Journal of Strategic Information Systems 27 (2018) 101–113

Contents lists available at ScienceDirect

Journal of Strategic Information Systems

journal homepage: www.elsevier.com/locate/jsis

Data analytics competency for improving firm decision making performance

Maryam Ghasemaghaei*, Sepideh Ebrahimi, Khaled Hassanein

DeGroote School of Business, McMaster University, Hamilton, Ontario, Canada

ARTICLE INFO

Keywords:
Data analytics competency
Data quality
Bigness of data
Analytical skills
Domain knowledge
Tools sophistication
Decision making performance

ABSTRACT

This study develops and validates the concept of Data Analytics Competency as a five multi-dimensional formative index (i.e., data quality, bigness of data, analytical skills, domain knowledge, and tools sophistication) and empirically examines its impact on firm decision making performance (i.e., decision quality and decision efficiency). The findings based on an empirical analysis of survey data from 151 Information Technology managers and data analysts demonstrate a large, significant, positive relationship between data analytics competency and firm decision making performance. The results reveal that all dimensions of data analytics competency significantly improve decision quality. Furthermore, interestingly, all dimensions, except bigness of data, significantly increase decision efficiency. This is the first known empirical study to conceptualize, operationalize and validate the concept of data analytics competency and to study its impact on decision making performance. The validity of the data analytics competency construct as conceived and operationalized, suggests the potential for future research evaluating its relationships with possible antecedents and consequences. For practitioners, the results provide important guidelines for increasing firm decision making performance through the use of data analytics.

1. Introduction

The availability of data with enormous volume, velocity, and variety has resulted in a Big Data revolution that has the potential to lead to improved firms' decision making performance with associated competitive advantages (Chen et al., 2012). To that end, data analytics is being increasingly leveraged by many firms to deal with the massive amounts of data they collect and fulfill their growing needs for better and faster decisions (Fernández et al., 2014; Loebbecke and Picot, 2015). Data analytics is a combination of processes and tools, including those based on predictive analytics, statistics, data mining, artificial intelligence, and natural language processing (Russom, 2011), often applied to large and possibly disperse datasets for gaining invaluable insights to improve firm decision making (Ertemel, 2015). Over the past two decades, data analytics has become a critical organizational Information Technology (IT) competency due to the increased amounts, speed of change, and types of data in business (Kambatla et al., 2014). Firms need to improve their *data analytics competency* (i.e., firm's ability to effectively deploy data analytics-based resources in combination with other related resources and capabilities) to make better, more informed, and faster decisions.

There is some evidence that using data analytics tools can help organizations improve their decision making performance. However, recent studies found that many firms that invested in data analytics could not take full advantage of using these tools. For example, in a recent report, only 25% of firms reported that analytics has "significantly" improved their organization's outcomes

* Corresponding author.
E-mail addresses: ghasemm@mcmaster.ca (M. Ghasemaghaei), s.ebrahimi@mcmaster.ca (S. Ebrahimi), hassank@mcmaster.ca (K. Hassanein).

http://dx.doi.org/10.1016/j.jsis.2017.10.001
Received 15 March 2017; Received in revised form 2 August 2017; Accepted 7 October 2017
Available online 22 October 2017
0963-8687/

(Deloite, 2013). Likewise, Colas et al. (2014) found that only 27% of firms that invested in data analytics reported their initiatives as successful. They argue that most firms could not take full advantage of using these tools due to a variety of reasons such as having low quality data, not using appropriate data analytics tools, and the lake of available analytical skills. Akter and Wamba (2016) also argue that while using data analytics has great potentials for improving firms' outcomes, organizations need to address various challenges in order to reap the benefits. Ghasemaghaei et al. (2017) argue that not all companies investing in data analytics can improve their decision making and different firm resources may play critical roles in successfully using these tools. In addition, Wu et al. (2016) argue that the high failure rate of successfully using data analytics could be due to the fact that often the necessary required conditions to generate insights from data analytics are neglected and most firms only focus on data aspects (e.g., data volume) to generate insight. Hence, it is critical to empirically explore the characteristics of more successful data analytics initiatives. Given the growth in the use of data analytics and the mixed outcomes it has obtained, this study focuses on conceptualizing, operationalizing and validating the concept of *data analytics competency* and strives to understand the impact of data analytics competency on firm decision making performance (i.e., decision quality, and decision efficiency). Defining and validating data analytics competency and its impact on firm decision making performance will offer new insights into the IT competency literature.

In recent years, firms are increasingly in possession of rich sets of data on their customers, businesses, markets and environments, which has been collectively called "Big data" (a concept indicating data that is high in volume, variety and velocity). Furthermore, the ultimate value and result of data analytics is greatly affected by the quality of the data used (Kwon et al., 2014). Without having high quality data (e.g., timely and relevant data) the improvement in firm decision making performance as a result of using data analytics could be impeded (Ghasemaghaei et al., 2016; Sukumar and Ferrell, 2013). Similarly, if users of data analytics do not have proper analytical skills and domain knowledge, decision making performance improvement within firms could be hampered (Waller and Fawcett, 2013). Likewise, if firms do not use sophisticated analytical tools that provide real-time insight, firm decision making performance could be impeded (Davenport, 2013). Hence, and consistent with Bharadwaj's (2000) framework which classified key IT-based resources as IT infrastructure, human IT resources, and IT-enabled intangibles, in the context of this study, data analytics competency is categorized as tools sophistication (IT infrastructure), employee analytical skills and domain knowledge (human IT resources), and data quality and bigness of data (IT-enabled intangibles). While it has been suggested that firms could improve their decision making performance through being competent in the use of data analytics, no studies have conceptualized, operationalized and validated the concept of data analytics competency.

In order to address the gaps identified above, we draw on Bharadwaj's (2000) key IT-based resources framework and Huber's (1990) theory of effects of advanced IT on decision making (Huber's theory) to pursue the following objectives: (i) define and validate the data analytics competency as a multidimensional formative index; and (ii) develop and validate a model to understand the impact of data analytics competency on firm decision making performance (i.e., decision quality, and decision efficiency). Both of the above endeavors are novel aspects not previously considered in the IS literature. Therefore, this study tries to address the following research questions: (1) what are the most critical firm resources that form data analytics competency? and (2) Whether and to what extent does data analytics competency impact firm decision making performance?

2. Deriving the theoretical framework

The Resource-Based View (RBV) of the firm posits that organizations compete on the basis of unique firm resources that are rare, difficult to imitate, and valuable (Barney, 1991). Grant (1991) classified firm resources into tangible, intangible, and personnel-based resources. Tangible resources encompass the physical assets of the firm, while intangible resources include assets such as product quality, and personnel-based resources include technical know-how and employee training. Firms create competitive advantage by assembling their resources to work together to generate organizational capabilities. Competency, thus, refers to a firm's ability to integrate, assemble, and deploy valued resources (Prahalad and Hamel, 2006).

Adopting the RBV of the firm, IS researchers have identified various IT related resources that serve as potential sources of competitive advantage (Seddon, 2014; Tallon, 2008; Wamba et al., 2017; Watjatrakul, 2005). Drawing on Grant's classification scheme for firm resources, Bharadwaj (2000) classified key IT-based resources as: (1) the tangible firm resource consisting the physical IT infrastructure components, (2) the human IT resources comprising the managerial and technical IT skills, and (3) the intangible IT-enabled resources such as knowledge assets.

The physical IT infrastructure, which comprises the computer and communication technologies, has been described as a critical firm resource in obtaining long-term competitive advantage (Bharadwaj, 2000). Viewed from the RBV perspective, the IT infrastructure is considered as a firm resource that makes feasible innovation within firms (Duncan, 1995; Venkatraman, 1991). Indeed, IT infrastructures that enable firms to obtain, process, and share information in real time, and detect opportunities and threats in a timely manner represent an invaluable firm resource (Reed and DeFillippi, 1990) that are central to the resource-based view.

Organizational human resources generally encompass the training, relationships, experience, and insights of its employees (Barney, 1991). Bharadwaj (2000) suggests that the critical dimensions of human IT resources include: (1) technical IT skills, such as systems analysis and design, programming, and competencies in emerging technologies, and (2) the managerial IT skills, which include abilities to effectively interact and coordinate with user community. Firms with strong human IT resources are able to work and communicate with business units more efficiently, integrate the business and IT planning processes more effectively, anticipate future business needs of the firm, and conceive of and develop cost effective and reliable applications that support the business needs of the firm faster than competitors (Bharadwaj, 2000). In addition, the successful use of IT systems and the managerial ability to coordinate the multifaceted activities have been found to be main distinguishing factors of successful firms (Sambamurthy and Zmud, 1992). Viewed from a resource-based perspective, human IT resources are difficult to obtain and complex to imitate; therefore, they serve as a source of competitive advantage. Indeed, the wide difference in economic benefits that companies gain from IT has been largely attributed to differences in their human IT resources (Mata et al., 1995).

