address the new threat landscape. To cope with new threats, therefore, organization are pivoting to proactive approach which requires significant improvement in cyber agility.

Njilla et al. [6] defined cyber agility as an attack avoidance technique designed to render cyberattacks ineffective. Contributing to the literature, Hult and Sivanesan [16] suggested a more agile approach to cybersecurity to ensure that information assets are properly protected. In their view, effective cyber agility is essential for organizations to quickly respond to and contain the devastating effects of cyberattacks.

Resource-based view (RBV)

Resource-based view (RBV) has been extensively studied in business literature and has been "widely acknowledged as one of the most prominent and powerful theories for describing, explaining, and predicting organizational relationships" across the business disciplines ([17], p. 1300). RBV considers the firm as a collection of tangible and intangible resources; however, only those that are valuable, rare, inimitable, and nonsubstitutable can generate competitive advantage [18]. Previous studies in management literature have proposed several theories to explain the way in which organizations develop and execute a business strategy to compete in their environments. Some of these include transaction cost economics, agency theory, network view, and institutional theory [19, 20]. Out of these theories, only RBV considers an organization as a collection of resources and presents a cogent framework to combine the disparate resources in a firm to generate competitive advantage [17]. This premise is especially relevant to this study because the model presents big data analytics capabilities as specific resources that an organization possesses and can leverage to provide better protection for firm assets.

Extant literature has also provided empirical evidence that demonstrates that RBV as one of the most compelling theories in information security (IS) and other business disciplines to explain the association between firm resources and their operational and market performance. For instance, Melville et al. [21] demonstrated RBV as a relevant framework to understand the value of implementing IS to improve firm performance. In particular, they documented that the business value of IT depends on internal and external factors including complementary organizational resources and developing relationships with trading partners. Gu and Jung [22] corroborated these findings and provided empirical support that shows RBV as a robust framework that could be used to identify and categorize IS resources; and to measure the impact of IT resources on a firm's drive for superior performance.

Further, the results of recent studies have confirmed the findings of early works and demonstrated empirically the ability of big data analytics to improve firm performance. For instance, drawing from RBV, Wang et al. [23] reported that big data analytics capability generated business benefits for deploying organizations, especially in the healthcare industry. In line with these findings, Loebbecke and Picot [24] posited that big data analytics, along with digital transformation are being leveraged not only to optimize existing business processes to increase the efficiency of business operations, they are also transforming how the society at large communicate and cooperate. Chae et al.[25] utilized RBV methodology to demonstrate that organizations which deployed and utilized the advanced analytic functionalities of big data analytics achieved greater operational and market performance than their peers without big data analytics capabilities. Overall, these studies have demonstrated, both theoretically and empirically, that RBV is a relevant paradigm to understand the association between organizational resources and organizational performance. Given that the main objective of this study is to propose big data analytics as a critical IT resource that can be leveraged to enhance organizational internal knowledge management and agility capabilities, the choice of RBV as a theoretical framework for this study seems appropriate.

Knowledge management

Extant literature has documented that to compete effectively in today's interconnected global economy, firms must leverage their existing knowledge and create new ones on a continuous basis. Thus, the importance of knowledge to an organization has led to extensive research into the way in which organizations acquire and utilize knowledge to gain and maintain favorable market position (Gold, Malhotra, & Segars). In its simplest form, knowledge management consists of processes that allow an organization to capture, store, transform, and transfer knowledge among the varying units [26, 27]. In his study, Spender [28] examined knowledge management processes and classified them into four broad dimensions—knowledge acquisition, conversion, application, and protection. In this paper, I propose three of the four dimensions as pertinent to organizational cyber agility and firm capability to effectively protect information assets.

Acquisition process

Acquisition-oriented knowledge management processes are oriented toward obtaining or acquiring new knowledge. Studies have demonstrated that the knowledge acquisition process is an important element in managing knowledge within an organization [29]. Specifically, researchers have identified two primary sources of acquiring new knowledge: acquire entirely new knowledge and create new knowledge out of existing knowledge through collaboration between individuals and between business partners [30, 31].

Conversion process

The second knowledge management process is knowledge conversion, which deals with organizational ability to make or convert existing knowledge into a useful form. Nahapiet and Ghoshal [32] argued that, for an organization to efficiently covert knowledge into its useful form, knowledge must be properly organized and structured. O'Dell and Grayson [33] posited that the knowledge-conversion process relies on a firm's ability to organize,

coordinate, combine, integrate and distribute knowledge. Similarly, the conversion process also includes mechanisms used to combine or integrate knowledge from various sources. This integration helps organizations reduce redundancy, thereby improving efficiency through the elimination of excess volume. Similarly, conversion also enables the organization to replace knowledge that has become outdated [33].

Application processes

Application-based processes are those oriented toward the actual use of the accumulated knowledge. These processes include storage, retrieval, and sharing of the acquired and converted knowledge [34]. An important element in the knowledge application is the storage and retrieval process that enables the organization to have an uninterrupted and quick access to knowledge. In addition to the storage and retrieval process, the application process includes the mechanisms used to share knowledge to enhance organizational capability [26]. Aiavi and Leidner [35] emphasized that an effective application of knowledge is instrumental for organizations to improve operational efficiency and reduce costs.

Development and research model

This paper adopts a big data analytics-enabled (BDET) approach to the resource-based view (RBV) of the firm and seeks to establish boundary conditions for the value of certain information technology (IT) capabilities. RBV is concerned with identifying the resources and capabilities that enable a firm to attain and maintain the superior performance that cannot be easily duplicated by competitors (e.g. [36]). Studies that examine the adoption of RBV as a framework to investigate the contributions of IT resources typically demonstrate positive outcomes relating to a firm's competitive advantage. For instance, Gupta and George [37] noted that RBV is based on the premise that firms can achieve a competitive advantage and improve organizational internal efficiency by combining IT resources with their other internal capabilities.

To explore the association between big data analytics and other IT capabilities, I first refine the conceptualization and measurement of knowledge management capability as a latent construct reflected in its three dimensions: acquisition, conversion, application, and security. In addition to the KM capabilities, the model leverages the National Institute of Standards and Technology (NIST) cybersecurity framework to conceptualize two types of cyber agility: pre-incidence and post-incidence agility. Last, the model examines the association among BDA capabilities, KM enablers, and two goals of effective firm cybersecurity—(1) improvement of cyber threat detection, and (2) reduction of business impacts from cyberattacks.

Explanatory variables: big data analytics capabilities

The first step in this model is to define the explanatory variable used in BDET methodology [23, 25, 37]. In line with the previous studies, I utilize big data analytics as the explanatory variable to examine the impact of its functionalities on firm performance. BDA has been described as a collection of aggregation analytics, and interpretation techniques that transform data into mechanisms useful in decision making [38]. For

this model, I examine two important dimensions of BDA: its architectural components and technological capabilities.

The limited literature in big data analytics has documented the three main architectural components of big data analytics: (1) data aggregation, (2) data analysis, and (3) data interpretation [23]. The first component is data aggregation. Ward et al. [39] described data aggregation of big data analytics as the tool to collect disparate data from multiple sources, both internal and external to the organizations, and transform them into a format that is easier to read and analyze. Data aggregation is made up of three components which are the acquisition, transformation, and storage [40].

