

BIG DATA IN THE BRAZILIAN PUBLIC HEALTH SECTOR: CONCEPT, CHARACTERISTICS, BENEFITS, AND CHALLENGES

Dayse Karenine de Oliveira Carneiro¹

Gislayne da Silva Goulart^{1,2}

Rafael Barreiros Porto¹

University of Brasilia (UnB), Brasília – DF, Brazil¹

Federal University of Mato Grosso do Sul (UFMS), Nova Andradina – MS, Brazil²

Big Data (BD) has arisen as a strategic way to manage and process varied and elevated volumes of information that are published these days. In this context, the purpose of this article is to investigate the adoption of BD for the management of financial resources related to Brazilian public health services and activities, in the perception of the actors involved. Using the multimethod approach, we propose a concept of BD in the public health sector; characterize the adoption of BD in an organization within this sector with attributes such as volume, velocity, variety, value, veracity, variability, and visualization; and finally, identify categories and variables in the literature which are related to the challenges and benefits that are present and absent in this phenomenon, as well as new variables related to two managerial benefits: sustainability and leadership support. The results offer an opportunity to understand the use of BD in the Brazilian public health sector, especially for the efficient management of the resources used to create collective value in public health.

Keywords: Big Data, public health sector, e-government, Brazil



BIG DATA NO SETOR PÚBLICO DE SAÚDE BRASILEIRO: CONCEITO, CARACTERÍSTICAS, BENEFÍCIOS E DESAFIOS

Big Data (BD) surge de forma estratégica para o gerenciamento e processamento de variado e elevado volume de informações geradas atualmente. Nesse contexto, o objetivo do artigo é investigar a adoção do BD na gestão dos recursos financeiros relacionados com as ações e serviços públicos de saúde, na percepção dos atores envolvidos. Utilizando-se de abordagem multimétodo, foi proposto conceito de BD no setor público de saúde brasileiro; o BD adotado em organização do referido setor foi caracterizado por meio dos atributos volume, velocidade, variedade, valor, veracidade, variabilidade e visualização; e por fim, foram identificadas categorias e variáveis da literatura relacionadas com os desafios e benefícios presentes e ausentes no fenômeno, bem como novas variáveis relacionadas com os benefícios gerenciais: sustentabilidade e apoio da alta gestão. Os resultados revelaram a oportunidade de compreensão do BD no setor público de saúde brasileiro, em especial para uma gestão eficiente dos recursos utilizados para criação de valor coletivo na saúde pública.

Palavras-chave: *Big Data*, setor público de saúde, governo eletrônico, Brasil

BIG DATA EN EL SECTOR DE LA SALUD PÚBLICA BRASILEÑA: CONCEPTO, CARACTERÍSTICAS, BENEFICIOS Y RETOS

Big Data (BD) surge de manera estratégica para la gestión y procesamiento de variado y elevado volumen de información generada actualmente. En este contexto, el objetivo del artículo es investigar la adopción del BD, para la gestión de los recursos financieros relacionados con las acciones y servicios públicos de salud, en la percepción de los actores implicados. Utilizando un enfoque multimétodo, se propuso un concepto de BD en el sector público de salud brasileño; el BD adoptado en la organización del referido sector fue caracterizado por medio de los atributos volumen, velocidad, variedad, valor, veracidad, variabilidad y visualización; y por último, se identificaron categorías y variables de la literatura relacionadas con los desafíos y beneficios presentes y ausentes en el fenómeno, así como nuevas variables relacionadas con los beneficios gerenciales: sostenibilidad y apoyo de la alta gestión. Los resultados demostraron la oportunidad de comprensión del BD en el sector público de salud, en especial para una gestión eficiente de los recursos utilizados para la creación de valor colectivo en la salud pública brasileña.

Palabras clave: *Big Data*, sector público de salud, gobierno electrónico, Brasil

1. INTRODUCTION

Given the advances provided by innovations in Information and Communications Technologies (ICTs), various daily societal activities generate immeasurable amounts of valuable data which organizations can use to transform into economic and social value (GÜNTHER *et al.*, 2017; KULKARNI *et al.*, 2020). Considering the quantity and complexity of this data, which cannot feasibly be analyzed by traditional tools, Big Data (BD) has been used in a strategic manner, including as a tool in the electronic government policies of many countries (TWIZEYIMANA; ANDERSSON, 2019), for the management of information that integrates the most varied types of data, involving statistical and informative methods of analyzing complex realities (RUMBOLD & PIERSCIONEK, 2017; GIACALONE *et al.*, 2018).

The health sector is similar in its production of countless data regarding patients, exams, procedures, policies, and the entire managerial sector of these organizations with the objectives of improving the quality of health care and minimizing its costs. Thus, BD can be utilized to effectively analyze the large volume of data generated by this sector to respond to the new challenges (ARNOLD-SCHMITT & TRISKA, 2014; WANG *et al.*, 2018; KULKARNI *et al.*, 2020).

It should be noted that important studies have already been conducted to analyze the adoption of BD in the health area, however, these studies have been scarce in relation to the public sector, which reflects the need to analyze the adoption of BD in the management of public health resources and their potential in the creation of social well-being (GÜNTHER *et al.*, 2017; SIVARAJAH *et al.*, 2017; MEHTA & PANDIT, 2018; SHASTRI & DESHPANDE, 2020).

In addition, the understanding of the benefits and challenges that accompany the adoption of BD in the context of the public health sector has become relevant given the indispensable nature of the efficient management of resources dedicated to the maintenance and investment in the public network, which cause grave harm to the health care of the system's users when absent. Given the relevance of the present study, we have elaborated the following research question: What are the concepts, characteristics, challenges, and benefits related to the adoption of BD for the management of financial resources for Brazilian public health services and activities?

Thus, the purpose of this study is to investigate the adoption of BD in the management of financial resources related to Brazilian public health services and activities, in the perception of the actors involved. We have decided to analyze the case of the Brazilian Ministry of Health and its adoption of BD in order to achieve our proposed objective.

Therefore, based on this objective, the present work is composed of sections which deal with the theoretical assumptions of BD, the characterization of BD, and the benefits and challenges of adopting BD. We will then present the methodological procedures utilized, followed by the results and a discussion of them based on an analysis of public health sector data from a developing country. Finally, we will present our conclusions and recommendations for future studies.

2. THEORETICAL REFERENCES

The use of the term BD began in the beginning of the twenty-first century, and became more prevalent in academic literature in 2012, which reflects the increasing growth of scientific production and reveals the importance of this theme in the private and public sectors. The term BD still does not possess a broad established definition, but is generally understood according to the lens that is being used to approach it. Studies (SHENG *et al.*, 2017; SIVARAJAH *et al.*, 2017) point out that this theme is still emerging and that it requires deeper conceptual and empirical investigations, such as the realization of case studies.

Chen *et al.* (2014) consider BD to be an abstract concept, generally considered to be a set of data that cannot be collected, processed, or managed by traditional Information Technology (IT) tools – software and/or hardware – that are usually used to treat data in a reasonable amount of time. BD is usually composed of a great quantity of data generated and collected from a variety of sources in various formats, such as internal organizational systems, the internet, social media, and intelligent devices, at a high rate of speed. The implementation of BD makes it possible for managers to broaden their knowledge of their businesses and the flexibility of their strategic decisions in such a way that it increases the effectiveness and efficiency of both (MCAFEE & BRYNJOLFSSON, 2012).

