Routledge
Taylor & Francis Group

Journal of Management Information Systems

ISSN: 0742-1222 (Print) 1557-928X (Online) Journal homepage: https://www.tandfonline.com/loi/mmis20

Creating Strategic Business Value from Big Data Analytics: A Research Framework

Varun Grover, Roger H.L. Chiang, Ting-Peng Liang & Dongsong Zhang

To cite this article: Varun Grover, Roger H.L. Chiang, Ting-Peng Liang & Dongsong Zhang (2018) Creating Strategic Business Value from Big Data Analytics: A Research Framework, Journal of Management Information Systems, 35:2, 388-423, DOI: 10.1080/07421222.2018.1451951

To link to this article: https://doi.org/10.1080/07421222.2018.1451951

Published online: 15 May 2018.

Check for updates

Creating Strategic Business Value from Big Data Analytics: A Research Framework

VARUN GROVER, ROGER H.L. CHIANG, TING-PENG LIANG, AND DONGSONG ZHANG

Varun Grover (vgrover@uark.edu) is the David D. Glass Endowed Chair and Distinguished Professor of Information Systems at the Walton School of Business, University of Arkansas. His work focuses on the impacts of digitalization on individuals and organizations. He has published extensively in the information systems (IS) field, with over 400 publications, 220 of which are in major refereed journals. He was recognized as one of 100 Highly Cited Scholars in Business by Thompson Reuters (2013). He has been a senior or associate editor of several major IS journals, a recipient of numerous awards for his research and teaching, and is a Fellow of the Association for Information Systems. He has been invited to give numerous keynote addresses and talks at various institutions and forums around the world.

Roger Chiang (roger.chiang@uc.edu) is a professor of information systems (IS) at the University of Cincinnati. He received his Ph.D. in computers and information systems from the University of Rochester. His research interests are in business intelligence and analytics, data and knowledge management, and intelligent systems. He has served as the senior or associate editor of numerous IS journals. He has published over 50 refereed articles in journals and conference proceedings, including *ACM Transactions on Database Systems, ACM Transactions on Management Information Systems, Communications of the ACM, Journal of Management Information Systems, Marketing Science, MIS Quarterly*, and others.

Ting-Peng Liang (tpliang@mail.nsysu.edu.tw; corresponding author) is a Life-time National Chair Professor at National Sun Yat-Sen University in Taiwan and Fellow, Leo Awardee and president-elect of the Association for Information Systems. He received his Ph.D. degree from the Wharton School of the University of Pennsylvania and has taught at the University of Illinois, Purdue University, Chinese University of Hong Kong and City University of Hong Kong. His research interests include electronic commerce, intelligent decision support systems, and strategic applications of information technologies. He has published extensively in *Journal of Management Information Systems, MIS Quarterly, Management Science, Operations Research, International Journal of Electronic Commerce*, and others. He is the editor in chief of the *Pacific Asia Journal of AIS* and has served on the editorial board of many other journals.

Dongsong Zhang (zhangd@umbc.edu; corresponding author) is a professor in the Department of Information Systems at the University of Maryland, Baltimore County and guest chair professor at the International Business School, Jinan University, China. He received his Ph.D. in management information systems from the University of Arizona. His research interests include social computing,

Journal of Management Information Systems / 2018, Vol. 35, No. 2, pp. 388–423.

ISSN 0742–1222 (print) / ISSN 1557–928X (online)
DOI: https://doi.org/10.1080/07421222.2018.1451951

mobile human–computer interaction, data analytics, health information technologies, and online communities. He has published over 140 papers in journals and conference proceedings, including *Journal of Management Information Systems, IEEE Transactions on Knowledge and Data Engineering, IEEE Transactions on Software Engineering*, and *MIS Quarterly*, etc. His research has been funded by the U.S. National Science Foundation, U.S. National Institutes of Health, U.S. Department of Education, National Natural Science Foundation of China, Chinese Academy of Sciences, and Google, Inc.

ABSTRACT: Despite the publicity regarding big data and analytics (BDA), the success rate of these projects and strategic value created from them are unclear. Most literature on BDA focuses on how it can be used to enhance tactical organizational capabilities, but very few studies examine its impact on organizational value. Further, we see limited framing of how BDA can create strategic value for the organization. After all, the ultimate success of any BDA project lies in realizing strategic business value, which gives firms a competitive advantage. In this study, we describe the value proposition of BDA by delineating its components. We offer a framing of BDA value by extending existing frameworks of information technology value, then illustrate the framework through BDA applications in practice. The framework is then discussed in terms of its ability to study constructs and relationships that focus on BDA value creation and realization. We also present a problem-oriented view of the framework—where problems in BDA components can give rise to targeted research questions and areas for future study. The framing in this study could help develop a significant research agenda for BDA that can better target research and practice based on effective use of data resources.

KEY WORDS AND PHRASES: big data, big data analytics, big data capabilities, big data infrastructure, research framework, strategic business value.

Digital networks now connect an increasing number of people, devices, and sensors, which transform the ways that businesses generate, communicate, share, access, and analyze data and adapt to environmental changes. Diverse data are generated not only internally, but also from public, proprietary, and purchased sources at unprecedented rates. This phenomenon is broadly known as big data, which encompasses not only structured data such as transactional records stored in traditional databases and data warehouses, but also unstructured data such as text documents, web content, videos, audio, images, and sensor data. Unstructured data constitute 95 percent of big data. Volume, velocity, and variety are the original three Vs to characterize big data. Volume refers to the ever-growing large magnitude of data. Velocity refers to the fact that data are generated and arriving continuously at an unparalleled speed and must be dealt with in a timely manner. Variety means that data are in diverse formats, ranging from structured data to unstructured data (e.g., text documents). In addition, veracity and value are also critical characteristics of big data [29]. Veracity refers to biases, noise, and abnormality in data. It is concerned with uncertainty, unreliability, or inaccuracy of data.

Big data is a new yet powerful source for potentially immense economic and social value and for gaining a competitive advantage on par with an organization's capital assets and human talent. According to a 2016 report from PromptCloud, big data has grown from a $6.8 billion industry to a whopping $32 billion industry in just three years. IDC forecasts the market of big data technology and services to grow at a 23.1 percent compound annual rate, reaching $48.6 billion in 2019.[1] With the volumes of organizational data moving past terabytes to tens or even hundreds of petabytes, businesses and information technology (IT) leaders are embracing unique opportunities to capitalize on big data to gain the competitive advantage. Companies are reported to spend more than 10 percent of their IT budget on data alone [19] and are undergoing a revolution by leveraging big data and analytics (BDA) as a strategic asset to guide their decision-making and improve business processes and outcomes [22].

BDA is the application of statistical, processing, and analytics techniques to big data for advancing business. BDA is becoming important to address unique customer requirements essential to developing and sustaining a competitive advantage [38]. Businesses are undertaking analytics initiatives to predict customers' propensity to buy certain new products in order to (1) make better and more personalized recommendations for future purchases or offer discounts; (2) determine root causes of failures, roadblocks, and defects in near real time, or even predict and fix potential failures before they happen; (3) understand consumers' experience with products or services through analysis of online consumer reviews or call center data for quality improvement and product innovation; (4) develop fast responses in time of crisis and develop anomaly detection; (5) fine-tune internal processes and pinpoint operational roadblocks within an enterprise; and so forth. Insights gained from analytics of streams of structured and unstructured data can answer questions that businesses have not even considered before. Arguably, no single business trend in the past decade has had as much potential impact on incumbent IT investments as BDA.

On the other hand, hype may create unfounded pressure on firms to adopt BDA. The ultimate success of any big data projects lies in realizing strategic business value, which gives firms a competitive advantage. A 2016 Gartner survey revealed that while big data investments continued to rise, there were signs of tapering. Many big data projects provide disappointing results. Gartner predicted that 60 percent of big data projects through 2017 would fail to go beyond piloting and experimentation and would be abandoned. In a recent survey of 199 technology executives, Gartner found that many companies had struggled to obtain insights that can make real differences, despite the fact that 48 percent of surveyed firms invested in big data projects in 2016, warning that the big data bubble could be about to burst. These concerns induce firms to rethink their big data strategies to have a stronger focus on the value-creating power and return on investment of big data initiatives. The focus is shifting from the glamour of big data itself to how it impacts specific business areas and metrics [24].

Although research on big data has received increasing attention in the past few years, research on strategic business value of big data remains scarce [14]. Recently, there have been debates on how to assess the value of big data (also called "data monetization" and "data valuation") to shed light on how the investment in BDA can yield tangible business value. If big data are viewed as a valuable asset, it is imperative that businesses determine their actual accounting, economic, financial, or strategic value. However, businesses face difficulty in determining the monetary value of data. First, this challenge emerges due to the unique nature of digital data, which are reusable, can be integrated, and are never really consumed. The common accounting practices in evaluating traditional organization assets cannot accurately quantify the monetary value of data. Data not only exhibit increasing returns but could also yield more value when they are integrated with data from other sources. Furthermore, data are only the input to generating knowledge and insights valuable for decision making. The value of data is revealed through the combination of insight generation and its actual use. It is an analogy to preparing a meal, in which many ingredients are mixed and cooked (integrated) together. It is necessary to have high-quality ingredients as well as a skillful chef to prepare a delicious meal. Correspondingly, firms need not only good-quality data but also appropriate information systems (IS), analytics tools, and human analytics talent to generate valuable knowledge and insights useful for decision making.

Therefore, the key big data challenge for all businesses and researchers [14] is "How to translate big data into commensurately valuable information and business insights via BDA that justify the requisite investments?" Firms must understand key elements of BDA to enable strategic planning, including anticipated returns, potential impacts on competitiveness and current operational ecosystems, as well as opportunity cost and risks for the investment [25].

This study aims to answer key questions regarding the relationships among IT investments in BDA, BDA capability building, and strategic value creation. To do this, we briefly describe the value proposition of big data analytics, and discuss strategic IT business value and our BDA value creation and realization framework. The proposed framework focuses on building BDA assets and capabilities and the realization of BDA capabilities to create strategic business value. We then discuss research and practical implications of the framework and issues with value manifestation. Our focus is to conceptualize the contribution of BDA to strategic business value, and frame research implications.

Value Proposition of Big Data Analytics

Businesses today have been eager to set up their BDA initiatives and strategies because they do not want to face potentially negative consequences of not doing so. However, the long-term strategic implications of BDA implementation are not well understood. Academics have not studied *how* BDA can lead to a competitive advantage and strategic business value. Success of big data strategies requires more than just the data asset, the techniques to collect and manage big data, and knowledge and implementation experience of analytics methods and tools. It also requires an understanding of the mediating process and mechanisms so that BDA can serve as a resource to harness strategic business value and keep firms competitive.

