CONSEQUENTIAL FACTORS OF BIG DATA'S ANALYTICS CAPABILITY FOR DIGITAL TRANSFORMATION

Artigo Científico Completo

Luis Hernan Contreras Pinochet	Guilherme de Camargo Belli Amorim	Durval Lucas dos Santos Júnior	Cesar Alexandre de Souza
Universidade Federal de São Paulo (UNIFESP)	*Universidade Federal de São Paulo (UNIFESP)*	*Universidade Federal de São Paulo (UNIFESP)*	*Universidade de São Paulo (USP)*
luis.hernan@unifesp.br	*amorim_camargo@hotmail.com*	*durval.lucas@unifesp.br*	*calesou@usp.br*

Abstract

The article aims to discover, by identifying the main models of its use - descriptive, diagnostic, predictive and prescriptive - the consequent factors of Big Data's Analytics Capability in companies that seek Digital Transformation. The research had a quantitative approach, using a survey of data from companies located in the state of São Paulo - Brazil. Structural Equation Modeling was used to validate the model. As a result, all hypotheses with direct and indirect relationships were accepted. Satisfaction was the construct that had the most explanatory power in the model. This study revealed that the transformation digital that occurred in companies is not limited to the area of technologies and systems, but also in the way professionals are being inserted in this context. It is necessary to invest more in analytical tools, as well as people trained in analysis of these models, in addition to a change of mindset, which will dictate the bias of the company's strategic decision making. It will also be necessary to invest in professional training, so that those involved can better handle Big Data tools to increase their influence in the organizational decision-making process.

FATORES CONSEQUENTES DA CAPACIDADE ANALÍTICA DE BIG DATA PARA TRANSFORMAÇÃO DIGITAL

Resumo

O artigo tem como objetivo investigar por meio da identificação dos principais modelos de sua utilização – descritiva, diagnóstica, preditivos e prescritivos – os fatores consequentes da Capacidade Analítica de Big Data em empresas que buscam a Transformação Digital. A pesquisa teve abordagem quantitativa, por meio de levantamento de dados junto a empresas localizadas no estado de São Paulo – Brasil. A Modelagem de Equações Estruturais foi usada para validar o modelo. Como resultado, todas as hipóteses com relações diretas e indiretas foram aceitas. Satisfação foi o construto que teve maior poder explicativo no modelo. Este estudo revelou que a transformação digital ocorrida nas empresas não se limita à área de tecnologia e sistemas, mas também na forma como os profissionais estão sendo inseridos nesse contexto. É necessário investir mais em ferramentas analíticas, bem como em pessoas treinadas na análise desses modelos, além de uma mudança de mentalidade, que ditará o viés da tomada de decisão estratégica da empresa. Também será necessário investir na capacitação profissional, para que os envolvidos manejem melhor as ferramentas de Big Data para aumentar sua influência no processo de tomada de decisão organizacional.

1 Introduction

In a world with continuous technological changes, innovation is a determining factor for the economic growth and survival of companies. Innovation has the ability to predict which technologies could potentially change an industry, create markets and increase accessibility and viability of the business (Nwankpa & Roumani, 2016). The organization must carefully analyze the environment to detect changes, disseminate knowledge and adjust according to what will create value for the customer. This implies not only adopting but managing new technologies and their results (Warner & Wäger, 2019), in a process known as Digital Transformation.

When considering Digital Transformation as a strategy for business development, there is high scientific relevance, and it has the potential to be revolutionary or even disruptive (Bhimani & Willcocks, 2014). Thus, considering Digital Transformation and its context, in which organizations are increasingly under pressure to have a quick response and constant adaptation related to a highly connected and dynamic global environment, it is essential that analytical skills are developed so that the results are a combination of management, learning and process reconfiguration processes.

The application of analytical science to Big Data can be defined as a set of methodologies used to capture, store and process a large volume of information from various sources (structured and unstructured data) to accelerate decision making and generate competitive advantages for the organization. Rapid data growth is due to advances in digital sensors, communications, computing and storage that created huge data sets (Halevi & Moed, 2012; Kacfah Emani et al., 2015).

