RAM Revista de Administração Mackenzie

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
Special Issue, doi:10.1590/1678-6971/eRAMD210063

THE ROLE OF BIG DATA STEWARDSHIP AND ANALYTICS AS ENABLERS OF CORPORATE PERFORMANCE MANAGEMENT

MAURICIUS M. MEDEIROS[1]
https://orcid.org/0000-0001-5552-4897

ANTONIO C. G. MAÇADA[1]
https://orcid.org/0000-0002-8849-0117

NORBERTO HOPPEN[2]
https://orcid.org/0000-0003-1827-8695

To cite this paper: Medeiros, M. M., Maçada, A. C. G., & Hoppen, N. (2021). The role of big data stewardship and analytics as enablers of corporate performance management. *Revista de Administração Mackenzie*, *22*(6), 1–32. doi:10.1590/1678-6971/eRAMD210063

Submission: Mar. 15, 2021. **Acceptance:** May 5 2021.

1 Federal University of Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil.

2 University of Vale do Rio dos Sinos (Unisinos), São Leopoldo, RS, Brazil.

ABSTRACT

Purpose: Digital transformation and big data (BD) have generated a real revolution in data-driven management. While BD improves corporate performance management (CPM), this also implies increasing exposure to risks at various BD's life cycle stages. As regulatory requirements and the need for database analysis in various business areas increase, the organization must establish definitions, policies, and processes to ensure data quality in order to protect and leverage its data to obtain a competitive advantage. Therefore, understanding data stewardship (DS) and business analytics (BA) is essential for business management. The purpose of this study is to analyze the role of BD, DS, and BA as enablers of CPM.

Originality/value: We contribute to the theory by conceptualizing, validating, and discussing the DS construct and by highlighting its role together with BA in the relationship between BD and CPM. The evidence in this study indicates that, in practice, DS and BA are critical paths for organizations to obtain better control over the effects that BD can have on business performance management.

Design/methodology/approach: A survey was conducted with 312 managers who use big data analytics (BDA) in Brazilian organizations. The data were analyzed through structural equation and mediation tests.

Findings: The findings suggest that DS and BA, both alone and jointly, can transmit the BD effect to CPM. However, a better level of model adjustment is obtained when there is a serialized multi-mediation in this relationship, being DS an antecedent to BA.

KEYWORDS

Big data. Business analytics. Data governance. Data stewardship. Corporate performance management.

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

1. INTRODUCTION

The advent of digital transformation and big data (BD) implies an opportunity for the adoption of business analytics (BA) to support the analytical resources needed to extract insights that can lead to better business management decisions (Chen, Chiang, & Storey, 2012; Kitchens, Dobolyi, Li, & Abbasi, 2018; Seddon, Constantinidis, Tamm, & Dod, 2017). This can change the way companies compete through better understanding, processing, and exploiting vast amounts of data coming from different internal and external sources and processes (Ferraris, Mazzoleni, Devalle, & Couturier, 2019). In addition, BD strategies are a significant resource for competitiveness, performance, agility, and innovation, which is strongly influencing strategy formulation due to the increase of data and analytical capabilities (Côrte-Real, Ruivo, & Oliveira, 2019; Dubey, Gunasekaran, & Childe, 2019). The main clusters described in the literature on big data analytics (BDA) are related to decision-making and performance management (Rialti, Marzi, Ciappei, & Busso, 2019).

However, the outdated IT infrastructure, complexity, and chaos inherent in BD, data management, data quality, data security, lack of data science skills in organizations, privacy concerns, and organizational cultures that do not lead to data-driven operations or data-driven decision-making are the main barriers and obstacles to the effective implementation of BD strategies. It is necessary to create a clear organizational vision related to BD, but if the senior management does not value data-driven decision-making, its behavior will affect decision patterns at all levels of the organization (Alharthi, Krotov, & Bowman, 2017; Tabesh, Mousavidin, & Hasani, 2019). This denotes the relevance of data-driven strategic management. Corporate performance management (CPM) combines management practices and Information Technology (IT) to enable the planning, measurement, and evaluation of the implementation and execution of organizational strategies to leverage business performance (Richards, Yeoh, Chong, & Popovič, 2019). CPM is data-driven (Acito & Khatri, 2014), which requires integrating a wide variety of external and internal data sources.

However, without adequate data control, organizations are more exposed to risks at various stages of the BD life cycle (Coyne, Coyne, & Walker, 2018). In this context, BA faces some data-related challenges, such as extractions, fluctuations, duplications, and security flaws, whose irregularities and inconsistencies need to be identified and corrected before decisions are made based on incorrect data (Fleckenstein & Fellows, 2018).

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

Organizations need a data governance program to balance value creation and risk exposure and achieve the effective coordination necessary to succeed and sustain competitive advantage (Coyne et al., 2018; Riggins & Klamm, 2017). However, the development of governance and data stewardship mechanisms and policies is one of the most complex contemporary organizational challenges (Khatri & Brown, 2010; Nielsen, 2017; Tallon, Ramirez, & Short, 2013).

Despite the importance of the theme and the existence of several conceptual studies, empirical research on governance and data management is scarce and fragmented, as pointed out in the literature review conducted by Abraham, Schneider, and vom Brocke (2019). Moreover, the existing empirical research on Data Stewardship (DS) addresses the issue in an exploratory way through qualitative research – see Nokkala, Salmela, and Toivonen (2019) and Plomp, Dintzner, Teperek, and Dunning (2019). Terms such as data governance, data management, and data stewardship are sometimes used interchangeably, although they are separate elements. Data governance refers to organizational policies and rules for solving problems or providing services to data stakeholders (Harrison et al., 2019; Lillie & Eybers, 2018). Data management deals with organizational processes related to the data life cycle – acquisition, pre-processing and treatment, distribution, exclusion, and disposal (Côrte-Real et al., 2019; Fleckenstein & Fellows, 2018). These policies, rules, and processes are implemented by the data stewards responsible for the exercise of DS, which is the link between data policy and the implementation of actions by users (Alhassan, Sammon, & Daly, 2018; Plomp et al., 2019).

Therefore, understanding and measuring the impact of management and data analytics is essential to manage business performance. As the guardian of data governance, data stewardship plays an essential role in the implementation of BA (Harrison et al., 2019). However, there is also a scarcity of empirical research on the implications of BDA use in organizations (Mikalef, Boura, Lekakos, & Krogstie, 2019), especially in the governance, management, and stewardship of BD (Abraham et al., 2019; Nokkala et al., 2019) and the relationship between BA and CPM (Richards et al., 2019). Although the findings by Richards et al. (2019) indicate that BA affects CPM positively and significantly, there are no studies that consider their interrelation with BD stewardship. Therefore, this study seeks to answer the following research questions:

- How much do BD, DS, and BA impact CPM?
- How does BD impact CPM, directly and indirectly, through data stewardship and business analytics?

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

Thus, the purpose of this study is to analyze the role of BD, data stewardship, and business analytics as enablers of CPM.

The findings suggest that DS and BA, alone and jointly, can transmit the BD effect to CPM; however, a better level of model adjustment is obtained when there is a serialized multiple mediation in this relationship, and DS is an antecedent of BA. Such findings are especially relevant because the use of BDA requires data to be collected and analyzed in a centralized manner, ensuring the application of standards, protocols, methods, and tools (Grover, Chiang, Liang, & Zhang, 2018). This reinforces the importance of DS, whose role is to ensure the effective application of definitions, policies, and procedures for the use of data throughout the organization (Koltay, 2016). On the one hand, there is a growing need for the democratization of BD analysis, as the training and use of data analysis tools (for example, PowerBI, Tableau, QlikView, SAS, and others) has created self-service BA activities (self-service BA) due to the new skills of business specialists with data handling skills, which can enable greater flexibility for managers in corporate management (Riggins & Klamm, 2017). However, this requires a common understanding of this data throughout the organization (Abraham et al., 2019). On the other hand, when dealing with personal data in business, it is necessary to ensure compliance with stricter regulatory requirements (Alharthi et al., 2017; Khatri & Brown, 2010; Plotkin, 2020; Thompson, Ravindran, & Nicosia, 2015), such as the General Data Protection Regulation (GPDR) in the European Union, or the California Consumer Privacy Act (CCPA) in the North-American state of California, or the General Data Protection Act – *Lei Geral de Proteção de Dados* (LGPD) in Brazil.

