BUSHOR-1561; No. of Pages 12
ARTICLE IN PRESS
Business Horizons (2019) xxx, xxx—xxx

Available online at www.sciencedirect.com
ScienceDirect

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

www.elsevier.com/locate/bushor

Implementing big data strategies: A managerial perspective

Pooya Tabesh [a,*], Elham Mousavidin [a], Sona Hasani [b]

[a] Cameron School of Business, The University of St. Thomas - Houston, 3800 Montrose Blvd., Houston, TX 77006, U.S.A.
[b] The University of Texas at Arlington, 500 UTA Blvd., Arlington, TX 76010, U.S.A.

KEYWORDS
Big data analytics;
Decision making;
Big data strategy;
Big data applications;
Strategy implementation;
Machine learning

Abstract Despite considerable recent advances in big data analytics, there is substantial evidence that many organizations have failed to incorporate them effectively in their own decision-making processes. Advancing the existing understandings, this article lays out the steps necessary to implement big data strategies successfully. To this end, we first explain how the big data analytics cycle can provide useful insights into the characteristics of the environments in which many organizations operate. Next, we review some common challenges faced by many organizations in their uses of big data analytics and offer specific recommendations for mitigating them. Among these recommendations, which are rooted in the findings of strategy implementation research, we emphasize managerial responsibilities in providing continued commitment and support, the effective communication and coordination of efforts, and the development of big data knowledge and expertise. Finally, in order to help managers obtain a fundamental knowledge of big data analytics, we provide an easy-to-understand explanation of important big data algorithms and illustrate their successful applications through a number of real-life examples.

1. The big data rush

After James W. Marshal discovered gold nuggets in the Sacramento Valley in 1848, thousands of prospective gold miners traveled by land and sea to California to seek their fortunes from large quantities of gold hidden in the riverbeds ("California Gold

* Corresponding author
E-mail addresses: tabeshp@stthom.edu (P. Tabesh), mousave@stthom.edu (E. Mousavidin), sona.hasani@mavs.uta.edu (S. Hasani)

https://doi.org/10.1016/j.bushor.2019.02.001

Rush," 2010). At the time, the process of mining gold (i.e., prospecting activities and extraction of gold from its ore) was laborious and not all of the fortune seekers were successful in their search for a road to quick riches (Rohrbough, 1998). Nevertheless, the ensuing gold rush was a significant event in the American 19th century history with many socio-economic consequences (Owens, 2002).

Today, the ability to collect and process massive amounts of data rapidly has created a new gold rush, and many organizations and their analysts have devoted considerable time, money, and effort to obtain profitable insights into the behavioral and structural characteristics of their environments. As was the case with the American gold prospectors of the 1840s, a new generation of dreamers have emerged bent on seeking fame and fortune, this time through collecting big data and extracting valuable business insights from them.

2. Big data and big data dreams

Big data refers to the large and complex data assets that require cost-effective management and analysis for extraction of insights from them (Gupta & George, 2016). Four specific features, also known as the 4Vs, characterize big data (Kietzmann, Paschen, & Treen, 2018; Sivarajah, Kamal, Irani, & Weerakkody, 2017):

1. *Volume* refers to the large scale of big data, which requires innovative tools for their collection, storage, and analysis.

2. *Velocity* refers to the rate at which the data are generated or updated, pointing to the real-time nature of big data.

3. *Variety* refers to the variation in types of data. Big data can come in diverse and dissimilar forms from multiple sources, such as texts, spreadsheets, audios, videos, and sensors. Big data are usually unstructured (e.g., text, audio) and are not organized in a structured manner in a relational database (e.g., tables, spreadsheets).

4. *Veracity* refers to the complex structures of big data assets that make them ambiguous, imprecise and inconsistent. For example, the data related to consumer opinions posted on social media can be biased, inaccurate, and ambiguous.

In the era of the big data rush, we are experiencing unprecedented changes in the business data analytics landscape that have generated growing interest in advanced big data tools and techniques. Some refer to these new developments as *data fetish* (Rasmussen & Ulrich, 2015), while others find the term *data deluge* ("The Data Deluge," 2010) a more accurate description. The extreme enthusiasm in big data analytics is legitimate as the digital data in the universe doubles every 2 years and the size of the useful business-related data continues to grow exponentially (Turner, Gantz, Reinsel, & Minton, 2014). Indeed, some 90% of the digital data available today have been generated in the past 2 years (Henke, Libarikian, & Wiseman, 2016).

The confluence of data proliferation, algorithmic advancement, and more powerful computing and storage facilities have opened new possibilities for transformation of data into business insights, decisions, and actions (Chui, Kamalnath, & McCarthy, 2018). In line with these trends, organizations are racing to keep up with the changes and leverage the benefits of useful information embedded within large volumes of data. In this respect, the adoption rate of popular big data analytics tools is rising in every major industry (Mayhew, Saleh, & Williams, 2016; Mazzei & Noble, 2017). In fact, based on International Data Corporation (IDC) estimates, the big data analytics market will surpass $203 billion in worldwide revenues by 2020 (Press, 2017).

This growth has led more and more organizations to seek their fortunes through new big data strategies. Nevertheless, and as it was the case with the process of extracting gold during the American gold rush, extracting valuable insights from big data and transforming them into useful actions is a task with which many organizations continue to struggle (Zeng & Glaister, 2017).

3. High failure rates of the big data strategies: What now?

In spite of the popularity of big data analytics as a game changer in revolutionizing the way organizations make decisions and operate, surveys show that around 80% of businesses have failed to implement their big data strategies successfully (Asay, 2017; Gartner, 2015). More than 65% of organizations have reported below average returns on their data management investments (Baldwin, 2015). While many organizations have jumped on the bandwagon to take advantage of big data opportunities (Ashayeri, 2016), only a small percentage of them have benefitted from their investments (Ross, Beath, & Quaadgras, 2013). Overall, the evidence suggests that, regardless of their efforts, many organizations cannot realize their "big data dreams" (Mazzei & Noble, 2017) in an effective manner.