The classic VRIO framework (valuable, rare, costly to imitate, organizationally embedded), based on a resource-based view (RBV) of firms [3], can be adopted to examine and develop BDA as a strategic resource to realize business value and sustain competitive advantage. When a firm considers and initiates a BDA strategy, it can ask the following questions adopted from the VRIO framework to assess whether BDA has potential to generate strategic business value:

1. Valuable: Does the BDA enable your organization to obtain valuable insights to exploit new business opportunities and/or neutralize competition threat?
2. Rare: Are your big data content, analytics capability, or the combination of them, rare? Can a few of your competitors acquire or possess them?
3. Costly to imitate (imitable): Do your competitors without a BDA capability face challenges or obstacles in obtaining or developing it? Is it difficult or almost impossible for your competitors to imitate what you can do with BDA?
4. Organizationally embedded: Do your organization's business strategies and culture support the exploitation of valuable, rare, and costly-to-imitate BDA resources?

Big data analytics can be valuable

While several dimensions of data quality, such as completeness, consistency, accuracy, and timeliness, have been discussed (e.g., [42]), the key question of data value stems from whether BDA can provide novel and valuable insights to exploit new business opportunities or defend competition threats. If the answer is yes, which is often difficult to determine a priori, then BDA can become a valuable resource for business. Insights generated from it can be used to create business value in many areas, such as business process improvement, product and service innovation, customer experience and market enhancement, organization performance improvement, and the creation of symbolic value such as business image and reputation.

For example, insights about business processes can be used to improve their efficiency, productivity, accessibility, and availability, and even to transform business processes information flows can be leveraged to create competitive advantage and dampen competition [10]. BDA obtains insights about products and services by analyzing both internally generated data and external user-generated content (e.g., online product/service reviews). Firms can use those insights to increase product/service differentiation, adjust the price of products/services, and develop innovative products and services. Likewise, insights about customers and markets can be used to improve customer satisfaction and loyalty, lock in customers and suppliers, and create a niche market. Finally, many firms have used BDA to obtain insights about their organization performance, which can enrich organizational intelligence for decision making, develop a dynamic organizational structure to respond to market and environmental changes, make better capacity utilization, and increase return on assets.

Big data analytics can be rare

BDA can be rare, meaning that few competitors can acquire or possess the same or similar capability. We can evaluate two key components in answering this question. First, by the content of big data itself, BDA may be imitable. For example, the user-generated content available at online social media becomes abundant and easy to collect and analyze. But it is strategically critical that firms own their internally generated data obtained from business interactions with employees, customers, suppliers, and so on, which are proprietary and almost impossible to be bought or obtained by other firms. Integration of internally generated data with externally obtainable data can make a big data asset rare. The second key component is the analytics capability. Although big data management and analytics tools, such as Hadoop and various data and text mining tools, can be purchased from vendors, analytics knowledge, human talent, expertise, and experience with advanced big data management and analytics tools can be rare, domain-specific, and unique to a business context. It can only be nurtured and grown through the actual implementation within the context. This ongoing process requires continuous assessment as rare BDA resource will soon become a competitive necessity when more competitors have similar capabilities.

Big data analytics can be inimitable

The third question is whether BDA is costly to imitate, indicating that competitors that do not have it cannot imitate, buy, or substitute it at a reasonable price. There are many reasons that BDA is difficult and costly to imitate or inimitable. First, the big data asset and analytics capability require a long time to develop and nurture. Firms often engage in "learning by doing" and evolve from exploratory use to a more institutionalized form of use. Second, a firm may develop its own analytics algorithms and methods, which makes the analytics capability proprietary and inimitable. Furthermore, the development and implementation of BDA capability are subject to a firm's IT maturity, the decision-making culture, and IT leadership, which makes BDA costly to imitate.

Big data analytics can be organizationally embedded

The last issue is whether a valuable, rare, and costly-to-imitate BDA can be organizationally embedded. The answer could be yes when this question is viewed as a classic alignment issue. The implementation of BDA is IT-oriented, but a BDA strategy must be business-oriented. Success of BDA depends on the inclusion of BDA in a firm's long-term business strategy, and the mechanisms in place to facilitate business alignment with this strategy. Such alignment involves processes, policies, procedures, organizational structure/governance, and corporate culture to leverage data for competitiveness.

BDA as a valuable resource may not be sufficient by itself to sustain competitive advantage. Without BDA, however, a firm can suffer competitive disadvantage. The combination of BDA with other organizational resources and capabilities provides a new way to sustain competitive advantage. Assessing BDA a priori is challenging, since many instantiations of BDA (i.e., how data come together and how insights are recognized and acted upon) can be unpredictable. However, firms can get better at this by taking advantage of what is learned through the BDA implementation to yield improved decisions, innovative products, or automated processes. The three gears, data, insights, and actions, need to move together, as firms improve their ability to generate strategic business value through BDA development and implementation.

Strategic Business Value: Functional Versus Symbolic

Business value of IT has been an ongoing IS research topic and is replete with frameworks. Chau et al. [8] argue that IT value is referred to as the value provided as a consequence of IT use and proposed a three-dimensional taxonomy, which includes four major levels—user satisfaction, individual impact, organizational impact, and societal impact. Tallon et al. [54] proposed a process-oriented approach to measure the business value of information technology. Seddon et al. [49] raised two important issues for measuring IT value: the stakeholder's perspective from which IT value is measured and the type of system being evaluated. Melville et al. [36] developed a model based on a resource perspective, which includes IT resources and complementary resources that affect business processes, which in turn influence organizational performance. Kohli and Grover [27] expanded the framework to define intermediate value (e.g., business process improvement), output value (e.g., better customer services), and competitive value (e.g., first to market). Schryen [48] reiterated the framework to examine internal versus external value and synthesized different levels to examine the economic impacts of IS as object of evaluation and time of evaluation.

The strategic value created by BDA can be functional and/or symbolic. Figure 1 classifies these two kinds of value. *Functional* value (e.g., market share, financial performance) refers to performance improvement directly resulting from adopting BDA, while *symbolic* value (e.g., positive brand image and reputation, mitigating environmental pressure) is largely derived through the "signaling effect" of investment in BDA. From the strategic fit perspective, we can think of functional value as the fit between technology and organizational tasks, and symbolic value as the fit

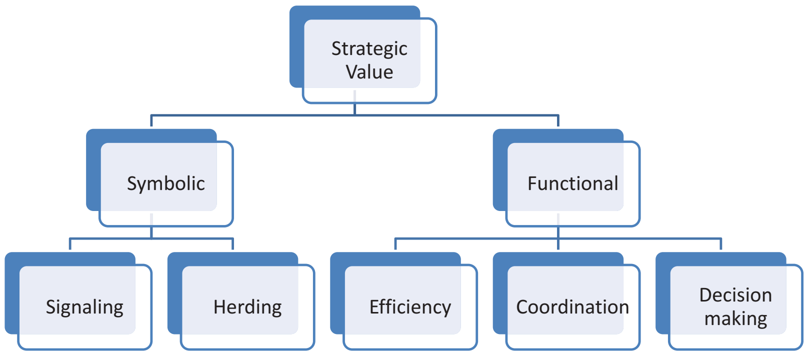

Figure 1. BDA Strategic Value

between technology and organizational environment. These two types of strategic value may not be mutually exclusive.

Functional value occurs through the chain of eventually converting assets to tangible (and intangible) value. This falls under the aegis of resource-based logic [2, 59], where BDA and other complementary resources can be configured synergistically to create unique capabilities that are targeted at value-creating entities (e.g., customers, decisions, processes). These effects have tangible manifestation in traditional productivity and financial metrics of a firm. They are usually achieved through improved efficiency, coordination, and decision making. The most commonly reported value of BDA, which should be assessed, includes performance improvement, cost and time reduction, product and service innovation, and improvement of business-consumer relationships.

While most previous literature focuses on functional value of IT, symbolic value of IT, which offers a clear signal to interested stakeholders, is also an important aspect of strategic value. Symbolic values may be observed from market reaction (e.g., real estate market reaction by Sun et al. [52]) or opportunity seeking (e.g., sensing innovation by Roberts et al. [44]). Two theories indicate the existence of symbolic values. First, adoption of BDA delivers a signal of organizational innovation. Signaling theory indicates that a signal or communication of private information by an agent reduces information asymmetry between the agent and the principal [51]. The principal, often represented by the shareholders of a company, may interpret the signal as value enhancing and boost the stock price (through their buying behavior) and the value of the company. With BDA investments in technologies, tools, skills, and leadership viewed at the forefront of innovation in social/ media discourse, a firm can gain substantial reputational effects via the signals made through such investments and other BDA initiatives [58].

In one of the few direct assessments of data value, McAfee and Brynjolfsson [34] interviewed executives of 330 public North American companies about their organizational and technology management practices and gathered performance data from their annual reports and independent sources. They reported that the more the companies characterized themselves as data-driven, the better they performed on objective measures of financial and operational performance. Specifically, companies in the top third of their respective industry in the use of data-driven decision making were approximately 5 percent more productive and 6 percent more profitable than their competitors.

Another theory underlying symbolic value is the herd behavior of following the crowd. To retain their customers or reputation, some organizations may adopt information technologies to show that they are in the mainstream and are competitive. In a longitudinal study, Sun [53] showed that herding behavior occurs "when adopting a new technology are provoked primarily by the observation of prior adoptions and perceptions of uncertainty regarding the adoption of new technology" (p.1013). Hence, the symbolic value should not be ignored when we examine the strategic implications of adopting BDA.

This functional-symbolic dichotomy provides a way to extend our understanding of the strategic role of BDA. For instance, Figure 2 shows a matrix that illustrates the strategic roles of BDA. If both functional and symbolic values are high, BDA may be a strategic transformer for the firm to enhance both internal values and image in the marketplace. When only high functional value is anticipated, BDA can be a good performance enhancer for increasing productivity. When only high symbolic value is anticipated, adopting BDA may be an image builder to generate positive signals to stakeholders. If neither value is high, then the firm may not actively seek BDA value but adopt a defensive stance.