Next, this study starts with the theoretical foundation for defining constructs and developing the research model. Subsequently, the method, analysis and discussion of the results, and final considerations are presented.

2. THEORETICAL FOUNDATIONS

2.1 Corporate performance management enabled by big data analytics

Business strategies are increasingly data-dependent, and BDA is redefining innovation, competition, and productivity (Côrte-Real et al., 2019). BDA supports the set of technologies, methods, and advanced applications for data storage, management, and analysis to improve decision-making

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

(Chen et al., 2012). However, BD and BA are distinct elements. Figure 2.1.1, below, presents these definitions.

Figure 2.1.1

BIG DATA AND BUSINESS ANALYTICS

Big data	• BD can be defined as the use of a large amount of data, from different sources and types, created quickly, which implies greater challenges to collect, manage and process them through traditional systems and resources to support decision making (Ghasemaghaei & Calic, 2020). • The essential characteristics of BD are volume, variety, and velocity, as well as variability (Chen, Mao, & Liu, 2014). • The term big data is used to describe the massive volume of digital data produced by human activity that is very difficult to manage using conventional data analysis tools, being characterized by the 3 Vs: volume, variety, and velocity (Alharthi et al., 2017).
Business analytics	• BA includes the use of models, formulas and algorithms to configure the set of rules or instructions elaborated to solve business problems, and is often subdivided into four dimensions: descriptive, diagnostic, predictive and prescriptive (Appelbaum, Kogan, Vasarhelyi, & Yan, 2017; Duan, Cao, & Edwards, 2020; Fleckenstein & Fellows, 2018). • Contributes to BD analysis, improving the understanding of performance patterns, the preparation of analysis of trends, in order to enable the projection of forecasts and potential risks and future results, in addition to the identification and adoption of the best strategies, in order to optimize the objectives, maximizing opportunities and potentialities, or minimizing risks and weaknesses (Acito & Khatri, 2014; Appelbaum et al., 2017; Duan et al., 2020).

Source: Elaborated by the authors.

BD is more technology-related, while BA supports decision-making at various levels within the organization. Therefore, organizations must first collect and store data and then apply BA processes. Thus, a competitive advantage can be created by enabling decisions to be taken more quickly or accurately (Dubey et al., 2019; Fernando & Engel, 2018). With strategic data management, the organization can protect and leverage its data (DalleMule & Davenport, 2017), making it more competitive to achieve goals and obtain value (Grover et al., 2018). Ferraris et al. (2019) found that companies that developed more BDA capabilities than others, both technological and managerial, increased their performances.

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

BA's rise is somewhat related to the notion of using performance measures as a way of testing business hypotheses (Acito & Khatri, 2014). Supported by scorecards and dashboards, CPM helps to analyze and visualize various performance metrics (Chen et al., 2012). CPM includes the activities, processes, methodologies, metrics, and technologies used by organizations to establish, implement, measure, monitor, and manage the performance of business strategies (Weeserik & Spruit, 2018). CPM seeks to understand how business performance can be measured and the determinants of business performance, combining management and IT practices, to enable organizational performance (Richards et al., 2019). However, CPM depends on integrating external and internal data from a wide variety of sources (BD) to, for example, outline a course for the organization, define performance indicators, collect and analyze performance data, and take corrective actions (Richards et al., 2019).

CPM can be observed through the development and learning of organizational skills, the optimization of the efficiency and effectiveness of internal processes, the improvement in the relationship management with customers, suppliers, and stakeholders, increasing the economic and financial added value for the business, which is essential for agility, innovation and competitive performance in contemporary business environments (Côrte-Real et al., 2019; Dubey et al., 2019; Mikalef & Pateli, 2017; Wamba & Akter, 2019; Mikalef, Krogstie, Pappas, & Pavlou, 2020). Therefore, CPM enabled by BDA facilitates the formulation of the strategy and the control and evaluation of the business performance.

2.2 Data stewardship: the link between data governance and data management

The terms data governance, data management, and data stewardship are generally used as synonyms, but although interdependent, they are different concepts (Koltay, 2016). Initially, it is important to clarify the difference between them. Figure 2.2.1 presents the concepts.

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

Figure 2.2.1

DATA GOVERNANCE, DATA STEWARDSHIP, AND DATA MANAGEMENT

Data governance	• It refers to the structuring of data management policies to ensure the quality, standardization and security of data, including the analysis of how information is distributed and administered; and, also, the definition of organizational structures and the agents responsible for managing the entire data life cycle (Khatri & Brown, 2010; Nielsen, 2017; Harrison et al., 2019; Abraham et al., 2019). • It specifies a multifunctional data management framework as a strategic asset of the organization, and decision-making rights and responsibilities, formalizing policies, standards and procedures, and monitoring data compliance (Abraham et al., 2019).
Data stewardship	• It denotes the responsibility of administrators to the data, as it concerns how administrators care for their data, in order to ensure that data-related work is carried out in accordance with policies and practices established by the governance and that the data is accurate, accessible, usable and current (Koltay, 2016; Brous, Janssen, & Vilminko-Heikkinen, 2016). • It is a collection of data management methods, which cover acquisition, storage, integration and procedures for maintenance, distribution and use, whose objective is to ensure accuracy, validity, quality, safety, management and data retention (Nokkala et al., 2019; Rosenbaum, 2010).
Data management	• It is intended to ensure that the most important information for the organization is well defined. It concerns how data is collected, stored, transformed, distributed and consumed throughout the organization throughout its life cycle, and this includes the rules that standardize structured formats, such as databases and archiving systems, data integration systems, and management processes that consume them (Fleckenstein & Fellows, 2018). • It consists of the processes necessary to manage their entire life cycle - acquisition, pre-processing and treatment, distribution, exclusion, and disposal (Côrte-Real et al., 2019). • A set of processes that includes data governance, data and document storage and architecture management, data quality and security, and data development through Business Intelligence/Analytics (Surbakti, Wang, Indulska, & Sadiq, 2020).

Source: Elaborated by the authors.

Based on the revised literature, data governance refers to the decisions that are made and who makes them, to ensure effective management and effective use of resources; in addition, data management involves implementing decisions as part of the daily implementation of data governance policies (Khatri & Brown, 2010; Alhassan et al., 2018). Although they are

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

distinct concepts, both are interdependent. Furthermore, data stewardship is defined as the function or set of activities in which data-related work is carried out, in accordance with the policies and practices established by governance, to ensure that the organization's information is reliable, healthy, of good quality, and preserved (Alhassan et al., 2018; Brous et al., 2016; Khatri & Brown, 2010; Nokkala et al., 2019; Plomp et al., 2019; Rosenbaum, 2010). It is considered that DS is responsible for operationalizing the data policy and for their implementation of user actions, being, therefore, the link between governance and data management.

