Association for Information Systems

AIS Electronic Library (AISeL)

CAPSI 2020 Proceedings

Portugal (CAPSI)

10-2020

Systematic Literature Review and Research Agenda

Josiane Caraffini

Antônio Maçada

Mauricius Medeiros

Follow this and additional works at: https://aisel.aisnet.org/capsi2020

Big Data Analytics in Financial Institutions:
Systematic Literature Review and Research Agenda

Josiane Piva Testolin da Silva Caraffini, Universidade Federal do Rio Grande do Sul - UFRGS, Brasil, josianetsc@gmail.com

Antônio Carlos Gastaud Maçada, Universidade Federal do Rio Grande do Sul - UFRGS, Brasil, acgmacada@ea.ufrgs.br

Mauricius Munhoz De Medeiros, Universidade Federal do Rio Grande do Sul - UFRGS, Brasil, mauricius.medeiros@gmail.com

Abstract

Although interest in big data analytics (BDA) has increased in recent years, studies on the subject in the context of financial institutions (FI) are still rare. Given this gap, the objective is to analyze how scientific research approaches BDA in the context of FI. A systematic review of the literature is carried out. The results show that the literature focuses on the themes: risk management, marketing, web and social media, technology and data analytics and consequences of BDA use. In addition to defining and identifying the themes of BDA application in FIs, this study contributes to the proposition of a framework that consolidates the research agenda and proposes directions for future studies on BDA in the context of FI, such as the use of BDA in FI, in particular, on the themes of financial risk management and marketing.

Keywords: Big Data Analytics; Financial Institutions; Banking; Systematic Literature Review.

1. INTRODUCTION

In the past, data were available in limited quantities, and nowadays, with the advancement of technology, they are available in a large volume, velocity, and variety of data (Dremel, Herterich, Wulf, & vom Brocke, 2020). In this context, big data (BD) stands out, which has the potential to enable new business opportunities (Anfer & Wamba, 2019). The digital technology of big data analytics (BDA) enables organizations to analyze large volumes of data from multiple sources and types to generate valuable insights and information that help improve performance and competitiveness (Dubey *et al*., 2019; Maroufkhani, Wagner, Ismail, Baroto, & Nourani, 2019; Mikalef, Boura, Lekako, & Krogstie, 2019; Dong & Yang, 2020).

A study by Capgemini Consulting (2018) highlights that in North America, 60% of financial institution leaders believe that BDA leads to a considerable competitive advantage. Thus, data collection and analysis have become increasingly important to organizations that want information about their business processes to improve their products and services and manage their business efficiently (Khan, 2019). A recent study by Gartner Inc. study (2019) found that data investments and data analytics favor efficiency in internal or operational costs, business transformation, customer service improvement, and business agility. In financial institutions, in addition to helping identify market trends, BDA can provide internal process optimization and risk reduction (Bhuvana, Thirumagal, & Vasantha, 2016).

20.ª Conferência da Associação Portuguesa de Sistemas de Informação (CAPSI'2020)
16 e 17 de outubro de 2020, Porto, Portugal
ISSN 2183-489X

1

In this context, it is noteworthy that recent studies have been developed to analyze BDA in the context of financial institutions, such as the research of Lehrer, Wieneke, Brock, Jung and Seidel (2018), Zhang, Zhu, Yan and Li (2018), Shirazi and Mohammadi (2019) and Hung, He and Shen (2020) who investigated BDA under aspects related to marketing. A survey by Li (2019) addressed BDA through a financial risk management system. Alzaidi (2018) and Elia, Polimeno, Solazzo and Passiante (2018) also analyzed, respectively, the role of BDA in decision-making and its value generation. Therefore, considering that BDA is an important research area (Côrte-Real, Ruivo, Oliveira, & Popovic, 2019), the aim of this study is to analyze how scientific research approaches BDA in the context of financial institutions. To achieve this objective, a systematic literature review (SLR) was defined as a research strategy.

Although the interest in BDA has increased in recent years, both among researchers and professionals (Mikalef, Papas, Krogstie, & Giannakos, 2017; Lin, Hou, & Hsiao, 2019; Elhoseny, Hassan, & Sing, 2020), BDA has not reached yet its full level of disruptive potential (Cusumano, Yoffie, & Gawer, 2020). Additionally, financial institutions have access to a large volume of their clients' data, however, there is no certainty that they are managing to convert this advantage into real benefits (Hung, He, & Shen, 2020). Thus, this study contributes to the theory with the identification of the specificities and the thematic areas of use of the BDA and with the proposition of a research agenda on BDA in the context of financial institutions. As a practical contribution, this research offers the definition of BDA applied to the context of financial institutions, and how this implies addressing the potential of its purpose and benefits, which can help organizations take advantage of this resource to generate results. It is also highlighted, as a practical contribution, the identification of the themes of BDA application in financial institutions. Thus, these organizations can evaluate the applications of BDA to meet the needs within the various areas of the organization. For instance, in risk management, which is considered an increasingly critical issue in financial institutions due to the increase in electronic transactions in the digital environment and the associated risks.

Next, this study is structured as follows: in section 2, a theoretical review on the concept of BDA is presented; in section 3 the research method is presented; in section 4, the results obtained in the SLR are analyzed and discussed; and section 5 contains the conclusions. Finally, this study concludes with references and appendices A and B.

2. BIG DATA ANALYTICS

Big data (BD) is related to a massive amount of data collected from various sources (Akter, Wamba, Gunasekaran, Dubay, & Childe, 2016; Anwar, Khan, & Shah, 2018). Chen, Preston and Swink (2015, p. 7) emphasize "three Vs" in their concept of BD "[...] defined as high-volume (large-scale), high-speed (move/transmission) and high-variety information assets (e.g., numeric, text, video, etc.) [...]". This characterization from the "Vs" is commonly emphasized in studies on BD and can be illustrated as follows: i) the volume is related to the total set of data with great and increasing magnitude; ii) velocity reflects how data are collected, updated and analyzed; (iii) the variety refers to the various sources and formats of the data; iv)

veracity is related to uncertainty, biases, noise and data abnormality; and v) the value is related to the insights and benefits generated by the data (Mikalef *et al*., 2017; Grover, Chiang, Liang, & Zhang, 2018; Khan, 2019).

Watson (2014) states that data alone does not create value for the company, so BDA stands out for the potential for capturing large amounts of data, using analytical techniques and demonstrating results in such a way that it allows value creation (Chen, Preston, & Swink, 2015; Maroufkhani *et al*., 2019). Figure 1 demonstrates the concepts used by the authors in recent studies to define the BDA.