Data analysis is the second architectural component of big data analytics. This functionality is used to process and perform analyses on data from disparate sources to discover information useful for decision making [39]. To better understand the data analysis functionality of big data analytics, Delen [41] identified and documented the three components of data analysis: descriptive, predictive, and prescriptive analytics. Each element is differentiated by the type of data processed and the purpose of the analysis.

The third architectural component of big data analytics is data interpretation. This component generates outputs such as reports and visual representations (charts, dashboards, etc.) leveraged by organizations in the decision-making process. In particular, big data analytics data interpretation layer has been used to produce real-time reports, critical business operation alerts, proactive notifications, and operational key performance indicators (KPIs) [40].

Elements of big data analytics capability

The second dimension of big data analytics examined in the extant literature is its technological capabilities. Gupta and George [37] documented four types of big data analytics capability: analytical capability, decision support capability, traceability, and predictive capability. These capabilities allow big data analytics to process, in parallel, large data volumes and visualize data in a real-time or near real-time basis. It is these capabilities that differentiate big data functionalities from traditional business intelligence systems. The first big data analytics capability is analytical, which enables organizations to improve process efficiency and deliver business value that might have been previously difficult or impossible to discover [38].

With its decision support capability, big data analytics provides critical information such as historical reports, statistical analyses, time series comparison, and executive summaries to managers and executives to facilitate better decision making [42]. The third capability is predictive, which is used to enhance models employed for forecasting and planning; and to predict future market trends and business opportunities. The last big data analytics capability is traceability. This capability allows organizations to track critical data from diverse IT systems such as transactional and business intelligence applications. Taken together, BDA provides organizations with the ability to discover undetected correlations, patterns, and trends between specific variables of interest across multiple dimensions.

IT-enabled transformation resources

In this section, I examine the role of IT-enabled transformation resources in the model. IT-enabled transformation resources are defined as organizational capabilities that leveraged BDA functionalities to improve operational performance. In this model, knowledge management and cyber agility are considered as the intermediate IT capabilities. The first organizational capability is Knowledge Management (KM). This paper focuses on three of KM elements that are pivotal for KM to deliver business value—acquisition, conversion, and application. Also, to enhance our understanding of organizational agility, I examine the effects of BDA on the two forms of cyber agility—pre-cyber incidence and post-cyber incidence agility.

Outcomes

To conceptualize the ultimate outcomes or the business value of the model, I utilize a two-dimensional view of cyber capabilities—to improve cyber threat detection and to reduce the business impacts of cyberattacks. In line with the literature, I regard improve cybersecurity capabilities as the extent to which a firm effectively leveraged in cyber capabilities to ensure better protection for valuable information assets. I propose that big data analytics can be leveraged to improve the two dimensions of cyber agility; thereby, it serves as the critical link between knowledge management enablers and business value. Big data analytics does this by enhancing the three processes of KM—acquisition, conversion, and application. It also improves the two types of cyber agility. This premise is aligned with that of Goldman et al. [14] and Volberda [43] in which they suggested that, for an organization to achieve improved agility, including in agility in cyber protection, it requires the capabilities to process, on timely basis, a large volume and variety of distributed information that are both internal and external to the organization. Taken together, this model proposes that big data analytics is an essential IT resource capable of improving knowledge management enablers, ultimately accelerating cyber agility and contributing to improvements in the protection of a firm's assets from cyber thefts.

As shown in Fig. 1, the research model adopted in this study follows a linear progress path that begins from the explanatory variables to practices, then moves on to the intermediate outcomes (knowledge management and cyber agility, which are considered "benefits"), and finally demonstrates examples of improved protection for a firm's valuable assets ("business value" in the model).

Hypothesis development

Big data analytics and knowledge management

Organizations have long recognized Information Technology (IT) as an enabler of superior firm performance. To leverage these capabilities, therefore, most large organizations have invested considerably in IT solutions, especially in building knowledge management infrastructures. While these firms typically have been able to leverage their KM capabilities in terms of accumulating a massive amount of data, many of them have been unable to effectively utilize their collection of information assets to derive desired business benefits. In response to the perceived limitations of current knowledge

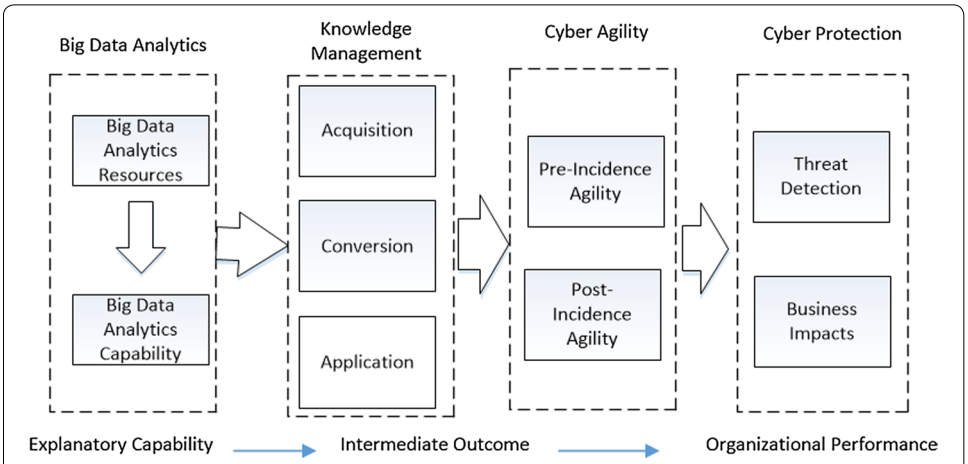

Fig. 1 Big data analytics-enabled transformation model

management, many firms have launched big data analytics initiatives to leverage their information assets to achieve competitive sustainability.

Using different types of analytic tools such as data visualization, natural language processing, data mining, and statistical analysis, big data analytics enables a firm to obtain new critical information about its competition and broader economic environment from the existing data repository [4]. In particular, big data analytics allows a firm to find new knowledge that is either internal or external to the firm, to effectively track sources of knowledge, and to create a catalog of internal organizational knowledge. Also, big data analytics capabilities enable an organization to use its existing knowledge more effectively to track and respond to demands from customers and protect valuable information. Cao et al. [38] noted that BDA is an essential IT resource that allows an organization to improve the organizational capability to create, transfer, and store knowledge from diverse sources. Therefore, I hypothesize that:

H1 BDA enhanced three elements of knowledge management in organizations—acquisition, conversion and application processes.

Big data analytics and cyber agility

A large stream of literature has asserted that IT can enable agility by speeding up decision making, facilitating communication, and responding quickly to changing conditions [44]. In this study, I extend the literature to examine how BDA enhances organization cyber agility, which in turn contributes to effectively protect critical information assets from cyberattacks.

The frequency, intensity, impact, and sophistication of cyberattacks continue to grow. And at the same time, the actual time to detect and respond to threats is increasing [1]. To effectively address these issues, it has become imperative for organizations to deploy IT solutions that could help to improve cyber response capabilities. Cyber agility is important because cyberattacks are notoriously quick to carry out, and the devastating operational and financial consequences noticeable in a matter of minutes. For instance,

Schiavone et al. [45] documented that 75% of organizations that had experienced cyber-attacks reported that their network infrastructures and systems were compromised in minutes from an attack to data exfiltration, regardless of the size and maturity of the organization or the amount of money invested in information security.