As presented in Figure 2.2.1, to define the construct DS, other concepts already consolidated in the literature were reviewed: "data governance" and "data management". This helped to differentiate the elements and delimit the specific characteristics of this new construct. Thus, "data governance" and "data management" are not constructs addressed in this study, but serve as a conceptual background to: 1. help define the context and problem under study; 2. enable conceptual differentiation to define the role of the DS construct; and 3. help, as a context, in discussing the implications of the study. In fact, data stewardship is intertwined with governance and data management (Nokkala et al., 2019) because it precisely contemplates the practical purpose of both, as it refers to the effective operationalization and application of data management actions, according to the principles established by data governance.

For Plotkin (2020), it is up to DS to improve data quality through the following initiatives: standardizing data at the corporate level; mediate the resolution of data-related problems (e.g., divergence in data quality rules and requirements); ensure alignment and communication of data governance objectives throughout the organization, communicating the rules to data users; collaborate with stakeholders in managing definitions, policies, procedures, and data-related issues; and promote its use as an asset to gain competitive advantage. In this context, data stewards are administrators and business operators responsible for translating data governance policy into the implementation of practices according to the recommendations and requirements established by it (Plomp et al., 2019); they are generally organized groups (committees, boards, user groups) that represent business stakeholders and are responsible for making decisions about the treatment of assets that contribute to the communication between users and storage, management and sharing of data. Thus, they act as the channel between IT and the business, whose challenge is to ensure that one of the most critical assets of the corporation – its data – is used to its maximum capability.

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

2.3 Development of the research model

The literature indicates that a volume and variety of data generated from a high speed (BD) is a relevant resource for agile and effective decision-making (Conboy, Mikalef, Dennehy, & Krogstie, 2020; Ghasemaghaei & Calic, 2020). Furthermore, BD solutions support data integration, enabling a considerable increase in the quantity and quality of information, making the impact on performance greater (Wamba et al., 2017). For Urbinati, Bogers, Chiesa, and Frattini (2019), the role of BD is to assist decision-making, that is, to support CPM. Thus, the first hypothesis suggests that dealing with BD influences CPM.

- H1: BD impacts CPM.

Integrating several external and internal BD sources can become uncontrollable (Appelbaum et al., 2017; Coyne et al., 2018), implying more significant challenges to control the veracity, vulnerability, volatility, and validity of data (Fleckenstein & Fellows, 2018). To manage BD and traditional data repositories, data lakes are needed – structured and unstructured data repositories distributed within the organization and online. Thus, dealing with BD implies more challenges for the DS, as proposed in the second hypothesis.

- H2: BD impacts DS.

BD value is unlocked only when used to drive decision-making through analytical processes (Mikalef et al., 2019). Organizations need to first collect and store data and then apply BA processes (Fernando & Engel, 2018). Therefore, the third hypothesis holds that BD affects BA.

- H3: BD impacts BA.

Governance and data management resources are essential for data analysis (Harrison et al., 2019; Lillie & Eybers, 2018). Considering that the role of DS is the application of policies, governance practices, and data management (Alhassan et al., 2018; Plomp et al., 2019; Rosenbaum, 2010; Koltay, 2016; Thompson et al., 2015), it is argued that DS is essential for BA and CPM.

- H4: DS impacts BA.
- H5: DS impacts CPM.

BA contributes to understanding performance standards, assessing the environment, analyzing trends and projections of future results, and formu-

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

lating the best strategies (Duan et al., 2020). This allows the organization to better understand its customers and drive process optimization to improve its efficiency and effectiveness (Grover et al., 2018; Mikalef et al., 2019; Wamba et al., 2017), increasing productivity and organizational performance, supported by the sixth hypothesis.

- H6: BA impacts CPM.

Data integration and management facilitate data sharing and establishing a corporate data view (Fleckenstein & Fellows, 2018; Grover et al., 2018; Seddon et al., 2017). This makes it possible to develop analyses in a shared way throughout the organization, stimulating new and interesting uses and problem-solving capacity . Therefore, integrating BD across the organization is essential to provide good quality information for analytical purposes (Harrison et al., 2019; Lillie & Eybers, 2018). It is noteworthy that, when dealing with BD, quality is considered fundamental for decision-making (Wamba & Akter, 2019). However, this requires management structures and DS to guarantee accuracy, validity, quality, security, management, and data conservation (Rosenbaum, 2010). Therefore, DS can influence the quality and efficiency of analytical results and management information.

- H7: DS mediates BD relationships with BA (H7a) and with CPM (H7b).

To overcome the complexity and overload caused by BD, it is crucial to generate insights (Seddon et al., 2017). Furthermore, to enable managers to find only the data capable of providing valuable insights, organizations must obtain a rich combination of information through them (Kitchens et al., 2018). Therefore, the hypothesis is that BA can mediate the relationship between BD and DS with CPM.

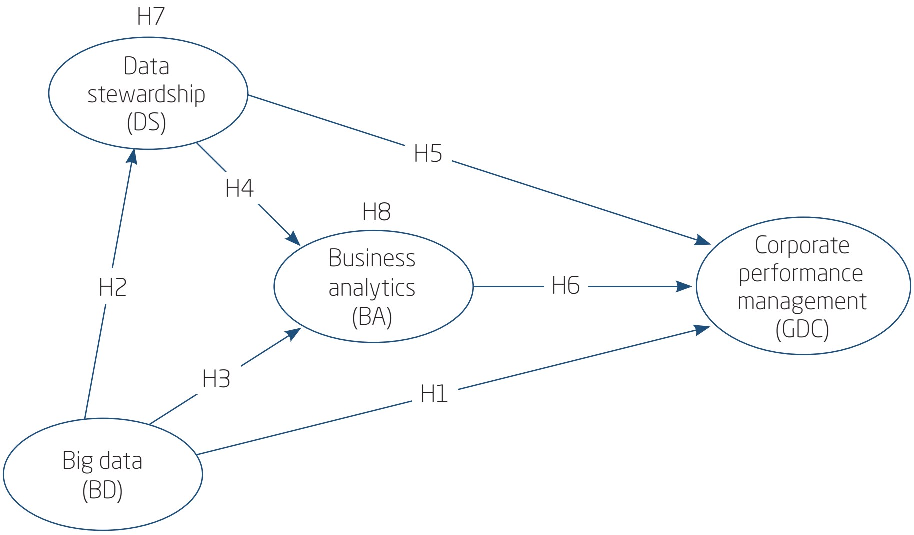
- H8: BA mediates the relationship between BD and CPM (H8a) and between DS and CPM (H8b).

Acito and Khatri (2014) indicate that extracting value from data requires alignment between resources and analytical capabilities with CPM, structuring and ordering the relationship between these elements. The structural framework for BA proposed by Acito and Khatri (2014) was considered a background for the organization of the relationships discussed in this study. It is understood that DS is at the operational and informational level, building the bridge between data (BD) and data analysis (BA). Based on the developed hypotheses, the research model was elaborated (Figure 2.3.1).

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

Figure 2.3.1

RESEARCH MODEL



Source: Elaborated by the authors.

This model fills a research gap, as it observes the DS construct, which, despite playing an important role in operationalizing governance and data management, is not found in the information systems literature; it also analyzes its relationship with the construct that deals with business data analysis (BA) processes to explain how BD affects CPM.

3. METHOD

A quantitative study was carried out for the empirical evaluation of the research model. For data collection, a survey was developed based on a structured questionnaire since the focus of the research is to measure the participants' perceptions about the use of BD and the role of DS and BA in CPM. In the sequence, the development of the instrument, the data collection procedures, the demographic data of the sample, and the statistical analyses performed are described.

3.1 Development of the research instrument

For the development of the research instrument, after conceptualizing the constructs, we used the existing literature to develop items that repre-

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

sent the definition of the constructs. In addition, the instrument's face and content validation was carried out with the support of five specialists in the area (two professionals, two masters in business administration, and a Ph.D. in information systems). During the face validation process of the instrument, adjustments were made to the structure of the questionnaire, such as the removal of items that contained some level of ambiguity in their definition, the integration of items with similar definitions, and the adequacy in the description of some items.