A main contribution of the resource-based view is its explicit recognition of the value of intangible firm resources (Bharadwaj, 2000). In general, firm-specific intangibles such as product quality, and knowledge assets, tend to be idiosyncratic, and tacit (Winter, 1998). While firm intangible resources serve as the basic units of analyses, firms create competitive advantage by assembling resources that work together to create firm competencies (Bharadwaj, 2000).

Adopting the Bharadwaj's (2000) key IT-based resources framework, in this study, data analytics competency is categorized as tools sophistication (IT infrastructure), employee analytical skills and domain knowledge (human IT resources), and data quality and bigness of data (IT intangibles). We define data analytics competency as a firm's ability to deploy and combine data analytics resources for rigorous and action-oriented analyses of data. In the following paragraphs, the identification of data analytics competency created by the interaction of bigness of data, data quality, analytical skills, domain knowledge, and tools sophistication are explicated.

Bigness of data refers to the increasing availability of data that provides the impetus for the use of data analytics[1] (Lycett, 2013). Considering the emerging nature of "big data", several definitions of the concept currently exist. Some practitioners and scholars use the notion of 'Vs' to define 'big Data'. For example, Douglas (2001) proposed a three fold definition of big data encompassing the "three Vs": volume, variety, and velocity, which has been supported by many other studies (e.g., Kwon and Sim, 2013; Lycett, 2013; McAfee and Brynjolfsson, 2012; Raghupathi and Raghupathi, 2014; Russom, 2011). Volume is defined as the amount of data, which is increasing significantly. The widespread use of smart devices and increasing digitization of content is resulting in increasing volumes of data (Newell and Marabelli, 2015; Rusitschka et al., 2014). It is estimated that by 2020, the data created will be 50 times more than the amount of data in 2011 (Ertemel, 2015). Variety refers to the many sources and types of data. Firms are now dealing with structured data (e.g., numbers, dates), semi-structured data (e.g., XML documents), and unstructured data (e.g., videos, social media data) from within and outside the organization (Abbasi et al., 2016; Li et al., 2008). Velocity refers to the speed at which the data is created. In the context of big data, data is created at an enormous speed which is almost in real-time. For example, in 2015, 100 h of video were uploaded to YouTube every minute (Ertemel, 2015).

In addition to the three V's, other dimensions of big data have also been mentioned. For example, value has emerged as another characteristic of big data, which focuses on the importance of extracting economic benefits from big data (Dijcks, 2012; Iview, 2012; Wamba et al., 2015). In addition, White (2012) and Wamba et al. (2015) suggested 'Veracity' as another characteristic of big data, which highlights the importance of quality of the data and the level of trust in various data sources. Other scholars added "Variability" as another dimension of big data which considers the inconsistencies of the data flow (Katal et al., 2013). However, most studies argue that the three Vs of big data (i.e., volume, variety, and velocity) constitute its primary characteristics (Chen et al., 2015; Russom, 2011; Ward and Barker, 2013). Lam et al. (2017) argue that while volume, variety, and velocity are the primary characteristics of big data, veracity and value could be considered as the endogenous variables of big data. Particularly, they state that as the volume, velocity, and variety of data flowing into firms increase, more knowledge about internal operations is needed to verify data accuracy. However, as most firms are unable to verify the accuracy of all their data, they employ sampling procedures and use only a small subsample of information for making decisions. Hence, data with high velocity, variety, and volume increases the chances of data inaccuracy. In addition, Lam et al. (2017) argue that having access to data that is high in terms of volume, variety, and velocity allows firms to make real-time adjustments to their offerings and interact with their customers on a continuous basis, which lead to increases in extracting economic benefits from the available data (i.e., data value).

De Mauro et al. (2015) suggest that although some studies extended the "3 V's" model and introduced multiple features of Big Data, such as Value (Dijcks, 2012), and Veracity (Schroeck et al., 2012), the nucleus of the concept of Big Data can be expressed by its volume, variety, and velocity. Therefore, in line with the aforementioned research (e.g., Lycett, 2013; McAfee and Brynjolfsson, 2012; Raghupathi and Raghupathi, 2014), in this study we define bigness of data as data that is high in volume, variety, and velocity.

Big data, as one of the critical organizational intangible resources, with its focus on volume, variety, and velocity is an influential strategic resource to discover unforeseen patterns and to develop sharper insights about businesses, customers, markets and environments to make better firm decisions (Fernández et al., 2014). Hence, based on Bharadwaj's (2000) key IT-based resources framework, bigness of data, an intangible IT-enabled resource, is considered as a dimension of data analytics competency.

Data quality is defined as the quality of raw facts that reflect the characteristics of an entity or event (Detlor et al., 2013). Wang et al. (1996) developed a framework conceptualizing the underlying aspects of data quality. This framework consists of four data quality categories, including intrinsic, contextual, representational, and accessibility, and the underlying dimensions for each category. The intrinsic category refers to the innate correctness of data regardless of the context in which it is being used such as accuracy and objectivity of data. The contextual category represents the quality of data that may vary according to the particular task at hand, such as completeness, timeliness, and relevancy of data. The representational category refers to the degree to which the data is presented in a clear manner, such as representational consistency, and ease of understanding. Finally, the accessibility category refers to the ease with which the data sought is obtained.

Notwithstanding the highly acknowledged role of data quality as a critical requirement for making effective decisions, recent report shows that data quality is in fact a main obstacle in having competent data analytics (Hazen et al., 2014). One reason for such concern about data quality in the big data era is the firms' desire for analysis of ever larger amounts of acquired data. Although

[1] Data analytics is a combination of processes and tools that is being leveraged by firms to deal with the massive amounts of data (i.e., big data) they collect and fulfill their growing needs for better decisions (Lam et al., 2017; Wamba et al., 2015).

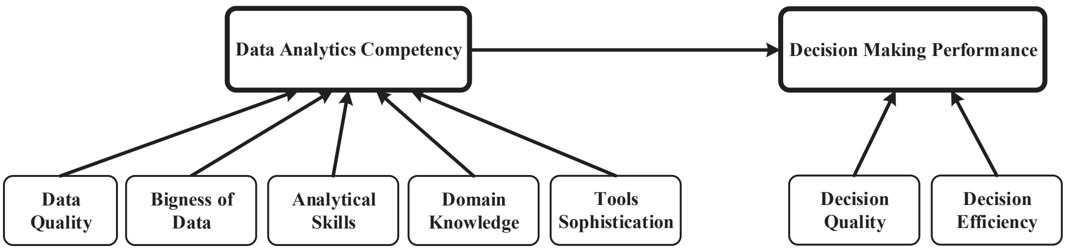

Fig. 1. Proposed research model.

advanced data analytics tools have the potential to identify useful information from data (Philip, 2008), the ultimate results and value of using data analytics will still be affected by the quality of the data used (Lycett, 2013; Popovič et al., 2014). As suggested by Lycett (2013), to obtain valuable business insights and improve firms' decision making performance by employing data analytics, organizations need to use data that is high in quality. Hence, based on Bharadwaj's (2000) key IT-based resources framework, data quality, an intangible IT-enabled resource, is considered as a dimension of data analytics competency.

Employee domain knowledge and analytical skills refer to a combination of knowledge and skills that gives an individual the potential for effectiveness in data analysis task performance (Draganidis and Mentzas, 2006). Having a wealth of data has little value unless it is used to improve firm performance. Many organizations have invested in data analytics that enable transformation of data into business insights, but the human component required for this process has not been given sufficient attention (Waller and Fawcett, 2013). According to RBV, to make firm resources difficult to imitate by competitors, attention to other related firm elements (e.g., human resources) is considered as a critical factor (Peteraf, 1993). Hence, in the context of this study, attention to individuals' competencies is a critical factor in improving decision making through the use of data analytics.

More specifically, possessing domain knowledge and applying it to the analysis of interest is a critical aspect of data integration and analysis. The required domain knowledge includes a deep understanding of the procedures, facts, and processes involved in a given firm/industry. Having sufficient domain knowledge enables the analyst to better identify the key attributes and thus to solve business problems of interest to the firm more effectively (Sukumar and Ferrell, 2013). Hence, based on Bharadwaj's (2000) key IT-based resources framework, domain knowledge, a human IT resource, is considered as a dimension of data analytics competency.