Gupta and George [37] confirmed this assertion in their studies in which they argued that traditional knowledge management tools are too slow and inefficient to allow organizations to adequately response on time to cyber incidences; therefore, they suggested big data analytic solutions to provide the information to aid organizations' agility capabilities to help shorten the amount of time and efforts required to respond and contain the cyberattacks.

Big data analytics provides an integrated platform that enforces standardization and integration of data and processes that are essential to enhance cyber agility. Also, the integration capability of BDA allows organizations to gather and share information in a timely manner. It also provides access to real-time, consistent, and comprehensive security information, which is essential for fast, efficient decision making [37]. The real-time access to pertinent information about the changing threat landscape allows organizations to respond rapidly to cyber incidences, which invariably contributes to the improvement of cyber agility. Further, Eastman et al. [1] argued that through the use of cyber analytics, organizations can predict unusual cyber activities including the ability to detect active insider and external threats. Taken together, the big data not only ensures the processing of detailed data, it also integrates diverse data types, delivered at various speeds and frequencies, all of which are essential to improve cyber agility [16]. Therefore, I hypothesize that:

H2 BDA has a positive impact on cyber agility.

Big data analytics and cybersecurity

The increased number of attack vectors and threat actors has resulted in exponential growth in the level of cybersecurity complexity for organizations of various sizes. While organizations have a wealth of existing or easy-to-access data that could support improved security, they lack the advanced analytic capability to analyze and effectively utilize these assets. In other words, current cybersecurity solutions are limited because of their general inability to efficiently analyze all data assets. Specifically, organizations are realizing that the traditional dump-and-analyze methodology has proven to be ineffective because it lacks the capability to store and analyze the needed data history in a timely fashion. Therefore, a new capability is required to leverage and evaluate data in a way that enhances cybersecurity technologies. Big data analytics offers the functionalities to assist organizations in achieving greater threat identification and remediation processes that are essential to mitigate cyber risks.

With big data analytics, organizations have the ability to store, process and analyze massive cybersecurity data sets relatively cheaply and quickly. A whole new area of opportunity has been unlocked in advanced analytics to enable business insights and improved decision making. In particular, the advanced analytics techniques such as data/text mining, machine learning, and pattern matching enhance the diagnosis and

predictive and automate data analysis needed to generate insights and answer complex security questions. Invariably, big data analytics provides better visibility into network activity and tools to proactively detect malicious behavior before a breach occurs. Therefore, I hypothesize that:

H3 Big data analytics is positively associated with superior cyber protection.

Research method

Sample and data collection

For this paper, I follow the scale development procedure developed by MacKenzie et al. [46]. He suggested collecting two sets of data: one to evaluate the instrument properties including multicollinearity, discriminant validity, and reliability, etc. The second dataset is to be used to reevaluate the scale properties and establish their nomological validity. Thus, two studies were conducted—study I and study II. Data collected during the pilot study (study I), was used to refine the scale items. Based on the results from the pilot study, data for the main (study II) were collected from a new sample with the scale properties re-examined and refined.

Pilot study

The first set of data for the pilot study was collected from IS and business executives who were members of one of the largest groups dedicated exclusively to professionals with big data experience—the big data and analytics group on LinkedIn. There were approximately 338,791 total members as of April 2018 when the data was collected. The survey was sent to 967 members who were identified as both business and IS executives based on their profiles (including job titles). In total, 176 responses (18.2%) were received. Respondents represented a diverse range of affiliated industries including energy, financial services, manufacturing, computing and information services, and healthcare.

For study II, two types of organizations were targeted—organizations that have implemented BDA and those that have not yet deployed big data analytics solutions. Two versions of the same survey were developed and distributed to two separate groups. The purpose was to provide a control group as a reference point to determine the effects of big data analytics on organizational knowledge management, organizational agility, and superior cyber protection. This control sample was also used to test various models aimed at distinguishing organizations that have implemented big data analytics capabilities from those that have yet to deploy such solutions.

To ensure adequate representation of both interest groups, two approaches were adopted to collect the data. First, a survey questionnaire was distributed through different LinkedIn groups and forums. Organizations that have deployed big data analytics solutions were reached through the two largest of such groups (by membership)—Chief Data Officer group and Data Science Central group. To gather data from other organizations (non-BDA), other IS and business professional groups in LinkedIn were used. Second, as an additional validation, a set of questions was included in the survey to determine whether an organization had implemented BDA. The questions specified key attributes of BDA, such as whether organizations have access to large data sets and

Table 1 Sample data by industry

Industry	Expected	Observed
Computer/consulting	55	53
Energy	66	62
Finance/banking	42	38
Manufacturing	47	43
Retail	83	77
Healthcare	91	86
Services	89	86
Others	39	34
Total	512	479
T-test	$p < 0.001$	

have adopted parallel computing approaches (e.g. Hadoop, etc.). These questions were adapted from an instrument developed by Gupta and George [37].

For study II, a new survey was created. The survey link was then sent to the group members by the owner/moderator of the CDO group. The survey was sent to 2415 members who were both business and IS executives. A total of 479 responses (19.8%) were received. I then examined the responses to check if there was anyone from the initial survey who also responded to this survey. No common respondents were found in the two studies.

Response rate and tests for non-response bias

As with studies that have utilized online surveys, non-response bias presented a potential problem for this study. Therefore, I utilized two methods to test for non-response bias. In the first test, I compared attributes of survey respondents with attributes of all individuals who received the survey using industry affiliations and firm size. Next, I compared the industry affiliation and firm size of respondents to industry and firm size from the complete list of survey recipients. Then, I compared the observed frequencies of the respondents with expected frequencies based on the surveyed respondents. As shown in Table 1, the t-test indicated no significant difference between the respondents and the full sample.

In the second non-response bias test, the surveys received were grouped into two "waves" based on the date returned, with later respondents serving as a surrogate for non-respondents. The first wave included those surveys received within 3 weeks of the initial online contact, while the second wave included surveys received after the follow-up email. The first survey, which was sent to both groups in April 2018, resulted in a total of 361 responses, out of which 55 incomplete responses were discarded, leaving 306 usable responses. Subsequent requests (and reminders) were sent 2 months later, which produced an additional 211 responses, out of which 38 were discarded with 173 usable responses. This led to a total of 479 usable responses.

To further test for non-response bias, I compared the characteristics of the early and late responses with each other. I conducted a test of differences of proportion by comparing responses from two waves using characteristics such as firm size, industry classification, IT dept size, and respondents' official position in the company. None of

the differences were significant. Also, Levene's test of homogeneity of variance showed no significant difference in variance among data sources (F = 1.412). While the above tests could not guarantee the absence of any non-response bias, they suggested that the respondents were representative of the population surveyed.

Operationalization of constructs

This study's variables were operationalized using multi-item reflective measures (on a seven-point scale). The measurement items were adopted from previous studies. The final instrument is presented in Appendix.

Knowledge management The three processes of knowledge management—acquisition, conversion, and application—were measured with items ranging from four to six. The measurement items were designed to capture the presence and the effects of knowledge management in an organization.

Cyber agility drawing on the NIST cybersecurity framework, I examined firm agility in relations to cyberthreats and attacks. I utilized two forms of cyber agility: pre-cyber incidence and post-cyber incidence agility.