In interpreting the relatively high failure rates of organizational big data strategies, previous research has mainly focused on a shortage of organizational resources (e.g., data, infrastructure, talent) required for the adoption of big data analytics. Surprisingly, while managers—as opposed to data scientists—should be at the forefront of this process, the roles and responsibilities of managers in implementing big data strategies have remained largely underexplored (Mikalef, Pappas, Krogstie, & Giannakos, 2018; Zeng & Glaister, 2017).

In this article, we aim to combine and extend the findings of information technology and strategy implementation research to delineate managerial responsibilities for successful big data strategy implementation. In doing so, we first introduce the big data analytics cycle to elaborate on the process of transforming big data into insights, decisions, and actions. Then, through an extensive review of the literature, we summarize the most important challenges that adversely affect the implementation process and offer specific recommendations toward effective implementation of big data strategies. Among these recommendations, which are rooted in the strategy implementation literature, we emphasize managerial responsibilities in providing continued commitment and support in facilitating effective communication and coordination of efforts and in developing the big data knowledge and expertise.

As an additional contribution, we provide managers with a better understanding of different big data analytics algorithms and their applications. As we discuss in detail, while a fundamental understanding of big data and their applications is crucial for successful implementation of big data strategies, important details related to different types of big data analytics techniques rarely have been provided to practitioners.

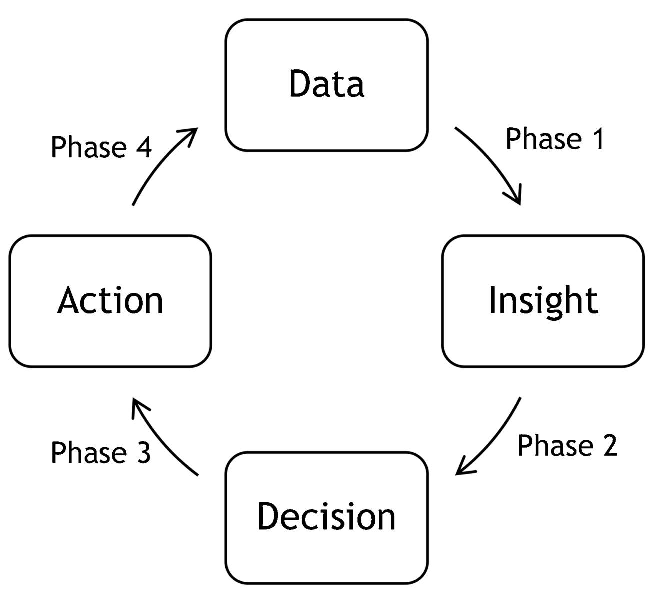
4. Big data analytics cycle

The goal of big data analytics is to enhance organizational decision making and decision execution processes. Informed decision making is one of the building blocks of organizational success, and the importance of comprehensive analysis of information before making operational and strategic decisions has been highlighted in the works of many organizational researchers and practitioners (e.g., Dean & Sharfman, 1996; Fredrickson, 1984). In making important decisions, managers collect data, generate several alternative strategies, and carefully evaluate these strategies and their outcomes before making final decisions. Once implemented, the realized outcomes of the decision will be evaluated to generate additional information that is cycled back into the subsequent decision making phases.

The process of big data analytics resembles the aforementioned comprehensive decision making scheme used to enhance decision quality and outcomes. The big data analytics cycle is comprised of four important phases depicted in Figure 1:

- *Phase 1*: Large, diverse, and usually unstructured data are collected from internal and external

Figure 1. Big data analytics cycle



sources and processed (i.e., cleaned and analyzed) using advanced analytics tools and algorithms in order to generate insights. These insights are then interpreted by decision makers and used in the process of decision making.

- *Phase 2*: Insights generated in Phase 1 are transformed into decisions. This is done by managers who contextualize the insights generated from their data analysis and attach meaning to them (Zeng & Glaister, 2017).

- *Phase 3*: The decisions are transformed into specific operational actions. In other words, decisions are executed.

- *Phase 4*: Transformation of decisions into actions generates additional outcomes (i.e., data points) which are cycled back into the process for future decision making efforts.

This way, a self-perpetuating cycle of big data analytics can significantly benefit organizational decision making processes and outcomes. Large volumes of both the internal operations data (e.g., inventory updates, employee performances, financial transactions, consumer behaviors, sales) and the data collected from external sources (e.g., customer ratings, e-commerce communications, social media) are gathered and transformed into actionable insights.

For example, Walmart Labs—the big data analytics arm of Walmart—via analysis of large volumes of real-time data related to online and in-store sales, can generate new insights about daily or weekly changes in needs and preferences of Walmart customers (i.e., generation of insights from data). Then, the generated insights that are indicative of previously unknown trends in consumer purchasing habits can lead managers to modify supply schedules or product offerings to address emerging customer needs (i.e., transformation of insights into decisions and actions). Subsequently, the implemented changes can influence demand and trigger additional changes in the purchasing behavior of customers that creates new data points that are collected and stored for future analysis (i.e., transformation of actions into new data).