Having the right talent and skills to analyze and interpret data are considered as important factors in generating business insights from the use of data analytics which would lead to higher firm decision making performance (Wong, 2012). When analysts have low proper analytical skills on how to perform data analyses, they may want to postpone their tasks, it might take them longer to execute needed analyses, they may make mistakes and, as such, they will not be able to solve the problems at hand (Ghasemaghaei et al., 2015). Hence, drawing on Bharadwaj's (2000) key IT-based resources framework, analytical skills, a human IT resource, is considered as a dimension of data analytics competency.

Tools sophistication refers to the maturity and complexity of the tools (Raymond and Paré, 1992) which captures the level of technological expertise within the organization (Chwelos et al., 2001). Based on how sophisticated the analytical tools are, the depth of analysis could vary in firms (Davenport, 2013). For example, sophisticated analytical tools are able to provide information about past or current events, help firms understand why something happened in the past, provide accurate projections of future happenings, and recommend one or more courses of action and show the likely outcome of each. Thus, sophisticated tools offer more possibilities for firms to generate business insights and to improve firm decision making performance (Petrini and Pozzebon, 2009; Cao and Duan, 2015; Gillon et al. 2012). Hence, drawing upon Bharadwaj's (2000) key IT-based resources framework, tools sophistication, an IT infrastructure resource, is considered as a dimension of data analytics competency.

3. Research model

Fig.1 depicts our proposed research model, illustrating the hypothesized relationship between data analytics competency and decision making performance; "*the higher the organization's level of data analytics competency, the higher the firm decision making performance will be*".

Consistent with the Bharadwaj's (2000) key IT-based resources framework, we argue that the five Data Analytics Competency dimensions (i.e. data quality, bigness of data, analytical skills, domain knowledge, and tools sophistication) are pivotal, whereby they all contribute to overall data analytics competency in the organization. Since each dimension makes a unique contribution to data analytics competency, together they are considered as dimensions forming *data analytics competency*. Data analytics competency is thus conceived and operationalized as a multidimensional, formative index[2]. Particularly, this construct is formed by its respective first-order constructs, which complement each other and combine to form and serve the purpose of data analytics competency. In addition, a change in one of the dimensions of data analytics competency (e.g., analytical skills) does not necessarily cause an equal

[2] A higher (or second)-order construct is either represented or constituted by its dimensions (i.e., lower (or first)-order constructs). Such constructs are modeled as reflective and formative respectively (Becker et al., 2012). Whereas in the former case, the general concept is "manifested" by its dimensions, in the latter case, the formative construct is an aggregate construct that is a composite of its dimensions (Edwards, 2001). In other words, the dimensions combine to produce the construct (Edwards, 2001; Wetzels et al., 2009).

change in other first-order constructs (e.g., bigness of data, tools sophistication). As a result, a formative model is deemed more appropriate than a reflective model (Pavlou and El Sawy 2006).

Huber (1990) developed a theory that proposes that, once adopted, new technology deeply alters the way decisions are made in an organization. Particularly, the theory postulates that advanced technologies will affect organizational intelligence, and decision making, with the final result of the adoption being the shortening of the firm decision making processes and the improvement of the quality of decisions (Carlson et al., 1999). According to Huber's theory, the proper use of IT leads to organizational intelligence that is more accurate, timely, comprehensive, and available with a more rapid identification of firm opportunities and problems (Huber, 1990). Previous studies found that IT capability leads to improved decision making in firms. For example, Lu and Ramamurthy (2011) argue that firms that have high levels of IT capability make more timely and accurate decisions compared to the firms with lower levels of IT capability. Liu et al. (2013) state that routinizing IT applications in firms' business processes enables them to have real-time analysis and insights which provide support for operational, and strategy decisions. Likewise, Roberts and Grover (2012) argue that competent IT can quickly route relevant information to firms' employees for making effective decisions in a timely manner. In the context of data analytics, Wamba et al. (2017) and Hagel (2015) argue that data analytics tools are increasingly becoming a critical component of firms' decision-making processes. Brown et al. (2011), Ghasemaghaei et al. (2017), and McAfee and Brynjolfsson (2012) argue that the importance of data analytics tools stems from their ability to help firms make better, more informed and faster decisions. Hence, as improved firm's decision making has been suggested as the eventual goal for data analytics (Ertemel, 2015), in this study, decision making performance is considered as the dependent variable which refers to the users' evaluations of decision quality and efficiency in their decision making process (Huber, 1990; Jarupathirun, 2007). Decision quality focuses on decision outcomes that are high in terms of accuracy, precision, and reliability, while decision efficiency concerns arriving at decisions quickly (Jarupathirun, 2007).

Based on RBV, we argue that data analytics provides strengths for firms, but other firms' elements (e.g., data, people) are also necessary to take full advantage of data analytics (Ghasemaghaei et al., 2015). Particularly, effectively combining firms' resources together could increase decision making performance, which helps firms provide better services and products rooted in an in-depth knowledge of its customers, markets, and environment resulting in a sustainable competitive advantage (Ferguson et al., 2005). Hence, in this study, data analytics competency is suggested to increase firm decision making performance.

4. Methodology

The proposed research model was tested using a survey of IT managers and data analysts, each representing a different organization. This sampling choice was made since IT managers and data analysts opinions should reasonably and meaningfully reflect relevant organizational-level constructs in both the technology and business domains (Carter et al., 2011) which are the main areas covered in the proposed research model.

4.1. Participants and incentives

A national market research firm administered the survey to IT managers and data analysts whose roles within firms were verified. Participants who reported using data analytics tools were allowed to take the survey. Participants were incentivized by a chance to win 1 of the 5 monthly $1000 prizes that were awarded by the research firm. The draft survey instrument was pilot tested with a sample of data analysts and IT managers. Feedback from the pilot round respondents resulted in minor modifications to survey items. The final survey yielded 151 valid questionnaires. Following Roldán and Sánchez-Franco (2012), the minimum sample size required to detect a medium effect size at a power of 0.80 and alpha of 0.05 would be 91 cases. Thus, our sample satisfies the sample size requirements for the proposed research model.

4.2. Measures

While some of the measurement scales for the constructs were selected from the extant literature, others were developed. Data quality was measured as a formative construct using a 6-item scale adapted from Wang et al. (1996). Domain Knowledge and analytical skills were measured as reflective constructs using 3-item scale adapted from Tippins and Sohi (2003) and a 4-item scale adapted from Bassellier and Benbasat (2004), respectively. Decision making performance was measured as a second order formative construct using 6-item and 2-item reflective scales for decision quality and decision efficiency, respectively, adapted from Jarupathirun (2007). Bigness of data and tools sophistication constructs were developed as formative constructs following Moore and Benbasat's (1991) methodology (see Appendix A), as no validated scales to measure these variables exist in the literature. The items and sources are outlined in Appendix B.

Two potentially relevant control variables were also included in the survey. First, we included firm size, operationalized with number of employees, because larger firms may have more resources than smaller ones (Chen et al., 2014) which may affect our results. Second, we included firm industry as another control variable because depending on the type of industry, the importance of improving firm decision making performance could vary in firms.

Table 1
Internal consistency and discriminant validity.

	CR	CA	Analytical skills	Domain knowledge	Decision quality	Decision efficiency
Analytical Skills	0.90	0.93	**0.83**			
Domain Knowledge	0.91	0.94	0.70**	**0.80**		
Decision Quality	0.93	0.95	0.73**	0.73**	**0.76**	
Decision Efficiency	0.97	0.98	0.65**	0.64**	0.55**	**0.97**

Note: Composite reliability = CR; Cronbach's alpha = CA
All measures were based on seven-point Likert scales ranging from "strongly disagree" (1) to "strongly agree" (7).

5. Data analysis and results

5.1. Sample

The sample consisted of 151 North American IT managers and data analysts representing firms that use data analytics. 10% of participants fell in the 20–30 age group, 49% of the participants fell in the 31–50 age group, 39% of participants fell in the 51–65 age group, and 2% of participants were above 65 years old. 87 (57%) of the participants were female. Participants worked in firms with different sizes ranging from less than 100 employees to more than 5000 employees. Specifically, 5% of the participants worked in firms with less than 100 employees, 42% worked in firms with a size of between 100–1000 employees, 28% worked in firms with a size of between 1000–5000 employees, and 34% of the participants worked in firms with more than 5000 employees. Regarding industry type, 30% of participants worked in manufacturing firms, 48% worked in services firms, 15% worked in financial firms, and 7% worked in utility firms. In the surveyed organizations analytics use range from "not much" to "very often", with a median of "quite often".