Chen, Preston, & Swink (2015)	• Process of using advanced technologies to examine big data in order to discover useful information (for example, hidden patterns, unknown correlations, etc.) to help improve decisions between business processes, between functions or companies.
Wamba et al. (2015), Anfer e Wamba (2019)	• It is a holistic approach to manage, process and analyze the dimensions related to "5 Vs" data (volume, variety, velocity, veracity and value) in order to create actionable insights for sustained value delivery, measure performance and establish competitive advantages.
Mikalef (2017)	• Processes and tools used to extract insights from big data. BDA encompasses not only data, but also elements of tools, infrastructure and ways of visualizing and presenting information.
Grover *et al*. (2018)	• It is the application of statistical techniques, processing and BD analysis to advance the business.
Dubey *et al*. (2019)	• An organizational capability with tools, techniques and processes that allow the company to process, organize, visualize and analyze data, thus producing insights that enable data-driven operational planning, decision-making and execution.
Maroufkhani et al. (2019)	• A collection of data and technology that accesses, integrates and reports all available data, filtering, correlating and reporting insights that are not achievable with previous data technologies.
Mikalef, Boura, Lekako e Krogstie (2019)	• A new generation of technologies and architectures, designed to economically extract value from large volumes of a wide variety of data, allowing high-speed capture, discovery and / or analysis.
Dong e Yang (2020)	• It is a specific application to manage, prioritize and analyze big data for business purposes. To manage and take advantage of big data, companies need to use BDA to manipulate the 5V's of big data, adopting a new way of understanding consumer behavior and developing data-driven marketing and innovation strategies.

Figure 1 - BDA Concepts

Based on the concepts presented by the authors, it can be said that "BDA is a set of technologies, processes, methods and techniques of analysis and interpretation applied to BD to extract information that generates value to the company and generates competitive advantage". Günther, Mehrizi, Huysman and Feldberg (2017) point out that the adoption of advanced analytics technologies provides great data to be used for the development of ideas, products, and services, and thereby generating value for the company.

Thus, the scope of BDA goes beyond the data, as it includes the elements of technology, processes, analysis and visualization of data (Mikalef *et al*., 2017) that can be used to generate marketing information, business processes, risk management, web and social media (Vassakis, Petrakis, & Kopanakis, 2018). In this context,

Grover *et al*. (2018) highlight the possibilities of using BDA for the prediction of customer propensity, in relation to the offer of products and services, and that by improving the customer experience the company would be developing and maintaining a competitive advantage. While Mikalef *et al*. (2017) highlight the complexity related to data-driven culture, and that instituting this would be a complex task, requiring the involvement of managers at various levels of the organization.

3. METHODOLOGY

3.1. *Strategy and research design*

A systematic literature review (SLR) was conducted based on the guidelines presented by Webster e Watson (2002). To ensure accuracy, the steps proposed by Wolfswinkel, Furtmueller and Wilderom (2013), composed of five stages, were adopted. The choice of these steps is justified by the difficulty related to the methods of structuring and presenting a SLR (Webster & Watson, 2002). Figure 2 illustrates the steps taken to perform this research.

Definition of inclusion and exclusion criteria for literature

Inclusion:
- Journals and conferences;
- Scientific Articles

Exclusion:
- Dissertations, case studies, books, etc.;
- Articles not written in English or Portuguese.

Search terms:
"big data analytics",
"big data analytics capabilities",
"banking",
"financial institution (s)" and
"financial service sector".

Literature search
(N=89)

Research date:
- March/2020.

Data Bases:
- AISeL (2), EBSCOhost (7), Science Direct (10), Scopus (52) and Web of Science (18).

Refinement of the Literature
(N=59)

Duplicate articles removal:
- 29 articles.

Removal of articles that did not develop the theme of this study:
- 30 articles.

Analysis of selected articles
(N=30)

Descriptive analysis.
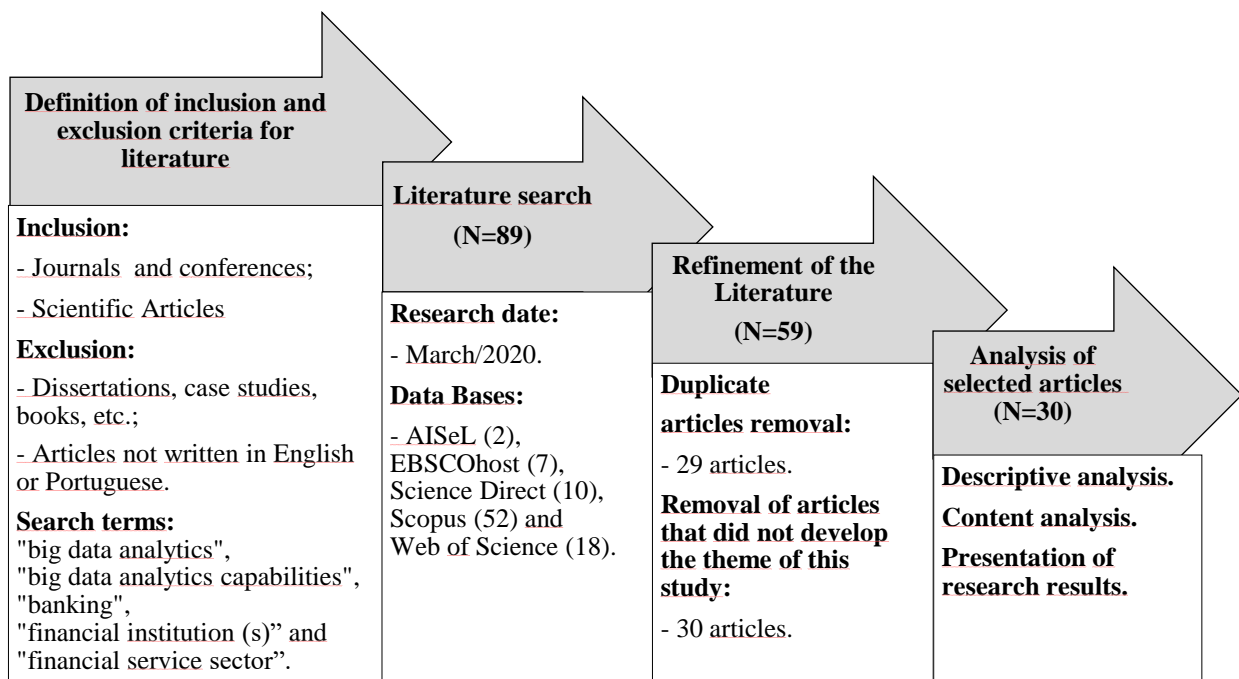**Content analysis.**
Presentation of research results.

Figure 2 - SLR Steps

The first stage consisted of the **definition of inclusion and exclusion criteria from the literature**. The inclusion criteria in the literature searched for articles from high quality sources. Thus, the recommendations of Webster and Watson (2002) were observed and only articles from journals and conferences were included at any date of publication. Dissertations, case studies, books, etc., as well as publications that were not written in English or Portuguese were excluded. For the selection, studies focused on the BDA topic in the financial

sector were defined as eligible. Therefore, the search terms defined were "big data analytics", "big data analytics capabilities", "banking", "financial institution" and "financial service sector".

The **literature research** was conducted in March 2020 in the following databases: AISeL, EBSCOhost, Science Direct, Scopus and Web of Science. These databases enclose a significant gamma of journals and publications of congresses of the Information Systems (IS) area (Webster & Watson, 2002). The research had been limited to the terms title, abstract and keywords. As a result of this stage, 89 articles were obtained.

The **refinement of the literature** was performed through the removal of duplicate articles (29 articles) and articles that did not meet the inclusion criteria (30). At this stage, the fonts were screened by the content of the title, abstract and main body. Articles that used search terms only in an exemplifying or illustrative way were discarded. At the end of this stage, 30 articles remained.