Pre-cyber incidence agility pre-cyber activities are those conducted by potential threat actors prior to launching an attack such as reconnaissance. For this form of cyber agility, I utilized three of the five NIST cybersecurity framework options: (1) identity agility was measured with three items designed to indicate the firm's ability to respond quickly identity potential cyber threats, (2) detect agility measured the speed and agility in detecting imminent threats, and (3) protect—concerns with proactive measures to protect information assets.

Post-cyber incidence Activities or countermeasures initiated during and after a cyberattack to contain its effects and to stop the attack. The proxies for this form of agility are the remaining two NIST cybersecurity framework functions: (1) *Respond agility*, which measures a firm's ability to respond to changes in attack vectors and to react quickly to cyberattacks, and (2) *Recover*, which measure reflects a firm's agility to recover from cyberattack and reduce the financial and operational impacts of such attacks.[4]

[4] The NIST cybersecurity framework consists of five functions—identity, detect, protect, respond and recover. These five functions were selected because they represent the five primary pillars for a successful and holistic cybersecurity program. They aid organizations in easily expressing their management of cybersecurity risk at a high level and enabling risk management decisions.
Identity—the identify function assists in developing an organizational understanding on how to better manage cybersecurity risk to systems, people, assets, data, and capabilities. Understanding the business context, the resources that support critical functions, and the related cybersecurity risks enable an organization to focus and prioritize its efforts, consistent with its risk management strategy and business needs.
Protect—the protect function outlines appropriate safeguards to ensure the delivery of critical infrastructure services. The Protect Function supports the ability to limit or contain the impact of a potential cybersecurity event.
Detect—the detect function defines the appropriate activities to identify the occurrence of a cybersecurity event. The Detect Function enables the timely discovery of cybersecurity events.
Respond—the respond unction includes appropriate activities to take action regarding a detected cybersecurity incident. The Respond Function supports the ability to contain the impact of a potential cybersecurity incident.
Recover—the recover function identifies appropriate activities to maintain plans for resilience and to restore any capabilities or services that were impaired due to a cybersecurity incident. The recover function supports timely recovery to normal operations to reduce the impact from a cybersecurity incident.

Table 2 Results of exploratory factor analysis: joint factor analysis

Item	Mean	Std. dev	Minimum	Maximum	Factor1	Factor2	Factor3	Factor4
BDA1	5.35	1.089	2	7	0.717	0.465	0.178	0.094
BDA2	5.10	1.165	1	7	0.717	0.303	0.244	0.150
BDA3	5.30	1.342	1	7	0.837	0.404	0.128	0.116
BDA4	5.45	1.146	1	7	0.741	0.134	0.382	0.183
KM1	5.55	1.395	1	7	0.434	0.768	0.325	0.191
KM2	5.95	0.887	1	7	0.142	0.818	0.004	0.032
KM3	5.85	1.089	1	7	0.493	0.775	0.040	0.040
KM4	5.80	0.834	1	7	0.457	0.795	0.001	0.025
CA1	4.75	1.118	1	7	0.444	0.379	0.875	0.270
CA2	4.80	1.196	1	7	0.114	0.470	0.822	0.117
CA3	5.00	1.257	1	7	0.248	0.033	0.798	0.231
CA4	5.25	1.372	2	7	0.144	0.194	0.839	0.094
CP1	4.80	1.281	1	7	0.264	0.279	0.020	0.746
CP2	5.60	0.995	1	7	0.139	0.378	0.134	0.840
CP3	5.70	1.081	1	7	0.465	0.282	0.441	0.764
CP4	5.60	1.046	1	7	0.137	1.000	0.384	0.822

Factor 1: big data analytics; Factor 2: knowledge management; Factor 3: cyber agility; Factor 4: cyber protection

Superior cyber protection The proxies to measure improved cyber protection are improved threat detection and a reduction in business impacts of breaches. Each of these dimensions is measured using three items. These items were designed to compare the cyber capabilities of organizations that have deployed big data analytics with those that have yet to deploy the same advanced analytics functionalities.

Control variables

Consistent with previous studies, four control variables were used in this study—firm asset size, employee size, IS size and industry affiliation. The firm asset was measured using firm revenue; firm size was the number of full-time employees (FTE); IS size was measured as the ratio of number of FTEs in the IS department to firm-wide FTEs; and finally, the industry sector was a binary variable with 1 for service firms and 0 for manufacturing firms.

Measurement model

Convergent and discriminant validity

I performed various tests to assess the construct validity and reliability of the instrument using multiple models. In each estimated model, items that cross load or demonstrate poor reliability were dropped and the model re-estimated. The purpose of this exercise was to ensure the strength of measurement at the item level such that estimates among constructs were not confounded.

Table 2 presents the results of the exploratory factor analysis (EFA). A four-factor structure emerged with all predefined indicators loading on to their respective constructs, thus affirming convergent validity. I also performed a confirmatory factor analysis (CFA) to assess convergent validity and reliability. As shown in Table 3, the indicators

Table 3 Results of confirmatory factor analysis: correlation and reliability of latent constructs

Constructs	Respondent	Item	Mean	Std dev	1	2	3	4	Range of factor loadings	Composite reliability	AVE
Big data analytics	IS executive	6	0.48	0.500	0.83				0.78–0.95	0.88	0.71
Knowledge management	IS executive	26	4.87	1.382	0.59	0.89			0.81–0.88	0.93	0.78
Cyber agility	Bus. executive	7	5.08	1.412	0.53	0.59	0.82		0.81–0.89	0.91	0.72
Cyber protection	Bus. executive	8	5.17	1.417	0.33	0.29	0.41	0.9	0.77–0.91	0.83	0.75

Model fit indices: χ^2 (df) = 124.18 (89), p = 0.0111; root mean square error of approximation (RMSEA) = 0.0623, normed fit index (NFI) = 0.89, goodness-of-fit index (GFI) = 0.87, adjusted goodness-of-fit index (AGFI) = 0.81, comparative fit index (CFI) = 0.95. Square-root of AVE values along the diagonal

loaded were high, greater than 0.73 on their respective constructs. Second, the fit indices of the measurement model were all within the normally specified thresholds. Third, composite reliability for each construct was greater than 0.7, and the average variance extracted (AVE) for each construct was above 0.5. Table 4 presents the correlation among all indicators. Together, these results confirmed the reliability, convergent validity, and discriminant validity of the measures.

Tests of common method bias and survey data

Several analyses were performed to test for common method bias in the survey data. First, data was collected from multiple respondents (business and IS executives) to minimize the threat of common method bias. The questionnaires were divided into two parts: A and B. Part A included questions relating to cyber agility and cyber protection variables designed for business executives who were assumed to be in a better position to assess the effects of big data analytics on cyber protection. Part B was designed to obtain responses from suitable IS executive to provide information about the firm's IT management practices and IT capabilities. In particular, they were asked to respond to questions that were IT-related (knowledge management capabilities). Unlike business executives, IS executives are more likely to be familiar with the impacts of big data analytics on knowledge management systems. Second, I conducted Harman's post hoc single-factor analysis to examine for method bias in the data.

Results

Table 5 presents the sample statistics. The sample firms were distributed across a wide range of industry sectors. On average, the sample firms had 1978 employees enterprise-wide. This shows that the sample firms were a good representation of the diversity of firms from small to large organizations. Their IT departments, on average, had 18 employees. The average budget of the sample was about 1.7% of sales revenue.