According to Hashem *et al.* (2015) BD, as the name itself suggests, deals with an excessive quantity of data that is so large that it cannot be categorized in regular relational databases due to the flows of data that are being generated, captured, and analyzed quickly. More recent studies (RUMBOLD & PIERSCIONEK, 2017; SHENG *et al.*, 2017) consider BD to be data that is excessively large, either structured, semi-structured or unstructured, produced continually by different sources in real time, and which influences decision making and the course of events. BD is complex by nature and requires advanced technologies and algorithms. In other words, the traditional statistical tools of Business Intelligence (BI) are no longer efficient in treating current BD.

Within this context, Big Data Analytics (BDA) has appeared and can be considered to be part of the process of extracting insights from *BD* (SIVARAJAH *et al.*, 2017) and an intensive architecture tool which makes possible the generation, collection, storage, advanced analysis, and visualization of data, as well as value generated in terms of decision making, performance measurement, and the establishment of competitive advantages (WAMBA *et al.*, 2015; SIVARAJAH *et al.*, 2017; SAGGI & JAIN, 2018). BDA addresses the techniques employed to process, examine, identify, and expose hidden underlying patterns and significant relationships among other perceptions related to the context of the application under investigation (IQBAL *et al.*, 2018).

These days, BD and BDA are treated as integral and relevant parts of various sectors and functions within the global economy because they can improve operational efficiency, drive new financial cash flow, and offer competitive advantages over competitors (SIVARAJAH *et al.*, 2017;

SAGGI & JAIN, 2018). In the public sector, for example, the adoption of BD and BDA can help improve public transport, health, education, avoid fraud, reduce costs, and increase the well-being of the population, thus contributing to the shaping of a more efficient modern society (KIM *et al.*, 2014; MEHTA & PANDIT, 2018; WANG *et al.*, 2018). In addition, it can contribute to the social and economic development of developing countries (IQBAL *et al.*, 2018). Despite these benefits, it should be noted that according to Fredriksson *et al.* (2017), the concept of BD and BDA is just beginning to be put into practice in the public sector and requires more studies to yield a more cohesive understanding of its theory and practice in this sector.

According to the digital transformation strategies, electronic government has been found to be a vital center of modern economies (GIL-GARCIA & FLORES-ZÚÑIGA, 2020). In this sense, the term electronic government or e-government can be understood as a set of actions that aim to improve the quality and increase the accessibility of information, transactions, and government services through the use of ICTs (CRIADO & GIL-GARCIA, 2019; DIAS, 2019).

The electronic transformation required by government services using ICTs is considered a complex, expensive task and is generally associated with the informatization of public services and business systems (JANSSEN & VAN DEN HOVEN, 2015). In addition, it needs to align with legislation and policies for improving quality, data accuracy, information security, and controlling the data flow of e-government (ABU-SHANAB, 2020).

Due to rapid and constant technological changes, e-government faces the challenge of managing the large volume of data generated and managed by public organizations, leading them to implement BD technologies (LÖFGREN & WEBSTER, 2020). Therefore, government can use BD to discover trends and patterns resulting from the data so that the government can reach goals, provide better services to society, and, therefore, generate public value (MORABITO, 2015; TWIZEYIMANA & ANDERSSON, 2019).

A trend towards consensus can be observed in terms of the characteristics of BD, even though there is no established broad definition in existence. BD was originally characterized in terms of the 3Vs: Volume – quantity of data; Variety – variety of data; and Velocity – frequency/time in which the data is generally generated and processed (MCAFEE & BRYNJOLFSSON, 2012, CHEN *et al.*, 2014). Later there were included the 4Vs and 5Vs: Volume, Variety, Velocity, Value – richness of the data in providing benefits, and Veracity – the reliability of the data (see CHEN *et al.*, 2014; CHEN & ZHANG, 2014, WAMBA *et al.* 2015). Current studies consider the 7Vs, with the addition of Variability – how dynamically data changes significance, and Visualization – which is how the output is presented (see SEDDON & CURRIE, 2017; SIVARAJAH *et al.*, 2017). The Vs represent the continual expansion of data in terms of its complexity and attributes. Other Vs may emerge with the adoption of BD according to the context of their use and/or data analysis since they are associated with Value (SEDDON & CURRIE, 2017). Studies (FREDRIKSSON *et al.*, 2017; SHENG *et al.*, 2017; SIVARAJAH *et al.*, 2017; IQBAL *et al.*, 2018) also relate a paucity in terms of

deeper investigations of the concept and characteristics of BD, which is why it is necessary to seek a more refined definition of them according to the scenario which is being adopted. Within this context, the current study presents a framework based on the literature and made up of the 7Vs, their descriptions, and key questions, which makes it possible better characterize BD, as can be seen in Box 1.

Box 1 – Framework for the characterization of Big Data

Components	Description	Key questions	Theoretical references
Volume	Volume refers to the quantity and/or size of data to be collected, processed and/or analyzed.	What is the volume of data collected? What is the volume of data processed?	Chen <i>et al.</i> (2014) Chen & Zhang (2014) McAfee & Brynjolfsson (2012) Wamba <i>et al.</i> (2015) Seddon & Currie (2017)
Velocity	Velocity of the data refers to the frequency with which the data is generated, transmitted, processed and/or updated.	How frequently is the data generated? How much time is necessary to transmit, store, process or structure new data? How frequently is the data delivered?	Chen <i>et al.</i> (2014) Chen & Zhang (2014) McAfee & Brynjolfsson (2012) Wamba <i>et al.</i> (2015) Seddon & Currie (2017)
Variety	Variety concerns the structural heterogeneity of the composition of Big Data, which can be structured, semi-structured or unstructured. In other words, the data can come from a variety of sources and formats and have different degrees of structure, or forms in which they are registered and significance.	Which types of data are collected? What is the structure of the data? What are the dimensions of the data? What is the format of the content? What is the format of the data? How is the data registered (human management, scientific instruments, or automatic sensor results, among others)?	Chen <i>et al.</i> (2014) Chen & Zhang (2014) McAfee & Brynjolfsson (2012) Wamba <i>et al.</i> (2015) Seddon & Currie (2017)
Value	Value refers to the possible benefits from the adoption/use of Big Data. Also known as the potential wealth of the data.	What are the possible benefits of the use of this tool? Of its results?	Chen <i>et al.</i> (2014) Günther <i>et al.</i> (2017) Seddon & Currie (2017) Sheng <i>et al.</i> (2017) Sivarajah <i>et al.</i> (2017)

Components	Description	Key questions	Theoretical references
Veracity	Veracity measures the accuracy of the data and its potential use in analysis. It deals with the inherent reliability, accuracy, quality, and integrity of the data.	<p>What is the origin of the data?</p> <p>Is there ambiguity in the information?</p> <p>Is there completeness in the data?</p> <p>How reliable is the data to generate these conclusions?</p>	<p>Vasarhelyi, Kogan & Tuttle (2015)</p> <p>Sivarajah <i>et al.</i> (2017)</p> <p>Seddon & Currie (2017)</p>
Variability	Variability configures how dynamic the storage, processing, and possible interpretations of the data are according to the data structure. It also refers to data whose meaning can change quickly and constantly.	<p>What are the various types of data that need to be processed?</p> <p>How should the different types of data be processed together?</p>	<p>Seddon & Currie (2017)</p> <p>Sivarajah <i>et al.</i> (2017)</p>
Visualization	Refers to the presentation of information in the most intuitive and effective manner, revealing the connectivity of a great quantity of data in graphic form.	<p>Is there connectivity between the data? What is the visual level of the interpretation of patterns and trends which are present in the data?</p>	<p>Chen & Zhang (2014)</p> <p>Seddon & Currie (2017)</p> <p>Sivarajah <i>et al.</i> (2017)</p>

Source: elaborated by the authors.