For the **analysis of the selected articles**, descriptive analysis and content analysis techniques were used to structuring it thematically (Creswell & Creswell, 2017). In this stage, the content was classified and grouped according to the categories defined from the literature, which are the themes addressed in this research, namely: Risk Management, Marketing, Web and Social media, Technology and Data Analytics and BDA Use. For data analysis processing, were used NVivo software and spreadsheets in Excel.

3.2. Characteristics of the sample

It is noteworthy that of the 30 selected articles, 18 of them were published in conferences and 12 of them in journals. The year with the highest number of publications was 2017 and, despite a reduction, it remained stable in the following two years, showing that the theme has been receiving more attention. This result follows the increase in interest related to BDA in recent years among academics and professionals (Mikalef *et al.*, 2017; Lin, Hou, & Hsiao, 2019; Elhoseny, Hassan, & Sing, 2020). It is reinforced by the context of financial institutions, which have access to a large amount of behavioral, transactional and demographic data of their clients, representing a promising field in the use of BDA (Hung, He, & Shen, 2020). Appendix "A" presents the main characteristics of the analyzed articles presented in the form of a timeline.

As can be seen in Appendix "A" there is no highlight, in terms of frequency, for the publication of authors and journals/conferences, considering that they do not present a frequency greater than two. Most studies are empirical in nature, representing 76% of the articles analyzed, especially the articles that used the quantitative approach that correspond to 57% of the articles in the sample. As technical research procedures, the documentary research that represents 33% of the articles stands out, this result characterizes the practical application of data in BD analyses. It is also noteworthy that of the 30 articles, only Hung, He and Shen's article (2020) presents a basic theory for conducting the study, namely the Resource-Based View.

Thus, in the next section is performed the description of the **analysis of the selected sources and discussion of the results.**

4.   DATA ANALYSIS AND RESULTS DISCUSSION

4.1.   *Evolution of the BDA concept in the context of Financial Institutions*

Considering the opportunities that financial institutions have in the use of BDA, through the access to a large amount of their customers' data (Hung, *et al*., 2020) and the disruptive potential of the BDA (Cusumano *et al*., 2020), a BDA definition applied to the context of financial institutions can help address its purpose and help these organizations turn this potential into results. To support the consolidation of this definition, we used the analysis of the frequency of the keywords of the selected articles (excluding the search terms) to elaborate a keyword cloud, which is shown in Figure 3.

Figure 3 - Sample keyword cloud

From the analysis of the keyword cloud and the context of each word, it is verified that the articles highlight the data (big data), the technological aspects (hadoop), the methods and techniques of analysis (regression, value tree, analytics), their use (fraud, decision, detection, risk, segmentation) and the context of research in financial institutions (financial, insurance). Subsequently, we analyzed the concepts related to BDA present in the sample, which are presented in Figure 4.

Bhuvana, Thirumagal e Vasantha (2016)	• Ability that supports banks in customer segmentation, combining a variety of data, such as customer behavior patterns, demographic variables and social media sentiment analysis.
Balasupramanian, Ephrem e Al-Barwani (2017)	• Process of collecting, structuring, analyzing and evaluating larger data sets to find patterns and other useful information. BDA helps to understand the information hidden in the data and, at the same time, helps to identify the data that is most important for business and for decision-making at critical moments.
Angée *et al.* (2018)	• It refers to performing analyses directly in a big data environment. This activity requires more technical talent and programming skills than traditional analysis.
Alzaidi (2018)	• It makes it possible to extract insights from big data, serving different departments of the financial institution, which can offer a variety of solutions in the form of fraud detection, risk management, monitoring customer behavior and understanding their needs. Thus, banks can develop an effective decision-making structure, which can offer a high level of efficiency.
Lehrer *et al.* (2018)	• It provides methods and tools capable of collecting, processing and analyzing large amounts of tracking data, allowing organizations to generate valuable information. This information has the potential to create competitive advantage and support customer-oriented service innovation.
Bürger (2019)	• It is an approach to use big data and advanced analytics to add value, improve efficiency and establish competitive advantage.
Dutta *et al.* (2019)	• It is an important and flexible tool available for analyzing data and information for decision-making.
Indriasari, Gaol e Matsuo (2019)	• It is the use of advanced analytics techniques for very large and diverse data sets, which include structured, semi-structured and unstructured data, from different sources and in different sizes.
Hung *et al.* (2020)	• Advanced methods of analyzing and including non-traditional data (such as non-financial data) in order to further improve credit assessment or dynamic monitoring and improve competitiveness.

Figure 4 - Evolution of the concept of BDA in financial institutions

In the definition of the BDA, some studies highlighted its potential in commercial analysis of customers and in processes related to fraud detection, risks and credit assessment (Bhuvana *et al*., 2016; Alzaidi, 2018; Hung *et al*., 2020). Similarly, Dong and Yang (2020) highlighted the use of BDA for commercial purposes, as well as Vassakis, Petrakis and Kopanakis (2018) highlighted the potential of BDA for generating information that is useful for marketing and risk management functions. Other research has focused the definition of BDA on its role in relation to better decision-making (Balasupramanian, Ephrem, & Al-Barwani, 2017; Alzaidi, 2018; Dutta, Rao, Basu, & Tiwari, 2019). For example, Chen et *al.* (2015) pointed out that the application of BDA helps in decision-making, both in relation to the business and in company functions. In this sense, Dubey et *al.* (2019) point out that the insights generated by the BDA enable the company to make better data-driven decisions.

The studies also address the definition of BDA as the competitive advantage gained through real-time data analytics (Lehrer, Wieneke, Brock, Jung, & Seidel, 2018; Bürger, 2019; Hung et *al*., 2020). Similarly, Wamba, Akter, Edwards, Chopin and Gnanzou (2015) and Anfer and Wamba (2019) pointed out that the insights created through BDA can lead to the establishment of competitive advantage. From the analyses carried out

from Figures 3 and 4, it can be inferred that: "in financial institutions, BDA consists of technologies, methods and techniques of analysis applied to BD to extract information related to customers, fraud, risks, among others, and with this, support in decision-making and provide better performance and competitiveness". Therefore, this is a definition of BDA in financial institutions considered in this study.

4.2. *Thematic Analysis of the Application of BDA in Financial Institutions*

BDA can be applied in various sectors and areas of the company. Thus, to analyze how scientific research approaches BDA in the context of financial institutions, the content was classified and grouped according to the categories defined from the literature (Risk Management, Marketing, Web and Social Media, Technology and Data Analytics and Consequences of BDA Use). These categories represent the themes of BDA for which subcategories that constitute the application of these themes in the context of financial institutions have been identified. Figure 5 presents the main results of this analysis.