The IS executives had on average 21 years of professional experience. Also, over 87% of the responding business executives held job titles such as vice president, president, director, and senior manager. On average, they had been involved in strategic planning development for 8.2 years. Similarly, about 91% of the IS executives were above the level of director, including chief information officers and vice presidents of IT. On average, they had been involved in corporate IT strategy formulation for 9.2 years. Taken together, these attributes indicated that the respondents were high ranking within their organization and highly competent to answer the questions of this study.

Hypotheses testing

Regression analysis was conducted to test the research hypotheses. The multi-item measures were transformed into summated scales. Similar to Cohen et al. (2003), to reduce any potential problems of multicollinearity, the variables were mean centered prior to forming the multiplicative product term. To test the hypotheses, six multivariate regressions were conducted. For all regressions, the independent variable is a dummy variable with a value of one if big data analytics technologies were deployed in an organization and zero otherwise. The control variables were drawn from previous studies and were mean centered to ensure easy interpretation of the coefficients.

Table 4 Summary statistics and correlations

Variable	Mean	Std. deviation	BDA	KM	Agility	Protection	Asset	Employees	IS size	Industry
BDA	0.48	0.500	1							
KM	4.87	1.382	0.596**	1						
Cyber agility	5.08	1.412	0.707**	0.719**	1					
Cyber protection	5.17	1.417	0.737**	0.657**	0.769**	1				
Asset	4.41	1.743	0.801**	0.535**	0.717**	0.712**	1			
Employees	4.61	1.658	0.737**	0.493**	0.692**	0.718**	0.701**	1		
IS size	4.91	1.472	0.721**	0.528**	0.758**	0.668**	0.691**	0.684**	1	
Industry	4.92	1.558	0.763**	0.599**	0.782**	0.700**	0.701**	0.651**	0.701**	1

*p < 0.10, **p < 0.05, ***p < 0.01

Table 5 Sample characteristics

	Pilot—study 1 (%) N = 176	Study 2 (%) N = 422
Industries		
Computer/consulting	11	11
Energy	14	13
Finance/banking	9	8
Manufacturing	12	9
Retail	13	16
Healthcare	17	18
Services	16	18
Others[a]	8	7
Total BDA experience		
Less than 3 years	51	42
3–6 years	42	47
More than 6 years	7	11
Number of employees		
Fewer than 1000	34	21
Between 1001 and 2500	19	29
Between 2501 and 5000	16	14
More than 5000	31	36

[a] Other industries include agriculture, utilities, real estate, government agency etc

Table 6 Results of analyses of research hypotheses

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Explanatory variable	0.036***	0.021***	0.043***	0.273	0.282	0.225***
Firm asset	0.103***	0.099***	0.110***	0.061	0.165	0.146*
Firm size	0.061***	0.051***	0.048***	0.079	0.245	0.223
IS size	0.038***	0.048***	0.038***	0.231	0.078	0.024
Industry	0.077	0.096	0.079	0.157	0.137	0.082
R-square	0.71	0.75	0.69	0.71	0.59	0.61

Model 1—BDA effects on cyber protection; Model 2—BDA impacts on KM; Model 3—BDA on agility, Model 4—KM on agility; Model 5—KM on cyber protection; Model 6—interaction of KM and agility on cyber protection

*p < 0.10, **p < 0.05, ***p < 0.01. One-tailed tests. All variables are mean centered for moderation analyses

In Hypothesis 1, I test to determine the effects of big data analytics on cybersecurity knowledge management capabilities. I perform both univariate and multivariate regressions. Untabulated results show that 69% of respondents reported significant improvements in their knowledge management with the deployment of big data analytics technologies. Similar to previous studies, the dependent variable is drawn from average the score from participant's responses to the knowledge management questions [35, 37, 47]. Table 6 presents the results of hierarchical regression analyses. As shown in Table 6, the results in Model 1 provide strong support for H1 as indicated by the significant positive coefficients of big data analytics on knowledge management construct ($b = 0.36$, $p < 0.01$), which demonstrates that big data analytics enhanced these processes. For the control variable, I found a significant positive effect on firm size, firm asset, and IS size. These results indicate that organizations which have deployed and leveraged big

data analytics capabilities are more likely to have larger IT departments, a higher number of total employees, and greater revenue than organizations which have yet to implement the big data analytics capabilities. The results are not surprising considering that large organizations are more likely to have the needs and resources to be early adopters of new technologies. And in some cases, large organizations are more likely to have a global footprint and are therefore more likely to face tougher competitions from rivals in multiple countries than small size or domestic companies. The indicator of the industry was not significant. I interpreted this finding to suggest that organizations in both the service industry and the manufacturing industry are as likely to deploy the big data analytics functionalities.

Hypothesis 2 is designed to test the effects of big data analytics on cyber agility. Similar to knowledge management capabilities results, the univariate results show that 72% of organizations that have implemented big data analytics reported significant improvements in their ability to detect and react quickly to cyber threats and attacks. The results of Model 2 confirmed the univariate results and provide strong support for H2 as demonstrated by the significant positive coefficients of BDA on cyber agility ($b = 0.21$, $p < 0.01$). These findings suggest that big data analytics functionalities enhance cyber agility, and thereby enable deploying organizations to response proactively to cyberattacks.

Hypothesis H3 posts in the alternative that organizations that deploy big data analytics functionalities are more likely to achieve superior cyber protection. The univariate results show that organizations that have implemented big data analytics functionalities are more likely to achieve greater cyber protection than organizations without big data analytic capabilities. In particular, the untabulated results indicate that, compared to 56% of organizations without big data analytics, 78% of companies with big data analytics considered their cyber protection capabilities to be adequate and robust to protect critical information assets. The multivariate regression confirmed this result. Specifically, the findings in Table 6, model 3 show a significant positive effect of big data analytics on cybersecurity capability ($b = 0.43$, $p < 0.01$) over and above the four control variables. This suggests that, if properly implemented and utilized in the decision-making process, big data analytics capability contributes to improved cyber protection.

To further enhance our understanding of the association between big data analytics and other organizational capabilities, I performed additional analyses. First, I test to determine relations between cyber knowledge management capabilities and cyber agility. I regressed the knowledge management process on cyber agility, controlling for the effects of firm size, IS size, and industry. In Table 6, Model 4, the results show a positive but insignificant effect of knowledge management capability on agility ($b = 0.273$, $p > 0.10$). These findings are consistent with previous studies that suggest knowledge management capability alone might not be sufficient for organizations to react effectively to changes in the marketplace [48].

In the second test, I examined the association between knowledge management and cybersecurity capability. The result in Model 5 shows an insignificant positive effect of knowledge management on firm effectiveness of cyber capabilities ($b = 0.282$, $p > 0.10$). Similar to agility, these results support the findings of previous studies that suggested that organizations need a moderating capability to achieve the desired benefits from

knowledge management initiatives [30]. In the final regression, the result in Table 6, Model 6 indicates a significantly positive effect of the interaction of knowledge management capability and cyber agility ($b = 0.225$, $p > 0.01$). This finding suggests that, when knowledge assets are effectively leveraged to enhance organizational agility, the combination of both capabilities can improve a firm's ability to effectively secure valuable assets from cyberattacks.