From the management perspective, the relevance of Big Data tools in organizational restructuring processes that require a "new process" of organizations is clear; demanding a digital performance, from logistics to communication. Additionally, organizations interested in using Big Data tools in their business should initially define what their goals are within a process, and these goals should be aligned with those of the organization itself. This guarantees that the choice of tools, as well as its attributes, is the most suitable for its analysis (Subramaniam et al., 2019).

This article aims to identify the consequent factors of Big Data Analytical Capacity in companies that seek Digital Transformation by identifying the main models of use (descriptive, diagnostic, predictive and prescriptive). To this end, in this quantitative research we conducted a survey (n=141) in companies located in the state of São Paulo - Brazil. To validate the model, we used Structural Equation Modeling (SEM).

The contribution of this article is to present a proposal to expand the research model based on the studies of the model by Hooi, Abu and Raim (2018) that analyzed the relationship between "Big Data Analytics Capability" and "Product Innovation Performance". The authors scrutinized the studies by Hannachi (2015) and Gunday, Ulusoy, Kilic and Alpkan (2011) to improve the construct "Big Data Analytics Capability"; the studies by Tippins and Sohi (2003) and Wang et al. (2012) for the "Product Innovation Performance" construct, and for the "Process-oriented dynamic capabilities" construct formulated from Kim et al. (2011). Furthermore, the authors examined the relationship between Big Data Analytical Capacity and the constructs "Satisfaction" and "Business Value", as observed in the literature by Akter et al. (2017).

2 LITERATURE REVIEW

2.1 Digital Transformation as a Strategy Digital

Transformation can be defined as a set of changes that are leveraged and designed based on technologies. For companies, digital transformation is understood as an organizational change based on the application of Big Data, analytics, cloud, mobile and social media platforms. The contemporary focus is for organizations to go digital and to generate value for consumers through personalized services (Schallmo, Williams & Boardman, 2017), and to create value for consumers in the digital age.

Being digital is not just about using digital technologies, rather this transformation requires the organization to cultivate an agile and entrepreneurial mindset (Day & Schoemaker, 2016). The digital transformation then becomes a path of no return, providing competitive advantages to organizations with plans aimed at accomplishing this (Nwankpa & Roumani, 2016).

The changes taking place in the market, in relation to digital technologies, are more comprehensive than products, business processes, sales channels or supply chains, as entire business models are being redesigned and constantly being deposed. The process of becoming a digital organization requires a transformation or reconfiguration, changing its business models through the use of new technologies.

In this sense, integrating and exploring new digital technologies is one of the biggest challenges that companies face in recent years (Hess et al., 2016). Every sector or company will feel the effects of digital transformation, and the mishandling of digital transformation can lead a company to failure (Ananyin et al., 2018; Cohen & Tripsas, 2018; Day & Schoemaker, 2016; Eggers & Park, 2018).

In view of this, analytics capability leads organizations to transform their routines for quick adaptation to changes in the external environment. They are organization-specific capability and can be a source of advantage in changing environments. They relate to the organization's ability to integrate, build and reconfigure internal and external competencies to adapt to rapidly changing environments (Andreeva & Chaika, 2006).

Therefore, the increased importance of innovation related to transformation in the current organizational environment is associated with the nature of Digital Transformation as a strategy. Technological innovations are elements of this continuous process and, due to the predominant patterns of imitation and invention, they have become the center of attention.

2.2 Big Data in business and Product Innovation Performance

Initiatives to transform Big Data into business indicators and business performance forecasts have revealed their difficulties (Xu et al., 2016). The internet was one of the great drivers of the creation of the term Big Data (Hilbert & Lopez, 2011). Big Data can be understood as the set of information that cannot be analyzed using traditional processes and tools.

This means that companies are increasingly seeing the need to deal with large volumes of useful data and information. The primary purpose is to benefit from the analysis of structured or unstructured data, which are continually emerging due to new devices, as well as social networks and corporate information systems (Kwon, Lee & Shinb, 2014). Considering that the main proposal of Big Data is to promote the inclusion of information from the web in the context of data analysis of the institutions, consequently managers start to define business strategies using data that evolve in real time in the digital environment (Hilbert & Lopez, 2011).

Companies can perform real-time monitoring based on data from a business performance dashboard resulting from large blocks of data in the initial entry of Big Data analysis. Big Data has become a trend in this era of data disruption, with an increasing number of applications across a range of varied businesses. With more people maintaining most of their lives online and contributing to word-of-mouth through online, Big Data Analytics is becoming an important agenda for companies when they consider launching new products.