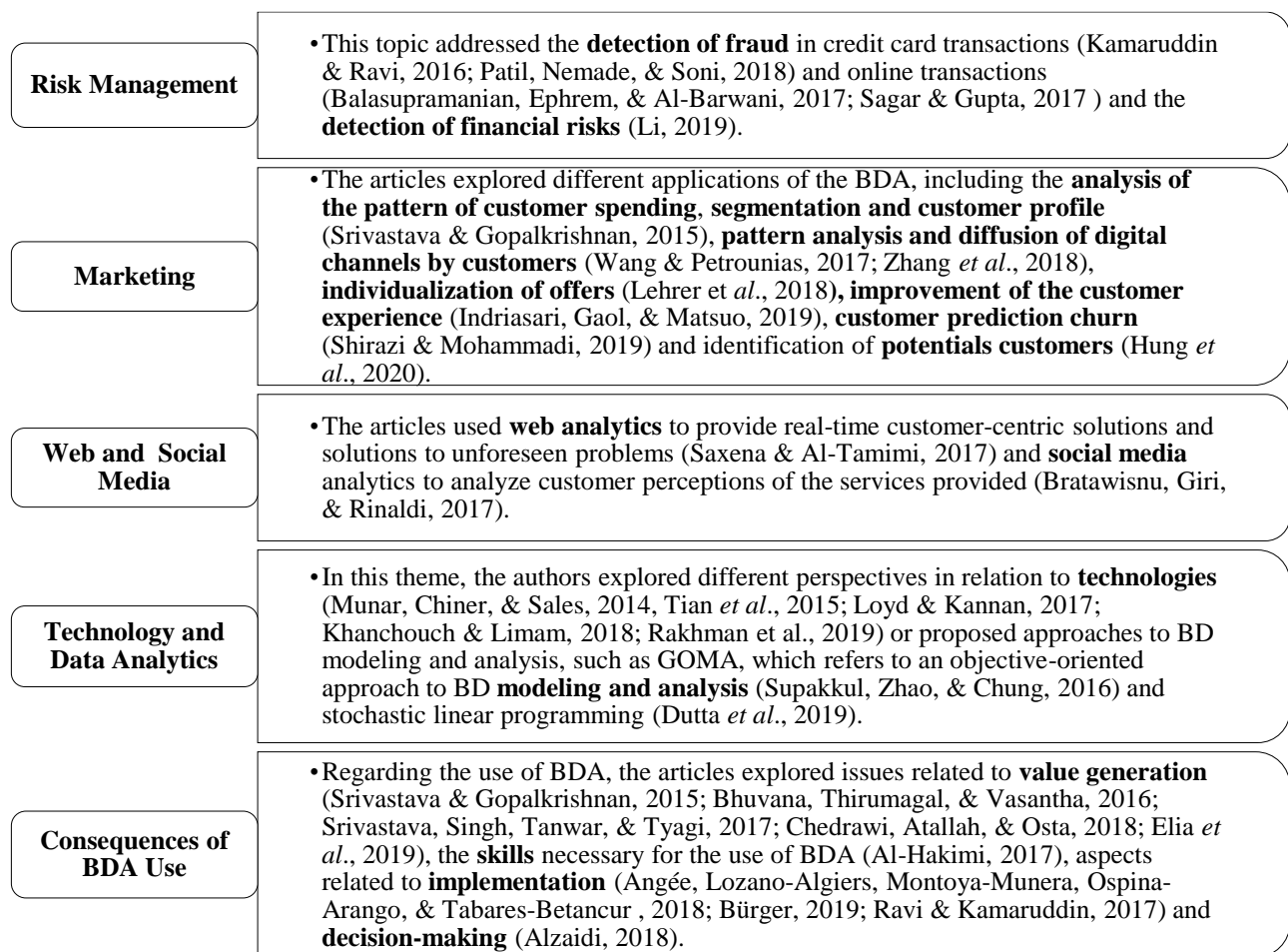| Risk Management	• This topic addressed the **detection of fraud** in credit card transactions (Kamaruddin & Ravi, 2016; Patil, Nemade, & Soni, 2018) and online transactions (Balasupramanian, Ephrem, & Al-Barwani, 2017; Sagar & Gupta, 2017) and the **detection of financial risks** (Li, 2019).
Marketing	• The articles explored different applications of the BDA, including the **analysis of the pattern of customer spending**, **segmentation and customer profile** (Srivastava & Gopalkrishnan, 2015), **pattern analysis and diffusion of digital channels by customers** (Wang & Petrounias, 2017; Zhang *et al.*, 2018), **individualization of offers** (Lehrer *et al.*, 2018)**, improvement of the customer experience** (Indriasari, Gaol, & Matsuo, 2019), **customer prediction churn** (Shirazi & Mohammadi, 2019) and identification of **potentials customers** (Hung *et al.*, 2020).
Web and Social Media	• The articles used **web analytics** to provide real-time customer-centric solutions and solutions to unforeseen problems (Saxena & Al-Tamimi, 2017) and **social media** analytics to analyze customer perceptions of the services provided (Bratawisnu, Giri, & Rinaldi, 2017).
Technology and Data Analytics	• In this theme, the authors explored different perspectives in relation to **technologies** (Munar, Chiner, & Sales, 2014, Tian *et al.*, 2015; Loyd & Kannan, 2017; Khanchouch & Limam, 2018; Rakhman et al., 2019) or proposed approaches to BD modeling and analysis, such as GOMA, which refers to an objective-oriented approach to BD **modeling and analysis** (Supakkul, Zhao, & Chung, 2016) and stochastic linear programming (Dutta *et al.*, 2019).
Consequences of BDA Use	• Regarding the use of BDA, the articles explored issues related to **value generation** (Srivastava & Gopalkrishnan, 2015; Bhuvana, Thirumagal, & Vasantha, 2016; Srivastava, Singh, Tanwar, & Tyagi, 2017; Chedrawi, Atallah, & Osta, 2018; Elia *et al.*, 2019), the **skills** necessary for the use of BDA (Al-Hakimi, 2017), aspects related to **implementation** (Àngée, Lozano-Algiers, Montoya-Munera, Ospina-Arango, & Tabares-Betancur , 2018; Bürger, 2019; Ravi & Kamaruddin, 2017) and **decision-making** (Alzaidi, 2018).

Figure 5 - Themes applied to BDA in financial institutions

In the financial sector, **risk management** is a topic of interest to both researchers and government regulators (Li, 2019). In which the **detection of fraud** in transactions carried out by digital channels and by debit and

credit cards stands out. For example, Patil, Nemade, and Soni (2018) addressed the BDA for high-volume data processing and for the use of various machine learning algorithms to detect fraud in credit card use. The study by Li (2019) addressed the **detection of financial risks**, with the projection of an information system for risk management, using BDA, to perform real-time monitoring and dynamic assessment of risk sources.

BDA's significant role in **marketing** has been recognized among companies in various sectors, especially in the financial sector (Hung *et al*., 2020). In the study by Srivastava and Gopalkrishnan (2015) the authors analyzed the application of BDA in performing transactional and sentimental analysis of customers and, with this, identifying the **customer's spending pattern**, the use of service channels and indicating the cross-selling of products based on the profile to increase the rate of correct answers. Wang and Petrounias (2017) applied the BDA to analyze the influence of demographic factors on the use of digital channels by customers, indicating improvements to be made to increase the use of these channels. Lehrer *et al*. (2018) analyzed the implementation of BDA for the provision of individualized service to customers of companies from four sectors (insurance, banking, telecommunications, and e-commerce). Indriasari, Gaol and Matsuo (2019) analyzed the application of BDA in Indonesia's and global banks to improve the customer experience, such as providing more relevant, correct, and appropriate content to the customer. Shirazi and Mohammadi (2019) built a **customer's turnover predictive model** using the BDA, including structured file data, integrated with unstructured data from sources such as online web pages, number of website visits, and phone conversation logs. Hung *et al*. (2020) applied the BDA to analyze three types of supply chain relationships at a bank to **identify potential corporate customers**. The results of this study indicate that the approach can significantly increase a customer's response rate to the marketing campaign and improve the approval rate.