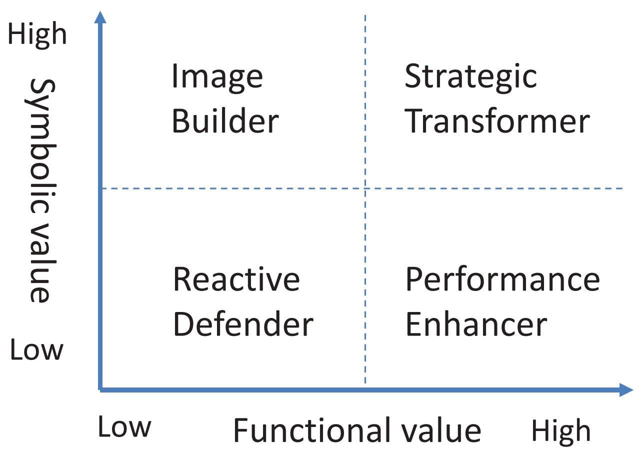

Figure 2. Strategic Roles of Big Data Analytics

How Value Is Created from BDA: A Conceptual Framework

A pivotal issue related to BDA is *how* to create strategic value. When examining successful cases of value realization, it is clear that BDA requires investments in data assets, technological assets, and human talent in order to generate something novel and valuable. Firms are motivated to plan, make, and manage changes to cope with the dynamics of the market. In their book *Managing Change for Competitive Success*, Pettigrew and Whipp [41] distinguished three dimensions of strategic change: (1) content, which is referred to what strategic changes should be made. It includes objectives, purpose, and strategic goals; (2) process, which concerns how changes should be made; and (3) context, which is about conditions through which changes can be made. Below we integrate content, process, and context in describing changes through BDA that create strategic business value.

To frame our understanding, we use (1) the general framing of dynamic capabilities indicating that in turbulent environments, companies engage in capability building and realization by building and reconfiguring internal and external resources in order to achieve supernormal performance [46, 55], and (2) IT-value models proposed by Soh and Markus [50] and Melville et al. [36], which describe how IT investments build assets/resources and create impacts on both process and variance representations. The value chain can be influenced by a variety of contextual factors. By adapting these source models based on specific instantiations of BDA in practice, we propose a conceptual framework for creating value from BDA, as shown in Figure 3, which integrates which integrates key constructs into two processes: capability building and capability realization. In the next two subsections, we describe these two processes.

Building Big Data Analytics Capabilities

Converting IT investment in BDA to valuable capabilities is a dynamic process that involves identification of where, how, and what value will be created. Capabilities include the ability to both manage and analyze data to create new insights. Firms need to develop a BDA strategy and be clear about how tangible (e.g., increased revenue or decreased cost) and/or intangible (e.g., increased customer satisfaction) value can be created from BDA [27].

Establishing a Big Data Analytics Infrastructure

Given the changing big data environment epitomized by the five Vs, the investment in BDA departs from traditional investments in structured, static, and deliberately collected data. To develop BDA capabilities, a firm must invest three necessary elements in its infrastructure: big data asset, analytics portfolio, and human talent. These infrastructural investments can generate business value [28].

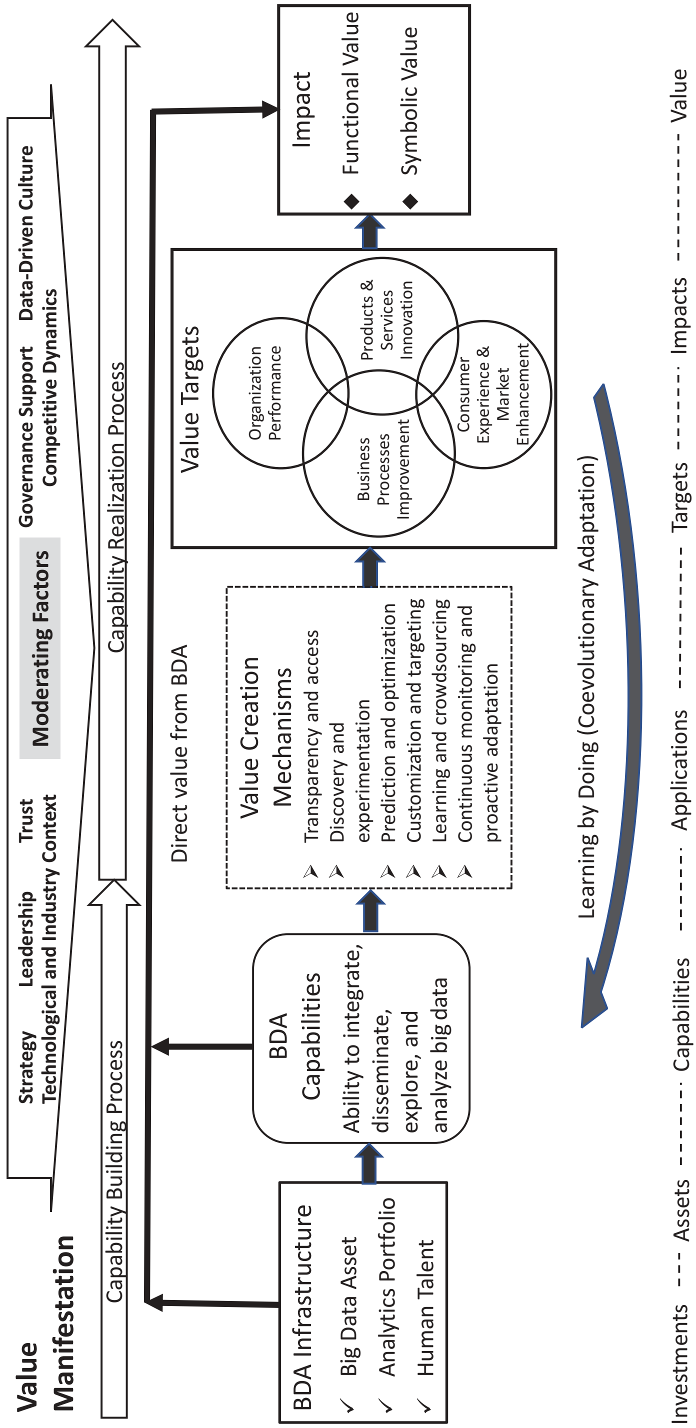

Figure 3. Value Creation by Big Data Analytics

A big data infrastructure includes data sources (e.g., transactional, clickstream, social media, user-generated, external databases) and a platform needed for collecting, integrating, sharing, processing, storing, and managing big data. There are several major open-source big data frameworks and initiatives using distributed storage, such as Apache Hadoop, Apache Mahout, Spark, and Storm. While most firms develop these infrastructures in-house, there is an increasing trend toward leveraging infrastructure, platform, and database services through cloud computing. Given the identified business goals, firms need to ensure that the data currently available are sufficient and suitable, or additional data need to be collected and incorporated. Therefore, an infrastructure should be capable of sharing data, collecting new types of data, or integrating new sources of data. Of course, firms must address any potential security, privacy, regulation compliance, or liability issues that might emerge, especially under the condition in which the third-party data and personnel are involved in data analytics.

Investment in a strong analytics infrastructure is required for analyzing prioritized data by exploring different ways to integrate and analyze data to create novel insights. Several decisions need to be made regarding data aggregation, transformation, distributed computing, tool selection, and analytics models that should be used. Potential analytical applications may include consumer sentiment analysis, financial risk modeling, marketing campaign analysis, cross-selling, fraud detection, recommendation improvement, and price and performance optimization.

To leverage investments in data and analytics, arguably the most critical element is the human talent infrastructure. Expertise and experience are needed to design and implement BDA strategies. Without the right group of skilled big data experts, it is impossible to develop and carry out a BDA strategy. This is actually one of the biggest challenges for firms. Big data professionals include data scientists, developers, programmers, analysts, and modelers that can serve significant roles in both managing and analyzing data, particularly the plethora of unstructured data in diverse formats. The most intensive use of people occurs during the input (design of BDA strategy) and output (interpretation of results) stages.

Developing Big Data Analytics Capabilities

Firms need to have strong capabilities of integrating, managing, sharing, and analyzing big data in diverse formats to support different value-creating needs. Today's BDA has evolved from the early era of business intelligence 1.0 characterized by structured content, OLAP, dashboards, data mining, and statistical analysis, to 2.0 characterized by unstructured online content, opinion mining, web analytics and intelligence, social media analytics, and social network analysis, to the current 3.0 era characterized by mobile and sensor-based content, context-relevant analysis, and mobile visualization and HCI [9].

Contemporary BDA handles big data that are more diverse, granular, real-time, and iterative. BDA capabilities need to economically generate value from data in a very large volume and variety through enabling high velocity capture, discovery, and/or analysis [61]. BDA provides a forward-looking view, enabling firms to anticipate and execute on future opportunities based on real-time insights discovered from high-volume streaming data sources, current events, and ongoing business processes. BDA includes all three types of analytics: (1) descriptive analysis that reports on the past; (2) predictive analysis that develops models based on past data for future prediction; and (3) prescriptive analysis that uses models to specify optimal behaviors and actions. BDA has an increasing emphasis on prescriptive analytics.

Core to developing BDA capabilities to leverage the analytics infrastructure is the portfolio, including text analytics (e.g., information extraction, text mining, sentiment analysis, topic modeling), predictive analysis (e.g., regression, survival analysis, time series analysis), audio analytics (e.g., automatic speech recognition, phonetic-based analysis, Interactive Voice Response), video analytics (e.g., motion and object detection, whole-to-part inductive analysis), social media analytics, geographic (location and spatial) analytics, streaming analytics, and graph analytics (e.g., graph partitioning, network analysis).

Realization of Big Data Analytics Capabilities

The amount of big data should not make it a big deal. It is the ability to actually do something meaningful and valuable with it [18]. How big data may create value for firms and public agencies is the utmost important question that requires further research [16]. Today, a new, more pragmatic view of big data is taking hold, dominated not by discussions about the volume, velocity, and variety of data, but by the value of data—the ability to generate actionable insights and apply them to business practices to accelerate innovation, drive optimization, and improve business performance. When used appropriately, BDA can yield previously unknown, valuable, and actionable insights to help refine business processes; develop business initiatives; discover service flaws or operational road blocks; streamline supply chains; better understand customers and predict market trends; and develop new products, services, and business models. The value creation mechanisms are key elements of sustainable business models.

Value Creation Mechanisms

A key question is how BDA capabilities create value. In every instantiation of BDA, the capabilities create results, which then must be turned into actions to impact targets. In other words, what are the value creation mechanisms of BDA to have a positive impact on decisions, customers, processes, or other targets? There are different ways to create value using BDA [31]. For instance, BDA can make information transparent throughout a firm, thereby improving its usability (e.g., for decisions) at a much higher frequency. In addition, organizations can collect more accurate and detailed performance data, which can be used to diagnose issues and boost performance. On the customer end, BDA allows ever-finer segmentation of customers, therefore enabling businesses to provide much more precisely tailored products or services, and substantially improve customer experience. BDA can be used to improve the development of the next generation of smart products and services.