5.2. Measurement model

Prior to testing the research model, we evaluated the validity of each of the two second-order formative constructs (i.e., data analytics competency and decision making performance). We first assessed the construct validity and reliability of the first-order reflective constructs (i.e., analytical skills, domain knowledge, decision quality, decision efficiency) by measuring the internal consistency and discriminant validity. To evaluate the measurement item reliability of reflective constructs, the loadings of each measurement item on its intended construct were assessed and compared with the recommended tolerance of 0.70 (Chin, 1998) (see Appendix C). As Table C1 shows, all indicators loaded most highly on their own theoretically assigned construct, and at a minimum threshold of 0.70. Gefen and Straub (2005) suggest that "loadings of the measurement items on their assigned latent variables should be an order of magnitude larger than any other loading" (p. 93) and the difference should be at least 0.10. As shown in Appendix C, this criterion was also met.

To demonstrate the internal consistency of the constructs, the composite reliability and Cronbach's alpha were calculated for each reflective construct. As shown in Table 1, all constructs met the recommended tolerance of being higher than 0.70 (Fornell and Larcker, 1981). In Table 1, the diagonal elements are the square roots of the average variance extracted (AVE) of variables, and the off-diagonal numbers represent the correlation between variables. According to Barclay et al. (1995), to have adequate discriminant validity, the square root of the AVE of a construct must be larger than the correlation between that construct and any other construct. As shown in Table 1, all variable pairs met this requirement. The means and standard deviations for all the constructs in our model are shown in Table B1 (see Appendix B).

For formative constructs, we used the Variance Inflation Factor (VIF) statistic to examine whether the formative measures are highly correlated (Petter et al., 2007). The VIF values of all first-order formative constructs (i.e., data quality, bigness of data, tools sophistication) were below the stringent threshold value 3.3 (Diamantopoulos and Siguaw, 2006). Thus, our measures do not have a multicollinearity problem. To evaluate the measurement properties for the second-order formative construct (i.e., data analytics competency and decision making performance), we followed the formulation suggested by Bagozzi and Fornell (1982). We first multiplied item values by their PLS weights and summed them up for each first-order indicator. Then, based on a weighted sum of the first-order indicators, the second-order variables were assessed by creating composite indices (Diamantopoulos and Winklhofer, 2001). The generated composite index values were used as the measures for data analytics competency and decision making performance. We also used the VIF statistic to examine whether the formative constructs are correlated too highly (Petter et al., 2007). The VIF values of data analytics competency and decision making performance dimensions were below the threshold value 3.3 (Diamantopoulos and Siguaw, 2006). Thus, no multicollinearity issue was observed in our second-order formative constructs.

We also measured the outer model weights, and outer model loadings to further establish convergent and discriminant validity of the model constructs (Gefen and Straub, 2005). Results showed that all the outer model loadings[3] were significant at the 0.05 alpha

[3] The outer loadings of data quality, bigness of data, analytical skills, domain knowledge, and tools sophistication were 0.82, 0.82, 0.91, 0.89, and 0.80, respectively, on data analytics competency. In addition, the outer loadings of decision quality and decision efficiency were 0.87 and 0.98, respectively, on decision making performance.

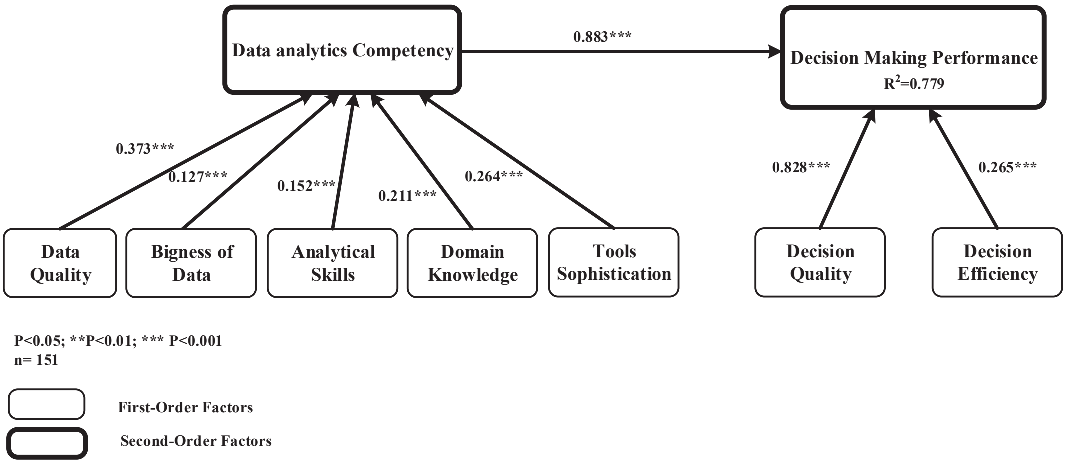

Fig. 2. Results of research model.

level. Moreover, construct reliability was measured by examining the loadings of the manifest variables on their respective dimension. A minimum loading cut-off is to accept dimensions with loadings of 0.70 or more, which indicates that there is more shared variance between the dimension and its manifest variable than error variance (Carmines and Zeller, 1979; Dwivedi et al., 2006; Hulland and Business, 1999). Results showed that all loadings were larger than 0.70, with each dimension contributing significantly to the formation of each construct. The results also showed that the outer model weights of each data analytics competency indexes[4] impacted significantly on this construct which indicates the relative importance of each index in forming data analytics competency.

Additional preliminary analyses were performed for assessing the potential effects of common method bias. These analyses are described in Appendix D. The results of these analyses indicate that there is a low likelihood of a meaningful common method variance component in our data. Hence, assessing the structural model was deemed to be appropriate.

5.3. *Structural model*

SmartPLS version 2.0 (Ringle et al., 2005) was used to analyze the structure model. Bootstrapping employing 500 re-samples was used for assessing significance levels. Supporting the main hypothesis, results (Fig. 2) show that data analytics competency is significantly associated with decision making performance (path coefficient (β) = 0.883, $p < .001$); the squared multiple correlation coefficient (R^2) of 0.779 indicates that data analytics competency explains a large amount of variance in the endogenous construct and, thus, is the most important antecedent of decision making performance. This strong relation is not surprising given the comprehensive nature of the first-order constructs forming our formative second-order data analytics competency construct.

The existence of a strong, positive relationship between data analytics competency and decision making performance indicates the validity of both constructs. If either construct is not valid we are unlikely to see this strong relationship (Edwards and Bagozzi, 2000; Diamantopoulos and Winklhofer, 2001; Sedera and Gable, 2010). This further evidence of construct validity is sometimes referred to as ‘identification through structural relations’ (Jarvis et al., 2003).

6. Post hoc analyses

6.1. *Impact of each data analytics competency index on decision quality and decision efficiency*

To better understand the impact of data analytics competency on firm decision making performance, a post hoc analysis was performed to observe the direct impact of the five data analytics competency indexes on decision quality and decision efficiency separately. As shown in Fig.3, results revealed strong and significant path coefficients for the impact of all data analytics competency indexes on decision quality. However, interestingly, all constructs, except bigness of data, impacted significantly on decision efficiency. This means that while data that is high in volume, variety, and velocity increase firm decision quality, it does not necessarily increase the speed of arriving at decisions.

6.2. *Control variable effects*

We examined whether the impact of control variables (i.e., firm size, and industry type[5]) on dependent variables were significant.

[4] The outer weight of data quality, bigness of data, analytical skills, domain knowledge, and tools sophistication were 0.2, 0.40, 0.36, 0.28, and 0.205, respectively, on data analytics competency. In addition, the outer weight of decision quality and decision efficiency were 0.19 and 0.51, respectively, on decision making performance.