Web and social media analytics with BDA application can provide intelligence in customer analytics for the offering of products and services (Vassakis *et al*., 2018). In the research of Saxena and Al-Tamimi (2017) they used **data extracted from the websites** of the banks studied, internet banking and Facebook pages to analyze the opportunities for creating strategies for the use of products and services from BDA. Bratawisnu, Giri and Rinaldi (2017) applied BDA to make the association of consumer perceptions about the services provided in internet banking, using data obtained by tracking on **social media,** such as Twitter. Therefore, it demonstrates that the analysis of these data can improve the bank's image through customer perception on social media. In the theme **technology and data analytics,** research addresses issues related to the necessary architectures and technologies for the proper use of BDA, as well as analytics models that can be used by financial institutions. For example, Dutta *et al*. (2019) created a BDA architecture design, contemplating business rules and model for cross-selling analysis in a bank. Another example is the research by Supakkul, Zhao, & Chung (2016) in which they proposed an objective-oriented modeling approach to the BDA, using as an example a scenario of cancelling accounts of retail bank customers.

Regarding the theme **consequences of BDA use**, Chedrawi, Atallah and Osta (2018) analyzed the role of the BDA in terms of value generation, with the reduction of the transactional cost of the two main Lebanese banks. Al-Hakimi (2017) analyzed the factors that affect the **skills** related to the use of BDA and proposed self-

efficacy and the level of analysis as a framework for a successful career in the banking sector. Bürger (2019) proposed a roadmap that helps structure the **adoption of BDA** in all areas of a bank. Alzaidi (2018) analysed the impact of BDA on **decision-making**, indicating that banks can develop an effective decision-making structure, which can offer a high level of efficiency.

4.3. Research Agenda - BDA in Financial Institutions

While the use of BDA is seen to add business value and generate competitive advantage, organizations still face some challenges in achieving the expected benefits (Akter *et al.*, 2016). Appendix 2 relates the opportunities for future research presented in the analyzed articles, as well as research questions that can contribute theoretically and practically to the BDA research in financial institutions.

The growing number of transactions carried out by customers through digital means brings benefits to financial institutions, such as cost reduction and scope of performance. However, at the same time, there is the increasing number of cyber-attacks that result in fraud and financial losses for these institutions (Patil. Nemade, & Soni, 2018). Therefore, as the volume of transactions increases, so does the challenge of financial institutions to identify fraudulent events before they occur (Kamaruddin & Ravi, 2016) and effectively manage their financial risks (Li, 2019), thus highlighting the potential of BDA to work with these large volumes of real-time data for risk management.

In the marketing theme, Srivastava and Gopalkrishnan (2015) point out that BDA in banks assists in offering better services to customers, by their ability to analyze the pattern of spending, segmentation, and customer profile. Indriasari, Gaol and Matsuo (2019) point out that banks still face customer experience challenges and that BDA enables banks to focus their strategies on the customer based on data-orientation. Similarly, Bratawisnu, Giri and Rinaldi (2017) and Saxena and Al-Tamimi (2017) point out that banks have not explored all the opportunities and challenges of using data from the web and social media yet, and that in this sense there is a great potential of the BDA to provide customer-centric and real-time solutions, as well as anticipating problems in relation to them.

For Loyd and Kannan (2017) BDA is an emerging theme that is expanding in several areas, highlighting in this sense the challenge of aspects related to technology, such as the definition of the best architecture and analytics models that enable financial institutions to make the best use of the BDA. Bürger (2019) considers BDA a game changer for its operational and strategic potential and highlights the importance of its adoption throughout the company to achieve better performance and competitiveness. In this sense, the challenges of its use are linked to how financial institutions can successfully implement the BDA (Angée *et al*., 2018) and how to harness its potential for decision-making (Alzaidi, 2018). Another important challenge of the use of BDA is related to skills, ranging from knowledge of analysis to data-driven culture (Akter *et al.*, 2016). By improving skills, financial institutions will be able to improve their business and take more advantages of BDA (Al-Hakimi, 2017).

Finally, value generation is considered a relevant topic through BDA (Elia et *al*., 2019). Günther *et al*. (2017) reinforce that IS research on the generation of value from BD is still limited to some empirical studies. Similarly, Mikalef *et al*. (2019) highlight the substantial investments made by companies in BD and reinforce the importance of conducting further studies that empirically investigate all aspects related to BDA, especially in relation to value generation. Thus, we conclude the proposition of this research agenda on BDA in financial institutions, obtained through the consolidation of suggestions presented in the literature along with the suggestion of specific research questions for each theme. Next, we present the framework that summarizes the findings of this study.

4.4. Framework - BDA in Financial Institutions

From the conduct of this SLR it was possible to identify relevant research areas to the phenomenon of the study, to identify which research methodologies and approaches the analyzed articles used to achieve the objectives of their respective researches, as well as the indications of further studies, both from the methodological point of view and from the point of view of the identified themes. Thus, based on the concept of the BDA, the themes and subcategories of BDA application, and research agenda introduced previously, a framework is proposed that summarizes the findings of this study. The framework is presented in Figure 6.
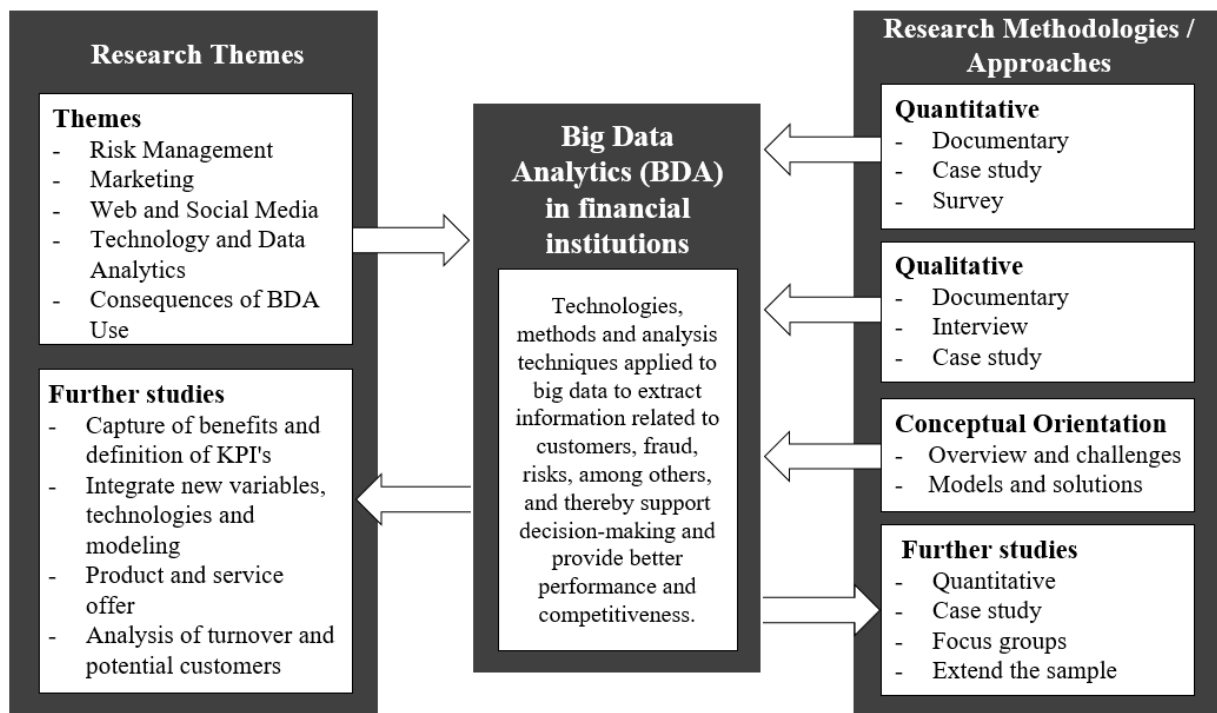