Discussion

Organizations today are faced with larger and broader changes to their business processes than ever before. One feature of this change is the exponential rate at which organizations are generating data from previously uncapturable data and acquiring valuable information obtained from new sources. Undoubtedly, the velocity and dynamic nature of the business environment have created a competitive incentive among organizations to consolidate and leverage their information assets as a means of creating value that is sustainable over time. Ironically, the importance of information assets to an organization also makes it a tempting target for hackers, criminals, nation-state entities, and others. Therefore, securing these kinds of valuable assets has become a top priority for virtually all firms.

This study examines the moderating effects of BDA on organization cyber capabilities. I proposed and tested two distinct dimensions of cybersecurity capabilities—improved threat detection and reduced business impacts of cyberattacks. These dimensions are empirically explored to better understand the impacts of BDA capability on organizational processes.

Linking big data capabilities with potential benefits

Studies have identified four capabilities of BDA—traceability, analytical, decision support, and predictive capabilities [37]. This paper finds some support for the mediator role of big data analytics functionalities on knowledge management and agility capability. Specifically, the results reveal that different big data analytics capabilities enhance various elements of knowledge management and the two forms of cyber agility, thereby improving organizational ability to compete in a competitive global environment. As shown in Fig. 2, an important benefit from big data analytics deployment is that organizations were able to leverage the predictive and analytical capabilities of big data analytic to create or discover new knowledge about their marketplace and its participants. Similarly, the decision-making capabilities of big data analytics are instrumental in improving organizational ability to effectively apply accumulated knowledge to create greater business value. Regarding BDA influence on cyber agility, the enhanced KM capabilities, especially new knowledge acquired about potential threats, allow an organization to react more promptly to the changing cyber landscape. Overall, the KM capabilities enhanced by BDA functionalities stimulate cyber agility, resulting in improved cyber protection for valuable information assets.

The survey instrument provided to the respondents gave them the opportunity to share additional comments about any major challenges prior to the decision to deploy BDA as well as the chance to highlight the benefits of implementing big data analytics to improve their company's cyber capabilities. Below are some of their comments:

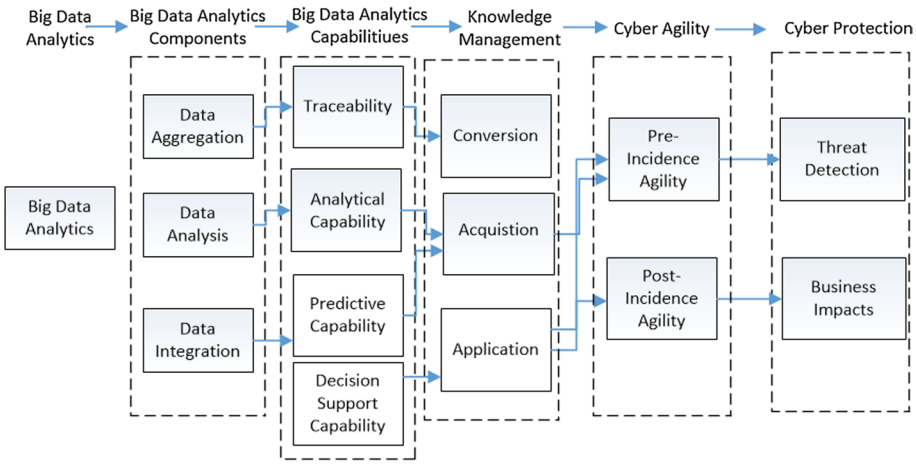

Fig. 2 Results of the big data analytics-enabled transformation model

Business benefits or improvements

The data we use for analytics purposes resides on a platform that can perform advanced modeling, machine learning, and visualization on the data in-place, as well as supporting traditional dashboarding and relational data representations.

Our big data platform provides us improved performance, enabling faster processing and analysis of data spanning long time periods or wide geographic areas.

All major sources of our security-relevant data are accessible via the analytics platform and robust processes are in place to ensure data quality.

Our data science/big data tools allow us to effectively cross-references security log data with relevant metadata from external systems such as asset inventories, company directory, etc. and can model the same entities (users, machines, etc.) across multiple data sets.

With the data analytics solutions, data is now being presented to analysts using advanced visualization tools such as Spotfire and modeled in more intuitive forms such as graph/network data representations rather than strictly tabular.

The advanced models utilizing machine learning is being applied to effectively identify quickly-changing modern cyberthreats that slip through traditional rules.

The predictive analytics/big data technologies are seamlessly integrated into the routine workflows for incident investigation, monitoring, and threat intelligence for a wide variety of use cases.

Challenges without BDA

Over 2 billion log events per day are being captured in our proprietary legacy systems based on relational databases, but the traditional knowledge management/

analytics capabilities are limited with out of the box dashboards, i.e. QRADAR.

Capacity issues—systems not performant with current data volumes.

Low accuracy and completeness of data in the SIEM limits the effectiveness of investigation and analysis.

Lack of data normalization and incorporation of reference data makes global and cross-system analyses of larger data sets infeasible.

Data is available to users in tabular format with limited visualization capabilities.

Modeling is limited to the application of deterministic rules and pattern matching.

Implication for practice

From the analytical results, there are four main streams of findings. First, the paper contributes to limited research in the capabilities of big data analytics to stimulate organizational ability to achieve competitiveness in today's global economy. While a significant number of studies on BDA has focused almost entirely on technological perspectives, this paper, drawing from PBV perceptive, provides some empirical support that suggests positive impacts of big data analytics capabilities towards organizational ability to effectively compete in the marketplace. The findings indicate that firms need to continuously nurture and develop superior firm-wide IT capability, such as big data analytics, to successfully manage and leverage organizational information assets to achieve expected business benefits.

Second, the paper contributes to the knowledge management literature by exploring the impacts of big data analytics' advanced analytics capabilities on knowledge management. To the best of my knowledge, this study is the first to investigate this perspective and, in doing so, contributes to and complements the current literature on knowledge management. The results of the analysis suggest that knowledge management alone might not be sufficient to deliver expected organizational capabilities. Therefore, the advanced analytical functionalities of big data analytics are necessary to enhance an organization's ability to leverage its knowledge assets to produce optimal market strategies and make quality decisions. Specifically, the results imply that big data analytics can influence the three knowledge management elements of acquisition, conversion, and application to deliver business value.

Third, this paper adds to the body of research that outlines the ways in which an organization can achieve meaningful improvements in organizational agility, especially relating to the ever-changing cyber landscape. To the best of my knowledge and ability, this author empirically investigated and demonstrated the benefits of implementing big data analytics to improve organizational cyber agility. The results demonstrate that organizations that deploy and leverage BDA achieve a greater capability to rapidly and proactively detect and response to cyber threats. The results also indicate that all dimensions of the NIST cyber framework—efforts to identify, detect, respond, protect and

recover valuable information assets—can be significantly improved with the deployment and use of big data analytics capabilities.

Last, this paper adds to the limited number of studies that have examined how organizations are leveraging big data analytics to improve cybersecurity capabilities. These studies find some evidence that the predictive and analytic capabilities of big data analytics are instrumental to firms' ability to effectively leverage cybersecurity technologies to protect valuable information assets.