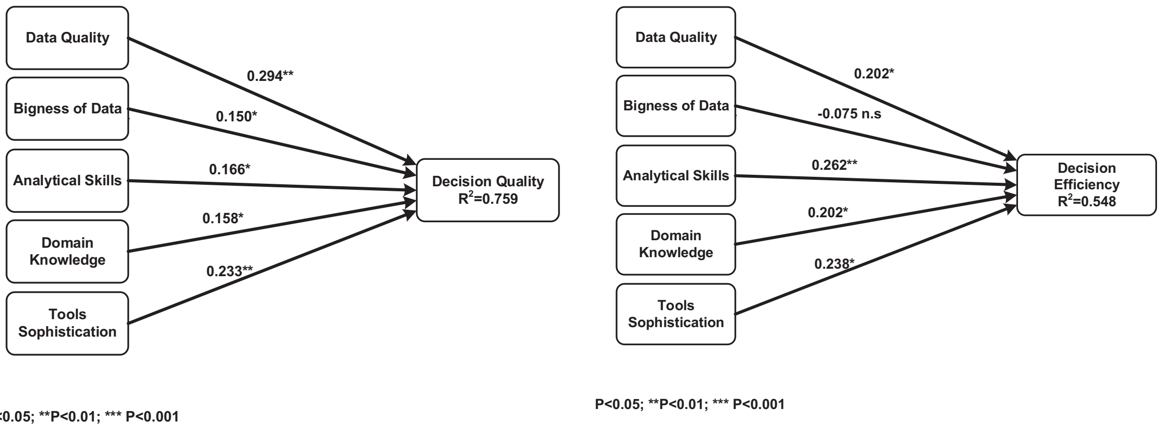

Fig. 3. Impact of data analytics competency indexes on decision quality and decision efficiency.

Table B1
Measurement items of the variables.

Construct Names	Measurement Items (7-point scale)	Mean	SD	Resources
Analytical Skills	7-point Likert scales ranging from "strongly disagree" to "strongly agree": ● Our data analytics users are knowledgeable when it comes to utilizing such tools. ● Our data analytics users possess a high degree of data analytics expertise. ● Our data analytics users are skilled at using data analytics tools.	4.6	1.1	Tippins and Sohi (2003)
Data Quality	7-point Likert scales ranging from "strongly disagree" to "strongly agree":In my organization, data used in data analytics: ● is reliable ● has an appropriate level of details ● is secure ● is timely ● is relevant to the task at hand ● is accurate	4.83	1.0	Wang et al. (1996)
Domain Knowledge	7-point Likert scales ranging from "strongly disagree" to "strongly agree":In my organization, there is a high level of knowledge of the ● External environment (e.g., government, competitors, suppliers, and customers) ● Organizational goals and objectives ● Core capabilities of the organization ● Key factors that must go right for the organization to succeed	4.8	1.2	Bassellier and Benbasat (2004)
Decision Quality	7-point Likert scales ranging from "strongly disagree" to "strongly agree":In my organization, decision outcomes are often ● accurate ● correct ● precise ● flawless ● error-free ● reliable	4.3	1.1	Jarupathirun (2007)
Decision Efficiency	7-point Likert scales ranging from "strongly disagree" to "strongly agree": ● In my organization, the time to arrive at decisions is fast. ● In my organization, the speed of arriving at decisions is high.	3.9	1.5	Jarupathirun (2007)
Bigness of Data	7-point Likert scales ranging from "strongly disagree" to "strongly agree":My organization processes ● High volumes of data. ● Real time data. ● Different types of data.	4.8	0.9	Developed
Tools Sophistication	7-point Likert scales ranging from "strongly disagree" to "strongly agree":In my organization, we use tools that ● Provide information processing and retrieval capabilities ● Perform modeling and simulation ● Perform natural language analytics (extracting information from unstructured sources such as social media) ● Provide real-time insight ● Identify problems ● Evaluate different alternatives	4.5	1.0	Developed

[5] Firm size was coded as a dummy variable (i.e., 1 for less than 100 employees, 2 for between 100–500 employees, 3 for between 501–1000 employees; 4 for

Table C1
Loading and Cross Loading of Measures.

	Analytical skills	Decision efficiency	Decision quality	Domain knowledge
AnalyticalSkills1	**0.90876**	0.59143	0.64969	0.61602
AnalyticalSkills2	**0.89334**	0.61179	0.67072	0.66885
AnalyticalSkills3	**0.93769**	0.58128	0.69010	0.66291
DecisionEfficiency1	0.62907	**0.98663**	0.52911	0.62154
DecisionEfficiency2	0.65590	**0.98777**	0.56714	0.65367
DecisionQuality1	0.66917	0.48977	**0.89047**	0.68713
DecisionQuality2	0.68492	0.50070	**0.88108**	0.66971
DecisionQuality3	0.66926	0.51011	**0.88795**	0.69198
DecisionQuality4	0.59027	0.43096	**0.85087**	0.55547
DecisionQuality5	0.55104	0.41162	**0.79447**	0.51792
DecisionQuality6	0.68562	0.54322	**0.90226**	0.70637
DomainKnowledge1	0.66826	0.54073	0.59383	**0.83757**
DomainKnowledge2	0.58082	0.59214	0.61579	**0.88630**
DomainKnowledge3	0.62615	0.60399	0.67845	**0.92073**
DomainKnowledge4	0.66995	0.57667	0.69001	**0.93201**

Results showed that firm size did not significantly influence firm decision making performance ($\beta = -0.084$; $p > .05$). Similarly, industry type did not significantly impact firm decision making performance ($\beta = -0.04$; $p > .05$). Thus, the effects of the control variables were deemed marginal.

7. Discussion

7.1. Theoretical contributions

There is some evidence that using data analytics tools can help organizations to improve their decision making performance. However, studies reported that many firms that invested in data analytics could not take full advantage of using these tools (Deloite, 2013; Ghasemaghaei et al., 2016). In addition, although the academic and practitioner literature emphasizes the impact of using data analytics tools on firm decision making performance, there is no study that quantitatively provides evidence for the impact of effective use of data analytics (data analytics competency) on firm decision making performance. Hence, we sought to examine what makes this effect work. As improving firms' decision making performance is the ultimate goal of data analysis in the world of data analytics, understanding the factors impacting it is a novel addition to the data analytics literature. In this study, we addressed this knowledge gap by relying on the Bharadwaj's (2000) key IT-based resources framework and Huber's (1990) theory of effects of advanced IT on decision making to define and validate data analytics competency for firms as a multidimensional formative index; and develop and validate a model to understand the impact of data analytics competency on firm decision making performance (i.e., decision quality, and decision efficiency). Both of the above endeavors are novel aspects not previously considered in IS literature.

The study conceptualizes, validates, and operationalizes the notion of data analytics competency as a second-order formative construct. Formative construct validation test results support our conceptualization of data analytics competency as a formative index comprised of the five factors: data quality, bigness of data, analytical skills, domain knowledge, and tools sophistication. As discussed earlier, evidence of its identification through measurement relations is very strong; the strong predicted relationship observed with decision making performance (identification through structural relations) further supports the validity of the data analytics competency construct. The high R^2 for the endogenous variable in our model indicates that data analytics competency is the most important antecedent of decision making performance.

In further analyses, we found that somewhat surprisingly, although the bigness of data significantly increases firm decision making quality, it does not significantly influence firm decision efficiency. This means that while bigness of data is a valuable resource within organizations to improve firm decision quality, it indeed does not enhance the speed of arriving at firm decisions quickly. This might be due to the complexities in collecting, managing and analyzing big data (Jagadish et al., 2014). All other first-order constructs (i.e., data quality, analytical skills, domain knowledge, and tools sophistication) significantly impacted both decision quality and decision efficiency.

In summary, this paper represents the first empirical study that examines various dimensions of data analytics competency by drawing upon Bharadwaj's key IT-based resources framework. The nomological validity of these dimensions was confirmed through an empirical study based on a survey approach involving a relevant sample of IT managers and data analysts. In addition, this research contributes to the growing body of research on IT competencies by exploring the most critical factors that form IT competency within firms. Particularly, the study focuses on data analytics competency as over the past two decades, data analytics has become a critical organizational IT competency owing to the increased amounts of data in business, advancements in data analytics

(*footnote continued*)
between 1001–5000 employees, and 5 for more than 5001 employees); industry type was coded as a dummy variable (i.e., 1: manufacturing firms; 2: services firms; 3: financial firms; 4 utility firms).

tools and techniques, and the growing need for better, more informed, and faster decisions. Moreover, this study also draws on Huber's (1990) theory of effects of advanced IT on decision making to advance the IT competency literature by understanding the most critical factors that impact decision effectiveness and decision efficiency. We believe that the results of our study offer useful guidance to researchers interested in the antecedents and consequences of data analytics competency.