Figure 6 - Framework - BDA financial institutions

The framework highlights the themes of BDA application in financial institutions classified into five categories, which were used to inform a research agenda and highlight research areas that justify additional studies to develop necessary knowledge about BDA in the context of financial institutions. Thus, further studies

may use the information summarized here as a starting point for the study of BDA in the context of financial institutions.

5. CONCLUSION

With the opportunities created by the advance of information technology, financial institutions can exploit the potential benefits of BDA. Thus, the objective of this research was to analyze how scientific research approaches the BDA in the context of financial institutions. Through the SLR it was possible to characterize the studies on BDA in financial institutions, highlighting the conceptualization of BDA in this context, the themes addressed and the research opportunities. With the analysis of the sources were identified the most relevant themes on the application of BDA in financial institutions, namely: risk management, marketing, web and social media, technology and data analytics and consequences of BDA use. An agenda with research directions for further studies was also offered in order to contribute to the expansion of knowledge from studied themes, such as the use of BDA and its value generation, or, even, they can address topics little explored until then, such as risk management and marketing with regard to the application of BDA for credit and market analysis, among others. As a theoretical contribution, this research presents the identification of the specificities and the thematic areas of use of BDA and the proposal of a research agenda on BDA in the context of the financial institutions. As a practical contribution, this research offers the definition of the BDA applied to the context of financial institutions, and how this implies addressing the potential of its purpose and benefits, which can help organizations take advantage of this resource to generate results. It is also highlighted as a practical contribution, the identification of themes of BDA application in financial institutions. Thus, these organizations can evaluate the applications of the BDA to meet the needs within the various areas of the organization. As, for example, in risk management, which is considered an increasingly critical issue in financial institutions due to the increase in electronic transactions in the digital environment and the associated risks. As a limit of this research, it is highlighted that this study is limited to the method used, and, consequently, to the search terms, databases and articles analyzed. It is suggested that future studies may analyze other databases and other types of publications, where complementary information can be found. It is also suggested that future studies seek to answer the research questions presented here through qualitative or quantitative investigations, to contribute to the theory with empirical evidence on such problems.

REFERENCES

Akter, S., Wamba, S. F., Gunasekaran, A., Dubay, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics* (182), pp. 113-131.

Al-Hakimi, A.A.A. (2017). Big data skills required for successful application implementation in the banking sector. In: Mohamed Ali M. et *al.* (Eds) *Modeling, Design and Simulation of Systems*. Singapore: Springer.

Alzaidi, A. A. (2018). Impact of use of big data in decision making in banking sector of Saudi Arabia. *International Journal of Computer Science and Network Security* (18:12).

Angée, S., Lozano-Argel, S. I., Montoya-Munera, E. N., Ospina-Arango, J. D., & Tabares-Betancur, M. S. (2018). Towards an improved ASUM-DM process methodology for cross-disciplinary multi-organization big data & analytics projects. *Proceedings of International Conference on Knowledge Management in Organizations*. Springer, Cham.

Anfer, O. & Wamba, S. F. (2019). Big data analytics and strategic marketing capabilities: impact on firm performance. In: Rocha Á. et *al*. (Eds) *New Knowledge in Information Systems and Technologies*. Springer, Cham.

Anwar, M.; Khan, S. Z.; Shah, S. Z. A. (2018). Big data capabilities and firm's performance: a mediating role of competitive advantage. *Journal of Information & Knowledge Management* (17:4).

Balasupramanian, N., Ephrem, B. G., & Al-Barwani, I. S. (2017). User pattern based online fraud detection and prevention using Big Data Analytics and self-organizing maps. *Proceedings of International Conference on Intelligent Computing, Instrumentation and Control Technologies*, Kannur, Kerala, India.

Bhuvana, M.; Thirumagal, P. G.; Vasantha, S. (2016). Big Data Analytics - A leveraging technology for Indian commercial banks. *Indian Journal of Science and Technology* (9:32).

Bratawisnu, M. K., Giri, R. R. W, & Rinaldi, R. (2017). Association perception customer feedback with text network analysis in social media (case study on internet banking BRI, BCA, Mandiri in Indonesia). *Proceedings of International Conference on Information and Communication Technology*, Malacca City, Malacca, Malaysian.

Bürger, O. (2019). How to structure a company-wide adoption of Big Data Analytics. *Proceedings of International Conference on Wirtschaftsinformatik*, Siegen, Germany.

Capgemini Consulting. (2018). *Big Data Alchemy: How can Banks Maximize the Value of their Customer Data?* available (at: https://www.capgemini.com/wp-content/uploads/2017/07/bigdatainbanking_2705_v5_1.pdf.)

Chedrawi, C., Atallah, Y., & Osta, S. (2018). Big data in the banking sector from a transactional cost theory (TCT) perspective - the case of top Lebanese banks. In: Baghdadi Y., Harfouche A., Musso M. (Eds) *ICT for an Inclusive World*. Lecture Notes in Information Systems and Organisation, Springer, Cham.

Chen, D.Q., Preston, D.S., & Swink, M. (2015). How the use of big data analytics affects value creation in supply chain management. *Journal of Management Information Systems* (32:4), pp. 4-39.

Côrte-Real, N., Ruivo, P., Oliveira, T., & Popovic. Al. (2019). Unlocking the drivers of big data analytics value in firms. *Journal of Business Research* (97), pp. 160-173.

Creswell, J.W. and Creswell, J.D. (2017). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, SAGE Publications, Thousand Oaks, CA.

Cusumano, M. A., Yoffie, D. B., Gawer, A. (2020). *The future of platforms*. available (at: https://sloanreview.mit.edu/download-disruption-2020-series/)

Dong, J. Q. & Yang, C. H. (2020). Business value of big data analytics: A systems-theoretic approach and empirical test. *Information & Management* (57), pp. 1-9.

Dremel, C., Herterich, M. M., Wulf, J., & vom Brocke, J. (2020). Actualizing big data analytics affordances: A revelatory case study. *Information & Management* (57), pp. 1-21.

Dubey, R. et *al*. (2019). Big data analytics and organizational culture as complements to swift trust and collaborative performance in the humanitarian supply chain. *International Journal of Production Economics* (120), pp. 120-136.