Practical implications

This research provides several practical implications for organizations as they build capabilities to compete effectively in the new economy. First, the findings demonstrate the importance of big data analytics to business competitiveness. The results demonstrate that firms that adopt and utilize big data analytics capabilities are more likely to achieve improved cybersecurity capability than firms without it. The findings also suggest that while knowledge management might contribute to superior business performance and improve cyber protection, the proper capabilities need to be present in an organization to achieve intended results. The paper considers big data analytics as one of such capabilities. In particular, these findings suggest to managers and executives that an absence of these critical IT infrastructures might lead to failure of the program to transform organizations through knowledge management capabilities. Although the results of this research are unable to address all of the potential obstacles that managers may face in their quest to create knowledge-based organizations, they do imply that certain firms that deploy big data analytics capabilities will have a better chance of achieving expected returns on investment regarding their IT infrastructures.

Second, the results offer practical guidance to business and IT executives who are engaged in implementing big data analytics. One of the benefits of big data analytics is that it provides a capability to improve the decision support processes. However, to effectively leverage this capability, managers and employees should possess essential analytical skills to accurately interpret the advanced reports and results generated by BDA models. An incorrect interpretation of the BDA outputs could lead to serious errors of judgment and questionable decisions. Thus, it is imperative for organizations to ensure that the employees possess the required skills and to provide analytical training for employees who are lacking in areas such as business analytics, data mining, and basic statistics.

Conclusion

The rapid rise of the Internet and the digital economy has fueled the exponential growth in data that are being captured and stored by firms, and organizations are struggling to effectively manage and analyze these increased volumes of information. The reason organizations are collecting and storing more data than ever before is because business objectives depend on maintaining a competitive edge. Big data analytics is, therefore, a way to extract value from these huge volumes of information, a process that has been demonstrated to be instrumental in discovering and entering new markets as well as maximizing customer retention [49].

The capabilities of IT to deliver expected business value has been an enduring research question. The elusive link between IT, specifically knowledge management enablers, and firm performance calls for further research into intermediate organizational variables through which IT may influence firm performance. The primary results suggest that big data analytics capability enhances knowledge management process, and together they enable both pre- and post-incidence cyber agility. The findings also reveal that organizations can leverage the combination of knowledge management and agility to better and quickly respond to changes in a global threat landscape. Using resource-based view as a methodological framework, this paper provides some evidence that organizations that have deployed big data analytics are more likely to improve cyber protection than organizations without big data analytics capability. In particular, the findings suggest that big data analytics functionalities appear to be instrumental to the improvement in KM capabilities in an organization, which in turn, enhances cyber ability, ultimately contributing to stronger overall asset protection.

Abbreviations
BDA: big data analytics; KM: knowledge management; RBV: resource-based view; IT: information technology.

Acknowledgements
Special thanks to Dr. Chang Koh, Dr. Bendavid Itzhak and Dr. René M. Stulz for their useful suggestions.

Authors' contributions
The author read and approved the final manuscript.

Authors' information
Oluseyi Peter Obitade, MBA, Ph.D. PMP, CISM is a visiting Assistant Professor of Finance and MIS at University of St. Thomas, Houston, Texas. He received his Ph.D. in Information Systems from the University of North Texas, Denton, Texas. His research interests include enterprise resource planning (ERP), digital transformation, cybersecurity, cloud computing, accounting information system, project management, and escalation of commitments. Currently working on three independent papers both in MIS and Finance while collaborating with two professors on two additional papers. He has presented academic papers in two major finance conferences and taught seven undergraduate and MBA courses in finance and management information systems (MIS). In addition to his academic accomplishments, he is also an innovative, result-oriented management professional with over 20 years of professional experience as IT Director, Senior Program Manager and Global Information Security Manager. He has managed medium to large IT systems deployment projects including a $54 million Cybersecurity and Insider threat optimization program and $41 million global ERP business system implementation project for Fortune 500 companies including Chevron Corp, Sony Pictures Entertainments, GE, Kellogg, IBM and Johnson & Johnson. He has also managed IT departments and projects in 9 countries—USA, Australia, Canada, UK, Brazil, Argentina, Mexico, Singapore and China. He is currently a Senior Information Technology Manager (IT) at Chevron Corporation, Houston, Texas.

Funding
No external funding.

Availability of data and materials
Due to anonymities and confidentiality granted to the participants in the survey, the paper material or raw data will not be provided.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The author declare that he has no competing interests.

Appendix

See Table 7.

Table 7 Construct and indicator items with the supporting research

	Construct and indicator items	Supporting research
	KM acquisition process	
ACQ1	Has processes for acquiring knowledge about our customers.	Gold, Malhotra & Segars, 2001
ACQ2	Has processes for acquiring knowledge about our suppliers.	Gold, Malhotra & Segars, 2002
ACQ3	Has processes for exchanging knowledge with our business partners.	Gold, Malhotra & Segars, 2004
ACQ4	Has processes for acquiring knowledge about competitors within our industry	Gold, Malhotra & Segars, 2006
	KM conversion	
COV1	Has processes for converting knowledge into the design of new products and services	Gold, Malhotra & Segars, 2011
COV2	Has processes for converting competitive intelligence into plans of action	Gold, Malhotra & Segars, 2012
COV3	Has processes for transferring organizational knowledge to individuals.	Gold, Malhotra & Segars, 2014
COV4	Has processes for absorbing knowledge from business partners into the organization	Gold, Malhotra & Segars, 2016
COV5	Has processes for distributing knowledge throughout the organization.	Gold, Malhotra & Segars, 2017
	KM application	
APP1	Has processes for using knowledge in development of new products/ service	Gold, Malhotra & Segars, 2022
APP2	Has processes for using knowledge to solve new problems.	Gold, Malhotra & Segars, 2023
APP3	Uses knowledge to adjust strategic direction.	Gold, Malhotra & Segars, 2025
APP4	Is able to locate and apply knowledge to changing competitive condition	Gold, Malhotra & Segars, 2026
APP5	Makes knowledge accessible to those who need it.	Gold, Malhotra & Segars, 2027
APP6	Quickly applies knowledge to critical competitive needs.	Gold, Malhotra & Segars, 2029
	Pre-cyber incidence agility	
PRC1	We identify and detect cyber anomalies faster than previous	
PRC2	The speed of identifying cyber threats has improved	
PRC3	We identify vulnerabilities faster	
PRC4	We are able to respond faster to changing threat landscape	
	Post-cyber incidence Agility	
POC1	Our speed of responding to potential threat has increased	
POC2	Our incidence recovery rate has improved	
POC3	The time to recover from an incidence has gone down	
	Improved cyber incidence detection and resolution	
CID1	We have reduced the amount of time to collect and analyze data that are relevant to investigations	
CID2	We have reduced amount of time it takes to resolve cyber incidences	
CID3	We have reduced risk of cybertheft by reducing the number of false positives	
CID4	Our threat detection rate has gone up	
	Reduced business impacts	
RBI1	Cyber Incidences has less impacts on our productivity	
RBI2	Financial loss from cyber incidence has reduced considerably	
RBI3	Risk to theft of intellectual property by cyber criminals have dropped	
RBI4	Overall the business impacts of breaches have reduced significantly	
	Big Data Analytics Adoption measurements	
BDA1	We have access to very large, unstructured, or fast-moving data for analysis	Davenport (2014)
BDA2	We integrate data from multiple internal sources into a data warehouse or mart for easy access	Davenport (2014)
BDA3	We have explored or adopted parallel computing approaches (e.g. Hadoop) to big data processing	Davenport (2014)
BDA4	We have explored or adopted different data visualization tools	Davenport (2014)