7.2. Practical contributions

The results of this study also have important implications for managers engaged in using data analytics to gain competitive advantage. Improving decision making performance through the use of data analytics is the main motivation for organizations making significant investments in these technologies. Hence, having explained a large amount of the variance in decision making performance, managers need to pay particular attention to improving data analytics competency dimensions if they are to enhance firm decision making performance as a result of using these tools. Without such competency, the implementation of analytical tools may fail to improve decision making performance within firms.

To improve firm decision making, firms could, for example, embark on training to improve employees' analytical skills; when employees have the skills needed to fulfill their job demands, the results of their work are improved (Ayyagari et al., 2011). Likewise, managers need to ensure that employees who are making firm decisions through the use of data analytics have sufficient domain knowledge in order to accurately use the tools and interpret the results (Waller and Fawcett, 2013). Furthermore, managers can employ thorough selection processes when acquiring data analytics tools (Goodhue, 1988) to ensure that the selected tools is sophisticated enough to support all data elements needed for executing known and future analytical tasks. To increase data analytics competency, managers also need to invest in data quality to reduce the information processing time and to improve firm decision making effectiveness (Katal et al., 2013). As discussed, the availability of data with enormous volume, velocity, and variety has resulted in a Big Data revolution that has the potential to lead to improved firms' decision making performance (Chen et al., 2012). According to the results of this study, while bigness of data improves firm decision quality, interestingly, it does not impact firm decision efficiency. This could be due to the fact that facing with such large amount of data could increase the time to arrive at decisions within firms. Nonetheless, to improve data analytics competency, managers should use the data that is big in nature. Hence, we can conclude that organizations should be aware of the fact that although employing big data (i.e., data that is high in volume, variety, and velocity) may not enhance their decision efficiency, it increases the quality of their decisions.

In summary, this study demonstrates the unique importance of all five dimensions of data analytics competency, including data quality, bigness of data, analytical skills, domain knowledge, and sophisticated tools as they all significantly contribute to building data analytics competency in firms. In addition, the observed strong positive relationship between data analytics competency and decision making performance and almost all dimensions of data analytics competency on decision making quality and efficiency point to the importance of studying the avenues to enhance the status of each of these dimension in organizations.

7.3. Limitations and future research

As with any research, some limitations should be acknowledged in this study. First, firm decision making performance can be affected by factors other than data analytics competency. Hence, the impact of other factors (e.g., organizational structure, business processes) on firm decision making performance warrants future research. Second, this study was conducted among North American data analysts and IT managers in organizations of different sizes. Given the potential impacts of culture on users' perceptions regarding information technology, future research should examine the associations we study here, in other cultures. Third, the research model is tested using cross-sectional data. To investigate its stability, future studies could re-examine the findings using panel data. Fourth, as the nucleus of the concept of 'big data' can be expressed by its volume, variety, and velocity (Lycett, 2013; Raghupathi and Raghupathi, 2014), in this study we defined the bigness of data as data that is high in terms of these three characteristics. As some studies extended the "3 V's" model and introduced other features of 'big data', like Value, and Veracity (e.g., Wamba et al., 2015), future studies could validate and operationalize the bigness of data construct considering other suggested big data characteristics and examine whether the inclusion of such other V's will significantly impact the results. Finally, this study was not specific in terms of the types of decisions being made. As the relative importance of data analytics competency and its various dimensions may vary by the nature of the decision being made, future research is warranted, focusing to test the influence of data analytics competency and each dimension of it on decision making outcomes in specific contexts (e.g., marketing promotions, recruitment processes).

8. Conclusion

This study addressed an important gap in Information Systems research by providing theoretical evidence of a statistical relationship between data analytics competency and firm decision making performance. We relied on Bharadwaj's (2000) key IT-based resources framework and Huber's (1990) theory of effects of advanced IT on decision making to define and validate the data analytics competency as a multidimensional formative index and to develop and validate a model to understand the impact of data analytics competency on firm decision making performance (i.e., decision quality, and decision efficiency). The results of this study demonstrate the unique importance of all of the five data analytics competency indexes (i.e., data quality, bigness of data, analytical skills, domain knowledge, and sophisticated tools), which make a distinct and significant contribution to data analytics competency. In addition, almost all of the aforementioned five dimensions have significant impact on both decision quality and decision efficiency (with the exception of the impact of bigness of data on decision making efficiency). Such significant impacts along with the observed strong positive relationship between data analytics competency and decision making performance, suggest that improvement in the data analytics competency indexes will result in improving firm decision making performance. These results have significant implications to the practice of using data analytics within organizations and preconditions required to gain advantages in the form of improved decision making performance as a result of such usage.

Acknowledgements

This research is generously supported by the Social Sciences and Humanities Research Council of Canada.

Appendix A

A.1. Construct development process

We followed Moore and Benbasat's (1991) methodology to develop measurement scales for bigness of data and tools sophistication constructs. After conducting the literature review regarding these two constructs, we generated candidate measurement items with the help of four participants (one graduate students, one professor, one associate professor, and one assistant professors in MIS). To further check content validity, the instrument was sent to a panel of graduate students majoring in MIS to obtain their opinions on appropriate items for inclusion. This procedure generated four items for perceived bigness of data and seven items for perceived tools sophistication. The newly created measures for both bigness of data and tools sophistication, together with items for the other constructs, were shuffled into a random order and were given to the judges. This exercise was conducted in two rounds. In the first round, five doctoral students were asked to sort the items into separate categories based on their differences and similarities and to label the underlying constructs for each of the categories. The average inter-judge raw agreement rate was 82 percent. The judges were also asked to provide comments on unclear or ambiguous items. In this process, only a single item of perceived bigness of data and a single item of perceived tools sophistication were identified to be too ambiguous. According to the comments of the judges, the rest of the items were refined and were retained for the next sorting round.

In the second round, another five doctoral students were asked to sort the retained items based on the construct definitions. To ensure that the judges were not forced to fit any item into a predefined category, a "too ambiguous/doesn't fit" category was also included. This round of sorting ended with an average raw agreement rate of 93 percent, indicating a very high reliability (Moore and Benbasat 1991). This process also helped in establishing the discriminant validity of the items. The resulting measures for perceived bigness of data and tools sophistication contained three and six items, respectively (see Appendix B). These measures were further validated in the measurement model assessment.

Appendix B

See Table B1.
Survey items and constructs

Appendix C

See Table C1.

Appendix D

D.1. Common method bias assessment

To measure the presence of Common Method Bias (CMB) the Harman's one factor test was conducted. The unrotated solution to the principal component analysis (PCA) suggested several factors none of which explains the majority of variance suggesting a low likelihood of a substantive common method variance component in our data. We also applied the marker variable technique to further test for CMB (Lindell and Whitney, 2001). A marker variable (gender) was implemented in the research model that was theoretically unrelated to at least one other variable in the study (decision making performance). CMB can be assessed based on the correlation between the marker variable and the theoretically unrelated variable (Malhotra et al., 2006). This value (0.01) was assumed as the method variance, was parceled out from the other correlations, and the analysis was rerun. The results showed no significant difference between the adjusted ones and the original correlation estimates. Given the results of the Harman's one factor test and the market variable test we conclude that common method bias is not substantial in our data and, therefore, is not likely contaminating the results.

References

Abbasi, A., Sarker, S., Chiang, R.H., 2016. Big data research in information systems: toward an inclusive research agenda. J. Assoc. Inform. Syst. 17 (2), 3.