As another example, companies in the automobile industry can leverage the benefits of the big data analytics cycle to explore existing patterns in driver behaviors in order to make important business decisions. In a competition organized by Kaggle, Google's recently acquired big data analytics crowdsourcing platform, participants were tasked with identifying the factors contributing to reliability of a car. After combining massive data sets from different sources (e.g., demographic data, maintenance information, customer complaints, car specifications), it was revealed that orange-colored cars of any model have significantly fewer after-purchase problems (Gage, 2014). Such a surprising discovery could be indicative of specific personality characteristics or patterns of behavior in people who purchase orange vehicles. In interpreting and contextualizing these insights, one could conclude that buyers who purchase orange-colored cars tend to take better cares of their automobiles. Car dealerships, insurance companies, and automobile service providers can transform similar insights into specific pricing decisions or advertising campaigns.

Each phase in the big data analytics cycle entails specific tasks, requires specific resources, and needs critical attention from specific organizational actors. In particular, managers and data scientists have important influences on different phases of the cycle. Data scientists are mainly responsible for technical tasks related to data collection and analysis (i.e., Phases 1 and 4) and can help the interpretation process by communicating technical findings to managers (Phase 3). Managers play a primary role in interpretation and execution phases of the cycle. From devising the right data acquisition and management systems to interpreting the insights generated from big data analysis, managerial engagement in the process is required. Beyond the roles they play in fulfilling specific tasks in each phase, managers are responsible for the orchestration of efforts in the entire cycle to implement organizational big data initiatives successfully.

Table 1 presents the information related to main tasks, key required resources, and critical actors in each phase of the big data analytics cycle. In the next section, we review the factors that can limit the effective implementations of big data strategies, followed by a discussion of the responsibilities managers have to minimize them.

5. Barriers to effective implementation of big data strategies

Academics and practitioners have enumerated a long list of barriers to the full realization of big data benefits in organizations. Here, we summarize our review of the literature by discussing the technological and cultural barriers as two major categories of common constraints faced by many organizations (Alharthi, Krotov, & Bowman, 2017; LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011).

Table 1. A summary of main tasks, key resources, and critical actors

	Four phases of the Big Data Analytics cycle			
	Phase 1: Data → Insight	**Phase 2: Insight → Decision**	**Phase 3: Decision → Action**	**Phase 4: Action → Data**
Main Task	Data Cleaning & Analysis	Interpretation of Insights	Decision Execution	Data Collection & Storage
Key Resource/ Capability	Technical Infrastructure & Capability	Managerial Capability to Contextualize	Decision Execution Capability	Technical Infrastructure & Capability
Critical Actors	Data Scientists Managers (all levels)	Managers (all levels) Data Scientists	Managers (all levels)	Data Scientists Managers (all levels)

5.1. Technological barriers

The specific characteristics of big data impose certain technological constraints for big data management and analysis. Technological barriers range from the costly infrastructures required for big data acquisition, storage, and analysis to the shortage of qualified data scientists and analysts (Alharthi et al., 2017; Sivarajah et al., 2017). According to John Bantleman, CEO of RainStor, which sells database software used for big data projects, big data analytics systems need revamped IT infrastructures, as the conventional data management tools are not scalable to keep up with the pace of new data creation (Savitz, 2012). Indeed, as IBM reports, although more than 70% of companies are aware of the importance of the data analytics infrastructures in gaining competitive advantage, only about 22% can effectively govern their IT infrastructure ("IBM Report, n.d.").

Technological challenges impact the entire big data analytics cycle. Implementation of big data initiatives calls for capital investment in building or buying new data management systems (e.g., Hadoop, NoSQL) to enable effective storing and analysis of big data. In addition, unstructured data require extensive processing (e.g., cleaning) before they can be used for analysis. Acquisition of big data-specific know-how poses another significant challenge. While the demand for data scientists and analysts is projected to grow 30% over the next 3 years (Miller & Hughes, 2017), a survey of 430 senior executives shows that 66% of organizations are currently unable to successfully fill their available data scientist positions with qualified candidates (Boulton, 2015). In the words of Hal Varian, Google's chief economist: "Data [are] so widely available and so strategically important that the scarce thing is the knowledge to extract wisdom from it" (KPMG, 2014). In December 2017 alone, *Forbes* reported about 1,829 open positions on LinkedIn for machine learning engineering, which is merely a subset of the plethora of big data analytics jobs (Columbos, 2017).

The list of technical challenges extends beyond the problems related to technical infrastructure and talent to also include lack of technical acumen by senior managers. In fact, research shows the majority of big data investments fail due to managerial misunderstandings about the process or due to their inability to incorporate the insights gleaned from big data into organizational decisions (Ross et al., 2013). As Kyle Evans, chief data officer at the Australian property information services business CoreLogic, stated (Collins, 2014):

> While there's a lack of big data skills on the ground—in the IT or analyst staff—there's also a lack of big data capabilities in management. This can manifest in misunderstanding what big data actually are, or how to get the most out of a big data project.

Increasing concerns regarding consumer data ownership and privacy pose another technological challenge to many organizations (Douglas, 2013; Hirsch, 2016). In a recently publicized case, the data from 50 million Facebook profiles harvested by a third-party organization caused many raised eyebrows and resulted in reputation damages for Facebook and its CEO, Mark Zuckerberg, who was questioned for more than 10 hours by U.S. senators and representatives regarding the company's privacy policies. While data privacy legislation is on the rise (e.g., EU General Data Protection Regulation), organizations need to devise technical procedures to ensure that their activities in the entire big data analytics cycle comply with the new regulations. Such requirements will incur major additional costs to the organizations that use big data (Kottasová, 2018).