Big Data is the product, but also a facilitator of this futuristic regime, which is incorporated in Hal Varian's notion of "nowcasting". The ways in which Big Data forms and transforms various arenas of human activity vary depending on the character of the activity in question (Ekbia & Suri, 2015).

Using Big Data techniques and technologies, we can now make more accurate predictions and potentially better decisions when dealing with health epidemics, natural disasters or social disturbances. However, at the same time, it is imperative to note the opacity of data-based approaches to science, social policy, cultural development, financial forecasting, advertising and marketing (Kallinikos, 2013). Technological advancement enthusiasts refuse to recognize this side of the phenomenon, whether they are business consultants, policy consultants, media gurus or scientists who design attractive visualization techniques. The right approach, in response to this unbridled enthusiasm, is not to deny the Big Data light side, but to devise techniques that balance human judgment and the technological capacity to support.

It is in this context that the diversity and heterogeneity of perspectives between theorists and professionals on the Big Data phenomenon is observed. This starts with the definitions and conceptualizations of the phenomenon. Looking at Big Data as a product, a process or a new phenomenon that challenges human cognitive abilities, professionals come to several conclusions about what the main problems are, what solutions are available and where the focus should be. When dealing with the cognitive challenge, some professionals look for the solution focusing on what to exclude (Waldrop, 2008), while others, observing the increasing capacity of computers storage at decreasing costs, suggest a different strategy: buy more space storage and maintain everything (Kraska, 2013).

The key benefits of Big Data technology for management are reduced costs and increased revenues; increased operational efficiency; improvement in decision making; improvement of products and services; and improvement in the processes of innovation and development of new products and markets (Leeflang et al., 2014). Therefore, organizations must prepare to generate business formed on the analysis of Big Data based on the information that, becomes assets of strategic value for organizations (Chen, Chiang & Storey, 2014).

Innovation in products and services results in changes to existing ones or the creation of new ones. This can often be achieved by offering a new central benefit or developing a new way of providing a central benefit (Berry et al., 2006). Product and service innovation takes multiple forms in various sectors (Miles, 2008) and is evident in traditional service industries and manufacturing industries (Baines et al., 2009). This sometimes results from the use of specific methods and sometimes requires organizational and cultural changes (Rothenberg, 2007). Innovation in sustainable products and services requires continuous interactions of development, operations and improvement. Stimulating innovation in products and services is a timely research topic that presents a big gap between importance and current knowledge, especially when we observe central phenomena in emerging technologies such as Big Data (Ostrom et al., 2015).

2.3 Big Data Analytics Capability

The amount of data generated in the market is increasing. Therefore, it is increasingly necessary that companies have some tools developed to absorb this change. The analytical aspect of Big Data occurs by combining the ability to work with data - from different sources, structured or unstructured, in huge quantities, in real time - with the possibility of developing predictive statistical models, conceived through information from Big Data. Big Data analysis is performed to support decision making based on information about the user's context and the planning, management and learning for the systematic use of data. There are models of Descriptive, Diagnostic, Predictive and Prescriptive analysis, in which each analysis capacity has separate services and purposes (Gim et al., 2018).

Descriptive analysis is this type of analysis is used to assist in the display of data, usually at the beginning of a study, to summarize the information and explore the behavior of the data, helping to understand the frequency of certain information being analyzed. This is a set of technologies and processes to understand and analyze utility results using data (Song et al., 2014). It is the most basic process when working with a database.

Diagnostic analysis is done to understand the purpose of the work itself. The reason, who should consume certain products, when and where, are the types of questions that this type of analysis proposes to clarify. It also helps to assess the impact and scope of any action taken. It is a way to assist in the development of strategies to drive the expected result, or to avoid surprises (Gim et al., 2018).

Predictive analysis is this type of analysis is used based on historical data, or known trends, thus making it possible to see future trends. It is also used when analyzed in conjunction with the context in which that data was entered, consequently being able to predict their behavior in a similar scenario. This method predicts results based on data entry using a variety of statistical and mathematical techniques. This method analyzes a variety of relationships based on predictions and suggests a model for maximizing business results (Song et al., 2014).