Akter, S., Wamba, S.F., 2016. Big data analytics in E-commerce: a systematic review and agenda for future research. Electron. Markets 26 (2), 173–194.
Ayyagari, R., Grover, V., Purvis, R., 2011. Technostress: technological antecedents and implications. MIS Quarterly 35 (4), 831–858.
Bagozzi, R.P., Fornell, C., 1982. Theoretical concepts, measurements, and meaning. A Sec. Gener. Multivariate Anal. 2 (2), 5–23.
Barclay, D., Higgins, C., Thompson, R., 1995. The partial least squares (PLS) approach to causal modeling: personal computer adoption and use as an illustration. Technol. Stud. 2 (2), 285–309.
Barney, J., 1991. Firm resources and sustained competitive advantage. J. Manage. 17 (1), 99–120.
Bassellier, G., Benbasat, I., 2004. Business competence of information technology professionals: conceptual development and influence on IT-business partnerships. MIS Quarterly 673–694.
Becker, J.-M., Klein, K., Wetzels, M., 2012. Hierarchical latent variable models in PLS-SEM: guidelines for using reflective-formative type models. Long Range Plan. 45 (5), 359–394.
Bharadwaj, A.S., 2000. A resource-based perspective on information technology capability and firm performance: an empirical investigation. MIS Quarterly 24 (1), 169–196.
Brown, B., Chui, M., Manyika, J., 2011. Are you ready for the era of 'big data'. McKinsey Quarterly 4, 24–35.
Cao, G., Duan, Y., 2015. The affordances of business analytics for strategic decision-making and their impact on organisational performance. In: Proceedings of the 19th Pacific Asia Conference on Information Systems (PACIS), Singapore. Retrieved from http://www.pacis-net.org/file/2015/3004.pdf.
Carlson, P.J., Kahn, B.K., Rowe, F., 1999. Organizational impacts of new communication technology: a comparison of cellular phone adoption in France and the United States. J. Global Inform. Manage. (JGIM) 7 (3), 19–29.
Carmines, E.G., Zeller, R.A., 1979. Reliability and validity assessment. Sage Publications Inc, Newbury Park, CA.
Carter, M., Grover, V., Thatcher, J.B., 2011. The emerging CIO role of business technology strategist. MIS Quarterly Executive 10 (1), 19–29.
Chen, D.Q., Preston, D.S., Swink, M., 2015. How the use of big data analytics affects value creation in supply chain management. J. Manage. Inform. Syst. 32 (4), 4–39.
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.
Chen, Y., Wang, Y., Nevo, S., Jin, J., Wang, L., Chow, W.S., 2014. IT capability and organizational performance: the roles of business process agility and environmental factors. Eur. J. Inform. Syst. 23 (3), 326–342.
Chin, W.W., 1998. The partial least squares approach to structural equation modeling. Modern Meth. Bus. Res. 295 (2), 295–336.
Chwelos, P., Benbasat, I., Dexter, A.S., 2001. Research report: empirical test of an EDI adoption model. Inform. Syst. Res. 12 (3), 304–321.
Colas, M., Finck, I., Buvat, J., Nambiar, R., Singh, R.R., 2014. Cracking the data conundrum: How successful companies make big data operational. Technical report, Capgemini consulting, 2014. https://www. capgemini-consulting. com/cracking-the-data-conundrum.
Davenport, T.H., 2013. Analytics 3.0. Harv. Bus. Rev. 91 (12), 64.
De Mauro, A., Greco, M., Grimaldi, M., 2015. What is big data? A consensual definition and a review of key research topics. In: AIP conference proceedings (vol. 1644, pp. 97–104). AIP. Retrieved from http://aip.scitation.org/doi/abs/10.1063/1.4907823.
Deloite, 2013. Tha analytics advantage: We're just getting started, Retrieved November 8, 2016, from https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Deloitte-Analytics/dttl-analytics-analytics-advantage-report-061913.pdf.
Detlor, B., Hupfer, M.E., Ruhi, U., Zhao, L., 2013. Information quality and community municipal portal use. Govern. Inform. Quarterly 30 (1), 23–32.
Diamantopoulos, A., Siguaw, J.A., 2006. Formative versus reflective indicators in organizational measure development: a comparison and empirical illustration. Br. J. Manag. 17 (4), 263–282.
Diamantopoulos, A., Winklhofer, H.M., 2001. Index construction with formative indicators: An alternative to scale development. J. Mark. Res. 38 (2), 269–277.
Dijcks, J.-P., 2012. Oracle: Big data for the enterprise. Oracle White Paper.
Douglas, L., 2001. 3d data management: Controlling data volume, velocity and variety. Gartner. from https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf.
Draganidis, F., Mentzas, G., 2006. Competency based management: a review of systems and approaches. Inform. Manage. Comp. Sc. 14 (1), 51–64.
Duncan, N.B., 1995. Capturing flexibility of information technology infrastructure: a study of resource characteristics and their measure. J. Manage. Inform. Syst. 12 (2), 37–57.
Dwivedi, Y.K., Choudrie, J., Brinkman, W.-P., 2006. Development of a survey instrument to examine consumer adoption of broadband. Indus. Manage. Data Syst. 106 (5), 700–718.
Edwards, J.R., 2001. Multidimensional constructs in organizational behavior research: an integrative analytical framework. Org. Res. Meth. 4 (2), 144–192.
Edwards, J.R., Bagozzi, R.P., 2000. On the nature and direction of relationships between constructs and measures. Psychol. Meth. 5 (2), 155.
Ertemel, A.V., 2015. Consumer insight as competitive advantage using big data and analytics. Int. J. Comm. Financ. 1 (1), 45–51.
Ferguson, G., Mathur, S., Shah, B., 2005. Evolving from information to insight. MIT Sloan Manage. Rev. 46 (2), 51–58.
Fernández, A., del Río, S., López, V., Bawakid, A., del Jesus, M.J., Benítez, J.M., Herrera, F., 2014. Big data with cloud computing: an insight on the computing environment, MapReduce, and programming frameworks. Wiley Interdiscipl. Rev.: Data Min. Knowl. Discov. 4 (5), 380–409.
Fornell, C., Larcker, D.F., 1981. Structural equation models with unobservable variables and measurement error: Algebra and statistics. J. Mark. Res. 382–388.
Gefen, D., Straub, D., 2005. A practical guide to factorial validity using PLS-Graph: tutorial and annotated example. Commun. Assoc. Inform. Syst. 16 (1), 5.
Ghasemaghaei, M., Ebrahimi, S., Hassanein, K., 2016. Generating valuable insights through data analytics: a moderating effects model. In: Proceedings of the 37th international conference on information systems (ICIS), Dublin, Ireland.
Ghasemaghaei, M., Hassanein, K., Turel, O., 2015. Impacts of Big Data Analytics on Organizations: A Resource Fit Perspective. In: Proceedings of the 21st americas conference on information systems (AMCIS), Puerto Rico.
Ghasemaghaei, M., Hassanein, K., Turel, O., 2017. Increasing firm agility through the use of data analytics: The role of fit. Decision Support Systems, Retrieved from http://www.sciencedirect.com/science/article/pii/S0167923617301161.
Gillon, K., Brynjolfsson, E., Mithas, S., Griffin, J., Gupta, M., 2012. Business analytics: Radical shift or incremental change? In: Proceedings of the 33th international conference on information systems (ICIS), Orlando, FL. Retrieved from http://aisel.aisnet.org/icis2012/proceedings/Panels/4/.
Goodhue, D., 1988. I/S attitudes: toward theoretical and definitional clarity. ACM SIGMIS Database 19 (3–4), 6–15.
Grant, R.M., 1991. The resource-based theory of competitive advantage: implications for strategy formulation. Knowl. Strategy 3–23.
Hagel, J., 2015. Bringing analytics to life. J. Accountancy 219 (2), 24–25.
Hazen, B.T., Boone, C.A., Ezell, J.D., Jones-Farmer, L.A., 2014. Data quality for data science, predictive analytics, and big data in supply chain management: an introduction to the problem and suggestions for research and applications. Int. J. Prod. Econ. 154, 72–80.
Huber, G.P., 1990. A theory of the effects of advanced information technologies on organizational design, intelligence, and decision making. Acad. Manag. Rev. 15 (1), 47–71.
Hulland, J., Business, R.I.S., 1999. Use of partial least squares (PLS) in strategic management research: A review of four recent studies. Strateg. Manag. J. 20 (2), 195–204.
Iview, I.D.C., 2012. The Digital Universe in 2020: Bag Data. Bigger Digital Shadows, and Biggest Growth in the Far East. IDC iView: IDC Analyze the future, 1–16.
Jagadish, H.V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J.M., Ramakrishnan, R., Shahabi, C., 2014. Big data and its technical challenges. Commun. ACM 57 (7), 86–94.
Jarupathirun, S., 2007. Exploring the influence of perceptual factors in the success of web-based spatial DSS. Decis. Support Syst. 43 (3), 933–951.
Jarvis, C.B., MacKenzie, S.B., Podsakoff, P.M., 2003. A critical review of construct indicators and measurement model misspecification in marketing and consumer research. J. Consum. Res. 30 (2), 199–218.
Kambatla, K., Kollias, G., Kumar, V., Grama, A., 2014. Trends in big data analytics. J. Parallel Distributed Comput. 74 (7), 2561–2573.
Katal, A., Wazid, M., Goudar, R.H., 2013. Big data: issues, challenges, tools and good practices. In: Proceedings of the 6th international conference on contemporary computing, Noida, India.. IEEE. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6612229.
Kwon, O., Lee, N., Shin, B., 2014. Data quality management, data usage experience and acquisition intention of big data analytics. Int. J. Inf. Manage. 34 (3), 387–394.
Kwon, O., Sim, J.M., 2013. Effects of data set features on the performances of classification algorithms. Expert Syst. Appl. 40 (5), 1847–1857.