5.2. Cultural barriers

Organizational culture is a set of values, beliefs, and attitudes shared by the members of an organization (Schein, 1990). In terms of cultural barriers to big data strategy implementation, the challenges related to developing a data-driven culture have been highlighted in the literature. A *data-driven culture* is defined as "the extent to which organizational members (including top-level executives, middle managers, and lower-level employees) make decisions based on the insights extracted from data" (Gupta & George, 2016, p. 5). According to Ross et al. (2013), the lack of data-driven culture is among the major reasons for the high failure rate of big data projects. Executives in many organizations rely heavily on their prior experiences or intuitive hunches instead of following evidence-based and data-driven decision processes (McAfee, Brynjolfsson, Davenport, Patil, & Barton, 2012). If top managers do not value data-driven decision making, their behavior will affect the decision patterns at all levels of the organization. For example, in organizations in which the highest paid person's opinions (HiPPOs) have the highest influence on the decision making process, data-driven decision making cannot thrive. HiPPOs are the antithesis to data-driven decisions. According to Anderson and Li (2017), the best data-driven insights generated by skillful data analysts cannot be executed "if the report sits unopened on a desk." When HiPPOs are in effect, "managers don't care what the data says, especially when it disagrees with their preconceived notions" (Anderson & Li, 2017).

Difficulties in creating a unified vision about the organizational big data strategy is another major roadblock in front of effective implementation (Rogers & Meehan, 2007). For instance, a survey of executives from 108 countries showed that many organizational decision makers lack a common understanding of what big data analytics is and what benefits it can generate for their business operations and outcomes (LaValle et al., 2011). McAfee et al. (2012), in their evaluation of big data revolution, interviewed executives from 330 public North American organizations. Their study revealed that not every manager was embracing data-driven decision making and highlighted the importance of "articulating a compelling vision" to create buy-in (McAfee et al., 2012, p. 5). As an example, Royal Bank of Scotland (RBS) could align its marketing staff with the new big data initiatives through refocusing of its marketing objectives. RBS developed a new analytics department to reach its customers and help them with account management. "We no longer use analytics to get products out of the door . . . . but to help customers get most out of what we offer," said Christian Nelissen RBS head of analytics (Marr, 2016). He further highlighted that the new approach could reduce resistance from marketing staff who were initially not open to the idea of a using data analytics tools, "They're at the point where they understand what the data is trying to do and feel it helps them have good conversations with the clients" (Marr, 2016).

Cultural barriers, hand in hand with the technological barriers, can dampen effective implementation of big data strategies. The important responsibilities of managers for addressing these barriers are discussed in the next section.

6. Big data strategy implementation: Managerial responsibilities

Implementation of business strategies is a complicated process and most of the formulated strategies cannot be executed effectively (Hrebiniak, 2006). When it comes to big data strategies, the implementation process is even more complicated due to the aforementioned technological and cultural challenges specific to this area. Key organizational decision makers play a central role in the success or failure of big data initiatives and are responsible for creating a unified vision regarding the approach to big data analytics in organizations (Rasmussen & Ulrich, 2015). They can create and sustain a data-driven culture that values evidence-based decision making and encourages transformation of data into insights, insights into decisions, and decisions into successful execution. In the words of Gupta and George (2016, p. 5): "The intelligence gleaned from data will be of little use to an organization if its managers fail to foresee the potential of newly extracted insights."

Research on successful strategy implementation reports three categories of managerial levers that can help mitigate implementation challenges (Crittenden & Crittenden, 2008; Noble, 1999; Tawse & Tabesh, 2017):

1. *Structural influences*: Structural influences refer to existing budgets, plans, and control mechanisms that support implementation of a specific strategy. Managerial structural support for the formulated strategies is an important prerequisite for successful implementation.

2. *Relational influences*: Relational influences refer to interpersonal mechanisms that influence communication of objectives and coordination of efforts at different levels of the organization.

Table 2. Recommendations for successful implementation of big data strategies

Challenges impeding implementation efforts	Managerial responsibilities for successful implementation	Recommendations
Technological (T) • Costly data management tools • Lack of managerial analytics knowledge • Technical misunderstandings between managers and data scientists • Inherent challenges related to big data (e.g., 4Vs) • Technical requirements in compliance with data ownership and privacy regulations Cultural (C) • Lack of a shared understanding of big data and its goals • Extensive reliance on intuitive or experiential decision-making approaches • Dominance of HiPPOs in the decision making process	Providing commitment and support	• Have a long-term orientation in your big data analytics investments (C, T) • Foster a data-driven decision-making culture (C) • Assign dedicated data analytics teams and provide them with resource, data access, and bureaucratic immunity (C, T)
	Effective communication and coordination	• Clearly communicate the business problem being addressed to technical staff (C, T) • Create a common understanding of big data goals (e.g., assign translators familiar with technical and business languages) (C, T) • Encourage cross-functional collaborations (C, T) • Disseminate the generated data-driven insights and share the newly generated data points with the rest of organization (C, T)
	Gaining managerial analytics acumen	• Obtain and maintain relevant analytics knowledge through various sources (e.g., practitioner and academic sources) (T) • Help staff gain a general understanding of the big data analytics cycle from big data to insights, decisions, and actions that contributes to generation of a data-driven culture (C, T) • Help staff gain a basic understanding of statistics (C, T) • Incentivize staff to gain big data analytics knowledge (T)

For successful implementation, strategic initiatives should be well understood across the organization.

3. *Knowledge influences*: Knowledge influences refer to the managerial strategy-specific expertise that contributes to successful implementation. For instance, effective implementation of international growth strategies requires managers at all levels of the organization to possess fundamental knowledge of international markets.

We build on this categorization and extend the three levers of strategy implementation to the case of big data strategies. In the next section, we introduce and discuss the three important managerial responsibilities for addressing the technological and cultural barriers and successfully implementing big data strategies (see Table 2).