Prescriptive analysis is this technique helps to better understand what the possible results would be when making a decision. It helps to direct and seeks to provide alternatives for the purpose to be achieved. While Predictive Analysis helps show how things are likely to happen, Prescriptive Analysis recommends what actions should be taken to avoid a particular situation. This approach determines several alternatives and guides based on the results extracted from the descriptive and predictive analysis (Song et al., 2014). Prescriptive analysis, as a type of data analysis technique, provides predictions and personalized context information. This technique is used to support more effective decision making based on various ideas when business decision makers, such as CTOs and CEOs, analyze and predict complex situations (Gim et al., 2018).

3 RESEARCH HYPOTHESIS AND RESEARCH MODEL

Big Data Analytics Capability is recognized as the competence to provide business insights using the capacity of data management, infrastructure (technology) and talent (personnel) to transform a business into a competitive force. Big Data Analytics Capability was selected to understand strategic integration by applying the capabilities of this technology's features to shape and support business strategy. Big Data Analytics Capability can be understood by Gunday, Ulusoy, Kilic and Alpkan (2011), Hannachi (2015) and Hooi, Abu and Raim (2018) as:

- Big Data Operations: the company's ability to identify sources where large volumes of various types of data flow at high speed, collecting, storing and analyzing Big Data to meet the company's strategic and operational goals.
- Updating IT infrastructure: applications, hardware, data and networks to enable Big Data Analytics Capability staff to rapidly develop, deploy and support the necessary system components for a company.
- Advanced Analytics: professional capacity of the Big Data Analytics Capability team, including, skills or knowledge to perform assigned tasks.
- Strategic Uses of Big Data: the ability to manipulate routines in a structured (rather than ad hoc) way to manage IT resources according to the needs and priorities of the business.

Product innovation may be the result of Big Data Analytics Capability. In the case of Product Innovation Performance, the measures for the subjective Product Innovation Performance were mostly adapted from Im, Bayus and Mason (2003). This makes it possible to understand the management's perception of the market performance of new products, through the company's performance analysis, facing the market, in three dimensions that were analyzed in this research: Market Performance; Business Strategy Alignment; and Big Data Operations. The following hypothesis was elaborated:

> **Hypothesis 1a:** Big Data Analytics Capability positively influences Product Innovation Performance.

Process-oriented dynamic capabilities help to understand how the company is facing methodologies and business processes, compared to the competition. Product development is an example of a dynamic resource oriented to processes and includes the basic elements of strategizing (practice, process and practitioners). Together with learning and isomorphic mechanisms present in the organizational field, they explain how the product development process occurs (Goes, 2014; McAfee & Brynjolfsson, 2012). We see Big Data Analytics Capability as an important organizational capacity, leading to a sustainable competitive advantage in the Big Data environment. It is necessary for organizations to have specialists with the knowledge and experience of working in a competitive and highly dynamic market, which requires the adoption of Big Data as a way to remain competitive (Eisenhardt & Martin, 2000; Helfat et al., 2007; Wamba et al., 2017), so the following hypothesis was elaborated:

Hypothesis 1b: Big Data Analytics Capability positively influences Process-oriented dynamic capabilities.

Big Data Analytics Capability evaluation results in an affective or emotional response, such as Business Value and Satisfaction. In view of this, the existing literature identifies that the analytical excellence of Big Data has a significant positive impact on the Business Value, which ends up leading to the Satisfaction of users of Big Data analysis. The literature suggests that the greater the perception of the novelty of products and services on the market, the greater the potential value for users involved in the company (Nambisan et al., 2017). Therefore, Digital Transformation can bring many benefits to the company, among them, a closer relationship with the client, in addition to the development of its analytics capability (Chen, Mao & Liu, 2014). Therefore, the significance of the association between Big Data Analytics Capability and Business Valor was highlighted by the existing literature (Lavalle et al., 2011; Wixom et al., 2013). Therefore, the following hypotheses were formulated:

Hypothesis 1c: Big Data Analytics Capability positively influences Satisfaction.

Hypothesis 1d: Big Data Analytics Capability positively influences Business Value.