Lam, S.K., Sleep, S., Hennig-Thurau, T., Sridhar, S., Saboo, A.R., 2017. Leveraging frontline employees' small data and firm-level big data in frontline management: An absorptive capacity perspective. J. Serv. Res. 20 (1), 12–28.
Li, G., Ooi, B. C., Feng, J., Wang, J., & Zhou, L. (2008). EASE: an effective 3-in-1 keyword search method for unstructured, semi-structured and structured data. In Proceedings of the 2008 ACM SIGMOD international conference on Management of data (pp. 903–914). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1376706.
Lindell, M.K., Whitney, D.J., 2001. Accounting for common method variance in cross-sectional research designs. J. Appl. Psychol. 86 (1), 114.
Liu, H., Ke, W., Wei, K.K., Hua, Z., 2013. The impact of IT capabilities on firm performance: the mediating roles of absorptive capacity and supply chain agility. Decis. Support Syst. 54 (3), 1452–1462.
Loebbecke, C., Picot, A., 2015. Reflections on societal and business model transformation arising from digitization and big data analytics: a research agenda. J. Strateg. Inf. Syst. 24 (3), 149–157.
Lu, Y., Ramamurthy, K. (Ram), 2011. Understanding the link between information technology capability and organizational agility: an empirical examination. MIS Quarterly 35 (4), 931–954.
Lycett, M., 2013. "Datafication": Making sense of (big) data in a complex world. Eur. J. Inform. Syst. 22 (4), 381–386.
Malhotra, N.K., Kim, S.S., Patil, A., 2006. Common method variance in IS research: a comparison of alternative approaches and a reanalysis of past research. Manage. Sci. 52 (12), 1865–1883.
Mata, F.J., Fuerst, W.L., Barney, J.B., 1995. Information technology and sustained competitive advantage: A resource-based analysis. MIS Quarterly 19 (4), 487–505.
McAfee, A., Brynjolfsson, E., 2012. Big data: the management revolution. Harv. Bus. Rev. 90, 60–69.
Moore, G.C., Benbasat, I., 1991. Development of an instrument to measure the perceptions of adopting an information technology innovation. Inform. Syst. Res. 2 (3), 192–222.
Newell, S., Marabelli, M., 2015. Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of 'datification'. J. Strateg. Inf. Syst. 24 (1), 3–14.
Pavlou, P.A., El Sawy, O.A., 2006. From IT leveraging competence to competitive advantage in turbulent environments: The case of new product development. Inform. Syst. Res. 17 (3), 198–227.
Peteraf, M.A., 1993. The cornerstones of competitive advantage: a resource-based view. Strateg. Manag. J. 14 (3), 179–191.
Petrini, M., Pozzebon, M., 2009. Managing sustainability with the support of business intelligence: integrating socio-environmental indicators and organisational context. J. Strateg. Inf. Syst. 18 (4), 178–191.
Petter, S., Straub, D., Rai, A., 2007. Specifying formative constructs in information systems research. Mis Quarterly 623–656.
Philip, R., 2008. Data Requirements for Advanced Analytics. TDWI Checklist Report.
Popovič, A., Hackney, R., Coelho, P.S., Jaklič, J., 2014. How information-sharing values influence the use of information systems: An investigation in the business intelligence systems context. J. Strateg. Inf. Syst. 23 (4), 270–283.
Prahalad, C.K., Hamel, G., 2006. The core competence of the corporation. In: Strategische unternehmungsplanung—strategische unternehmungsführung (pp. 275–292). Springer Berlin Heidelberg.
Raghupathi, W., Raghupathi, V., 2014. Big data analytics in healthcare: promise and potential. Health Inform. Sci. Syst. 2 (1), 3.
Raymond, L., Paré, G., 1992. Measurement of information technology sophistication in small manufacturing businesses. Inform. Resour. Manage. J. (IRMJ) 5 (2), 4–16.
Reed, R., DeFillippi, R.J., 1990. Causal ambiguity, barriers to imitation, and sustainable competitive advantage. Acad. Manag. Rev. 15 (1), 88–102.
Ringle, C.M., Wende, S., Will, A., 2005. SmartPLS (Release 2.0 M3) http://www. smartpls. de. University of Hamburg. Hamburg: Germany.
Roberts, N., Grover, V., 2012. Investigating firm's customer agility and firm performance: the importance of aligning sense and respond capabilities. J. Business Res. 65 (5), 579–585.
Roldán, J.L., Sánchez-Franco, M.J., 2012. Variance-Based Structural Equation Modeling: Guidelines for Using Partial Least Squares. Research Methodologies, Innovations and Philosophies in Software Systems Engineering and Information Systems, pp. 193.
Rusitschka, S., Ramirez, A., UIO, G. V. G., Soylu, A., 2014. Big Data Technologies and Infrastructures. Retrieved from http://byte-project.eu/wp-content/uploads/2014/09/BYTE_D1-4_BigDataTechnologiesInfrastructures_FINAL.compressed.pdf.
Russom, P., 2011. Big data analytics. TDWI Best Practices Report, Fourth Quarter, pp. 1–35.
Sambamurthy, V., Zmud, R.W., 1992. Managing IT for success: The empowering business partnership. Financial Executives Research Foundation.
Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., Tufano, P., 2012. Analytics: the real-world use of big data. IBM Global Bus. Serv. 1–20.
Seddon, P.B., 2014. Implications for strategic IS research of the resource-based theory of the firm: a reflection. J. Strateg. Inf. Syst. 23 (4), 257–269.
Sedera, D., Gable, G.G., 2010. Knowledge management competence for enterprise system success. J. Strateg. Inf. Syst. 19 (4), 296–306.
Sukumar, S.R., Ferrell, R.K., 2013. "Big Data"collaboration: exploring, recording and sharing enterprise knowledge. Inform. Serv. Use 33 (3–4), 257–270.
Tallon, P.P., 2008. Inside the adaptive enterprise: an information technology capabilities perspective on business process agility. Inf. Technol. Manage. 9 (1), 21–36.
Tippins, M.J., Sohi, R.S., 2003. IT competency and firm performance: is organizational learning a missing link? Strateg. Manag. J. 24 (8), 745–761.
Venkatraman, N., 1991. IT-induced business reconfiguration. *The Corporation of the 1990s: Information Technology and Organizational*. Transformation 122–158.
Waller, M.A., Fawcett, S.E., 2013. Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. J. Bus. Logistics 34 (2), 77–84.
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. Int. J. Prod. Econ. 165, 234–246.
Wamba, S.F., Gunasekaran, A., Akter, S., Ren, S.J., Dubey, R., Childe, S.J., 2017. Big data analytics and firm performance: Effects of dynamic capabilities. J. Bus. Res. 70, 356–365.
Wang, R.Y., Strong, D.M., Guarascio, L.M., 1996. Beyond accuracy: What data quality means to data consumers. J. Manage. Inform. Syst. 12 (4), 5–33.
Ward, J.S., Barker, A., 2013. Undefined by data: a survey of big data definitions. ArXiv Preprint ArXiv:1309.5821. Retrieved from https://arxiv.org/abs/1309.5821.
Watjatrakul, B., 2005. Determinants of IS sourcing decisions: a comparative study of transaction cost theory versus the resource-based view. J. Strateg. Inf. Syst. 14 (4), 389–415.
Wetzels, M., Odekerken-Schröder, G., Van Oppen, C., 2009. Using PLS path modeling for assessing hierarchical construct models: guidelines and empirical illustration. MIS Quarterly 177–195.
White, M., 2012. Digital workplaces: Vision and reality. Bus. Inform. Rev. 29 (4), 205–214.
Winter, S., 1998. Knowledge and competence as strategic assets. Butterworth-Heinemann, The Strategic Management of Intellectual Capital, pp. 165–187.
Wong, D., 2012. Data is the Next Frontier, Analytics the New Tool. London: Big Innovation Centre, November. Available at: http://www. biginnovationcentre. com/Publications/21/Data-is-the-next-frontier-Analytics-the-new-tool.
Wu, C., Buyya, R., Ramamohanarao, K., 2016. Big Data Analytics = Machine Learning + Cloud Computing. ArXiv Preprint ArXiv:1601.03115. Retrieved from http://arxiv.org/abs/1601.03115.