6.1. Structural influences: Providing continued commitment and support

Commitment of company leadership to the formulated strategies and providing financial and structural support throughout the implementation process are important prerequisites for successful adoption of new organizational initiatives (Crittenden & Crittenden, 2008). The same principle applies to implementation of big data strategies, including the roles managers should play in providing support to overcome the technological and cultural barriers related to big data (Mikalef et al., in press). According to an article published in *McKinsey Quarterly*, full exploitation of big data analytics is impossible without senior and middle management's involvement in the process (Mayhew et al., 2016). Furthermore, time must be considered as an important resource along with capital investment in planning and implementation of big data strategies. In other words, managers should not expect their investments in technology to generate immediate returns as this process takes time to pay off (Mata, Fuerst, & Barney, 1995). This limitation is specifically significant when it comes to implementing advanced technology-related strategies, which are known to require "years rather than months" to generate results (Vidgen, Shaw, & Grant, 2017, p. 632).

Managers should stay committed to data-driven decision making and be persistent in providing structural support for big data initiatives. For example, research by Bain & Company revealed that managers in successful data-driven firms create dedicated data-insights task forces that are provided with adequate resources, bureaucratic immunity, and data access. These teams consistently explore business processes in different departments and identify problems for which big data analytics can offer solutions (Wegener & Sinha, 2013).

Overall, managerial commitment and support can significantly mitigate the cultural and technological barriers to big data strategies. For instance, managerial commitment to big data projects contributes to the generation of a data-driven culture by sending the right signals to everyone in the organization (e.g., Adrian, Abdullah, Atan, & Jusoh, 2018). Providing ample financial support for activities such as talent acquisition, data acquisition, and data management systems will address the aforementioned technological needs.

6.2. Relational influences: Effective communication and coordination of efforts

Relational mechanisms aimed at communicating strategic objectives and coordinating organizational implementation efforts are central to the process of strategy implementation (Noble, 1999; Peng & Litteljohn, 2001). Establishing a common understanding of big data goals among managers and their technical teams is an important step in addressing the implementation barriers and realizing big data dreams. While managers define business goals of big data initiatives, technical staff (e.g., data scientists) are responsible for data collection, cleaning, and analysis to generate insights from data (Davenport & Patil, 2012). Thus, for effective transformation of big data into meaningful insights and decisions, business goals of big data analytics should be communicated effectively to the technical staff.

At the early stages of the big data analytics cycle, managers play a critical role in communicating the business problems or the why of business analytics to the technical staff, who are responsible for technical aspects such as collecting and analyzing big data (e.g., Mayhew et al., 2016). Toward the end of the big data analytics cycle, managers should disseminate the data-driven insights to the middle and functional managers at the forefront of implementation. At all phases throughout the cycle, the communication channels between managers and data scientists should remain open. According to a 2014 report by KPMG, effective interaction between data scientists and managers is a critical success factor for big data analytics initiatives. Since, on many occasions, scientists and managers come from different backgrounds, effective communication is required to build trust and minimize misunderstandings. Eddie Short, partner and lead for data and analytics at KPMG, highlighted an example of a common misunderstanding between managers and scientists: “C-level executives scream for more data, seeing it as a panacea. Showing them that not all data is business critical—some is utterly useless—can be difficult” (KPMG, 2014).

The coordination of data-driven decision making processes at all levels of an organization is another important responsibility of managers in mitigating the challenges related to big data. In recent years, this importance necessitated new executive positions such as Chief Data/Digital Officer (CDO) to streamline the process and facilitate communication (Aiken & Gorman, 2013). In order to encourage cross-functional collaboration in the implementation of big data analytics strategies, the importance of building multiskilled teams has been emphasized. Managers are advised to form big data analytics teams comprised of data scientists and engineers with technical knowledge and translators who are familiar with both technical and business languages. These translators facilitate the communication between managers and technical staff and can help managers interpret the generated insights before transforming them into business decisions (Mayhew et al., 2016). This practice, over time, can create a rich culture of open communication that is instrumental in addressing big data analytics hurdles.

In sum, effective communication of big data goals can mitigate the cultural and technological barriers to successful implementation of big data strategies. Open organizational communication channels not only contribute to the generation of a shared vision about big data strategies (Chen, Chiang, & Storey, 2012) but also facilitate a coordinated effort in addressing the challenges related to data collection, storage, management, and analysis.

6.3. Knowledge influences: Gaining fundamental managerial analytics acumen

Managerial fundamental knowledge related to the strategic problem at hand is required for effective implementation of business strategies (Ethiraj, Kale, Krishnan, & Singh, 2005). In the context of big data strategies, managerial analytics acumen is a prerequisite for effective creation of value out of big data investments. Big data analytics tools, like other organizational resources, should be deployed properly in order to create desired outcomes. Thus, it is imperative for managers at different levels of the organization to be familiar with

Table 3. An overview of important big data analytics algorithms and applications

Categories	Algorithms Main Applications	Real-world examples
Descriptive Tools	**Clustering** Segmentation (e.g., customers, employees, products, services)	JPMorgan Chase uses clustering to segment millions of customers into different groups with different spending habits based on a combination of characteristics such as transaction types and account balances. This can lead to devising effective strategies for delivering the best set of products to its customers.
	Association Rule Discovery Bundling (e.g., products, services, customers) for different purposes	Large supermarkets use this technique. In this regard, association rule discovery can reveal that, if customers buy onions, there is an 80% probability that they will also buy tomatoes. Such insight, derived from the analysis of billions of data points, enables Walmart store managers to quickly determine how to stock store shelves.
	Sequential Patterns Discovery Recommendation systems	Pharmacies use this technique to identify the temporal relationships between prescribed medications. The insights derived from such approach can trigger patient-specific recommendations. As another example, Netflix and Amazon Prime can analyze a user's previous choices of movies to recommend a set of movies and sort them based on the predicted preference of the user.
Predictive Tools	**Classification** Prediction of the class or group to which a new entity (e.g., customer or product) belongs	Large banks can rely on a classification model to predict if an existing customer is likely to open a new savings account. Such a predictive model, built through the analysis of behaviors (e.g., habits and transactions) and characteristics (e.g., age and gender) of customers, may enable the bank to identify easy-to-convert customers and engage in targeted advertising activities to attract them.
	Regression Prediction of a variable of interest (e.g., sales price, consumer behavior)	Regression is heavily used in many industries. In the retail industry, for example, regression models predict future sales based on market conditions and other relevant historical data of customers and competitors. Similarly, sophisticated regression models can be used to estimate real estate prices based on analysis of various data points such as time of the year and interest rates.
	Anomaly Detection Identifying outliers in data	Deutsche Bank utilizes an anomaly detection tool to discover fraudulent activity at the point of transaction. As another example, Intel uses image-based anomaly detection tools for the purpose of manufacturing quality control in its microchip production and assembly lines.