In our framework, we propose six distinct mechanisms that *mediate* the linkage between BDA capabilities and value targets. Each value creation mechanism indicates the fundamental source of value being pursued. First, value can be created through transparency and access. The ability to generate descriptive data and disseminate them widely across a firm not only allows consistency in viewing the data but also facilitates a more complete visibility of the firm's business processes and outcomes. For instance, dashboards can provide real-time access to the well-being of various activity systems in a firm. Second, value could occur through discovery and experimentation. Discovery is often the most emphasized aspect of BDA—digging into data for both deep and pragmatic insights can yield important outcomes for various BDA targets. Similarly, in an increasingly digital world, big data can involve many small experiments. For instance, running longitudinal experiments can provide insights into causality that may have strong implications for the delivery of customer services, among other targets.

Third, value can be created through prediction and optimization mechanisms. In the former, the value is in determining probabilistic outcomes for the future on which present-day action can be taken. In the latter, leveraging big data and powerful analytics can determine the "best path" forward—something that was largely suboptimized and dependent on managerial judgment in the recent past. Fourth, BDA can facilitate customization of products and services, as well as targeting different market segments with digitally versioned products. Such mechanisms greatly bolster customer retention and other customer-related outcomes. Fifth, value can be obtained through the mechanisms of learning and crowdsourcing. Machine learning has been applied to many different contexts from online education to automated vehicles, while crowdsourcing is being used for predictions and leveraging innovative talent. Finally, the ability to monitor situations and adapt rapidly allows preemption of future problems. This is particularly prevalent with the explosion of data from the Internet of things, which can be used to monitor, warn, and adjust for situational abnormalities.

Targets of Value Creation

It is critical to integrate the views and priorities of stakeholders and get them on board when determining value targets. In our framework shown in Figure 3, we identify four distinct targets of BDA value creation: (1) organization performance such as the quality of decision making, (2) business process improvement (e.g., the greater efficiency of business processes through automation), (3) product and service innovation, and (4) customer experience and market enhancement (e.g., improved consumer satisfaction, retention, and customer-firm relationships). These targets are interrelated, as many initiatives involve multiple targets. For instance, automated analytics of business processes can improve customer service innovation.

First, BDA can create value by improving organizational decision making. This can be accomplished by providing broad and consistent access to data across an organization complemented with empowerment structures to act on the data, or through decision models that augment human decision making or are built into business processes. For instance, analyzing streaming data, such as real-time performance data and just-in-time inventory status, can have significant impacts on organizational performance (e.g., situational awareness, data breach, and fraud detection).

Second, BDA can create value by improving the effectiveness, efficiency, and productivity of business processes, which leads to better execution and less time spent on process breakdowns. Continuous improvement of business processes is a challenging task that requires complex and robust supporting systems. The three types of business process analysis are validation, verification, and performance, all requiring large volumes of process and event data [57]. These business process analyses can explicitly benefit from the results of BDA, such as process mining (e.g., inductive mining, event detection and analysis, bottleneck analysis, deviation analysis). BDA can help identify the strengths and weaknesses of a business process. What used to be a gut-feeling decision or a crudely automated decision can now be empirically supported with large-scale analyses.

Third, BDA can create value for product and/or service innovation. Data analysis based on customer clickstreams or purchasing patterns can enable customized web interfaces or promotional strategies. For example, user-generated content available on social media has become increasingly analyzed for gaining business insights. Customers often share their experiences with services and opinions on products and influence each other through online product reviews and ratings. By analyzing those online consumer product reviews and/or consumer discussion forums, a firm can identify the frequently reported product flaws/issues or consumer desired features of certain products or services, which provide insights for product/service innovation. Some companies nurture crowdsourcing innovation platforms to facilitate this. Monitoring of products at customers' sites (as GE does for its power turbines) can also facilitate both product and service innovation.

Fourth, BDA can also deliver a better customer experience and more competitive services, resulting in higher customer satisfaction and retention. Firms may identify competitive intelligence or acquire a significant number of new customers through social media analytics. They can identify potential competitors through online comments and link social activities, profiles, and historical purchase history of existing customers to create a better and more comprehensive understanding of their customers through BDA. For example, customers are increasingly looking for more specific answers to questions such as "what do other customers like me think of this product or service?" Answering such questions requires not only an understanding of what customers are looking for but also identification of "similar customers" and what exactly they have stated about a product or service. That is an opportunity for big data value creation.

Eventually, functional and/or symbolic value can be created. In the literature, many measures have been developed to assess functional value of information technologies, such as financial returns, market share, and so on. Symbolic value may also be assessed through perceptions of a company's innovative image, industry leadership, market influence, discourse and media discussion, satisfaction from diverse stakeholders (customers, government, and other related entities), and other reactions to BDA initiatives.

Contextual Enablers

A number of contextual enablers (i.e., moderators) can catalyze the conversion of BDA investments to value. A BDA strategy that depicts the objectives and approach to BDA is the glue that holds disparate BDA initiatives together. Further, alignment of a strategy with the organizational infrastructure requires a strong leadership in a firm to promulgate a data-driven culture. Lack of such a culture could be fundamentally detrimental to identifying and generating potential value of BDA. Emotional factors of decision makers may also have effect in real option IT investment decisions [40]. Becoming a data-centric organization will involve organizational and cultural changes and innovation. Strong data-driven cultures can yield predictions that are instrumental in determining where a company is going. A widely read 2011 McKinsey report [31] on big data suggested that a company's "data-driven mind-set" would be a key indicator of big data's value to companies.[2] The report gauged corporate cultures of fact-based decision making as an important indicator of big data's value potential. Firms should also set up processes, governance structures, and teams with complementary data skills. Active data governance refers to the overall management of the availability, usability, integrity, and security of data. A sound data governance in an enterprise should include a governing body or council, a set of data governance procedures, and a plan to follow and execute those procedures. When business questions arise, a data provenance assessment can help prioritize data that can offer new insights into these questions. A strategy roadmap can then harness existing and new data assets. BDA initiatives without clear business goals and strategies will fail.

The paths from investments to assets to capabilities to mechanisms to targets and finally to value largely represents functional value creation. We recognize that companies can create diverse datasets and strong analytics capabilities and directly "sell" them in the marketplace to create value. Alternatively, assets and complementary resources may provide symbolic value to a firm through their propensity to signal that they are innovative and at the forefront of technological development. These patterns are indicated by the direct (nonmediated) path in the framework.

A second note regarding the framework is recognition of the temporal aspect. Capability building and realization processes occur through a series of interrelated decisions that facilitate organizational learning. This learning allows for coevolutionary adaptation (i.e., the reverse arrow), which is a virtuous feedback cycle that enhances the ability of a firm to build and realize future BDA capabilities through experience, successes, and failures.

Below, we discuss the implications, particularly for research, of the framework in three parts. The first part is descriptive and illustrates how applications of BDA can be examined through the framework. The second part describes how the framework can be used by researchers to study BDA value, including its theoretical framing. The third part focuses on practical challenges in moving through the framework and how these challenges can frame key questions for a research agenda.

Face Validity of the Framework: Illustrating Paths

Based on the framework, we identify several value paths from capabilities to mechanisms to targets to impacts. Below, we illustrate the framework by using real-world cases from the practice of realizing strategic value through BDA (Table 1) . At the minimum, the ability to trace paths through the framework based on successful BDA applications provides face validity for the framing.

1. BDA at eBay

eBay is one of the world's largest online marketplaces. In 2012, the total value of goods sold on eBay was $75.4 billion. eBay serves over 164 million active users and approximately 1 billion live listings of items for sale.[3] One of the keys to eBay's extraordinary success is its ability to turn its big data, with more than 250 terabytes (TBs) in volume and high streaming data velocity with 6 billion writes and 5 billion reads generated daily, into useful business insights that benefit its customers directly. eBay has invested in DataStax Enterprise with integrated search and real-time big data analytics deployed across multidata centers. With substantial data access and analysis capabilities in place, eBay emphasizes customization to enhance its customers' digital experience in order to drive sales and satisfaction [13]. For example, eBay models personalization based on structured (e.g., sale item listings and purchases) and unstructured (e.g., behavioral activity synopsis and word clouds) data. Merchandising is improved by using machine learning to recommend similar items. eBay also develops predictive machine learning models for fraud detection, account takeover, and buyer/seller risk prediction. In doing so, eBay traverses the value foci on assets, capabilities, and multiple mechanisms to target consumers and generate strategic impact on both financial growth and its reputation for industry leadership.

2. BDA at CancerLinQ

The health-care industry has been one of the most focused application domains of BDA. The main assets in health big data include genetic data (e.g., gene expression,

Table 1. Paths to Value Creation of BDA in Real-World Organizations

	BDA infrastructure	BDA capabilities	Value creation mechanisms	Value targets	Impact	Moderators
E-Bay	DataStax Enterprise Structured/ unstructured data Streaming data	Models and real-time BDA	Customization Prediction Machine learning	Customers (Enhanced digital experience)	Functional (financial growth) and Symbolic value (reputation)	These companies have strong leadership and data-driven cultures. They foster BDA strategy. They also face competitive pressure to innovate with data.
CancerLinQ	Tracked real-time data on oncology patients and treatments, including genomic and social data.	Statistical, interactive data visualization tools, text mining	Access Discovery Experimentation Prediction Crowdsourcing	Customers (Physicians) (Improved patient outcomes)	Functional (life span) and Symbolic (satisfaction) value	
Walmart	Hadoop cluster, Data Café Telecom data, sales data, online clickstreams, social media data, economic data, meteorological data, Nielsen data, gas prices, and local events databases	Models and real-time BDA; social media analytics; data visualization; trend analysis; market basket analysis	Discovery Prediction Optimization Customization Continuous monitoring	Business outcome (e.g., sales) Customer experience and loyalty Operational efficiency	Functional (revenue) and Symbolic (reputation, leadership)	
Deutsche Bank	Hadoop, Data Lab Transactions, loans, consumer data, mortgages, bank accounts, and so on	Risk analysis Fraud detection Security Anomaly detection	Discovery Prediction Continuous monitoring	Business process and outcome Organizational performance	Functional (customer retention, financial performance)	
UPS	Platform for sensor data in vehicles (ORION)	In-house algorithms Spatial data maps	Prediction Optimization Experimentation	Business process improvement Service experience	Functional (cost reduction)	

sequencing data), payer-provider big data (e.g., electronic health records, clinicians' notes, drug prescription) [9], biomedical sensor data, online social media data, and so on. The pressure on turning to big data is not new in the U.S. health sector due to the soaring health-care expenses in the past two decades and rising interest in evidence-based medicine and patient care [23]. The Mckinsey Global Institute suggested that if U.S. health care were to use BDA creatively and effectively, this sector would create more than $300 billion in value every year, with two-thirds of the value coming from reducing U.S. health-care expenditure [31]. Advances in scientific medical knowledge, rapid introduction to new drugs and treatments, and the escalation of health-care costs require hospitals and physicians to assimilate new information and deploy new advances, including BDA, in order to improve processes, decisions, services, and outcomes in the form of high-quality and personalized care. BDA contributes to increasingly complex treatment decision making by potentially enabling more personalized medicine at the bedside and the rapid pace of scientific discovery.