The authors developed the instrument by operationalizing the constructs in variables according to the references present in the literature, as indicated in Appendix A. A Likert-type scale of five points of agreement was adopted to measure the variables, from 1 (I do not agree) to 5 (I fully agree). In addition to these items, demographic data were requested, and the frequency of BDA use was asked – on a five-point scale, from "rarely" to "very often". Finally, for the qualification of the sample, two questions were established to screen the participants with answers ("yes" or "no"). The first question asked if the respondent was a BDA user in the organization where he worked. For this, a definition and examples of BDA tools were presented. The second questioned whether the respondent had professional experience.

3.2 Collection procedure and sample treatment

The minimum sample size was estimated using the G*Power V3.1.9.4 software (Faul, Erdfelder, Buchner, & Lang, 2009). Then, in accordance with the recommendations of Cohen (1988) and Hair, Hult, Ringle, and Sarstedt (2017), using the values of the test power as 0.80, effect size (f^2) median (equal to 0.15), and considering that the latent variable CPM has three predictors, it was calculated that the minimum sample is 77 respondents. Therefore, the minimum sample was overcome since the sample is of 312 cases.

The desired population was managers of Brazilian organizations from various economic sectors who work with BD. Data were collected through a survey instrument, through an electronic form, answered by 366 people. For the qualification of the sample, two questions for screening the participants were established. Thus, 312 qualified observations were collected from experienced professionals and managers who confirmed the use of BDA in the organization in which they worked. More than 70% of the sample has a graduate degree. Approximately 75% of the participants are trained in business management, IT, and statistics, which is compatible with data science users. Most participants (61%) work in organizations in the service sector, especially in the areas of IT (18%), business management, marketing and

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

market intelligence (18%), and financial/banking services (11%). Each informant who answered the survey corresponds to an organization, so the sample represents 173 Brazilian organizations. There is a concentration of organizations from the South and Southeast regions (86 percent of the total), which are the most economically active regions in the country. Figure 3.2.1 shows the profile of respondents and organizations.

Figure 3.2.1

PROFILE OF RESPONDENTS

Total informants/organizations (n = 312)					
Age (years)	(%)	Professional experience (years)	(%)	Frequency of BDA use	(%)
≤ 25	8	≤ 2	3	1 – Rarely	1
26 ≤ x ≤ 35	44	2 ≤ x ≤5	11	2	6
36 ≤ x ≤ 45	32	6 ≤ x ≤ 10	31	3	19
46 ≤ x ≤ 55	12	11 ≤ x ≤ 15	25	4	26
x > 55	4	x >15	30	5 – Very often	48
Company size	(%)	Economic sector	(%)	Geographic region	(%)
Small	24	Services	61	North	1
Medium	31	Industry	14	Northeast	6
Large	45	Trade	6	Midwest	8
		Government	12	Southeast	39
		Others	7	South	46

Source: Elaborated by the authors

As this is primary data, it was necessary to ensure that no systematic bias influenced the information collected. Thus, Harman's single factor test (1976) was performed, whose result was 39.99%, lower than the minimum of 50%. Furthermore, the AFC test in the SPSS software, with varimax rotation and autovalue equal to 1.0, indicated the existence of the five components predicted in the research model for a total variation, explained by 66.47%, confirming all the dimensions predicted in the model. In addition, the analysis of the bias of "non-respondents" was performed, according to Armstrong and Overton (1977). When carrying out these tests, it was found that both the standard method bias and the "non-respondents" bias are not a significant problem.

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

4. DATA ANALYSIS

For data analysis, the IBM SPSS Statistics V20 and SmartPLS V3 tools were adopted. The measurement model was evaluated, and structural equations were modeled by partial least squares structural equation modeling (PLS-SEM). The PLS-SEM method was adopted because it allows working with complex models, being preferred for theoretical development and explanation of the construct variance (Hair et al., 2017; Hair, Risher, Sarstedt, & Ringle, 2019), recommended for management research, and widely used in information systems studies (Mikalef & Pateli, 2017).

4.1 Measurement model

The model deals with reflective constructs, supported by the SmartPLS software, internal consistency, composite reliability, convergent validity, and discriminant validity (Hair et al., 2019). The results are shown in Figure 4.1.1.

Figure 4.1.1

MEASUREMENT MODEL ASSESSMENT

Variable[1]	Indicators[2]			HTMT Criterion			
	α	CR	AVE	BA	BD	CPM	DS
BA	.875	.914	.728				
BD	.811	.876	.638	.518			
CPM	.910	.930	.690	.578	.332		
DS	.863	.897	.593	.639	.540	.561	

[1] Big data (BD); business analytics (BA); data stewardship (DS); corporate performance management (CPM).
[2] Cronbach's Alpha (α); composite reliability (CR); average variance extracted (AVE).

Source: Elaborated by the authors.

For all constructs, Cronbach's Alpha and composite reliability (CR) indicators are more significant than 0.7 (Hair et al., 2017), suggesting, respectively, the internal consistency and reliability of the constructs. The convergent validity, calculated using the average variance extracted (AVE) of each factor, indicates how much a given composition of observable variables represents a single latent variable. The AVE indicator of each of them was higher than the recommended – of 0.5 (Hair et al., 2017). Additionally, the

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

external loads of each item in its respective construct (see Appendix A) are greater than 0.7 (Hair et al., 2017). Therefore, it is concluded that the constructs have convergent validity.

Discriminant validity indicates how much one construct differs from the others. The most modern criterion is Henseler's Heterotrait-Monotrait Ratio (HTMT), by Henseler, Ringle, and Sarstedt (2015), which according to Mikalef et al. (2020) is better than Fornell-Larcker's, through which the average correlation of indicators between constructs measuring different aspects of the model is compared in relation to the average of correlations of indicators within the same construct, whose results should reach the limit of 0.85 (Henseler et al., 2015; Hair et al., 2019). Therefore, it is observed that there is convergent validity.

4.2 Structural model

The evaluation of the structural model is performed using the path coefficients, the level of significance of the relationships, the effect size (f^2), the Pearson's correlation coefficients (R^2) and predictive validity (Q^2) and the standardized root mean square residual (SRMR), as recommended by Hair et al. (2019). Preliminarily, the collinearity between the constructs was analyzed through the variance inflation factor (VIF), noting that collinearity is not a problem because all values are lower than 3 (Hair et al., 2019). Next, the bootstrapping procedure (5,000 samples) was used to test whether the hypotheses are significant, obtaining $p < 0.05$ (Hair et al., 2017). Except for H1, all other hypotheses were supported with a significance of less than 0.1%.

Next, we started to evaluate the variance portion of endogenous variables, which is explained by the structural model by Pearson's correlation – coefficient of determination (R^2). The BA and CPM constructs have large effects, as they have $R^2 > 26\%$, and the DS construct has a medium effect, as it has $R^2 > 13\%$ (Cohen, 1988). Figures 4.2.1 and 4.2.2 contain the results of the structural model.