general concepts and applications related to these techniques (Gupta & George, 2016). In other words, mitigating big data implementation challenges is not possible without senior and middle management's general understanding of the applications. These managers, as the most influential decision makers, are the closest to the very internal and external business challenges big data analytics attempts to tackle. Therefore, gaining a holistic understanding of the technical realities related to the big data analytics cycle is crucial for executives and can help them better envision the possible benefits of big data and orchestrate organizational efforts in this regard. Such understanding will enable managers to ask the 'right' questions for which big data analytics can generate valuable insights. Unfortunately, the lack of general understanding regarding the application of these methods has contributed to isolated organizational efforts to leverage big data in different functions but without benefiting from a unifying theme or purpose (Rasmussen & Ulrich, 2015).

Although a wide range of data analytics tools is available to organizations, organizational resources are limited and a hasty choice in acquiring costly tools, applications, or human knowledge may not always guarantee success (Sivarajah et al., 2017). That is, the effectiveness of these tools depends on how well they fit the problem domain at hand and how much they empower the organizations in addressing their most pressing needs (Mayhew et al., 2016). To this end, managers should learn the basics of data analytics to integrate them effectively into existing decision processes. In this respect, many organizations have started to educate their managers and staff with working knowledge of data science (Zettelmeyer, 2015). In addition to providing incentives for employees to gain knowledge from external sources (e.g., degrees, certificates), organizations can facilitate internal knowledge diffusion. At Utah's Loveland Living Planet Aquarium, a group of managers and data scientists organizes regular lunch-and-learn sessions to refresh employees' statistical knowledge, to familiarize them with the terminology, and to instill in them analytical thinking abilities (Dykes, 2017).

Overall, managerial analytics acumen can reduce the technological and cultural challenges facing big data implementation efforts. Familiarity of organizational decision makers with the fundamentals of big data analytics leads to a shared understanding of big data goals that is essential for the creation of a data-driven culture. In addition, increased organizational data literacy contributes to the technical aspects related to the big data analytics cycles. As explained before, managerial misunderstandings about big data is a major cause for big data strategy failures (Ross et al., 2013). In Table 2, we summarize the barriers to implementation of big data strategies and provide specific recommendations for successful implementation.

Given the importance of fundamental analytics knowledge for effective implementation of big data strategies, the following section briefly reviews some of the basic big data analytics algorithms and their applications. Existing research on fundamental methodological concepts of big data are extremely technical, making them inaccessible to most general managers. Thus, a simple and non-technical introduction to big data tools and their applications may prove useful.

7. Fundamentals of big data analytics tools and applications

To present a brief and yet comprehensive account of big data analytics techniques, we classify these techniques into two broad categories: descriptive and prescriptive analytics tools (Sivarajah et al., 2017). For each category, we introduce several important artificial intelligence-based algorithms to clarify their applications. Table 3 summarizes the applications for each of the algorithms and provides specific real-world examples for each of them.

7.1. Descriptive tools

The first category of big data analytics tools helps managers learn about the current state of their business based on data from the past. These tools address the 'what happened' question by uncovering existing states or patterns at an aggregate level (Sivarajah et al., 2017). Descriptive analytics tools can help uncover hidden and potentially useful information related to business processes. For example, consider a large bank that needs to understand different segments of its current customers better in order to provide specialized services for them. By using tools developed based on a clustering algorithm, the bank might be able to uncover distinct categories or clusters of customers with similar characteristics that are otherwise unknown to decision makers. In addition, the patterns of events or associations between two specific variables could be discovered using descriptive data analytics techniques (Li & Musings, 2017). In this regard, an association rule discovery algorithm can help supermarkets such as Walmart effectively estimate what products are more likely to be purchased together so that they could be placed near each other. One final algorithm used in descriptive analytics is sequential patterns discovery, which is designed to identify and reveal the unknown sequence of events in the past. For instance, this algorithm can be used to identify the sequence of purchases made by customers and unveil unknown customer behaviors useful for future advertising campaigns. Table 3 provides additional real-life examples related to these techniques.

7.2. Predictive tools

The second category of big data analytics tools helps managers predict probable future states, patterns, or outcomes based on analysis of existing data. These tools address the questions such as 'what is likely to happen?' and 'what should we do next?' among others (Sivarajah et al., 2017).