Dutta, G., Rao, H. V., Basu, S., & Tiwari, M. K. (2019). Asset liability management model with decision support system for life insurance companies: computational results. *Computers & Industrial Engineering* (128), pp. 985-998.

Elhoseny, M., Hassan, M. K, & Sing, A. K. (2020). Special issue on cognitive big data analytics for business intelligence applications: Towards performance improvement. *International Journal of Information Management* (50), pp. 413-415.

Elia, G., Polimeno, G., Solazzo, G., & Passiante, G. (2019). A multi-dimension framework for value creation through big data. *Industrial Marketing Management*.

Gartner Inc. (2018). *Adopt SMART Information Principles for Effective Data and Analytics Governance*. Available (at: https://www.gartner.com/doc/3881670/adopt-smart-information-principles-effective)

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

Günther, W. A., Mehrizi, M. H. R., Huysman, M., & Feldberg, F. (2017). Debating big data: a literature review on realizing value from big data. *Journal of Strategic Information Systems*.

Hung, J. L., He, W., & Shen, J. (2020). Big data analytics for supply chain relationship in banking. *Industrial Marketing Management* (86), pp. 144-153.

Indriasari, E, Gaol, F. L., & Matsuo, T. (2019). Digital banking transformation: application of artificial intelligence and big data analytics for leveraging customer experience in the Indonesia banking sector. *Proceedings of International Congress on Advanced Applied Informatics*, Toyama, Japan.

Khan, M. (2019). Challenges with big data analytics in service supply chains in the UAE. *Management Decision*.

Kamaruddin, S. & Ravi, V. (2016). Credit card fraud detection using Big Data Analytics: use of PSOAANN based one-class classification. *Proceedings of International Conference on Informatics and Analytics*, Pondicherry, India.

Khanchouch, I. & Limam, M. (2018). Adapting a Multi-SOM clustering algorithm to large banking data. *In*: Rocha Á., Adeli H., Reis L.P., Costanzo S. (Eds) *Trends and Advances in Information Systems and Technologies*. Springer, Cham.

Lehrer, C., Wieneke, A., vom Brocke, J., Jung, R., & Seidel, S. (2018). How Big Data Analytics Enables Service Innovation: Materiality, Affordance, and the Individualization of Service. *Journal of Management Information Systems*.

Li, Y. (2019). Design a management information system for financial risk control. *Cluster Computing* (22), pp. S8783–S8791.

Lin, J. L., Hou, T. C. T., & Hsio, Y. L. (2019). Introduction to the special issue of "Big data analytics: Using financial and non-financial information. *Asia Pacific Management Review* (24), pp. 10.

Loyd, B. D. & Kannan, D. Identifying design patterns for risk management system using Big Data Analytics. *Proceedings of International Conference on Trends in Electronics and Informatics (ICEI)*, Tirunelveli, India.

Maroufkhani, P. Wagner, R., Ismail, K. W. K, Baroto, M. B., & Nourani, M. (2019). Big data analytics and firm performance: a systematic review. *Information* (10:7), pp. 2-21.

Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2017). Big data analytics capabilities: A systematic literature review and research agenda. *Information Systems and e-Business Management*, pp.1-32.

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), pp. 261-276.

Munar, A., Chiner, E., & Sales, I. A big data financial information management architecture for global banking. *Proceedings of International Conference on Future Internet of Things and Cloud*, Barcelona, Espanha.

Patil, S., Nemade, V., & Soni, P. (2018). Predictive modelling for credit card fraud detection using data analytics. *Procedia Computer Science* (132), pp. 385-395.

Rakhman, R. A., Widiastuti, R. Y., Legowo, N., & Kaburuan, E. R. (2019). Big data analytics implementation in banking industry –case study cross selling activity in Indonesia's commercial bank. *International Journal of Scientific & Technology Research* (8:9).

Ravi V. & Kamaruddin S. (2017). Big data analytics enabled smart financial services: opportunities and challenges. *In*: Reddy P., Sureka A., Chakravarthy S., Bhalla S. (Eds) *Big Data Analytics*. Springer, Cham.

Sagar, R. K. & Gupta, Y. (2017). Demonetization: How big data analytics can unearth the black money of India. *Proceedings of International Conference on Intelligent Computing and Control Systems*, Madurai, India.

Saxena, S. & Al-Tamimi, T. A. S. M. (2017). Big data and internet of things (IoT) technologies in Omani banks: a case study. *Foresight* (19:4), pp. 409-420.

Shirazi, F. & Mohammadi, M. (2019). A big data analytics model for customer churn prediction in the retiree segment. *International Journal of Information Management* (48), pp. 238-253.

Srivastava, A., Singh, S. K., Tanwar, S., & Tyagi, S. (2017). Suitability of big data analytics in Indian banking sector to increase revenue and profitability. *Proceedings of International Conference on Advances in Computing, Communication & Automation*, Dehradun, India.

Srivastava, U. & Gopalkrishnan, S. (2015). Impact of Big Data Analytics on Banking Sector: Learning for Indian Banks. *Procedia Computer Science* (50), pp. 643-652.

Supakkul, S., Zhao, L., & Chung, L. (2016). GOMA: supporting big data analytics with a goal-oriented approach. *Proceedings of IEEE International Congress on Big Data.*

Tian, X., Han, R., Wang, L., Lu, G., & Zhan, J. (2015). Latency critical big data computing in finance. *The Journal of Finance and Data Science* (1), pp. 33-41.

Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data' can make big impact: findings froma systematic review and alongitudinal case study. *International Journal Production Economics* (165), pp. 234-246.

Wang, S. & Petrounias, I. (2017). Big data analysis on demographic characteristics of chinese mobile banking users. *Proceedings of IEEE Conference on Business Informatics*, Thessaloniki, Greece.

Webster, J. & Watsom, R.T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly* (26:2), pp.13-23.

Wolfswinkel, J. F., Furtmueller, E., & Wilderom, C.P. M. (2013). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems* (22), pp. 45–55.

Vassakis, K., Petrakis, E., & Kopanakis I. (2018). *Big data analytics: applications, prospects and challenges*. In: Mobile big data. Springer.

Zhang, J. M., Zhu, S. Z., Yan, W., & Li, Z. P. (2018). The construction and simulation of internet financial product diffusion model based on complex network and consumer decision-making mechanism. *Information Systems and e-Business Management*.