Figure 4.2.1

SIGNIFICANCE OF STRUCTURAL MODEL PATH COEFFICIENTS

Hypothesis	Path	Coefficient	T-statistics	p-values	Effect size (f^2)	Cohen's f^2 analysis	Empirical evidence
H1	BD -> CPM	-.006	0.009	.927 ns	.000	Not relevant	Not supported
H2	BD -> DS	.456	9.707	.000***	.263	Medium	Supported

(continue)

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

Figure 4.2.1 (conclusion)

SIGNIFICANCE OF STRUCTURAL MODEL PATH COEFFICIENTS

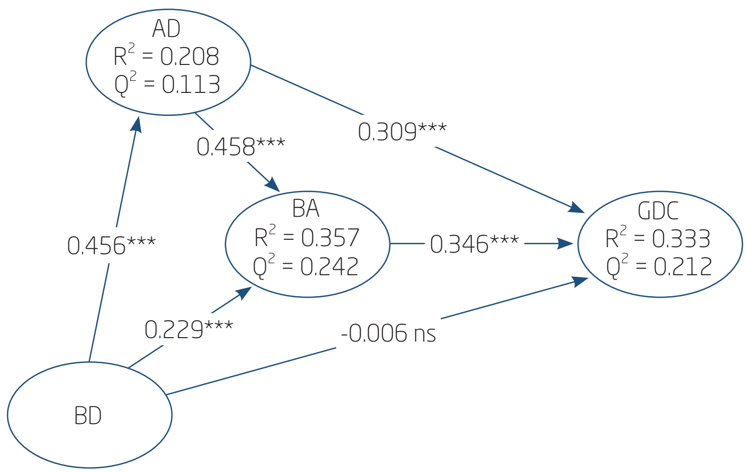
Hypothesis	Path	Coefficient	T-statistics	p-values	Effect size (f^2)	Cohen's f^2 analysis	Empirical evidence
H3	BD -> BA	.229	4.097	.000***	.064	Small	Supported
H4	DS -> BA	.458	9.062	.000***	.258	Medium	Supported
H5	DS -> CPM	.309	5.284	.000***	.090	Small	Supported
H6	BA -> CPM	.346	5.131	.000***	.115	Small	Supported

*** p < 0.001, ns – not significant.

Source: Elaborated by the authors.

Figure 4.2.2

RESEARCH MODEL WITH HYPOTHESIS TEST RESULTS



*** p < 0.001, ns – not significant.

Source: Elaborated by the authors.

To verify the portion that each exogenous variable represents in explaining the model's endogenous variables, the effect size was evaluated. It is observed that, except for the relationship between BD -> CPM, all other effects (f^2) of the Cohen's coefficient are above 0.02, which shows adequate results for latent factors (Henseler, Ringle, & Sinkovics, 2009). According to Cohen (1988) and Hair et al. (2017), $f^2 > 0.02$ represents a small-sized effect, while $f^2 > 0.15$, a medium-sized effect, and $f^2 > 0.35$, a large-sized effect. As shown in Figure 4.2.1, medium-size effects on the relationship between BD -> DS

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

and DS -> BA are evidenced, and also small-size effects in the relationship between BD -> BA, DS -> CPM, and BA -> CPM. Predictive relevance, on the other hand, is measured by the Stone-Geisser indicator (Q^2). It is observed that all endogenous variables have $Q^2 > 0$, which suggests significant predictive relevance (Hair et al., 2019).

The only recommended model adjustment criterion for partial least squares structural equation modeling (PLS) is the SRMR (Hu & Bentler, 1999). It is noteworthy that the SRMR index (0.059) meets the parameters recommended in the literature for its validation, a value lower than 0.08 (Hair et al., 2019; Hu & Bentler, 1999).

4.3 Analysis of mediations

For mediation, it is necessary that the independent variable significantly affects the dependent variable by removing the mediating variable (Zhao, Lynch, & Chen, 2010). This assumption is met because all relationships have significant direct effects when analyzed without the intervention of the mediating variables. Figure 4.3.1 presents the results of the mediation analysis.

Figure 4.3.1

MEDIATION TEST

Model	Hypothesis	SRMR	Direct effect	Indirect effect	Total effect	Mediation type
Simple mediations	H7a: BD -> DS -> CPM	.066	.073 ns	.215***	.288***	Complete
	H7b: BD -> DS -> BA	.072	.229***	.209***	.438***	Partial
	H8a: BD -> BA -> CPM	.054	.074 ns	.213***	.287***	Complete
	H8b: DS -> BA -> CPM	.062	.307***	.194***	.501***	Partial
Multiple type A	H7a, H8a: BD -> |DS+BA| -> CPM	.106	-.007 ns	.294***	.287***	Complete
Multiple type B	H7b, H8b: BD -> DS->BA -> CPM	.059	-.006 ns	.293***	.287***	Complete

*** p < 0.001, ns – not significant.

Source: Elaborated by the authors.

As for the analysis of simple mediations, the results show that both DS and BA completely mediate the relationship between BD and CPM

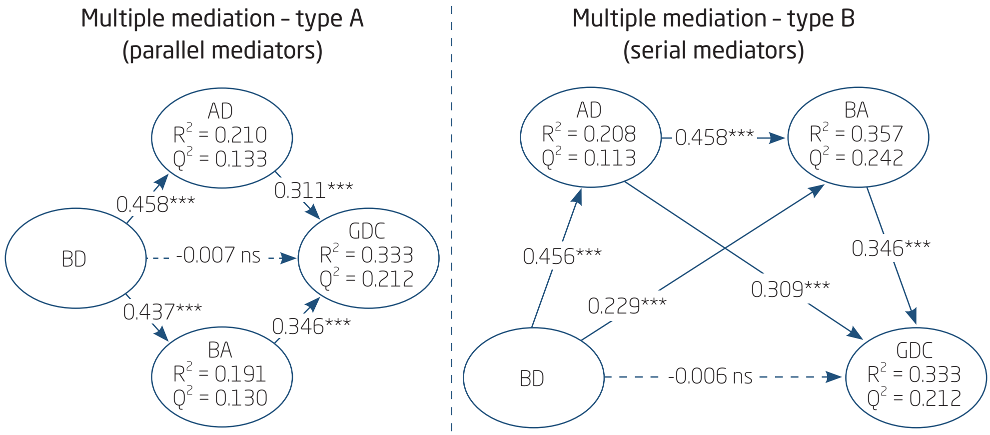
ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

(supported H7 and H8). Complete mediations mean that DS and BA constructs can fully convey the effects from BD to CPM. In addition, DS partially measures the relationship between BD and BA, and BA partially measures the relationship between DS and CPM. This means that DS can transmit a portion of the BD effect to BA and that BA transmits part of the effect from DS to CPM.

In addition to simple mediations, according to Zhang, Zyphur, and Preacher (2009), two different types of multiple mediation of mediating variables were analyzed: "type A", a multiple mediation that considers mediators with causality in parallel, and "type B", a multiple mediation that considers the serialized causality relationship between the mediating variables, starting from the premise that DS is antecedent to BA, as hypothesized in the research model. This procedure aims to assess which model of multiple mediations is the one with the most significant explanatory power. Figure 4.3.2 shows the results of the analyzed multiple mediation models.

Figure 4.3.2

MULTIPLE MEDIATION MODELS



*** $p < 0.001$, ns – not significant.

Source: Elaborated by the authors.

Regarding the analysis of multiple mediation, it is observed that the effect of BD in CPM is completely transmitted through explanatory variables, both when the mediators occur with causality in parallel – multiple mediation model "type A" – and when they occur in a serial way – multiple mediation model "type B". However, when they occur in a serial way, where mediator DS precedes mediator BA, although no differences in the

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

total effect are observed, important differences are observed in the model adjustment. First, the multiple mediation model "type B" presents higher coefficients of determination (R^2) and predictive relevance (Q^2) of the endogenous variable BA. Second, a considerable difference was observed in the model's adjustment indicator since the SRMR of the estimated model "type B" is lower (0.059) than that of "type A" (0.106). Therefore, it is verified that, in fact, the multiple mediation effect of these variables behaves in a serialized way (type B) because, for a good level of adjustment, the SRMR indicator must be less than 0.08 (Hair et al., 2017; Hu & Bentler, 1999).