According to Wamba *et al.* (2015), BD can be considered as a wholistic approach to generate, process, and analyze the Vs in order to create insights for the delivery of sustainable value, measure performance and establish competitive advantages. The authors add that the Vs make it possible to guide the development of more sophisticated tools to meet the needs of dealing with a sea of data which is constantly generated.

The adoption of BD presents organizations with legal, ethical, regulatory, and governance challenges related to the data due to its complexity. According to several studies (SIVARAJAH *et al.*, 2017; SAGGI & JAIN, 2018), the challenges of adopting BD may be divided into three broad categories, considering the data life cycle: data challenges – which are associated with the characteristics of the data itself; process challenges – which are correlated with technical issues (the capture, integration and transformation of data, the selection of models, and the providing of results); and overall managerial challenges (ex. cybersecurity). The challenges of BD, as well as its performance, can be investigated initially through the Vs, making it possible to understand the strategies and/or improvements that should be adopted to ensure that the implementation and management of BD initiatives are successful. In Figure 1, we can verify the

primary challenges identified by the BD literature.

Figure 1 – Challenges of Big Data

Categories	Variables	Authors
Data challenges	Large volume of data (Volume) Heterogeneities, various sources, unstructured data, a variety of data (Variety) Integrity, quality, origin, precision, inconsistency and pollution of the data (Veracity) Large inflow of data, high degree of scalability and rate of transmission (Velocity) Distinct meanings for the same data and rapid change in data (Variability) Visualization of groups of complex data, Presentation of results (Visualization) Generation of Value, connectivity and the associations of the data (Value) Absence of data	Sivarajah <i>et al.</i> (2017) Akhavan-Hejazi & Mohsenian-Rad (2018)
Process challenges	Availability of the data Selection Acquisition Preparation Storage Analysis, interpretation and presentation of results Monitoring and control in real time Data mining and cleaning Data integration Data modeling Data recovery Slow and onerous data processing	Blazquez & Domench (2018) Hadi <i>et al.</i> (2018) Mehta & Pandit (2018) Saggi & Jain (2018) Wang <i>et al.</i> (2018) Dash <i>et al.</i> (2019) Kumar <i>et al.</i> (2020)
Managerial challenges	Privacy and Security Data governance Transparency Data ownership Costs Culture of data sharing Legal issues Policies Ethics Training A lack of understanding of the results Rapid and reliable communication Cybersecurity	Shastri & Deshpande (2020)

Source: elaborated by the authors.

On the other hand, the implementation of BD and BDA offers organizations the potential benefits mentioned above. Kumar *et al.* (2020) describe that the function of BD in the health area is to extract knowledge from structured and unstructured data and points out that the main benefits of it are: improvement in the quality of health information; offers a big picture view of treatments provided to meet future needs; optimization of business growth-related decisions; immediate access for clinical data analysis; shortens the time of diagnostic tests; reduces surgery-related hospitalizations; explores new avenues of research; and, creates very competitive healthcare services.

Wang *et al.* (2018) assessed twenty-six cases of the implementation of BD in the health area, mapped the potential benefits generated by the BD and BDA, and identified that these benefits can be classified into five categories: IT infrastructure benefits; operational benefits – those obtained by improvements in operational activities; organizational benefits – which can be considered the results of the focus, cohesion, learning and execution of the chosen strategies; managerial benefits – which are obtained by business management activities that involve the allocation and control of resources; and finally, strategic benefits – which are obtained from strategic activities that involve long-term planning. We can verify the main benefits identified by these authors in Figure 2.

Figure 2 – Potential benefits of Big Data

Categories	Variables
IT infrastructure benefits	<ul style="list-style-type: none"> Reduces system redundancy Avoids unnecessary IT costs Transfers data quickly between healthcare IT systems Better use of healthcare systems Process standardization among various healthcare IT systems Reduces IT maintenance costs in terms of data storage
Operational benefits	<ul style="list-style-type: none"> Improves the quality and accuracy of clinical decisions Processes a large number of healthcare records in seconds Reduces patient travel time Immediate access for clinical data analysis Shortens the time of diagnostic tests Reduces surgery-related hospitalizations Explores new avenues of research
Managerial benefits	<ul style="list-style-type: none"> Offers quick insights about changing healthcare trends in the market Provides board members and department heads with sound information to support their decision making in a daily clinical setting Optimization of business growth-related decisions
Strategic benefits	<ul style="list-style-type: none"> Offers a big picture view of treatments provided to meet future needs Creates very competitive healthcare services

Categories	Variables
Organizational benefits	<p>Detects interoperability problems much more quickly than traditional manual methods</p> <p>Improves cross-functional communication and collaboration among administrative staff, researchers, clinicians and IT staff</p> <p>Enables the sharing of data with other institutions and adds new services, sources of content and research partners</p>

Source: adapted from Wang *et al.* (2018).

The review of existing literature (MEHTA & PANDIT, 2018; KULKARNI *et al.*, 2020; KUMAR *et al.*, 2020) indicates a gap in research on the benefits, challenges, and case studies related to the adoption of BD in the public sector of developing countries, such as Brazil, which reinforces the relevance of the present research.

3. METHODOLOGICAL PROCEDURES

To analyze the adoption of the BD phenomenon in the public sector, the study integrates qualitative and quantitative approaches to take advantage of the reconciliation of these methods to examine the same research problem, which permits their mutual corroboration, and thus increases the confidence in the results (BRYMAN, 2003).

The selection of the Ministry of Health case study, through the Fundo Nacional de Saúde (FNS), as the manager responsible for financial transfers of approximately 113 billion BRL in 2020 to the Brazil's Public Health Care System, known as SUS, was made due to the strategic importance of this area for Brazilian public health. In addition, the FNS's mission is to contribute to the strengthening of citizenship through the continuous improvement of the financing of health actions, and seeks daily to create mechanisms, such as BD, to make information available to the whole society regarding costing, investments, and financing for the SUS.

The governmental body's implementation of BD on February 16th. 2017, was demonstrated in the excerpt of Contract N°. 11/2017, published in the Official Diary of the Union, designed to consolidate the large volume of data coming from the processing of 12 transactional systems administered by the FNS and SUS's IT Department (DATASUS). In this way, the adoption of a BD solution incorporated the administrative flow of information related to the transfer of financial resources dedicated to financing the current and capital expenses of the Ministry of Health, including institutions and entities directly and indirectly subordinate to it which are part of the SUS.

In terms of data collection, we used a triangulation based on four techniques: bibliographic research, which referred to the collection, reading and interpretation of scientific articles for the elaboration of the big data characteristics analysis framework, potential benefits, and adoption challenges; semi-structured individual interviews to learn the perceptions and gather

information from public servants who have a systemic vision of BD; documental research of reports related to public health financial resources (FUNDO NACIONAL DE SAÚDE, [s. d.]) from the years 2002 to 2018 in order to characterize BD in the public sector and provide a basis for our interview script; and finally, participant observation with the IT professionals involved in the project to acquire knowledge of the context of the studied phenomenon which was duly registered on index cards.

We used the techniques of document analysis and the index cards of participant observation to understand the use of BD in the FNS and to build the structuring elements of the interview script with the project team, namely: BD concept; BD characterization; and benefits and challenges the use of BD by the public organization. To survey the information (Characterization, Benefits, and Challenges of BD) in the index cards, five meetings were held in March 2018 between one of the researchers and the team of IT professionals responsible for BD in the FNS meeting room. To analyze the information from the documents and participant observation, content analysis was used. Box 2 presents the interview script used and the sources used for its elaboration.