Predictive models enable decision makers to predict the estimated value of a variable of interest using existing data. A variety of techniques such as data mining, statistical modeling, and machine learning are used to predict events or outcomes that are otherwise unknown to decision makers (Chui et al., 2018). For instance, a large bank can rely on a classification algorithm to analyze data to predict whether an existing customer is likely to open a new savings account. Such a predictive tool built through the analysis of behaviors (e.g., habits, transactions) and characteristics (e.g., age, gender) of customers may enable the bank to identify easy-to-convert customers and engage in targeted advertising activities to attract them. Tools based on a regression analysis algorithm form another important type of predictive model that is used for predicting the value of continuous numeric variables. For example, regression models can predict future sales based on market conditions and other relevant historical data of customers and competitors. Another class of popular predictive tools use anomaly detection algorithm. Tools developed based on this algorithm enable decision makers to identify outliers and unusual patterns in events or behaviors. As an example, anomaly detection is widely used by financial institutions for fraud detection by identifying and preventing fraudulent activity in credit card transactions. That is, transaction patterns of millions of customers are analyzed to enable the algorithm to learn the normal patterns and detect unusual activities. Specific, real-world examples for each of the algorithms are provided in Table 3.

It should be noted that descriptive and predictive tools are mainly utilized by data scientists in the process of transforming big data into meaningful insights (i.e., the first phase in the big data analytics cycle). Therefore, as mentioned before, the basic managerial knowledge of these tools enriches manager-scientist interactions regarding big data goals and contributes to successful implementation of big data strategies.

8. Concluding remarks

Recent developments in the field of big data analytics have generated a new gold rush for extracting business value from big data. While success during the California gold rush was mainly a result of luck (e.g., having access to the right parcel of land), the realization of big data dreams is much more a matter of successful implementation. In this regard, after discussing the many challenges faced by big data dreamers, we showed that the road to riches passes through effective implementation of big data strategies. This, in turn, requires managers to gain a better understanding of the basic concepts of big data analytics, to stay committed to their big data initiatives, and to promote communication of big data goals. Only then can they effectively orchestrate activities to transform big data into insights, decisions, and actions, and strike it rich with big data.

Acknowledgment

We would like to thank Associate Editor Dr. Jan Kietzmann for his excellent suggestions and guidance. We also thank Dr. Ehsan Kakharizadi, Dr. Hassan Shirvani, and Dr. Samuel Condic for their feedback on earlier versions of this article.

References

Adrian, C., Abdullah, R., Atan, R., & Jusoh, Y. Y. (2018). Conceptual model development of big data analytics implementation assessment effect on decision-making. *International Journal of Interactive Multimedia and Artificial Intelligence, 5*(1), 101–106.

Aiken, P., & Gorman, M. M. (2013). *The case for the chief data officer: Recasting the C-suite to leverage your most valuable asset*. Waltham, MA: Morgan Kaufman.

Alharthi, A., Krotov, V., & Bowman, M. (2017). Addressing barriers to big data. *Business Horizons, 60*(3), 285–292.

Anderson, C., & Li, M. (2017, June 23). Five building blocks of a data-driven culture. *Tech Crunch*. Available at http://www.techcrunch.com/2017/06/23/five-building-blocks-of-a-data-driven-culture/

Asay, M. (2017, November 10). 85% of big data projects fail, but your developers can help yours succeed. *TechRepublic*. Available at http://www.techrepublic.com/article/85-of-big-data-projects-fail-but-your-developers- can-help-yours-succeed/

Ashayeri, J. (2016, August 17). Five common mistakes in big data analytics. *TIAS*. Available at http://www.tias.edu/en/knowledgeareas/area/marketing/article/5-common-mistakes-in-big-data-analytics

Baldwin, H. (2015, January 22). When big data projects go wrong. *Forbes*. Available at http://www.forbes.com/sites/howardbaldwin/2015/01/22/when-big-data-projects-go-wrong/#671a28642736

Boulton, C. (2015, December 9). Lack of big data talent hampers corporate analytics. *CIO*. Available at http://www.cio.com/article/3013566/analytics/lack-of-big-data-talent-hampers-corporate-analytics.html

California Gold Rush. (2010, April 6). *History.com*. Available at https://www.history.com/topics/westward-expansion/gold-rush-of-1849

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

Chui, M., Kamalnath, V., & McCarthy, B. (2018). An executive's guide to AI. *McKinsey & Company*. Available at http://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/an-executives-guide-to-ai

Collins, A. (2014, October 27). Big data skills to pay the big data bills. *Technology Decisions*. Available at http://www.technologydecisions.com.au/content/it-management/article/big-data-skills-to-pay-the-big-data-bills-378997873

Columbos, L. (2017, December 11). LinkedIn's fastest-growing jobs today are in data science and machine learning. *Forbes*. Available at http://www.forbes.com/sites/louiscolumbus/2017/12/11/linkedins-fastest-growing-jobs-today-are-in-data-science-machine-learning/#7b8472e851bd

Crittenden, V. L., & Crittenden, W. F. (2008). Building a capable organization: The eight levers of strategy implementation. *Business Horizons, 51*(4), 301–309.

The data deluge. (2010, February 25). *The Economist*. Available at https://www.economist.com/leaders/2010/02/25/the-data-deluge

Davenport, T. H., & Patil, D. (2012). Data scientist. *Harvard Business Review, 90*(5), 70–76.

Dean, J. W., & Sharfman, M. P. (1996). Does decision process matter? A study of strategic decision-making effectiveness. *Academy of Management Journal, 39*(2), 368–392.

Douglas, M. (2013). Big data raises big questions. *Government Technology, 26*(4), 12–16.

Dykes, B. (2017, March 9). Why companies must close the data literacy divide. *Forbes*. Available at http://www.forbes.com/sites/brentdykes/2017/03/09/why-companies-must-close-the-data-literacy-divide/#700648f2369d

Ethiraj, S. K., Kale, P., Krishnan, M. S., & Singh, J. V. (2005). Where do capabilities come from and how do they matter? A study in the software services industry. *Strategic Management Journal, 26*(1), 25–45.

Fredrickson, J. W. (1984). The comprehensiveness of strategic decision processes: Extension, observations, future directions. *Academy of Management Journal, 27*(3), 445–466.