5. DISCUSSION

The world and the data within it are in a continuous cycle of change. Therefore, organizations that can recognize change and react quickly and intelligently to it will have a greater competitive advantage; hence, the incorporation of the BDA strategy has decisive implications for driving and directing business strategies (Fleckenstein & Fellows, 2018). The ability to manage, analyze and act constitutes a "data-driven" decision system, characterized as a significant resource for competitiveness, performance, and innovation (Tabesh et al., 2019). The strategic use of BDA implies the transformation of organizations' management processes and decision-making culture (Frisk & Bannister, 2017).

Previous studies have indicated that the greater the volume, velocity, variety, and variability of data (BD), the greater the risks and opportunities for CPM (Appelbaum et al., 2017; Conboy et al., 2020; Coyne et al., 2018; Ghasemaghaei & Calic, 2020; Riggins & Klamm, 2017; Urbinati et al., 2019). BDA guarantees that data may be transformed into business information and knowledge, useful for efficient decision-making processes, thereby improving performance (Ferraris et al., 2019). This study confirms this idea by identifying that the relationship between BD and CPM has a significant effect when the relationship is analyzed in isolation, that is, without the intervention of other latent variables. In addition, the findings of this study expand the knowledge about this phenomenon by describing how other constructs can explain the transmission from BD to CPM.

In the presence of the mediating variables DS and BA, the hypothesis that BD influences CPM (H1) was not sustained, resulting in discovering the complete mediation of these intervening variables. In addition, it was found that BD influences both DS and BA (H2 and H3, supported), and this

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

indicates that, when dealing with BD, the role of DS and BA is even more relevant, since a greater volume, variety, and variability of data, generated with greater speed, requires more sophisticated governance, management and analytical treatments (Abraham et al., 2019; Fleckenstein & Fellows, 2018). This means that, when dealing with data from different business domains, which is generated, handled, analyzed, and interpreted by different organizational agents, it is essential that organizational data is managed and analyzed effectively (Alhassan et al., 2018; Coyne et al., 2018; Kitchens et al., 2018).

It was found that DS plays a central role in the model, as it significantly influences BA (supported H4), as documented by Harrison et al. (2019), and CPM (supported H5), confirming the hypothesized relationships in this study. In addition, BA directly influences CPM (supported H6), which is consistent with the findings by Richards et al. (2019). Thus, a complete mediation effect of DS and BA was found in the relationship between BD and CPM (supported H7 and H8), which suggests that both mediators, alone and in combination, can transmit the BD effect to CPM. There was also a partial mediation of BA in the relationship between DS and CPM (H9 and H10). However, in the presence of the other intervening variables, although the DS -> CPM path is significant ($p < -0.000$), it has a small size effect on the analysis of the f^2. Given these findings, another type of multiple mediation of DS and BA in the model was investigated, considering DS as an antecedent of BA. This new configuration of the multiple mediation model proved to be more efficient in transmitting the effect of BD to CPM. This confirms that BA plays a crucial role in managing corporate performance (Richards et al., 2019; Kitchens et al., 2018), as data analysis can contribute to the generation of insights for the identification of patterns, trend projection, scenario study, and process optimization (Duan et al., 2020; Seddon et al., 2017).

The findings show that DS is an effective approach to the execution and supervision of plans, policies, programs, and practices that control, protect, deliver and enhance the value of data assets (Koltay, 2016; Thompson et al., 2015). DS can guarantee the operationalization and application of the data governance policies and mechanisms that are important both to protect them from irregularities, inconsistencies, fraud, or security breaches, which guarantees their veracity and prevents vulnerability, as well as leverage them by providing the integration and control of the entire flow in the organization, serving to deal with the variability and volatility of the data, in order to guarantee the supply of good quality information for BA use, which ensures

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

one of the most critical assets of the corporation – its data – to be used to its maximum capability (DalleMule & Davenport, 2017; Fleckenstein & Fellows, 2018; Lillie & Eybers, 2018; Rosenbaum, 2010).

6. FINAL CONSIDERATIONS

This study analyzed the multiple mediation of DS and BA in the relationship between BD and CPM. This model fills a research gap, as it observes the DS construct, which despite playing an important role in operationalizing governance and data management, is not found in the information systems literature; and analyzes its relationship with the construct that deals with business data analysis (BA) processes to explain how BD affects CPM. The contributions, implications, limitations, and proposals for future studies are described below.

6.1 Theoretical contributions and managerial implications

Unlike other studies on the subject, the research design is based on a quantitative method. A survey was conducted with 312 managers who are users of BDA in Brazilian organizations from different economic segments and geographic regions. The data were analyzed through structural equation modeling and mediation tests. This made it possible to deliver valuable empirical evidence for management theory and practice.

First, it contributes to the theory by conceptualizing, validating, and discussing the role of the DS construct in corporate management. Another relevant contribution is related to the measurement of how much BD, DS, and BA directly impact CPM. It was also discovered that DS plays a fundamental role, as it affects CPM directly and indirectly through BA; it also mediates the relationship between BD and CPM. Based on these findings, progress was made in the analysis of multiple mediations to investigate whether it could be mediated in parallel (type a), both starting from the same antecedent (BD), or in a serial way (type b), having data stewardship as a precedent for BA. Thus, it was found that “type B” multi-mediation shows a better level of adjustment of the model, showing how the BD effect is transmitted to CPM through the multi-mediation of DS and BA.

The evidence in this study indicates that investing in the development of DS and BA capabilities is an assertive way to better control the positive and negative effects that BD can have on CPM (Dubey et al., 2019; Harrison et al., 2019; Richards et al., 2019; Nokkala et al., 2019; Plomp et al., 2019;

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

Rosenbaum, 2010). Moreover, it is important to emphasize that DS is beneficial even for organizations that do not deal with BD, as it is able to significantly influence CPM, both directly and indirectly, through BA. Therefore, the implications of this study are considerably broad, although some organizations do not deal with BD and others are still building their BA capability; in fact, every organization has DS in some way, because even in the absence of a governance program or standardized management processes, there will always be someone in the organization to whom people turn to with questions about the meaning of the data.

Using BD and BA can support organizations to change their decision-making culture, resulting in better and more effective decisions (Frisk & Bannister, 2017). Many organizations have adopted BA self-service to facilitate the expansion of data use in the organization (Riggins & Klamm, 2017). However, the excess of flexibility and accessibility to data can imply disorganization in the use of data and analytical tools, which can compromise the quality of the data and generate misinterpretations and wrong decisions (Appelbaum et al., 2017).

Data stewardship as an instrument of data governance (Koltay, 2016; Rosenbaum, 2010) needs to assess and ensure the balance between the level of accessibility to data in a distributed manner in the business areas and the availability of data science skills, as well as it needs to ensure compliance with policies, processes, and standards of data use, in order to avoid chaos in the use of data, guarantee its protection, and leverage the acquisition of insights and value. Data integrity, along with data protection and utilization, is the foundation of future excellence. Therefore, DS imposes basic preparedness in facing a threat against data integrity and ensures maximum leverage of data investments (Keywell, 2020).

Naturally, privacy concerns often hinder the adoption and use of BDA in organizations. The establishment of new regulatory requirements of GPDR, or CCPA, or LGPD, despite being barriers, can be used as a great opportunity for organizations. Thus, organizations need to accommodate this legislation and incorporate general best practices for handling sensitive customer data into their policies and operations (Alharthi et al., 2017). For example, organizations can benefit from good data management and business analysis practices to develop an ethical and legal framework for data sharing (Zarkadakis, 2020) in order to deal with the treatment of the high volume and variety of data generated at high velocity (BD) on the web and in business operations, available in organizational data lakes, open data or other sources, and also in order to establish strategic management to protect and

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

leverage data, both organization data and customer and stakeholder data (DalleMule & Davenport, 2017; Thompson et al., 2015).