Examples of ambitious and exciting investments in BDA include PCORNet (the National Patient-Centered Clinical Research Network) [19], the U.S. Food and Drug Administration (FDA)'s Mini-Sentinel, and the American Society of Clinical Oncology's CancerLinQ [47]. For example, BDA is transforming the fight against cancer. By using sophisticated analytics tools on powerful computing platforms and data assets in the form of genomic datasets of many patients, along with patient medical records and reference genome data, novel discoveries about genes involved in DNA damage repair and genome instability in cancers can be generated [26]. These healthcare BDA applications can provide insights for making effective and personalized treatment decisions. In addition, social media have also become an important data source for discovering new health-related information and knowledge, such as identification and surveillance of the trends of diseases (e.g., influenza) [11] and discovery of unknown adverse drug reaction and side effects [37] through structural and text mining of online user-generated content (e.g., Twitter). Here, assets and capabilities allow discovery and experimentation as well as prediction to improve patient care outcomes. BDA is targeted at physicians and patients with cancer to gain both functional value of increased life span and symbolic value of high satisfaction of physicians and patients.

3. BDA at Walmart

With more than 245 million customers visiting 10,900 stores,10 active websites across the globe, 2 million associates hired, and $36 million sales from over 4,000 stores in the United States on a daily basis, Walmart is a giant in the retailing industry. It was the world's largest retailer in 2014. Walmart collects 2.5 petabytes of data from 1 million customers every hour. In 2012, Walmart moved from a 10-node Hadoop cluster to a 250-node Hadoop cluster so that all the unstructured data generated at 10 different websites could be combined and collected into a new Hadoop cluster. Since then, Walmart has been speeding in BDA to provide the best e-commerce technologies with a motive to deliver the best customer shopping experience and improve its operational efficiency. It uses Hadoop and NOSQL technologies to provide internal customers with access to real-time, centralized data collected from different sources. Walmart creates the world's largest private cloud and analytics hub known as Data Café, which pulls information from 200 diverse sources. Walmart's analytic algorithms are designed to scan through the data in microseconds to come up with a real-time solution for a particular problem.

An example of early applications stemmed from the Hadoop data at Walmart is Savings Catcher, which alerts consumers who have already bought a particular product, about the item's price reduction by a competitor and sends a gift voucher to compensate the price difference. Data analytics helps Walmart to find purchase patterns in sales data and provide product recommendations to consumers based on which products were bought together or bought before the purchase of another particular product. The big data team at Walmart analyzes every clickable action on Walmart.com about what individual consumers buy, what is trending on Twitter, how local weather deviations or seasons affect buying patterns, and so on.

By leveraging social media data, such as Facebook comments, Pinterest pins, Twitter Tweets, and LinkedIn shares, Walmart can discover the trending products to be introduced to stores across the world. Social Genome is a big data analytics solution developed by WalmartLabs that analyzes the combined public data from the web, social media data, and proprietary data like contact information and e-mail addresses. It analyzes Facebook messages, tweets, YouTube videos, blog postings, and the like to enable Walmart to reach out to customers or friends of customers who tweet or mention something about Walmart products on social media to inform them about products that they may be interested in and provide them with special discounts.

In addition, Walmart uses the predictive analytics platform acquired from Inkiru that incorporates machine learning for targeted marketing and fraud prevention and reduces overstock. The value of the strategic changes that Walmart made is identified by analyzing the sales before and after leveraging big data analytics.[4] Walmart observes a significant 10–15 percent increase in online sales for $1 billion as well as reputational benefits due to being on the forefront of BDA.

4. BDA at the Deutsche Bank

The financial industry is one of the most data-driven industries. For decades, financial institutions have leveraged their internal insights on their customers to manage risk and fraud, as well as to improve product development, marketing, and customer communications. Advanced information technologies coupled with voluminous structured and unstructured big data, however, allow more real-time information analysis and better, faster decisions. Big data provide financial institutions with unprecedented ability and tools to digest digital and physical channel interactions, customer data, graph data, geolocation data, and so on, aiming to provide better customer experience and enhance fraud detection capabilities. Deutsche Bank currently has multiple Hadoop platforms.[5] Its Data Lab sits horizontally across the bank's businesses, providing internal data and analytics services. BDA projects in the Deutsche Bank have built risk platforms to mine and process data and analyze risk. Other data analytics used by the bank include the matching algorithm that enables the bank to monitor its performance and identify abnormal information through rule-based algorithms so that errors can be quickly flagged.

Wells Fargo, Bank of America, and Discover have used BDA to understand their customers and improve customer relationship management. They leverage their infrastructure and capabilities to monitor customers' "journeys" through the tangle of websites (e.g., website clicks), call centers (e.g., voice recordings), tellers (e.g., transaction records), and other branch personnel (e.g., bankers' notes) to understand the paths that customers follow through a bank, and how those paths affect attrition or the purchase of financial services [1]. Such BDA applications can use discovery and experimentation as well as prediction software to improve the quality of bank-customer interactions, identify reasons for customer attrition, and discover customer opportunities and problems.

5. BDA at UPS

In many BDA applications, reducing cost and improving efficiency are the primary foci. Since 2010, UPS has invested over $1 billion annually on BDA. Much of this is in a substantial platform for sensor data that includes installation of telematics sensors on over 50,000 delivery trucks of the company. As drivers do their shipment routes, the sensors on their trucks capture streams of data consisting of more than 200 elements, such as speed, direction, braking, oil pressure, seat-belt use, number of times that a truck is placed in reverse, and idle time, and send them in real time to servers for analysis. At the end of the day, data are uploaded and sent to the UPS data center. These streaming sensor data are not only used to monitor daily performance but also analyzed to discover problems and insights and enable UPS to optimize drivers' routes and identify when drivers can make certain adjustments, and to help reduce fuel consumption, emissions, and maintenance costs while improving customer service and driver safety. In 2011, UPS reduced 85 million miles in daily routes because of those data analytics, saving more than 8.4 million gallons of fuel. It is estimated that reducing one daily mile driven per driver could save $30 million in a year for the company [1]. The efficiency path followed through the framework involves investment in assets and capabilities that allow for the capture of data that can be used for optimization and prediction to improve business processes and their asset utilization.

Framing Research

Because BDA is relatively new, there has been little empirical research on how BDA can harness business value and linkage among BDA assets, impacts, and business performance [5, 7]. We need to have a better understanding of how, why, and when

BDA can be valuable for companies to gain a competitive advantage [6]. While the usefulness of BDA is known, the path toward strategic value creation may not always be clear and remains relatively unexplored.

Below, we offer two approaches to frame BDA research using the framework: (1) to develop constructs and theory, and (2) to deal with problematization in BDA value creation.

Framing Research on BDA Value: Constructs and Theory

To better explain and predict BDA value creation, it is important to identify constructs in the framework and provide a theoretical basis for their relationships. The goal of BDA research should not only test correlations and establish plausible causality but also draw strong conclusions by converging evidence from multiple, independent, and heterogeneous sources [20].

The basic framework can be represented by the following broad propositions:

> *Proposition 1:* Increasing BDA infrastructure investments in the quality and quantity of data and data management assets, analytics portfolio, and analytical skills will enhance BDA capabilities (the ability to integrate, disseminate, explore and analyze big data).
>
> *Proposition 2:* BDA capabilities will enable organizations to determine value targets (i.e., better decisions, process improvement, product innovation, and customer experience) mediated by one or more value creation mechanisms (i.e., transparency, access, discovery, experimentation, prediction, optimization, customization, targeting, learning, monitoring, adaptation).
>
> *Proposition 3:* Different value targets of BDA will result in different organizational impacts (functional and symbolic value).
>
> *Proposition 4:* Contextual factors (i.e., strategy, culture, governance, leadership, competition) will moderate the relationships in Propositions 1, 2, and 3.

Our proposed value creation framework is broad and widely applicable. Given the abstract nature of the propositions, specific relationships can be studied by identifying constructs in each category. Table 2 illustrates sample constructs that can be used by researchers to conceptualize and test the framework. Therefore, exploring relationships, for example, between integration of structured and unstructured data and its impact on service innovation through discovery, moderated by competitive intensity, could yield insights on how companies can gain value from customer-facing social media.

However, from a research perspective, discovering successful value creation paths according to the proposed framework might exhibit certain patterns that could be useful in explaining and predicting value outcomes. For instance, types and content of big data might vary and have different mechanisms that are more effective for

Table 2. Potential Constructs of BDA Research

BDA infrastructure	BDA capabilities	Value creation mechanisms
• Dollar investment in BDA assets • Diversity of data sources • Volume, variety, and velocity of data • Quality of data analysis skills • Type of analytical toolset	• Ability to integrate databases • Ability to disseminate structured and unstructured data • Ability to handle data volume • Ability to handle data variety • Ability to integrate external data • Ability to design data analysis • Ability to interpret data outputs	• Transparency • Access • Discovery • Experimentation • Prediction • Optimization • Customization • Targeting • Learning • Crowdsourcing • Monitoring • Proactive adaptation • Number of mechanisms triggered
Value targets	**Impact**	**Contextual**
• Quality of decisions • Business process efficiency • Product innovation • Service innovation • Customer experience • Market expansion	• Profitability • Return on assets • Equity price • Symbolic value (image, satisfaction, industry leadership) • Fraud detection • Market forecasting	• BDA strategy • Data-driven culture • Leadership • Data governance • Competitiveness • Environmental uncertainty

different targets. These patterns require underlying valuation logic to generate hypotheses under the broad BDA propositions listed above. OuYang [39] proposed three different views for performance generation from knowledge management capabilities and found different results from empirical evaluation.

In this study, we summarize five theoretical logics underlying the value creation of BDA, as described in Table 3, with each based on theoretical perspectives that have been shown to be useful in understanding "how" value is created in organizations. The logic of resources stems from RBV, the logic of alignment from much of the strategy-IT alignment literature, the logic of options from real options thinking, the logic of dynamics from dynamic capabilities, and the logic of absorptive capacity from management literature on exploitable innovation.