Box 2 – Interview Script

Category	Question	Source
Concept of BD	What do you mean by BD?	Bibliographic research
Concept of BD	What is the purpose of using BD in the FNS?	Bibliographic research
Characterization of BD	What is the importance of using BD in Public Administration, especially in the context of the FNS?	Document analysis; Participant observation
Benefits of BD	What are the benefits of using the tool?	Bibliographic research; Document analysis; Participant observation
Challenges of BD	What are the main challenges encountered during the use of BD in FNS?	Bibliographic research; Document analysis; Participant observation
Characterization of BD	What are the main results presented so far with the use of the BD by the FNS?	Document analysis; Participant observation
Characterization of BD	How are the analyzed data being presented? For whom? How are deliveries being used?	Document analysis; Participant observation

Source: elaborated by the authors.

We conducted the primary data collection through semi-structured individual interviews with FNS managers and IT professionals from FNS and DATASUS. The sample was defined

based on simple random criteria according to the agenda of these professionals and considering the universe of 10 public servants directly involved in the adoption of the BD project in the Ministry of Health. We conducted the interviews with three of the five managers, and four of the five IT professionals, during the months of April and May 2018, and they were recorded in digital audio files for transcription and data collection purposes. Box 3 presents the data for the individuals interviewed.

Box 3 – Sample Characteristics

Interviewee	Area
E1	IT Support Coordination
E2	IT Support Coordination
E3	IT Support Coordination
E4	IT Support Information
E5	Overall Coordination of Investment Analysis and Realization
E6	IT Support Coordination
E7	Administrative Board

Source: elaborated by the authors.

In this manner, in selecting the sample of those interviewed, we have sought to include professionals who occupy various roles involved in the adoption and utilization of BD in the Ministry of Health, assuming that these individuals would be the most qualified to provide a systemic vision of the financial and budgetary management of the funds dedicated to IT management in public health.

We emphasize that, before starting each interview, the objective of the research was explained to the interviewees. Furthermore, we informed them that we would guarantee the anonymity of their answers and that the research would follow all the principles of ethics. Also, the respondents were informed that they had the right to leave the interview at any time and should read and sign the Informed Consent Form.

In terms of the data analysis related to the concept of BD in the public health sector, interview transcriptions were analyzed using the IRaMuTeQ 0.7 program, which is known as the French acronym "*Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires*" (R interface for multidimensional analysis of texts and questionnaires). We chose this software because it is an open-source solution, expanding the accessibility to the study's methodological replicability. Iramuteq is based on the R software, python language, and allows researchers to analyze both unstructured textual data of different sizes (e.g., interviews, documents) and structured textual data (e.g., categorical variables). It has proven useful as it

provides statistical robustness, allowing to perform analyses that retrieve the semantic context to which the words belong (JUSTO & CAMARGO, 2014; RATINAUD, 2014).

A similarity analysis based on graphic theory permitted the identification of co-occurrence between words in the form of a connected maximum tree graph without cycles (MARCHAND & RATINAUD, 2012). This analysis made it possible to identify the portions of the interviewee statements that occur in several instances, as well as provided a graphic representation of the lexical context. A few words were grouped together to achieve a better congruence in terms of the semantic context. The central variables in capital letters should be considered the most representative variables in the graph, and the strongest correlations (X) are those that have thicker connections with others (GOULART, VIANA & LUCCHESI-CHEUNG, 2021; MARCHAND & RATINAUD, 2012).

It should be noted that Qualitative Data Analysis Software's (QDAS) have the potential to support a wide variety of mixed and qualitative research approaches and have been used by researchers in different knowledge fields (CHANDRA & SHANG, 2017). IRaMuTeQ 0.7 is not a method, and its outputs are not analytics per se. Thus, the graphical representations were subjected to analytical interpretation by the researchers (JUSTO & CAMARGO, 2014; SANTOS et al., 2017; BARDIN; 2011), allowing, their content to be understood.

To characterize BD through its architecture, as well as its main characteristics, we performed a documental analysis associated with the elements collected by the participatory observations, which describe the main attributes associated with its adoption in the public health sector. To identify the presence or absence of the categories and variables related to the benefits and challenges involved in adopting BD, we have opted to proceed with content analysis of the interviews, as defined by Bardin (2011) as a group of techniques used to analyze verbal communication in order to obtain inferences of knowledge through systematic procedures.

In this sense, the reconciliation of the collection and analysis techniques cited above are justified by the chance to take advantage of a rich methodology which offers a complementary nature which can broaden our understanding of the BD phenomenon in the Brazilian public health sector, as well as present opportunities for the improvement of the methodologies used in these studies.

4. RESULTS

The Brazilian Constitution of 1988 established in the Fundamental Rights and Guarantees, in Article 6, that health is the right of all citizens and the duty of the State, guaranteed through social policies and universal and equal access to actions and services, for its promotion, protection, and recovery (BRASIL, 1988). Hence the SUS, which provides for the organization in a regionalized and hierarchical public health care system, must conform to the following main

guidelines: comprehensive care, decentralization with autonomy in each sphere of Brazilian government, and community participation (PAIM, 2018).

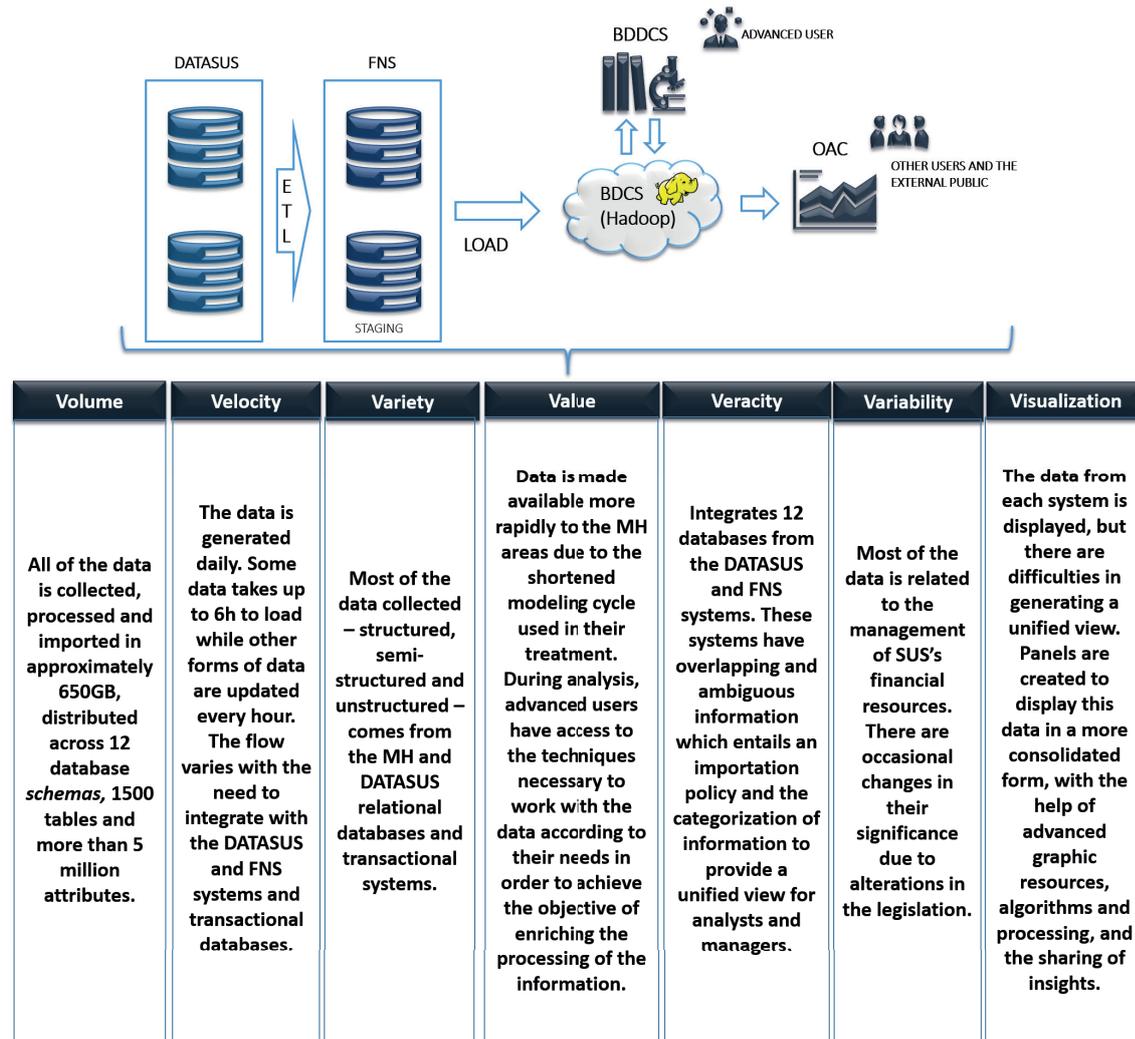
In order to guarantee the capillarity of the SUS, it is necessary to manage a large volume of resources to finance current expenses and investments in 5,570 municipalities, 26 states, and the Federal District, as well as the organizations and entities of the direct and indirect Public Administration members of the Brazilian public health care system. In view of the challenges that arise for the management of the resources, the BD was adopted by the FNS, from the Ministry of Health, in the search for continuous improvement in the financing of health actions.