Table 7 (continued)

	Construct and indicator items	Supporting research
BDA5	We have explored or adopted new forms of databases such as Not Only SQL (NoSQL) for storing data	Gordon (2007)
BDA6	We have explored or adopted open source software for big data analytics	Davenport (2014)

Received: 26 April 2019 Accepted: 16 July 2019
Published online: 03 August 2019

References

1. Eastman R, Versace M, Webber A. Big data and predictive analytics: on the cybersecurity front line. IDC Whitepaper, February. 2015.
2. Beyer MA, Laney D. The importance of 'Big Data': a definition. Stanford: Gartner; 2012.
3. El Sawy OA. Personal information systems for strategic scanning in turbulent environments: can the CEO go on-line? MIS Q. 1985;9(1):53–60.
4. Boyd D, Crawford K. Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. Inf Commun Soc. 2012;15(5):662–79.
5. Villars RL, Olofson CW, Eastwood M. Big data: what it is and why you should care. IDC White Paper. Framingham: IDC; 2011.
6. Njilla LL, Kamhoua CA, Kwiat KA, Hurley P, Pissinou N. Cyber security resource allocation: a Markov decision process approach. In: 2017 IEEE 18th international symposium on high assurance systems engineering (HASE). New York: IEEE; 2017. p. 49-52.
7. Cox M, Ellsworth D. Managing big data for scientific visualization. ACM Siggraph. 1997;97:21–38.
8. Gandomi A, Haider M. Beyond the hype: big data concepts, methods, and analytics. Int J Inf Manag. 2015;35(2):137–44.
9. Davenport T. Big data at work: dispelling the myths, uncovering the opportunities. Brighton: Harvard Business Review Press; 2014.
10. Power D. Decisions support systems: concepts and resources for managers. Westport: Quorum Books; 2002.
11. Michael K, Michael M. The social and behavioral implications of location-based services. J Locat Based Serv. 2011;5:121–37.
12. Choi H, Varian H. Predicting initial claims for unemployment benefits. Google Inc. 2009. p. 1–5
13. Prahalad CK. In volatile times, agility rules, September 21, 2009. p. 80.
14. Goldman SL, Nagel RN, Preiss K. Agile competitors and virtual organizations: strategies for enriching the customer. New York: Van Nostrand Reinhold; 1995.
15. Mahmood T, Afzal U. Security analytics: Big data analytics for cybersecurity: a review of trends, techniques and tools. In: 2013 2nd national conference on information assurance. New York: IEEE; p. 129–34.
16. Hult F, Sivanesan G. What good cyber resilience looks like. J Bus Contin Emerg Plan. 2014;7(2):112–25.
17. Barney JB, Ketchen DJ, Wright M. The future of resource-based theory revitalization or decline? J Manag. 2011;37:1299–315.
18. Barney J. Firm resources and sustained competitive advantage. J Manag. 1991;17(1):99–120.
19. Palmatier RW, Dant RP, Grewal D. A comparative longitudinal analysis of theoretical perspectives of interorganizational relationship performance. J Market. 2007;71:172–94.
20. Makadok R. Toward a synthesis of the resource-based and dynamic-capability views of rent creation. Strateg Manag J. 2001;22:387–401.
21. Melville N, Kraemer K, Gurbaxani V. Information technology and organizational performance: an integrative model of IT business value. MIS Q. 2004;28(2):283–322.
22. Gu JW, Jung HW. The effects of IS resources, capabilities, and qualities on organizational performance: an integrated approach. Inf Manag. 2013;50:87–97.
23. Wang Y, Kung L, Wang WYC, Cegielski CG. An integrated big data analytics-enabled transformation model: application to health care. Inf Manag. 2018;55(1):64–79.
24. Loebbecke C, Picot A. Reflections on societal and business model transformation arising from digitization and big data analytics: a research agenda. J Strateg Inf Syst. 2015;24(3):149–57.
25. Chae B, Sheu C, Yang C, Olson D. The impact of advanced analytics and data accuracy on operational performance: a contingent resource-based theory (RBT) perspective. Decis Support Syst. 2014;59:119–26.
26. Appleyard MM. How does knowledge flow? Interfirm patterns in the semiconductor industry. Strateg Manag J. 1996;17:137–54.
27. Grant R. A knowledge-based theory of inter-firm collaboration. In: Academy of management best paper proceedings. 1995. p. 17–21.
28. Spender JC. Making knowledge the basis of a dynamic theory of the firm. Strateg Manag J. 1996;17:45–62.
29. Inkpen A. Creating knowledge through collaboration. Calif Manag Rev. 1996;39(1):123–41.
30. Leonard D. Wellsprings of knowledge: building and sustaining the source of innovation. Boston: Harvard Business School Press; 1995.
31. Nonaka I, Takeuchi H. The knowledge creating company: how Japanese companies create the dynamics of innovation. New York: Oxford University Press; 1995.

32. Nahapiet J, Ghoshal S. Social capital, intellectual capital, and the organizational advantage. Acad Manag Rev. 1998;23(2):242–58.
33. O'Dell C, Grayson C. If only we knew what we know: identification and transfer of internal best practices. Calif Manag Rev. 1998;40(3):154–74.
34. Davenport T, DeLong D, Beers M. Successful knowledge management projects. Sloan Manag Rev. 1998;39:43–57.
35. Aiavi M, Leidner DE. Review: knowledge management and knowledge management systems: conceptual foundations and research issues. MIS Q. 2001;25:107–36.
36. Bharadwaj AS. A resource-based perspective on information technology capability and firm performance: an empirical investigation. MIS Q. 2000;24:169–96.
37. Gupta M, George JF. Toward the development of a big data analytics capability. Inf Manag. 2016;53(8):1049–64.
38. Cao G, Duan Y, Li G. Linking business analytics to decision making effectiveness: a path model analysis. IEEE Trans Eng Manag. 2015;62(3):384–95.
39. Ward MJ, Marsolo KA, Froehle CM. Applications of business analytics in healthcare. Bus Horiz. 2014;57(5):571–82.
40. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst. 2014;2(1):3.
41. Delen M. Real-world data mining: applied business analytics and decision making. Upper Saddle River: FT Press; 2014.
42. Marr B. Big data: using SMART big data. analytics and metrics to make better decisions and improve performance. Hoboken: Wiley; 2015.
43. Volberda HW. Toward the flexible form: how to remain vital in hypercompetitive environments. Organ Sci. 1996;7(4):359–74.
44. Lucas HC Jr, Olson M. The impact of information technology on organizational flexibility. J Organ Comput Electron Commer. 1994;4:155–76.
45. Schiavone S, Garg L, Summers K. Ontology of information security in enterprises. Electron J Inf Syst Eval. 2014;17(1):71.
46. MacKenzie SB, Podsakoff PM, Podsakoff NP. Construct measurement and validation procedures in MIS and behavioral research: integrating new and existing techniques. MIS Q. 2011;35:293–334.
47. Gold AH, Malhotra A, Segars AH. Knowledge management: an organizational capabilities perspective. J Manag Inf Syst. 2001;18(1):185–214.
48. Kelly D, Amburgey TL. Organizational inertia and momentum: a dynamic model of strategic change. Acad Manag J. 1991;34(3):591–612.
49. Athey S. How big data changes business management, Stanford business, 2013.

Publisher's Note