Each of these logics has an instrumental concept for value creation. They can help create subframes from the proposed BDA framework that are tied together by this concept, creating testable theory-based models. For instance, RBV indicates that BDA interaction in the framework creates heterogeneity and immobility for value. Options logic indicates that BDA can have higher value as it provides options from greater learning and opening follow-on options. Alignment logic indicates efficiency and effectiveness from BDA can be increased through consistency among strategy,

Table 3. Five Theoretical Logics of Value Creation

Logics	Explanation	Relevance to BDA value framework	Illustrative hypotheses
Resources [33]	Heterogeneous and immobile resources generate higher value.	Having or creating heterogeneity and immobility by integrating elements of the BDA value framework can enhance value.	• *BDA initiatives that trigger more value creation mechanisms will result in higher value.* • *Integration of proprietary data into public databases will result in greater value than without integration.*
Alignment [21]	Creating consistency between goals and structures of organizational components leads to greater efficiency (low resource wastage) and effectiveness (goal orientation).	Consistency between the various elements of the BDA framework will yield more positive value outcomes.	• *Alignment between BDA strategy and BDA initiatives will result in greater value.* • *Alignment between IT and data governance will result in fewer BDA failures.*
Real options [17, 35]	Investments that are subject to high uncertainty, irreversibility, and flexibility should be valued more highly because they provide options to expand the investment under better conditions (e.g., when the investor has more experience).	BDA can benefit from options logic since these investments are uncertain and every "analysis" can yield new options.	• *Companies that cast a broad net by using diverse data sources generate greater value than companies that case a narrow net.* • *Diverse BDA analytical skills can provide more options for deeper analysis resulting in greater value.*
Dynamics [12, 55]	Based on Schumpeterian dynamics of disruption, value is created through exploitation of market opportunities through flexibly combining resources and learning by doing.	Flexibility of BDA infrastructure in using value creation mechanisms to target value allows firms to take advantage of opportunities and increase value through BDA.	• *Market dynamism moderates the relationship between BDA capabilities and value.* • *Ability to integrate diverse data sources is positively related to organizational value.*

(continues)

Table 3. Continued

Logics	Explanation	Relevance to BDA value framework	Illustrative hypotheses
Absorptive capacity [45]	Value is based on the ability to identify valuable external knowledge, assimilate or transform this knowledge into the firm's knowledge base, and apply this new knowledge through innovation and competitive actions.	BDA capabilities and organizational resources that can enhance absorptive capacity will create greater value for the firm through innovation.	• *The ability to integrate external databases with internal databases is positively related to value creation.* • *The positive relationship between BDA capabilities and product innovation is moderated by the extent of integration between internal and customer data.*

resources, and capabilities. Logic of dynamics suggests that BDA value increases when opportunities can be identified and exploited through flexible resource configuration. Finally, absorptive capacity logic indicates that the BDA value occurs when effective data integration occurs to create innovation. These logics can be applied to frame relationships derivable from the BDA framework. Table 3 illustrates possible hypotheses relevant to BDA value that could be generated from the different theoretical logic. Each of the above logics provides a valuable paradigm for designing research on the adoption and strategic value creation of BDA. Importantly, the logics can be combined in meaningful ways to provide a more granular understanding of the root of value creation in BDA.

Framing Research on BDA Value Based on BDA Problems

An alternate way to examine BDA and frame a research agenda is to look at the challenges facing companies in creating strategic value. A recent survey of executives at more than 400 companies, each with over $1 billion in revenue,[6] indicated that only 4 percent of companies were really good at analytics; 56 percent did not have appropriate systems to capture the data that they needed or did not collect useful data; and 66 percent lacked the right technology to store and access data.

Given these challenges, an alternate framing of a research agenda for BDA value creation can be done though problematization or identifying problems associated with creating and realizing BDA value and framing research around these problems. Problem-oriented research can deliver guidelines or explore solutions that are directed toward alleviating impediments to BDA value. Markus [32] describes "theories of the problem" research as dealing with problems related to "IT design, implementation and use" [32, p. 343], which would explain how and why a problem occurs, setting the stage for solutions that can alleviate the problem situation.

In this vein, based on the framework shown in Figure 3, we present (a) problems in building BDA capabilities based on big data asset, analytics portfolio, and human talent; (b) management problems in catalyzing movement through the framework, and (c) problems with assessing impact. Each of these is discussed below, followed by key research areas/questions that focus on the problems.

Big Data Asset

In building BDA capabilities, core data assets present many challenges. Three of these stand out—data quality, data integration and data security.

Data Quality

Pervasive data quality problems add a further layer of complexity to the real-time and actionable use of big data. There are several common data quality challenges:

1. Data are often noisy, erroneous, or even missing. For example, there are many jargons, misspelled words, and incorrect grammars in user-generated textual social media content, which poses significant technical challenges for computational linguistic analysis; the data captured by mobile and wearable devices and sensors can be noisy. There is a big chunk of data cleaning preprocessing that needs to be done prior to any analytics.
2. As data continue growing exponentially, it has become increasingly difficult for firms to ensure that their source of data and information is trustworthy. Veracity of big data, which is an issue of data validity, is a much bigger challenge than volume, velocity, and variety in BDA. For example, there has been increasing interest in leveraging online consumer product reviews for product/service evolution or innovations in the past few years [4, 16]. However, it is estimated that approximately 20–25 percent of online consumer reviews are fake [43]. Therefore, data hygiene through data cleaning, filtering, and selection to detect and remove noise and abnormality from data automatically becomes essential. The more consolidated, cleaned, accurate, and consistent data a firm has, the more likely it can make better decisions. However, this cumbersome activity involves meticulous statistical and analytical processes. Developing effective ways to detect and remove unauthentic data becomes critical to ensuring adequate trustworthiness of data.
3. There are many partially or even unlabeled data. Traditional supervised machine learning techniques require access to many labeled training samples (e.g., positive vs. negative instances of a target class). In some real-world big data (e.g., financial fraud detection and online fake review detection) projects, data are often unlabeled or only partially labeled. Partially labeled data

can be considered as a type of noise/inaccuracy in labeling a target class [22]. Such a lack of labeled data causes significant problems when using machine learning algorithms for model building. Therefore, we need more research on machine learning with partially labeled data, such as semisupervised machine learning (e.g., [60]).

4. There is a need to cope with imbalanced data, which occurs when training a classification model with data samples that do not represent different classes equally [43,62].

Due to these data quality challenges, there are important research questions that need to be further explored in addressing the quality issues of big data asset, such as:

RQ Set 1: Questions that focus on the development of uniform, well-accepted data quality standards and metrics for BDA that address a variety of data quality dimensions (e.g., accuracy, accessibility, credibility, consistency, completeness, integrity, auditability, interpretability, and timeliness).

RQ Set 2: Questions that focus on the development of more effective, automated techniques for data preprocessing that can detect and adjust for data anomalies effectively.

Data Integration

It is critical to establish a robust data management architecture that can support a variety of analyses of big data and deliver on-demand, real-time business insights to business users, enabling them to rapidly access what they need when they need it. To perform BDA effectively, integrating heterogeneous big data from diverse sources presents significant challenges to realizing value from BDA.[7] The true value in incorporating a big data initiative comes from integrating and unifying diverse sources of existing (e.g., data stored in legacy systems) and new data, both structured and unstructured. The variety dimension of big data determines the high complexity of data integration, especially when heterogeneous data come from external sources. Independent data sources may provide data collected at different time intervals and with different levels of granularity and aggregation.

In general, there are two possible approaches to tackling such data integration challenges. One is the loose mediator-based integration, in which data are kept at their original sources. A data management system acquires and combines data relevant to answering users' questions. Although this approach avoids data replication and leverages the storage and maintenance of the source data by its respective owners, it may not always be practical for many reasons. Also, it needs to develop resolution strategies to cope with conflicting conclusions drawn from different data sources. Another method is the tight integration approach that extracts, collects, and then replicates all relevant data from individual sources a priori into one large integrated big data repository. Such a process involves data extraction, cleansing, transformation, and loading. Then analytics tasks can be conducted directly on this dedicated data repository. In this approach, all the preprocessing is done in advance and offline, which can shorten the time needed for analytics and decision making. However, it must cope with issues of up-to-date replication and data consistency. Further research is needed on more advanced methodologies and techniques for data integration:

RQ Set 1: Questions that focus on data resolution strategies for handling inconsistent and even contradictory data from diverse sources.

RQ Set 2: Questions that focus on trade-offs between loose (via application program interfaces [API]) and tight integration (via replication) across platforms for integrating different types of data.

Data Security

Big data must be protected to prevent malicious data breaches. The growing cyber-security threats imply that organizations should expect and prepare for data breach, and design and implement measures to detect such breaches in a timely manner to minimize their negative impacts and make data secure and data management systems resilient. Organizations are expected to comply with data security regulations and report to both the regulator and individual consumers affected when a data breach takes place. As data volume and computing infrastructures get large, traditional methods of data security mechanisms for securing small-scale data and infrastructure are becoming inadequate [29].

Implementing pervasive and scalable data analytics would naturally and gradually rely on the cloud infrastructure [18]. As a result, effective security mechanisms are needed at different layers to provide data security. Kune et al. [29] presented an onion model of defense for big data security. At the distributed computing infrastructure layer, big data setup needs to be confined to an enterprise or several enterprises. Needed at the distributed data layer are privacy preserving mechanisms, data encryption techniques, data access control mechanisms, and security of data models. At the analytics security level, companies need to develop secured frameworks that allow them to use analytics securely based on authentication mechanisms. At the user level, confidentiality, integrity, and user authentication mechanisms must be established to validate users. Such needs are even more salient with mobile context-aware applications that involve mobile and wearable devices that keep track of users' location, activities, interaction with others, and so forth through sensors (e.g., GPS, Gyroscope, and Accelerometer).

In addition, potential technical and operational risks associated with exploiting, maintaining, and operating a large volume of heterogeneous and potentially sensitive data need to be clearly identified.

RQ Set 1: Questions that examine the scalability of traditional security tools for newer data contexts (such as social media, mobile, etc.).

RQ Set 2: Questions around managing "loss of control" of data in cloud environments and decentralized access that typify big data environments.

Analytics

LaValle et al. [30] identified three levels of analytics capability of firms (from low to high), including aspirational, experienced, and transformed. Organizations with aspirational analytics capability, which often focuses on efficiency or automation of existing processes and looks for ways to cut costs, have the lowest analytics capability and are the farthest from achieving their analytical goals. They lack some essential elements (e.g., people, processes, tools) for collecting, understanding, or acting on analytics insights. Organizations with experienced analytics capability have gained some analytics experience and often intend to go beyond cost savings. They are interested in developing effective methods to collect, incorporate, and act on analytics for business optimization. Organizations with transformed analytics capability have substantial experience with using analytics and with organizing people, processes, and analytics tools. They mostly focus on driving profitability and making targeted investments in niche analytics. LaValle et al. [30] reported that transformed organizations were three times more likely to outperform their competitors substantially than aspirational organizations.