We observe in Figure 3, based on the similarity analysis of the interviews, that the central variable in terms of BD in the public health sector is 'information.' The concept of BD is related to: technology, a large volume of data, the world of data, structured and unstructured data, existing public data, and the relationship with vs. the adoption of BD offers the ability to visualize results in dashboards, reports, and specific cross-references with activities and the budget. In addition, it may be noted that it has been verified that BD is very important in decision making and that the BD tool BDA, associated with BI, which is related to its capacity in terms of treatment and results, assists this governmental body (FNS) in managing this area and improving its strategy, as well as the velocity, quality, and transparency of its information.

Given the absence of a BD concept, especially in the context of the Brazilian public health sector (MEHTA & PANDIT, 2018; SHASTRI & DESHPANDE, 2020), we propose the following definition for the concept of BD, based on the theoretical framework and the words presented in the similarity analysis carried out in Figure 3: 'A large volume of information from structured, semi-structured or unstructured data related to public health that is continuously generated from various sources, which is used to influence decision making and provide transparency as well as generate public value and social well-being.'

permits the distributed processing of large groups of data on various machines and their storage in the BDCS, which is a platform to execute Hadoop system loads and develop BD applications. As a result, the processed data is made available through a BD Discovery (BDDCS) cloud storage service consisting of a catalogue for the exploring and visualization of this data, which is available in two forms – in its entirety for advanced users or in a limited manner for public consumption.

Figure 4 – The architecture and characterization of Big Data in the public health sector



Source: elaborated by the authors.

The interview data is based on content analysis, which makes it possible to profit from this material through its categorization procedures which identify the categories and variables related to the challenges and benefits of implementing BD in the public health sector presented in Figures 1 and 2 of the Theoretical References Section above.

Based on this categorization, two analysis vectors were defined: (a) Vector 1 corresponds to the declarations that reflect the presence of the analysis categories and the variables previously identified by the authors in the theoretical references, while (b), Vector 2, groups the analysis categories and the newly observed variables of this phenomenon.

Box 4 presents the main systematized data analysis results from the interviews. The related categories appear within the brackets, followed by the related variables that have previously been identified in the literature. The numbers between the parentheses represent the number of the interviewee who signaled his or her agreement or verbally agreed with the transcribed declarations.

Box 4 – Results of the content analysis of the challenges and benefits of the adoption of Big Data in the public health sector according to managers and IT professionals

Vector 1	Vector 2
<p>“Health has historically dealt with an enormous amount of financial data which has been accumulating over the past 20 years, especially the last eight years of this data which we’re trying to qualify to generate information and through this guarantee the transparency of financial resources, an increase in social control and a number of other important things.” [Data challenges – Large volume of data - Volume] - (I7)</p>	<p>“I would add that this project became viable due to the strong support of the Director of the FNS because there have been so many setbacks with upper management during this journey that otherwise it would have been abandoned.” [Managerial benefits – Leadership Support] - (I1)</p>
<p>“(…)We need to improve our qualifying of the information being generated within the MH so that quality structured data can be combined with unstructured data generated by society as a whole to aggregate more information which will inform the work of the MH as a whole.” [Data challenges – Quality – Veracity] - (I4)</p>	<p>“So, with Big Data we don’t have that rigid way of creating reports, which has an ecological aspect as well because generally all of our reports are printed, and the advantage of dashboards is that they don’t use paper. It’s something that you can share with various people and can use them in your presentations in a totally interactive manner.” [Managerial benefits – Sustainability] - (I2).</p>

Vector 1	Vector 2
<p>“So, this breaks the paradigm because logically we aren’t able to simply take several databases and throw them into a data lake and simply cross-reference them. That’s why they talk a lot about, 80% of the effort of Big Data in health goes into normalization analysis and cleaning the data.” [Process challenges – Data mining and cleaning] - (I2)</p>	<p>“Since it’s a very new experience, it’s the IT team that’s doing this, but none of it would be possible without the participation of the business managers and the executives and overall coordinators within the FNS”. [Managerial benefits – Leadership Support] - (I2)</p>
<p>“The idea is that this tool should not be restricted to just being used by FNS managers, because IT has spent a lot of time extracting this data to meet the needs of various areas, so a request sometimes takes almost 15 days because it’s very difficult to put the data in the format that the manager wants initially.” [Process challenges – Slow and onerous data processing] - (I1)</p>	<p>“And the FNS managers and division chiefs are also participating a lot in this process, because as I mentioned before there’s a strategic, tactical and operational side that’s fundamental in this process.” [Managerial benefits – Leadership Support] - (I2)</p>
<p>“I think that today the biggest challenge that we face is really the issue of the source and mainly its security from the point of view of whether this information is precise or not. There are a lot of things that are still being loaded manually and we know what can happen – a simple typing error and an extra 0 can make a big difference.” [Managerial challenges – Cybersecurity] - (I5)</p>	

Vector 1	Vector 2
<p>“(…) DATASUS understood this need and we managed to adhere to an agreement that registered the price of consulting for Oracle tools, and Big Data was included in this package.” [Managerial challenges – Legal issues] - (I1)</p> <p>“(…) What is attractive about it is the velocity with which it generates results, the quality with which it presents the results, and the velocity with which it generates this type of information, which is something that wasn’t even expected. You create a result that can make you change direction.” [Managerial benefits – Decision optimization related to business growth] - (I1)</p> <p>“The sum of all the information dashboards generated by BD ended up generating a larger dashboard for the Ministry of Health where they have a vision of the progress of all the financial adjustments, what has been spent, what has been paid for, what portion of a contract should be passed on, hospital agreements, all of this is mounted onto a large dashboard for the Ministry.” [Strategic benefits – Offers a big picture view of the treatments provided to meet future needs] - (I1)</p> <p>“We’re in the middle of a transition in which the federal government produces a lot of information, but it’s in small silos within various governmental bodies, institutions and agencies which all work together in a way, however this information is not disseminated in the most linear fashion. With the coming of Big Data, you can treat structured and unstructured data where you can have various databases that don’t communicate with each other within the structured model, but you can make correlations.” [Organizational benefits – Enables the sharing of data with other institutions and adds new services, sources of content and research partners] - (I2)</p> <p>“Another great difficulty is the issue of the people themselves, because we are very used to working with systems and adding our famous reports to these systems. So, there are various types of reports, but it’s difficult for the technology team to create specific reports for every need. What’s the advantage of the dashboards of Big Data? It’s that each person can create scenarios and filters and thus can obtain the same answer in a simpler manner that is more intuitive because these are interactive graphics.” [IT infrastructure benefits – Avoids unnecessary IT costs] - (I2)</p> <p>“Today the visibility that this tool is bringing is very important as well as the intermediate management of the overall coordinators, where they can see where there are errors in IT and where we need to organize ourselves.” [Operational benefits – Improves the quality and precision of decision making] - (I3)</p>	

Source: elaborated by the authors.