APPENDIX "A" - ANALYSIS OF THE TIMELINE OF PUBLICATIONS ON THE TOPIC

Year	SC[1]	SJ[2]	TT[3]	N	Authors	Method	Technique	Journal/Conference
2014	1	-	1	1	Munar, Chiner e Sales	Quantitative	Case study	Int. Conf. on Future Internet of Things and Cloud (FiCloud)
2015	1	1	2	2	Srivastava e Gopalkrishnan	Quantitative	Documentary	Int. Symp. on Big Data and Cloud Computing (ISBCC)
				3	Tian *et al.*	Theoretical		The Journal of Finance and Data Science
2016	2	1	3	4	Bhuvana, Thirumagal, & Vasantha	Theoretical		Indian Journal of Science and Technology
				5	Kamaruddin e Ravi	Theoretical		Int. Conf. on Informatics and Analytics (ICIA)
				6	Supakkul, Zhao e Chung	Quantitative	Case study	Int. Cong. on Big Data (Big Data Congress)
2017	8	1	9	7	Al-Hakimi	Quantitative	Survey	Asian Simulation Conference (AsiaSim)
				8	Balasupramanian, Ephrem, & Al-Barwani	Quantitative	Documentary	Int. Conf. on Intelligent Comp., Instrumentation and Control Tech. (ICICICT)
				9	Bratawisnu, Giri & Rinaldi	Quantitative	Documentary	Int. Conf. on Information and Communication Tech. (ICoICT)
				10	Loyd e Kannan	Theoretical		Int. Conf. on Big Data Analytics (BDA)
				11	Ravi e Kamaruddin	Theoretical		Int. Conf. on Big Data Analytics (BDA)
				12	Sagar e Gupta	Theoretical		Int. Conf. on Intelligent Comp. and Control Systems (ICICCS)
				13	Saxena e Al-Tamimi	Quantitative	Documentary	Foresight
				14	Srivastava *et al.*	Theoretical		Int. Conf. on Adv. in Comp., Com. & Automation (ICACCA)
				15	Wang e Petrounias	Quantitative	Survey	Conf. on Business Informatics (CBI)
2018	4	3	7	16	Alzaidi	Quantitative	Survey	Int. Journal of Computer Science and Network Security
				17	Angée *et al.*	Qualitative	Case study	Int. Conf. on Knowledge Management in Organizations (KMO)
				18	Chedrawi, Atallah, & Osta	Qualitative	Interview	Information and Com. Tech. in Organ. and Society (ICTO)
				19	Khanchouch e Limam	Quantitative	Documentary	World Conf. on Information Systems and Tech. (WorldCist)
				20	Lehrer *et al.*	Qualitative	Case study	Journal of Management Information Systems
				21	Patil, Nemade e Soni	Quantitative	Documentary	Int. Conf. on Comp. Intelligence and Data Science (ICCIDS)
				22	Zhang *et al.*	Quantitative	Documentary	Information Systems and e-Business Management
2019	2	5	7	23	Bürger	Qualitative	Case study	Int. Conf. on Wirtschaftsinformatik (WI)
				24	Dutta *et al.*	Quantitative	Documentary	Computers & Industrial Engineering
				25	Elia *et al.*	Qualitative	Case study	Industrial Marketing Management
				26	Indriasari *et al.*	Qualitative	Case study	Int. Cong. on Advanced Applied Informatics (IIAI-AAI)
				27	Li	Quantitative	Documentary	Cluster Computing
				28	Rakhman *et al.*	Quantitative	Case study	Int. Journal of Scientific & Technology Research
				29	Shirazi e Mohammadi	Quantitative	Documentary	Int. Journal of Information Management
2020	-	1	1	30	Hung *et al.*	Quantitative	Case study	Industrial Marketing Management

Notes: [1]Scientific Conference (SC); [2]Scientific Journal (SJ); [3]Total (TT).

20.ª Conferência da Associação Portuguesa de Sistemas de Informação (CAPSI'2020)   15
16 e 17 de outubro de 2020, Porto, Portugal
ISSN 2183-489X

APPENDIX "B"- RESEARCH AGENDA ON BDA IN FINANCIAL INSTITUTIONS

Theme	Suggestions identified in the reviewed literature	Research questions proposed based on SLR
Risk Management	• **Fraud Detection**: Extend and compare the proposed model with other machine learning tools (Kamaruddin & Ravi, 2016). Adjust the proposed model related to the overfitting of the decision tree and the detection of transactions in real time (Patil, Nemade & Soni, 2018).	How to apply machine learning approaches in the big data paradigm to detect fraud in bank transactions?
	• **Detection of financial risks**: no indication of future studies in the reviewed literature.	What BDA approaches are adopted by banks to detect financial risks?
Marketing	• **Analysis of customer spending patterns, segmentation, and customer profile**: Extend the analysis to cover the various data mining techniques that can be used by banks to improve the quality of the analysis (Srivastava & Gopalkrishnan, 2015).	How do financial institutions use BDA to identify offers of banking products and services that generate the most profit and which customer segments are most attractive?
	• **Pattern analysis and diffusion of digital channels by customers**: Add demographic characteristics, other functions for the mobile application and users' perception of these functions (Lehrer *et al.*, 2018).	Which BDA approaches are adopted by banks for the analysis of the pattern and diffusion of digital channels by customers?
	• **Individualization of offers**: Investigate additional uses of BDA based on comparable samples (Shirazi & Mohammadi, 2019).	How can banks use BDA to individualize offers to customers?
	• **Customer Experience**: Expand the sample to other banks (Indriasari *et al.*, 2019).	How does BDA influence strategies relate to customer experience in financial institutions?
	• **Customer prediction churn**: Consider sensitivity analysis to explore the correct receptive point of customers (Shirazi & Mohammadi, 2019).	What analyzes can be used by financial institutions to forecast customer turnover?
	• **Potential customers**: Consider other customer filtering criteria in the analysis (Hung *et al.*, 2020).	What criteria and analyzes can be carried out to identify potential customers in financial institutions?
Web and Social Medial	• **Web analytics**: no indication of future studies in the reviewed literature.	How can web analytics using BDA identify the perceptions of bank customers?
	• **Analysis of social media**: no indication of future studies in the reviewed literature.	What BDA approaches can be used to analyze social media in financial institutions?
Technology and Data Analytics	• **Technology**: Develop a data management system that supports different types of big data analytics and uses memory and other hardware resources effectively to provide high computing speed (Tian *et al.*, 2015). Apply the integrated cluster approach to other Hadoop tools, such as Flume and HBase, and compare clustering methods with the multi-SOM method using Hadoop and R (Khanchouch & Limam, 2018).	How can financial institutions leverage big data sources and cloud-based architecture to generate insight and business value?
	• **Modeling and analysis**: Conducting case studies to validate the benefits of the model (Supakkul, Zhao, & Chung, 2016). Extend the analysis using a statistical distribution different from that used in the research for the scenario generation activity (Dutta *et al.*, 2019).	Which BDA modeling and analysis approaches can be used to generate scenarios to support banks in decision-making?
Consequences of BDA Use	• **Value generation**: Quantify financial and non-financial benefits after the implementation of BDA and predict improvements in the bank's financial statements (Srivastava & Gopalkrishnan, 2015). Estender a pesquisa no mesmo contexto em outros bancos (Chedrawi, Atallah, & Osta, 2018). Extend the search in the same context in other banks (Chedrawi, Atallah, & Osta, 2018).	What is the role of BDA in the performance systems of financial institutions?
	• **Skills**: no indication of future studies in the reviewed literature.	How can financial institutions develop data literacy?
	• **Implementation**: Extend the research with the use of focus groups and projects like those used in the study (Angée *et al.*, 2018). Conducting additional case studies, considering holistic models of BDA capabilities, and considering research into successful and failed BDA adoption projects (Bürger, 2019).	What are the critical success and failure factors in the adoption of BDA in financial institutions?
	• **Decision-making**: no indication of future studies in the reviewed literature.	How can financial institutions develop a data-driven culture using BDA?