Gage, D. (2014, March 23). Big data uncovers some weird correlations. *The Wall Street Journal*. Available at http://www.wsj.com/articles/big-data-helps-companies-find-some-surprising-correlations-1395168255

Gartner. (2015, September 15). *Gartner says business intelligence and analytics leaders must focus on mindsets and culture to kick start advanced analytics*. Available at http://www.gartner.com/newsroom/id/3130017

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information and Management, 53*(8), 1049–1064.

Henke, N., Libarikian, A., & Wiseman, B. (2016). Straight talk about big data. *McKinsey Quarterly, 10*(1), 1–7.

Hirsch, P. B. (2016). The caliphate of numbers. *Journal of Business Strategy, 37*(6), 51–55.

Hrebiniak, L. G. (2006). Obstacles to effective strategy implementation. *Organizational Dynamics, 35*(1), 12–31.

IBM report: Most companies lack IT infrastructure strategy. (n.d.). *Analytics*. Available at http://analytics-magazine.org/ibm-report-most-companies-lack-it-infrastructure-strategy/

Kietzmann, J., Paschen, J., & Treen, E. (2018). Artificial intelligence in advertising: How marketers can leverage artificial intelligence along the consumer journey. *Journal of Advertising Research, 58*(3), 263–267.

Kottasová, I. (2018, May 11). These companies are getting killed by GDPR. *CNN*. Available at http://money.cnn.com/2018/05/11/technology/gdpr-tech-companies-losers/index.html

KPMG. (2014, February 17). *How can big data make a big difference to your performance*? Available at https://home.kpmg.com/jm/en/home/insights/2014/02/analytics-big-data-make-big-difference.html

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

Li, H., & Musings, S. (2017, April 12). Which machine learning algorithm should I use? *SAS*. Available at http://blogs.sas.com/content/subconsciousmusings/2017/04/12/machine-learning-algorithm-use

Marr, B. (2016, April 13). The wonderful big data strategy at Royal Bank of Scotland. *Forbes*. Available at http://www.forbes.com/sites/bernardmarr/2016/04/13/the-wonderful-big-data- strategy-at-royal-bank-of-scotland/#70b94652788b

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

Mayhew, H., Saleh, T., & Williams, S. (2016, October). Making data analytics work for you—Instead of the other way around. *McKinsey Quarterly*. Available at https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/making-data-analytics-work-for-you-instead-of-the- other-way-around

Mazzei, M. J., & Noble, D. (2017). Big data dreams: A framework for corporate strategy. *Business Horizons, 60*(3), 405–414.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D., & Barton, D. (2012). Big data: The management revolution. *Harvard Business Review, 90*(10), 60–68.

Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). *Big data analytics capabilities: A systematic literature review and research agenda*. Information Systems and e-Business Management.

Miller, S., & Hughes, D. (2017). *The quant crunch: How the demand for data science skills is disrupting the job market*. Boston, MA: Burning Glass Technologies.

Noble, C. H. (1999). The eclectic roots of strategy implementation research. *Journal of Business Research, 45*(2), 119–134.

Owens, K. N. (2002). *Riches for all: The California gold rush and the world*. Lincoln, NE: University of Nebraska Press.

Peng, W., & Litteljohn, D. (2001). Organisational communication and strategy implementation: A primary inquiry. *International Journal of Contemporary Hospitality Management, 13*(7), 360–363.

Press, G. (2017, January 20). 6 predictions for the $203 billion big data analytics market. *Forbes*. Available at https://www.forbes.com/sites/gilpress/2017/01/20/6-predictions-for-the-203-billion-big-data-analytics-market/#3d689e682083

Rasmussen, T., & Ulrich, D. (2015). Learning from practice: How HR analytics avoids being a management fad. *Organizational Dynamics, 44*(3), 236–242.

Rogers, P., & Meehan, P. (2007). Building a winning culture. *Business Strategy Series, 8*(4), 254–261.

Rohrbough, M. J. (1998). *Days of gold: The California gold rush and the American nation*. Berkeley, CA: University of California Press.

Ross, J. W., Beath, C. M., & Quaadgras, A. (2013). You may not need big data after all. *Harvard Business Review, 91*(12), 90–98.

Savitz, E. (2012). The big cost of big data. *Forbes*. Available at http://www.forbes.com/sites/ciocentral/2012/04/16/the-big-cost-of-big-data/#1ae45fe65a3b

Schein, E. H. (1990). Organizational culture: What it is and how to change it. In P. Evans, Y. Doz, & A. Laurent (Eds.), *Human resource management in international firms* (pp. 56–82). New York, NY: Springer.

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of big data challenges and analytical methods. *Journal of Business Research, 70*, 263–286.

Tawse, A., & Tabesh, P. (2017). Strategy implementation: A comprehensive framework. In *Academy of Management annual meeting proceedings*. Briarcliff, NY: Academy of Management.

Turner, V., Gantz, J. F., Reinsel, D., & Minton, S. (2014). The digital universe of opportunities: Rich data and the increasing value of the Internet of Things. *EMC²*. Available at https://uk.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf

Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research, 261*(2), 626–639.

Wegener, R., & Sinha, V. (2013, September 17). The value of big data: How analytics differentiates winners. *Bain & Company*. Available at http://www.bain.com/publications/articles/the-value-of-big-data.aspx

Zeng, J., & Glaister, K. W. (2017). Value creation from big data: Looking inside the black box. *Strategic Organization, 16*(2), 105–140.

Zettelmeyer, F. (2015). A leader's guide to data analytics. *Kellogg Insight*. Available at http://www.insight.kellogg.northwestern.edu/article/a-leaders-guide-to-data-analytics