The results of the content analysis performed on the interviews demonstrated that the defined categories – data challenges, process challenges and managerial challenges – were present in our interviewees' discourses. Box 5 systematizes the three categories, 42 variables related with challenges which were observed in the BD phenomenon in the public health sector, and five variables that were not validated in the survey.

Box 5 – Categories and variables related to the challenges of adoption Big Data in the public health sector

Categories	Validated Variables	Non-validated variables
Data challenges	Large volume of data (Volume)	
	Heterogeneities, various sources, unstructured data, a variety of data (Variety)	
	Integrity, quality, origin, inconsistency, pollution of the data (Veracity)	
	Large inflow of data, high degree of scalability and rate of transmission (Velocity)	Precision and
	Distinct meanings for the same data and rapid changes in data (Variability)	transmission of data
	Visualization of groups of complex data, Presentation of the results (Visualization)	
	Generation of Value, connectivity and the association of data (Value)	
Absence of data		
Process challenges	Availability of data	
	Selection	
	Acquisition	
	Preparation	
	Storage	
	Analysis, interpretation, and presentation of results	Data recovery
	Monitoring and control in real time	
	Data mining and cleaning	
	Data integration	
Data modeling		
Slow and onerous data processing		
Managerial challenges	Data governance	
	Transparency	
	Data ownership	
	Costs	Privacy
	Culture of data sharing	Ethics
	Legal issues	Rapid and reliable
	Policies	communication
	Training	
A lack of understanding of the results		
Cybersecurity		

Source: elaborated by the authors.

The variables which were not observed as challenges of this phenomenon were: precision and transmission of data, which is classified under data challenges; data recovery, which is classified under process challenges; and finally, privacy, ethics, and rapid and reliable communication which are classified under managerial challenges. All the other variables were mentioned by the interviewees.

In terms of the benefits of BD for the public health sector, five categories were identified based on the interview content: IT infrastructure benefits, operational benefits, managerial benefits, strategic benefits, and organizational benefits. Of the 21 variables defined as possible benefits, only three were not mentioned by the interviewees, namely: reduces patient travel time; shortens the time of diagnostic tests; and finally, reduces surgery-related hospitalizations. Box 6 presents the categories and variables observed for the studied BD phenomenon.

Box 6 – Categories and Variables related to the benefits of adoption Big Data in the public health sector

Categories	Variables from literature	New variables arising from the research
IT infrastructure benefits	<ul style="list-style-type: none"> Reduces system redundancy Avoids unnecessary IT costs Transfers data quickly between healthcare IT systems Better use of public health systems Process standardization among various healthcare IT systems Reduces IT maintenance costs in terms of data storage 	
Operational benefits	<ul style="list-style-type: none"> Improves the quality and accuracy of clinical decisions Processes a large number of health records in seconds Immediate access for clinical data analysis Explores new avenues of research 	
Managerial benefits	<ul style="list-style-type: none"> Offers quick insights about changing healthcare trends in the market Provides board members and department heads with sound information to support decision making in a daily clinical setting Optimization of business growth-related decisions 	<ul style="list-style-type: none"> Leadership Support Sustainability

Categories	Variables from literature	New variables arising from the research
Strategic benefits	Offers a big picture view of treatments provided to meet future needs Creates very competitive healthcare services	
Organizational benefits	Detects interoperability problems much more quickly than traditional manual methods Improves cross-functional communication and collaboration among administrative staff, researchers, clinicians, and IT staff Enables the sharing of data with other institutions and adds new services, sources of content, and research partners	

Source: elaborated by the authors.

Furthermore, the new variables that arose from the interviews – sustainability and leadership support – are in consonance with two overall benefits of BD for Public Administration: the use of technological innovations to promote sustainable development; and the importance of providing support to upper management in the implementation of these innovations in the public sector (DE VRIES, BEKKERS & TUMMERS, 2016).

5. DISCUSSION

The concept of BD in the public health sector that has been proposed based on these theoretical references and empirical data has provided us with an opportunity to strengthen an operational definition of the public health sector environment, which corroborates a need cited in the review of the literature performed by Mehta and Pandit (2018). In addition, it highlights that BD, as an innovation related to e-government, should aim at generating public value for the organization that adopts it, as well as the society that benefits from public health resources, which confirms the understanding of Twizeyimana and Andersson (2019) and Criado and Gil-Garcia (2019).

In terms of characterizing BD through its architecture, the seven analyses and observation components defined by the proposed framework - volume, velocity, variety, value, veracity, variability, and visualization – are attributes that make this phenomenon tangible in the public health sector, because as the descriptions and key issues proposed make it possible to measure this architecture and ratify the propositions related by the authors in Box 1.

The content analysis demonstrates that the challenges and benefits related to the implementation of BD in the public health sector are in consonance with findings in other contexts, according to the wholistic vision presented by the authors Sivarajah *et al.* (2017) and Wang *et al.* (2018).

The most mentioned data challenges identified were the large volume of data generated by these systems, the storage, and the low quality of data produced and the variety of this data, which were pointed out by Sivarajah *et al.* (2017), Blazquez and Domench (2018), Wang *et al.* (2018), Dash *et al.* (2019), and Shastri and Deshpande (2020).

In terms of the process challenges of BD, the interviewees reported that the analysis, interpretation, and presentation of the data used were the greatest problems encountered within the public health management context, followed by the integration and preparation of data, which corroborates the results presented by Wang *et al.* (2018), Akhavan-Hejazi and Mohsenian-Rad (2018) and Blazquez and Domench (2018).

Also, the managerial challenges of BD that were most frequently mentioned by managers and IT professionals included training, data governance, and a culture of sharing data, which were mentioned by the authors Mehta and Pandit (2018), Saggi and Jain (2018) and Wang *et al.* (2018).

Moreover, we also analyzed the challenges related to the transparency of information related to the large volume of financial resources for health care that the FNS faced, and that the BD provided the projection of this information in a unique database towards digital transparency in the public sector health, in line with proposed by the authors Matheus, Janssen and Janowski (2021) and Janssen and Van Den Hoven (2015).

Additionally, we highlight four variables not observed in the BD phenomenon in the Brazilian public health sector: precision and transmission of data; data recovery; privacy; ethics; rapid and reliable communication. The lack of validation by the interviewees of ethical and information security aspects with the use of BD refutes what was proposed by Dash *et al.* (2019) and reinforces the need for the application in Brazil of the Law 13.709, known as the General Law of Data Protection (BRASIL, 2018).

In terms of the benefits of IT infrastructure, the most mentioned were avoiding unnecessary IT costs, improving the usage of health systems and the standardization of processes among various IT systems. In addition, the operational benefits highlighted included processing a large number of health records and improved quality in decision making, according to Kumar *et al.* (2020). The most notable managerial benefit of BD was offering department heads solid information to support their decision making. In terms of the strategic benefits of BD, the interviewees cited the offering of a big picture view of treatments provided to meet future needs. Finally, in terms of organizational benefits, improvements in cross-functional communication and collaboration between administrative teams were highlighted by the interviewees. All the benefits mentioned corroborate the research of Wang *et al.* (2018).