RQ Set: Questions around identification of analytical capabilities for people and processes and the adoption of IT tools to realize them. This would include organizational-level issues of conceptualizing BDA maturity.

Human Talent

For some business organizations, the biggest challenge in deploying BDA may not be the technology itself, but how to cultivate human capital and organizational culture to support such efforts. BDA is a strategy and operational activity. An organization needs to build a team of people with sufficient BDA skills and talent to capitalize on the promise of big data. The reality, however, is that the lack of big data expertise (e.g., data scientist or big data professionals) in organizations is one of the major issues. In 2015, IBM, Cisco, and Oracle together had 26,488 open positions that required big data expertise. In Bain & Company's survey, 56 percent of executives said that their companies lacked the capabilities to develop deep, data-driven insights. At Walmart, when a new member joins the analytics team, he or she has to take part in the analytics rotation program, in which he or she will spend time with different departments to understand how BDA is being leveraged across the company. It is important for an enterprise to develop a dynamic, data-driven business culture.

RQ Set: Questions focused on delineating the distinct capability sets for an effective data scientist, the interrelationships between these sets, and implications for research and pedagogy.

Management (Contextual) Challenges

Management challenges that may prevent companies from succeeding in BDA projects include leadership and strategy [34]. Companies that have leadership teams with clear BDA strategies that have clear goals and can articulate the business case would likely succeed. Those teams can define what success is and ask the right questions. Correspondingly, company culture and governance are critical to creating the right environment for BDA. A culture that embraces data- and evidence-driven approaches to business decisions, and governance that delineates responsibility and accountability for data, are both catalysts for BDA value creation. While creating value, data deluge presents privacy concerns that may stir a regulatory backlash dampening the data economy and stifling innovation [56]. Data-sharing agreements among multiple parties involved in a BDA project need to be reached for data protection and privacy, including anonymization for open data, access control, rights management, and data usage control. Also, data sharing and privacy should be included as integrate data sharing and privacy as a part of research methodology [20]. Thus, there is an obvious need to address some of the most fundamental concepts of privacy law, including the definition of "personally identifiable information," the role of individual control, and the principles of data minimization and purpose limitation [56].

Without appropriate organizational structures and governance frameworks in place, it is impossible to collect and analyze data across an enterprise and deliver insights to where they are most needed. BDA requires data to be collected and analyzed with centralized governance, which ensures that all big data projects within an organization apply the common standards, protocols, methods, and tools. Meanwhile, organizations can also benefit from having a local, federated model for delivering BDA projects, which can improve the speed of analytics and ensure that gained insight is available to decision makers. Therefore, an organization needs to establish a governance framework that standardizes the operations of BDA across diverse operational areas while also supporting federated project delivery. Currently, only 27 percent of companies combine both [15].

Important research questions related to management challenges in BDA may include but are not limited to:

RQ Set 1: Questions that focus on the composition of (a) BDA strategy, (b) BDA culture, and (c) BDA governance to facilitate the value creation process.

RQ Set 2: Questions that focus on effective approaches or structures to protect user privacy in BDA initiatives.

Value Assessment

Before knowing how to measure strategic business value of BDA, businesses may want to measure functional value through return on investment (ROI). A Gartner survey of IT and business leaders conducted in June 2016 reported that when asked about ROI for big data efforts, a large proportion of companies (43 percent of those planning to invest and 38 percent of those that have already invested) did not know if their ROI would be positive or negative [24]. This uncertainty highlights the challenge in determining and measuring the business value of BDA projects.

According to Lisa Kart, the research director at Gartner, "While the perennial challenge of understanding value (of big data) remains, the practical challenges of skills, governance, funding and ROI come to the fore." Investment in big data involves the cost of various processes described in Figure 3: collecting or licensing data of sufficient quality, cost of required infrastructure (hardware and software) for transferring, cleansing, storing, transforming, integrating, and analyzing data across multiple sources, cost of training professionals and skill sets needed to develop and maintain models, timing of data and result availability, and the salience, explainability, and marginal business value of the results [22]. Very few companies have taken steps to rigorously quantify ROI for their BDA efforts. They need to get more precise in measuring ROI from BDA and examine the relationship between BDA and business outcomes. Practically, however, this is very difficult because there are so many factors or variables that must be considered, and these transcend dollars represented by revenue or cost savings. It requires mapping data, analytics, and business processes to desired business outcomes, and then measuring the impact or success of the outcomes.

Research questions around the challenge of BDA value assessment are important. These could include:

RQ Set 1: Questions around the development of objective short-term and long-term metrics that effectively measure BDA value.

RQ Set 2: Questions around delineation of the sources of tangible and intangible value realized by a company through BDA.

In closing, whether we study the creation of strategic business value of BDA from a pragmatic, process, variance, theory-based, or problem-based perspective, the framework and research questions provided in this study can help guide this important research agenda.

Conclusion

Despite the popularity of BDA in industry, many organizations have failed to reach their strategic goals after investing substantial resources. Very little has been written about why this is the case, and how strategic value from BDA can be fostered. In this study we offer a framework of value creation that has implications for both research and practice. Through this framing we argue that leveraging big data successfully to achieve its strategic business value requires a significant investment in not only data infrastructure and analytic technologies but also skilled analysts and strategic positioning. Businesses need to assess the strategic role of big data analytics and to invest in quality data, state-of-the-art tools, and data-savvy people who understand relevant technologies and data-drive business opportunities.

Further, building BDA capabilities involves developing both data and analytical capabilities, which poses significant challenges. For instance, integrating diverse sources of data to create a platform for analytics, in an environment characterized by four Vs is not easy, as is the ability to frame the right analysis, and discover and act on outputs. Coevolutionary adaptation through "learning by doing" cycles is important for strengthening BDA capabilities. Also, mediating the chain between capabilities and their realization are a variety of value creation mechanisms, which if identified can be nurtured to improve the various targets of value. Successful BDA yields strategic value —in both functional and symbolic forms. Successful initiatives are usually championed through an integrative BDA strategy and strong leadership. BDA requires a strong data-driven culture along with good structures for the governance of data. Importantly, the value realized from BDA is subject to weak links in the value creation process. For instance, excellent tools and data scientists will not help if the data are of poor quality, and big data experts may leave if a company is not committed to provide necessary resources to support analytics and deploy and benefit from the insights discovered from big data.

We apply our framework to a number of use cases on BDA to provide face validity. Further, we offer some core propositions, and five different theoretical logics to frame the root of value creation through BDA. As a complementary perspective, we articulate the research agenda from a "problematization" perspective. By articulating key challenges in realizing strategic value from BDA, we offer important research questions that address BDA challenges.

On a final note, we observe that much of the proliferating research on BDA is tactical, and argues for the novelty of datasets or analysis in addressing narrow business problems. The larger theoretical and strategic impact of the work seems to be under-emphasized. At a minimum, we hope the framing in this study provides a structure for broadening research in this important area and setting up an ambitious research agenda. We encourage future work to expand this framing to value creation and cocreation for entities outside an organization, such as producers, markets, and society.

Funding

This work was partially funded by the National Science Foundation (Award #CNS 1704800 and SES 1527684), National Natural Science Foundation of China (Award #71772075 and 71672074 to Dongsong), and Ministry of Science and Technology of Taiwan (MOST 106-2420-H-110-014 to Ting-Peng). Any opinions, findings, or recommendations expressed here are those of the authors and not necessarily those of the sponsors of this research.

Notes

1. http://www.idc.com/getdoc.jsp?containerId=prUS40560115.
2. A 2017 survey by McKinsey, also emphasized the importance of leadership and culture. See https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/fueling-growth-through-data-monetization.
3. https://www.forbes.com/sites/maribellopez/2016/10/04/how-ebay-uses-big-data-and-machine-learning-to-drive-business-value/#1e915d801f35.
4. https://www.dezyre.com/article/how-big-data-analysis-helped-increase-walmarts-sales-turnover/109.
5. http://cib.db.com/docs_new/GTB_Big_Data_Whitepaper_(DB0324)_v2.pdf.
6. http://www.bain.com/publications/articles/the-value-of-big-data.aspx.
7. A recent survey on "Big Data Trends for 2017" by Tableau, indicates that data variety, not volume or velocity presents the biggest challenge and is the new frontier for BDA value. See https://www.scribd.com/document/339346218/Whitepaper-Top-10-Big-Data-Trends-2017.

References

1. Abbasi, A.; Sarker, S.; and Chiang, R.H. Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, *17*, 2 (2016), ii–xxxii.
2. Barney, J.B. Resource-based theories of competitive advantage: A ten-year retrospective on the resource-based view. *Journal of Management*, *27*, 6 (2001), 643–650.
3. Barney, J.B., and Hesterly, W.S. *Strategic Management and Competitive Advantage: Concepts*. New Jersey, USA: Pearson, 2014.
4. Brun, C., and Hagege, C. Suggestion mining: Detecting suggestions for improvement in users' comments. *Research in Computing Science, 70* (2013), 199–209.
5. Côrte-Real, N.; Oliveira, T.; and Ruivo, P. Understanding the hidden value of business intelligence and analytics (BI&A). Paper presented at the Twentieth American Conference of Information Systems, Association of Information Systems, Savannah, Georgia, August 7–10, 2014.
6. Côrte-Real, N., Oliveira, T., and Ruivo, P. Assessing business value of big data analytics in European firms. *Journal of Business Research, 70* (2017), 379–390.
7. Carter, P. Big Data analytics: Future architectures, skills and roadmaps for the CIO, 2011. IDC. https://triangleinformationmanagement.com/wp-content/uploads/2013/12/bigdata-idc-wp.pdf (accessed on March 26, 2018).
8. Chau, P.Y.; Kuan, K.K.; and Liang, T.-P. Research on IT value: What we have done in Asia and Europe. *European Journal of Information Systems, 16* (2007), 196–201.
9. Chen, H.; Chiang, R.H.; and Storey, V.C. Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, *36*, 4 (2012), 1165–1188.
10. Clemons, E.K.; Dewan, R.M.; Kauffman, R.J.; and Weber, T.A. Understanding the information-based transformation of strategy and society. *Journal of Management Information Systems*, *34*, 2 (2017), 425–456.

11. Corley, C.D.; Cook, D.J.; Mikler, A.R.; and Singh, K.P. Text and structural data mining of influenza mentions in web and social media. *International Journal of Environmental Research and Public Health*, *7*, 2 (2010), 596–615.