6.2 Limitations and future studies

In the study, especially two limitations must be recognized. First, the confirmation of the partial mediation of BA in the relationship between DS and CPM suggests that there may be other intervening variables not considered in this model. Second, it is worth mentioning the geographical limitation, as the study was applied only in the Brazilian context. Although the mentioned limitations do not compromise the results obtained, the generalization of the results should be evaluated according to specific characteristics of the context and the sample analyzed.

Despite the relevance of DS for the success of data governance, there is a scarcity of empirical studies on it (Abraham et al., 2019; Nokkala et al., 2019). Therefore, future studies can deepen the investigation of the phenomenon through qualitative approaches and comparative case studies to understand the implications of management models and DS practices. It is also suggested that future studies seek to evaluate the influence of other intervening organizational factors that are promising in the context of BD and data science, such as data strategy, data-driven culture, data literacy, data quality, among others analytical capabilities.

O PAPEL DA ADMINISTRAÇÃO E ANÁLISE DE *BIG DATA* COMO HABILITADORAS DA GESTÃO DO DESEMPENHO CORPORATIVO

RESUMO

Objetivo: A transformação digital e o *big data* (BD) geraram uma verdadeira revolução no gerenciamento orientado a dados. Embora o BD melhore a gestão do desempenho corporativo (GDC), isso também implica aumentar a exposição a riscos em vários estágios do ciclo de vida do BD. À medida que aumentam os requisitos regulatórios e a necessidade de análise do banco de dados em diversas áreas de negócios, é necessário que a organização estabeleça definições, políticas e processos para garantir a qualidade dos dados, a fim de proteger e potencializar

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

seus dados para obter vantagem competitiva. Portanto, compreender a administração de dados (AD) e a *business analytics* (BA) é essencial para o gerenciamento dos negócios. O objetivo deste estudo é analisar o papel do BD, da AD e da BA como habilitadores da GDC.

Originalidade/valor: Contribuímos para a teoria ao conceituarmos, validarmos e discutirmos o construto AD e ao destacarmos o papel dela com a BA na relação entre BD e GDC. As evidências deste estudo indicam que na prática a AD e a BA são um caminho crítico para as organizações obterem um melhor controle dos efeitos que o BD pode ter na GDC.

Design/metodologia/abordagem: Realizou-se uma *survey* com 312 gestores que utilizam *big data analytics* (BDA) em organizações brasileiras. Os dados foram analisados por meio de equações estruturais e testes de mediação.

Resultados: Os resultados sugerem que a administração e a analítica de dados de negócios, tanto isoladamente quanto em conjunto, podem transmitir o efeito BD para a GDC. No entanto, um melhor nível de ajuste do modelo é obtido quando há uma multimediação serializada nesse relacionamento, sendo a AD um antecedente para a BA.

PALAVRAS-CHAVE

Big data. *Business analytics*. Administração de dados. Governança de dados. Gestão do desempenho corporativo.

REFERENCES

Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, *49*, 424–438.

Acito, F., & Khatri, V. (2014). Business analytics: Why now and what next? *Business Horizons*, 57(5), 565–570.

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

Alhassan, I., Sammon, D., & Daly, M. (2018). Data governance activities: A comparison between scientific and practice-oriented literature. *Journal of Enterprise Information Management*, *31*(2), 300–316.

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

Appelbaum, D., Kogan, A., Vasarhelyi, M., & Yan, Z. (2017). Impact of business analytics and enterprise systems on managerial accounting. *International Journal of Accounting Information Systems*, *25*, 29–44.

Armstrong, J. S., & Overton, T. S. (1977). Estimating nonresponse bias in mail surveys. *Journal of Marketing Research*, *14*(3), 396–402.

Brous, P., Janssen, M., & Vilminko-Heikkinen, R. (2016). Coordinating decision-making in data management activities: A systematic review of data governance principles. In H. Scholl et al. (eds.), Electronic Government. EGOV 2016. *Lecture Notes in Computer Science* (pp. 115–125). Cham: Springer.

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

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, *19*(2), 171–209.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: L. Erlbaum Associates.

Conboy, K., Mikalef, P., Dennehy, D., & Krogstie, J. (2020). Using business analytics to enhance dynamic capabilities in operations research: A case analysis and research agenda. *European Journal of Operational Research*, *281*(3), 656–672.

Coyne, E. M., Coyne, J. G., & Walker, K. B. (2018). Big data information governance by accountants. *International Journal of Accounting & Information Management*, *26*(1), 153–170.

Côrte-Real, N., Ruivo, P., & Oliveira, T. (2019). Leveraging internet of things and big data analytics initiatives in European and American firms: Is data quality a way to extract business value? *Information & Management*, *57*(1), 103–141.

DalleMule, L., & Davenport, T. H. (2017). What's your data strategy. *Harvard Business Review*, *95*(3), 112–121.

Delen, D., & Zolbanin, H. M. (2018). The analytics paradigm in business research. *Journal of Business Research*, *90*, 186–195.

Duan, Y., Cao, G., & Edwards, J. S. (2020). Understanding the impact of business analytics on innovation. *European Journal of Operational Research*, *281*(3), 673–686.

Dubey, R., Gunasekaran, A., & Childe, S. J. (2019). Big data analytics capability in supply chain agility: The moderating effect of organizational flexibility. *Management Decision*, *57*(8), 2092–2112.

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Method*, *41*(4), 1149–1160.

Fernando, F., & Engel, T. (2018). Big data and business analytic concepts: A literature review. *Americas Conference on Information Systems*, New Orleans, USA, 24.

Ferraris, A., Mazzoleni, A., Devalle, A., & Couturier, J. (2019). Big data analytics capabilities and knowledge management: Impact on firm performance. *Management Decision*, *57*(8), 1923–1936.

Fleckenstein, M., & Fellows, L. (2018). Implementing a data strategy. In M. Fleckenstein & L. Fellows, *Modern data strategy* (pp. 35–54). Cham: Springer.

Frisk, J. E., & Bannister, F. (2017). Improving the use of analytics and big data by changing the decision-making culture: A design approach. *Management Decision*, *55*(10), 2074–2088.

Ghasemaghaei, M., & Calic, G. (2020). Assessing the impact of big data on firm innovation performance: Big data is not always better data. *Journal of Business Research*, *108*, 147–162.

Grover, V., Chiang, R. H., Liang, T. P., & Zhang, D. (2018). Creating strategic business value from big data analytics: A research framework. *Journal of Management Information Systems*, *35*(2), 388–423.

Hair, J. F., Hult, G., Ringle, C., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). Los Angeles: Sage.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24.

Harman, H. H. (1976). *Modern factor analysis*. Chicago: University of Chicago Press.

Harrison, T. M., Luna-Reyes, L. F., Pardo, T., De Paula, N., Najafabadi, M., & Palmer, J. (2019, June). The data firehose and AI in government: Why data management is a key to value and ethics. *Annual International Conference on Digital Government Research, Dubai, United Arab Emirates,* (pp. 171–176). ACM, New York, NY, USA, 20.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135.

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. In R. R. Sinkovics & P. N. Ghauri (Eds.), *New challenges to international marketing* (pp. 277–319). Bingley, UK: Emerald Group Publishing.

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55.

Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, *53*(1), 148–152.

Keywell, B. (2020, October 21). Your board needs a data-integrity committee. *Harvard Business Review*. Retrieved from https://hbr.org/2020/10/your-board-needs-a-data-integrity-committee

Kitchens, B., Dobolyi, D., Li, J., & Abbasi, A. (2018). Advanced customer analytics: Strategic value through integration of relationship-oriented big data. *Journal of Management Information Systems*, *35*(2), 540–574.