6. CONCLUDING REMARKS

This article investigated the adoption of BD for the management of financial resources related to public health service activities and services, with its locus being the Brazilian Ministry of Health. To perform this analysis, we have used a multi-method approach through a triangulation of data collection and analysis techniques as a way to add conceptually and empirically to the relative understanding of this chosen phenomenon (BRYMAN, 2003).

Based on these techniques, we have sought a better understanding of the BD phenomenon, especially within the context of managing financial resources in Brazilian public health care system. The use of these analyses makes it possible to reflect on the categories and variables involved in the benefits and challenges related to the implementation of BD, which can be used for posterior intervention in these practices within a public organization context.

Given the absence of an established definition, we have proposed a concept of BD for the public health sector which consists of an elevated volume of information from structured, semi-structured and unstructured data related to public health, which is generated continually by various sources, designed to influence decision making, offers transparency, generates public value, and social well-being.

In terms of the characterization of the BD phenomenon in the public health sector, we have accomplished this by using a framework which presents the key issues related to the main attributes which measure its concept and architecture, namely the 7 Vs: volume, velocity, variety, value, veracity, variability, and visualization.

To identify the categories and variables of the challenges and benefits of the implementation of BD in the studied phenomenon, content analysis was performed on the seven interviews we conducted with managers and IT professionals, and through these stakeholders we have been able to identify five categories and 19 variables for benefits, three categories and 42 variables for challenges, and finally, two new variables related to the managerial benefits category for the implementation of BD: sustainability and leadership support.

It should be noted that the analysis of BD from the point of view of the presented results has some limitations, since the data presented refers exclusively to the reality of Brazil, the vision of professionals directly involved with the implementation, and therefore does not reflect the particular characteristics of managing public health resources in other countries.

Due to a growing interest in the implementation of BD, we suggest conducting the multimethod studies to explore and validate the categories and variables identified, which are considered to be secondary data generated by BD. Finally, we suggest studies that investigate and propose performance measures for the evaluation of the results of the implementation itself.

Moreover, it should be noted that e-government and innovative practices in public organizations, especially concerning the implementation of BD, should consist of organized searches for changes which can affect a planned proposal, and with this a culture of efficiently managing organizational and technological resources can be institutionalized as a way of creating collective value for society.

REFERENCES

- ABU-SHANAB, E. A. (2020). E-government contribution to better performance by public sector. In: *Open Government: Concepts, Methodologies, Tools, and Applications*. IGI Global, 2020. p. 1-17. DOI: 10.4018/978-1-5225-9860-2.ch001
- AKHAVAN-HEJAZI, H.; MOHSENIAN-RAD, H. (2018). Power systems big data analytics: An assessment of paradigm shift barriers and prospects. *Energy Reports*, v. 4, p. 91-100. DOI: <https://doi.org/10.1016/j.egy.2017.11.002>
- ARCHENAA, J.; ANITA, E. M. (2015). A survey of Big Data analytics in healthcare and government. *Procedia Computer Science*, v. 50, p. 408-413. DOI: <https://doi.org/10.1016/j.procs.2015.04.021>
- ARNOLD-SCHMITT, L. E.; TRISKA, R. (2014). Informação na área da saúde em tempos de comunicação móvel, *Big Data e computação cognitiva. Razón y palabra*, v. 18, n. 88.
- BARDIN, L. (2011). *Análise de conteúdo*. 3 reimp. Lisboa: Edições, 70.
- BLAZQUEZ, D.; DOMENCH, J. (2018). Big Data sources and methods for social and economic analyses. *Technological Forecasting and Social Change*, v. 130, p. 99-113. DOI: <https://doi.org/10.1016/j.techfore.2017.07.027>.
- BRASIL. (1988). Constituição da República Federativa do Brasil. Brasil. Disponível em: <http://www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm>. Acesso em: 8 abr. 2021.
- BRASIL. (2018). Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados. Disponível em: <>. Acesso em: 11 abr. 2021.
- BRYMAN, A. (2003). *Quantity and quality in social research*. Routledge.
- CHANDRA, Y.; SHANG, L. (2017). An RQDA-based constructivist methodology for qualitative research. *Qualitative Market Research*, v. 20, n. 1, p. 90-112. DOI: <https://doi.org/10.1108/QMR-02-2016-0014>
- CHEN, C. L. P.; ZHANG, C. Y. (2014). Data-intensive applications, challenges, techniques, and technologies: A survey on Big Data. *Information Sciences*, v. 275, p. 314-347. DOI: 10.1016/j.ins.2014.01.015
- CHEN, M.; MAO, S.; LIU, Y. (2014). Big Data: A Survey. *Mobile Networks and Applications*, v. 19, n. 2, p. 171–209. DOI: 10.1007/s11036-013-0489-0
- CRIADO, J. I.; GIL-GARCIA, J. R. (2019). Creating public value through smart technologies and strategies. *International Journal of Public Sector Management*. DOI: <https://doi.org/10.1108/IJPSM-07-2019-0178>
- DASH, S. et al. (2019). Big data in healthcare: management, analysis and future prospects. *Journal of Big Data*, v. 6, n. 54. DOI: <https://doi.org/10.1186/s40537-019-0217-0>

- DE VRIES, H.; BEKKERS, V.; TUMMERS, L. (2016). Innovation in the public sector: A systematic review and future research agenda. *Public administration*, v. 94, n. 1, p. 146-166. DOI: <https://doi.org/10.1111/padm.12209>
- DIAS, G. P. (2019). Fifteen years of e-government research in Ibero-America: A bibliometric analysis. *Government Information Quarterly*, v. 36, n. 3, p. 400-411. DOI: <https://doi.org/10.1016/j.giq.2019.05.008>
- FREDRIKSSON, C. (2017). et al. Big data in the public sector: A systematic literature review. *Scandinavian Journal of Public Administration*, v. 21, n. 3, p. 39-62.
- FUNDO NACIONAL DE SAÚDE. (s.d.). *Arquivos de Repasse Anual Fundo a Fundo: Repasse FAF com População – 2002 a 2018*. Disponível em: < <https://portalfns.saude.gov.br/publicacoes/>. Acesso em: 11 abr. 2019.
- GIACALONE, M.; CUSATELLI, C.; SANTARCANGELO, V. (2018). Big Data Compliance for Innovative Clinical Models. *Big Data Research*. DOI: <https://doi.org/10.1016/j.BDr.2018.02.001>
- GIL-GARCIA, J. R.; FLORES-ZÚÑIGA, M. Á. (2020). Towards a comprehensive understanding of digital government success: Integrating implementation and adoption factors. *Government Information Quarterly*, v. 37, n. 4, p. 101518. DOI: <https://doi.org/10.1016/j.giq.2020.101518>
- GOULART, G. DA S.; VIANA, M. M. & LUCCHESI-CHEUNG, T. (2021). Consumer perception towards familiar and innovative foods: the case of a Brazilian product. *British Food Journal*, v. 123, n. 1, p. 125-142. DOI: <https://doi.org/10.1108/BFJ-02-2020-0160>
- GÜNTHER, W. A. et al. (2017). Debating Big Data: a literature review on realizing value from Big Data. *Journal of Strategic Information Systems*, v. 26, p. 191–209. DOI: <http://dx.doi.org/10.1016/j.jsis.2017.07.003>
- HASHEM, I. A. T. et al. (2015). The rise of “Big Data” on cloud computing: review and open research issues. *Inf. Syst.*, v. 47, p. 98-115. DOI: <https://doi.org/10.1016/j.is.2014.07.006>
- HADI, M. S. et al. (2018). Big Data analytics for wireless and wired network design: A survey. *Computer Networks*, v. 132, p. 180-199. DOI: <https://doi.org/10.1016/j.comnet.2018.01.016>
- IQBAL, M. et al. (2018). A study of big data for business growth in SMEs: Opportunities & challenges. In: *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*. IEEE. p. 1-7. DOI: 10.1109/ICOMET.2018.8346368.
- JANSSEN, M. & VAN DEN HOVEN, J. (2015). Big and Open Linked Data (BOLD) in government: A challenge to transparency and privacy?. DOI: <https://doi.org/10.1016/j.giq.2015.11.007>
- JUSTO, A.M., & CAMARGO, B.V. (2014). Estudos qualitativos e o uso de softwares para análises lexicais. In Caderno de artigos: X SIAT & II Serpro (p. 37–54). UNIGRANRIO. DOI: <<https://lageres.wordpress.com/>>
- KIM, G-H, TRIMI, S. & CHUNG, J-H. (2014). Big-Data Applications in the Government Sector. *Communications of the ACM*, v. 57, n. 3, p. 78-85. DOI: 10.1145/2500873
- KULKARNI, A. J. et al. (Ed.). (2020). *Big Data Analytics in Healthcare*. Springer, Cham. DOI: <https://doi.org/10.1007/978-3-030-31672-3>