12. D'Aveni, R.A. *Hypercompetition: Managing the Dynamics of Strategic Maneuvering*. New York: Free Press, 1994.

13. Datastax. eBay engages customers with personalized recommendations, 2012. http://www.datastax.com/wp-content/uploads/2012/12/DataStax-CS-eBay.pdf (accessed on March 26, 2018).

14. de Vries, A.; Chituc, C.-M.; and Pommeé, F. Towards identifying the business value of big data in a digital business ecosystem: A case study from the financial services industry. In W. Abramowicz, R. Alt, and B. Franczyk (eds.), *International Conference on Business Information Systems*. Springer, 2016, pp. 28–40.

15. EY. Becoming an analytics-driven organization to create value: A report in collaboration with Nimbus Ninety, 2015. http://www.ey.com/Publication/vwLUAssets/EY-global-becoming-an-analytics-driven-organization/$FILE/ey-global-becoming-an-analytics-driven-organization.pdf (accessed on March 26, 2018).

16. Ferrando-Llopis, R.; Lopez-Berzosa, D.; and Mulligan, C. Advancing value creation and value capture in data-intensive contexts. In *Proceedings of 2013 IEEE International Conference on Big Data*. Santa Clara, CA: IEEE, October 6–9, 2013, pp. 5–9.

17. Fichman, R.G. Real options and IT platform adoption: Implications for theory and practice. *Information Systems Research*, *15*, 2 (2004), 132–154.

18. Fisher, D.; DeLine, R.; Czerwinski, M.; and Drucker, S. Interactions with big data analytics. *Interactions*, *19*, 3 (2012), 50–59.

19. Fleurence, R.L.; Curtis, L.H.; Califf, R.M.; Platt, R.; Selby, J.V.; and Brown, J.S. Launching PCORnet, a national patient-centered clinical research network. *Journal of the American Medical Informatics Association*, *21*, 4 (2014), 578–582.

20. George, G.; Haas, M.R.; and Pentland, A. Big data and management. *Academy of Management Journal*, *57*, 2 (2014), 321–326.

21. Gerow, J.E.; Grover, V.; Thatcher, J.B.; and Roth, P.L. Looking toward the future of IT-business strategic alignment through the past: A meta-analysis. *MIS Quarterly*, *38*, 4 (2014), 1059–1085.

22. Gopalkrishnan, V.; Steier, D.; Lewis, H.; and Guszcza, J. Big data, big business: Bridging the gap. BigMine '12: Proceedings of the 1st International Workshop on Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications, Beijing, August 12–16. New York: ACM, 2012, pp. 7–11.

23. Groves, P.; Kayyali, B.; Knott, D.; and Van Kuiken, S. The "big data" revolution in healthcare: Accelerating value and innovation, 2013. http://www.pharmatalents.es/assets/files/Big_Data_Revolution.pdf (accessed on March 26, 2018).

24. Heudecker, N., and Hare, J. Survey analysis: Big data investments begin tapering in 2016. Gartner Report. 2016. https://www.gartner.com/doc/3446724/survey-analysis-big-data-investments (accessed on March 26, 2018).

25. ISACA. Generating value from big data analytics. 2014. http://www.isaca.org/Knowledge-Center/Research/Documents/Generating-Value-from-Big-Data-Analytics_whp_Eng_0114.pdf?regnum=371545 (accessed on Jan 12, 2018).

26. Jiang, P., and Liu, X.S. Big data mining yields novel insights on cancer. *Nature Genetics*, *47*, 2 (2015), 103–104.

27. Kohli, R., and Grover, V. Business value of IT: An essay on expanding research directions to keep up with the times. *Journal of the Association for Information Systems*, *9*, 1 (2008), 23–39.

28. Kumar, R.L. A framework for assessing the business value of information technology infrastructures. *Journal of Management Information Systems*, *21*, 2 (2004), 11–32.

29. Kune, R.; Konugurthi, P.K.; Agarwal, A.; Chillarige, R.R.; and Buyya, R. The anatomy of big data computing. *Software: Practice and Experience*, *46*, 1 (2016), 79–105.

30. LaValle, S.; Lesser, E.; Shockley, R.; Hopkins, M.S.; and Kruschwitz, N. Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, *52*, 2 (2011), 21–32.

31. Manyika, J.; Chui, M.; Brown, B.; Bughin, J.; Dobbs, R.; Roxburgh, C.; and Byers, A. H. Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute. 2011.

32. Markus, M.L. Maybe not the king, but an invaluable subordinate: A commentary on Avison and Malaurent's advocacy of "theory light" IS research. *Journal of Information Technology*, *29*, 4 (2014), 341–345.

33. Mata, F.J.; Fuerst, W.L.; and Barney, J.B. Information technology and sustained competitive advantage: A resource-based analysis. *MIS Quarterly*, *19*, 4 (1995), 487–505.

34. McAfee, A., and Brynjolfsson, E. Big data: The management revolution. *Harvard Business Review*, 2012, 60–68.

35. McGrath, R.G. A real options logic for initiating technology positioning investments. *Academy of Management Review*, *22*, 4 (1997), 974–996.

36. Melville, N.; Kraemer, K.; and Gurbaxani, V. Review: Information technology and organizational performance: An integrative model of IT business value. *MIS Quarterly*, *28*, 2 (2004), 283–322.

37. Nguyen, T.; Larsen, M.E.; O'Dea, B.; Phung, D.; Venkatesh, S.; and Christensen, H. Estimation of the prevalence of adverse drug reactions from social media. *International Journal of Medical Informatics, 102* (2017), 130–137.

38. Nicola, S.; Ferreira, E.P.; and Ferreira, J.J.P. A quantitative model for decomposing and assessing the value for the customer. *Journal of Innovation Management*, *2*, 1 (2014), 104–138.

39. OuYang, Y.-C. Information system capabilities and organizational performance: Comparing three models. *Pacific Asia Journal of the Association for Information Systems*, *9*, 1 (2017), 1–28.

40. Park, E.H.; Ramesh, B.; and Cao, L. Emotion in IT investment decision making with a real options perspective: The intertwining of cognition and regret. *Journal of Management Information Systems*, *33*, 3 (2016), 652–683.

41. Pettigrew, A., and Whipp, R. *Managing Change for Competitive Success*. Oxford: Basil Blackwell, 1991.

42. Pipino, L.L.; Lee, Y.W.; and Wang, R.Y. Data quality assessment. *Communications of the ACM*, *45*, 4 (2002), 211–218.

43. Qiao, Z.; Zhang, X.; Zhou, M.; Wang, G.A.; and Fan, W. A domain oriented LDA model for mining product defects from online customer reviews. *50th Hawaii International Conference on Systems Sciences*. Waikoloa, Hawaii, 2017, pp. 1821–1830.

44. Roberts, N.; Campbell, D.E.; and Vijayasarathy, L.R. Using information systems to sense opportunities for innovation: Integrating postadoptive use behaviors with the dynamic managerial capability perspective. *Journal of Management Information Systems*, *33*, 1 (2016), 45–69.

45. Roberts, N.; Galluch, P.S.; Dinger, M.; and Grover, V. Absorptive capacity and information systems research: Review, synthesis, and directions for future research. *MIS Quarterly*, *36*, 2 (2012), 625–648.

46. Sambamurthy, V.; Bharadwaj, A.; and Grover, V. Shaping agility through digital options: Reconceptualizing the role of information technology in firms. *MIS Quarterly*, *27*, 2 (2003), 237–263.

47. Schilsky, R.L.; Michels, D.L.; Kearbey, A.H.; Yu, P.P.; and Hudis, C.A. Building a rapid learning health care system for oncology: The regulatory framework of CancerLinQ. *Journal of Clinical Oncology*, *32*, 22 (2014), 2373–2379.

48. Schryen, G. Revisiting IS business value research: What we already know, what we still need to know, and how we can get there. *European Journal of Information Systems*, *22*, 2 (2013), 139–169.

49. Seddon, P.B.; Staples, S.; Patnayakuni, R.; and Bowtell, M. Dimensions of information systems success. *Communications of the Association for Information Systems*, *2*, 3 (1999), 1–61.

50. Soh, C., and Markus, M.L. How IT creates business value: A process theory synthesis. *1995 International Conference on Information Systems, Amsterdam*, NL, 1995, pp. 29–41.

51. Spence, M. Job market signaling. *The Quarterly Journal of Economics*, 87, 3 (1973), 355–374.

52. Sun, D.; Du, Y.; Xu, W.; Zuo, M.; Zhang, C.; and Zhou, J. Combining online news articles and web search to predict the fluctuation of real estate market in big data context. *Pacific Asia Journal of the Association for Information Systems*, *6*, 4 (2014), 19–37.

53. Sun, H. A longitudinal study of herd behavior in the adoption and continued use of technology. *MIS Quarterly*, *37*, 4 (2013), 1013–1041.

54. Tallon, P.P.; Kraemer, K.L.; and Gurbaxani, V. Executives' perceptions of the business value of information technology: a process-oriented approach. *Journal of Management Information Systems*, *16*, 4 (2000), 145–173.

55. Teece, D.J., and Pisano, G. Dynamic capabilities and strategic management. *Strategic Management Journal*, *18*, 7 (1997), 509–533.

56. Tene, O., and Polonetsky, J. Big data for all: Privacy and user control in the age of analytics. *Northwestern Journal of Technology and Intellectual Property*, *11*, 5 (2012), 239–273.

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

58. Wang, P. Chasing the hottest IT: Effects of information technology fashion on organizations. *MIS Quarterly*, *34*, 1 (2010), 63–85.

59. Wernerfelt, B. The resource-based view of the firm: Ten years after. *Strategic Management Journal*, *16*, 3 (1995), 171–174.

60. Wu, L., and Zhang, M.-L. Multi-Label classification with unlabeled data: An inductive approach. In C. S. Ong and T. B. Ho (eds.), *Asian Conference on Machine Learning*, Canberra, November 13–15, 2013, pp. 197–212.

61. Zhang, M.; Ranjan, R.; Haller, A.; Georgakopoulos, D.; and Strazdins, P. Investigating decision support techniques for automating cloud service selection. *In Proceedings of the 4th IEEE International Conference on Cloud Computing Technology and Science (CloudCom)*, Taipei, December 3–6. Taipei: IEEE, 2012, pp. 759–764.

62. Zhang, D., Zhou, L., Kehoe, L. J., and Kilic, I. What Online Reviewer Behaviors Really Matter? A Study of Effects of Verbal and Nonverbal Behaviors on Online Fake Review Detection. *Journal of Management Information Systems*. *33*, 2 (2016), pp. 456–481.