Koltay, T. (2016). Data governance, data literacy and the management of data quality. *IFLA Journal*, *42*(4), 303–312.

Lillie, T., & Eybers, S. (2018, August). Identifying the constructs and agile capabilities of data governance and data management: A review of the literature. In K. Krauss, Turpin, M. Naude F. (eds.), Locally Relevant ICT Research. IDIA 2018. *Communications in Computer and Information Science* (pp. 313–326). Cham: Springer.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, *98*, 261–276.

Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. *Information & Management*, *57*(2), 103–169.

Mikalef, P., & Pateli, A. (2017). Information technology-enabled dynamic capabilities and their indirect effect on competitive performance: Findings from PLS-SEM and fsQCA. *Journal of Business Research*, *70*, 1–16.

Nielsen, O. B. (2017). A comprehensive review of data governance literature. *IRIS: Selected Papers of the Information Systems Research Seminar in Scandinavia*, (8), 120–133.

Nokkala, T., Salmela, H., & Toivonen, J. (2019). Data governance in digital platforms. *Americas Conference on Information Systems*, Cancun, México, 25.

Plomp, E., Dintzner, N., Teperek, M., & Dunning, A. (2019). Cultural obstacles to research data management and sharing at TU Delft. *Insights*, *32*(1), 1–11.

Plotkin, D. (2020). *Data stewardship: An actionable guide to effective data management and data governance*. London, UK: Elsevier.

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

Rialti, R., Marzi, G., Ciappei, C., & Busso, D. (2019). Big data and dynamic capabilities: A bibliometric analysis and systematic literature review. *Management Decision*, *57*(8), 2052–2068.

Richards, G., Yeoh, W., Chong, A. Y. L., & Popovič, A. (2019). Business intelligence effectiveness and corporate performance management: An empirical analysis. *Journal of Computer Information Systems*, *59*(2), 188–196.

Riggins, F. J., & Klamm, B. K. (2017). Data governance case at KrauseMcMahon LLP in an era of self-service BI and big data. *Journal of Accounting Education*, *38*, 23–36.

Rosenbaum, S. (2010). Data governance and stewardship: Designing data stewardship entities and advancing data access. *Health Services Research*, *45*(5), 1442–1455.

Seddon, P. B., Constantinidis, D., Tamm, T., & Dod, H. (2017). How does business analytics contribute to business value? *Information Systems Journal*, *27*(3), 237–269.

Surbakti, F. P. S., Wang, W., Indulska, M., & Sadiq, S. (2020). Factors influencing effective use of big data: A research framework. *Information & Management*, *57*(1), 103–146.

Tabesh, P., Mousavidin, E. & Hasani, S. (2019). Implementing big data strategies: A managerial perspective. *Business Horizons*, *62*(3), 347–358.

Tallon, P. P., Ramirez, R. V., & Short, J. E. (2013). The information artifact in IT governance: Toward a theory of information governance. *Journal of Management Information Systems*, *30*(3), 141–178.

Thompson, N., Ravindran, R., & Nicosia, S. (2015). Government data does not mean data governance: Lessons learned from a public sector application audit. *Government Information Quarterly*, *32*(3), 316–322.

Urbinati, A., Bogers, M., Chiesa, V., & Frattini, F. (2019). Creating and capturing value from big data: A multiple-case study analysis of provider companies. *Technovation*, *84*, 21–36.

Wamba, S. F., & Akter, S. (2019). Understanding supply chain analytics capabilities and agility for data-rich environments. *International Journal of Operations & Production Management*, 39(6–7–8), 887–912.

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356–365.

Weeserik, B. P., & Spruit, M. (2018). Improving operational risk management using business performance management technologies. *Sustainability*, *10*(3), 640–659.

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

Zarkadakis, G. (2020, November 10). "Data trusts" could be the key to better AI. *Harvard Business Review*. Retrieved from https://hbr.org/2020/11/data-trusts-could-be-the-key-to-better-ai

Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using hierarchical linear models. Problems and solutions. *Organizational Research Methods*, *12*(4), 695–719.

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, *37*, 197–206.

Appendix A

OPERATIONALIZATION OF CONSTRUCTS

Construct (references)	Item	Load	Item description
Big data Alharthi et al. (2017) and (Chen et al. (2014).	BD-01	.756	In the organization's business environment.... .... handles a large volume of data.
	BD-02	.820	... data is generated (collected, created) at high speed.
	BD-03	.835	... it deals with a wide variety (types and formats) of data.
	BD-04	.782	... deals with great variability of data.
Business analytics Delen and Zolbanin (2018) and Duan et al. (2020).	BA-01	.799	Business data are .... analyzed using descriptive techniques (categorization, consolidation, and classification) to convert them into useful information for managers.
	BA-02	.847	... explored to find out the causes of a specific problem.
	BA-03	.863	... to identify behaviors and predict trends.
	BA-04	.900	... to identify the best alternatives and optimize objectives.
Data stewardship Fleckenstein and Fellows (2018), Lillie and Eybers (2018), Khatri and Brown (2010), Koltay (2016), Plotkin (2020), Rosenbaum (2010), Surbakti et al. (2020), and Tallon et al. (2013).	DS-01	.725	In the organization it is adopted... ... governance policies and data management formally established.
	DS-02	.790	... a data architecture that makes it possible to identify what data is collected and in which processes it is used.
	DS-03	.775	... an organizational structure or management committee responsible for data quality and guidance to BDA users.

(continue)

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

Construct (references)	Item	Load	Item description
Data stewardship Fleckenstein and Fellows (2018), Lillie and Eybers (2018), Khatri and Brown (2010), Koltay (2016), Plotkin (2020), Rosenbaum (2010), Surbakti et al. (2020), and Tallon et al. (2013).	DS-04	.721	... BDA in an accessible corporate environment, in a shared way among business areas.
	DS-05	.799	... practices to ensure the protection of the data used in BA in the management or operation of the business.
	DS-06	.807	... data sharing to reduce "islands" or "silos" of data in business areas.
Corporate performance management Acito and Khatri (2014), Chen et al. (2012), Mikalef et al. (2019), Richards et al. (2019), and Weeserik and Spruit (2018).			Big data analytics...
	CPM-01	.843	... assists in the formulation and planning of the business strategy.
	CPM-02	.792	... assists in the control and evaluation of business performance.
	CPM-03	.844	... contributes to add economic and financial value to the business.
	CPM-04	.842	... helps optimize the efficiency and effectiveness of internal business processes.
	CPM-05	.818	... favors the management of relationships with customers, suppliers, and stakeholders.
	CPM-06	.845	... favors the development and learning of other organizational capabilities.

AUTHOR NOTES

Mauricius M. Medeiros, master from the Graduate Program in Administration (PPGAdm), University of Vale do Rio dos Sinos (Unisinos); **Antonio C. G. Maçada**, Ph.D. from the Graduate Program in Administration (PPGA), Federal University of Rio Grande do Sul (UFRGS); **Norberto Hoppen**, Ph.D. from the Administration Department, Université Pierre Mendès France (UPMF).

Mauricius M. Medeiros is now a Ph.D. student at PPGA-UFRGS; Antonio C. G. Maçada is now a professor at the PPGA-UFRGS; Norberto Hoppen is now a professor at the PPGAdm-Unisinos.

Correspondence concerning this article should be addressed to Mauricius M. Medeiros, Rua Washington Luiz, 855, Centro Histórico, Porto Alegre, Rio Grande do Sul, Brazil, CEP 90010-460. E-mail: mauricius.medeiros@gmail.com

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063

ISSN 1678-6971 (electronic version) • RAM, São Paulo, *22*(6), eRAMD210063, 2021
doi:10.1590/1678-6971/eRAMD210063