- KUMAR, Y. et al. (2020). Big Data analytics and its benefits in healthcare. In KULKARNI, A. J. et al. *Big Data Analytics in Healthcare*. Springer, Cham, p. 3-21. DOI: https://doi.org/10.1007/978-3-030-31672-3_1
- LÖFGREN, KARL, WEBSTER, C. William R. (2020). The value of Big Data in government: The case of 'smart cities'. *Big Data & Society*, v. 7, n. 1, p. 2053951720912775.
- MARCHAND, P.& RATINAUD, P. (2012). L'analyse de similitude appliqué aux corpus textuelles: les primaires socialistes pour l'élection présidentielle française. In: *Actes des 11eme Journées internationales d'Analyse statistique des Données Textuelles. JADT*, p. 687-699. Presented at the 11eme Journées internationales d'Analyse statistique des Données Textuelles, Liège, Belgique.
- MATHEUS, R., JANSSEN, M.& JANOWSKI, T. (2021). Design principles for creating digital transparency in government. *Government Information Quarterly*, v. 38, n. 1, p. 101550. DOI: <https://doi.org/10.1016/j.giq.2020.101550>
- MCAFEE, A., BRYNJOLFSSON, E. (2012). Big Data: the management revolution. *Harvard Business Review*, 1 ed.
- MEHTA, N.; PANDIT, A. (2018). Concurrence of Big Data analytics and healthcare: a systematic review. *International Journal of Medical Informatics*, v. 114, p. 57-65. DOI: <https://doi.org/10.1016/j.ijmedinf.2018.03.013>
- MORABITO, V. (2015). Big data and analytics for government innovation. In: *Big data and analytics*. Springer, Cham, p. 23-45. DOI: <https://doi.org/10.1007/978-3-319-10665-6-2>
- PAIM, J. S. (2018). Sistema Único de Saúde (SUS) aos 30 anos. *Ciência & Saúde Coletiva*, v. 23, n. 6, p. 1723-1728. DOI: <http://dx.doi.org/10.1590/1413-81232018236.09172018>
- RATINAUD, P. (2014). IRAMUTEQ - Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires (Versão 0.7 alpha 2). [Computer software]. DOI: <http://www.iramuteq.org>
- RUMBOLD, J. M. & PIERSCIONEK, B. K. (2017). What Are Data? A Categorization of the Data Sensitivity Spectrum. *Big Data Research*. DOI: <https://doi.org/10.1016/j.BDr.2017.11.001>
- SAGGI, M. K. & JAIN, S. (2018). A survey towards an integration of Big Data analytics to big insights for value-creation. *Information Processing & Management*. DOI: <https://doi.org/10.1016/j.ipm.2018.01.010>
- SEDDON, J. J. & CURRIE, W. L. (2017). A model for unpacking Big Data analytics in high frequency trading. *Journal of Business Research*, v. 70, p. 300–307. DOI: <https://doi.org/10.1016/j.jbusres.2016.08.003>
- SHASTRI, A. & DESHPANDE, M. (2020) A Review of Big Data and Its Applications in Healthcare and Public Sector. In KULKARNI, A. J. et al. *Big Data Analytics in Healthcare*. Springer, Cham, p. 55-66. DOI: <https://doi.org/10.1007/978-3-030-31672-3-4>
- SHENG, J., AMANKWAH-AMOA, J. & WANG, X. (2017). A multidisciplinary perspective of Big Data in management research. *International Journal of Production Economics*, v. 191, p. 97–112. DOI: <https://doi.org/10.1016/j.ijpe.2017.06.006>
- SIVARAJAH, U. et al. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, v. 70, p. 263–286. DOI: <https://doi.org/10.1016/j.jbusres.2016.08.001>
- SANTOS, V., SALVADOR, P., GOMES, A., RODRIGUES, C., TAVARES, F., ALVES, K. & BEZERRIL, M. (2017). Visualização de IRAMUTEQ nas pesquisas qualitativas brasileiras da área da saúde: Scoping review.

CIAIQ2017 - Investigación Cualitativa en Salud, 2. <https://Proceedings.Ciaiq.Org/Index.Php/Ciaiq2017/Article/View/1230/1191>

TWIZEYIMANA, J. D. & ANDERSSON, A. (2019). The public value of E-Government—A literature review. *Government Information Quarterly*, v. 36, n. 2, p. 167-178. DOI: <https://doi.org/10.1016/j.giq.2019.01.001>

VASARHELYI, M. A., KOGAN, A. & TUTTLE, B. M. (2015). Big Data in accounting: An overview. *Accounting Horizons*, v. 29, n. 2, p. 381–396. DOI: <https://doi.org/10.2308/acch-51071>

WAMBA, S. F. *et al.* (2015). How ‘Big Data’ can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, v. 165, p. 234-246. DOI: <https://doi.org/10.1016/j.ijpe.2014.12.031>

WANG, Y., KUNG, L. & BYRD, T. A. (2018). Big Data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change*, v. 126, p. 3-13. DOI: <https://doi.org/10.1016/j.techfore.2015.12.019>

WHITE, M. (2012). Digital workplaces: vision and reality. *Business Information Review*, v. 29, n. 4, p. 205-214. DOI: <https://doi.org/10.1177/0266382112470412>

Dayse Karenine de Oliveira Carneiro

 <https://orcid.org/0000-0002-6456-8557>

Doutora em Administração e Mestre em Administração Pública pela Universidade de Brasília (UnB). Bacharel em Administração pela Universidade Potiguar (UnP) e em Ciências Econômicas pela Universidade Federal do Rio Grande do Norte (UFRN). Analista Técnico de Políticas Sociais no Ministério da Saúde.
daysekoc@hotmail.com

Gislayne da Silva Goulart

 <https://orcid.org/0000-0002-0188-0965>

Professora da Universidade Federal de Mato Grosso do Sul (UFMS). Doutora em Administração pela Universidade de Brasília (UnB). Mestre e Bacharel em Administração pela UFMS.
gislayne.goulart@ufms.br

Rafael Barreiros Porto

 <https://orcid.org/0000-0003-2210-7098>

Pós-doutor em Economia Comportamental aplicada ao Marketing pela Cardiff Business School (UK). Doutor em Ciências do Comportamento pela Universidade de Brasília (UnB). Professor Associado da Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas Públicas (FACE) da Universidade de Brasília (UnB).
rafaelporto@unb.br