Competitive advantage is achieved by deploying and using distinct, valuable and inimitable resources and capabilities (Bhatt & Grover, 2005). The concept of Information Technology capacity is based on the assumption that, although resources can be easily replicated, a distinct set of resources mobilized by a company is not easy to replicate and will lead to competitive advantages (Santhanam & Hartono, 2003).

As part of the incorporation of higher-order Big Data Analytics Capability into a causal network, the study models it with criteria variables, such as Business Value and Satisfaction (Akter et al., 2017). We define Satisfaction as the general attitude response of Big Data analysts towards Big Data Analytics Capability. Value is seen as the degree of perceived benefits for the organization at a strategic level, such as competitive advantage. The impact of this analytics capability on the Business Value is a dominant concern in the Big Data environment and Business Value is one of the central items of Big Data, as it is directly influenced by its Analytical Capacity (Wixom et al., 2013). Business Value is identified as a mediator in the study because, first, Big Data Analytics Capability influences Business Value; next Business Value influences Satisfaction and, finally, the Big Data Analytics Capability influences Satisfaction (Baron & Kenny, 1986). In this sense, the following hypotheses were elaborated:

Hypothesis 1e: Big Data Analytics Capability positively influences Product Innovation Performance when it is mediated by Process-oriented dynamic capabilities.

Hypothesis 1f: Big Data Analytics Capability positively influences Satisfaction when it is mediated by Business Value.

Hypothesis 2: Process-oriented dynamic capabilities influence Product Innovation Performance.

Hypothesis 3: Business Value positively influences Satisfaction.

The elaborated research model measures the opportunities in the Big Data Analytics Capability in business and is illustrated in Figure 1. The circles represent the latent variables and the arrows, the causal relationships - hypotheses:

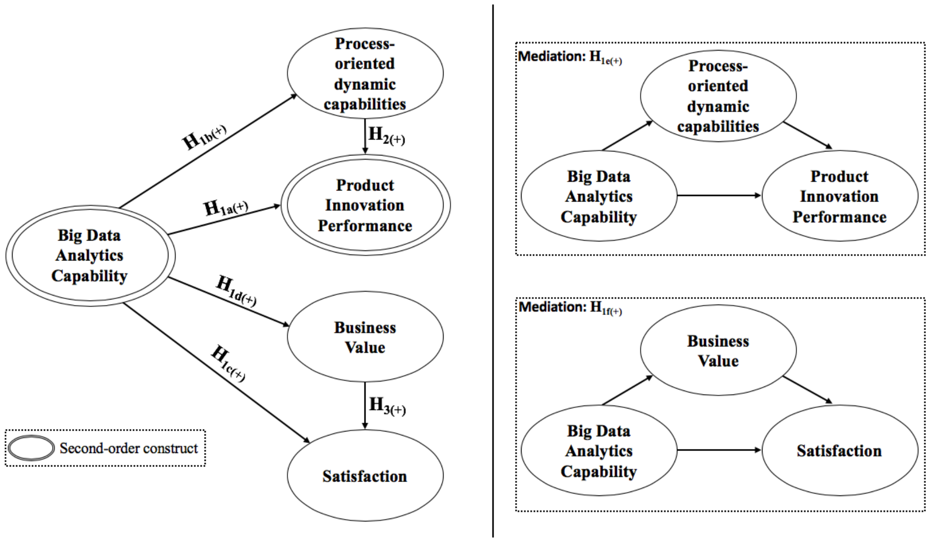

Figure 1: Proposed research model
Source: authors' elaboration.

4 METHOD

The research was conducted with individuals classified as managers and professionals who are in highly specialized positions with Big Data technologies and data analysis. The companies were identified through a database obtained by the Higher Education Institution through a partnership. The representation of the sample was considered adequate due to the analyzed position being predominantly analytical than executive. All selected companies are located in the state of São Paulo, Brazil. In addition, it is important to inform that the industry in São Paulo is the most modern and diversified in Latin America and is supported by a robust technological base, generating products with high added value in various economic sectors. In sum, the State of São Paulo concentrates 36% of Brazilian industrial production.

For this study, a pre-test was performed with 25 individuals to verify there was an understanding of the research instrument (Hair et al., 2010). The questionnaire was made available by email and by LinkedIn, to facilitate access to managers and professionals from the companies that participated in this research. Around 450 companies were contacted to participate in the survey. However, a total of 141 valid questionnaires were obtained, which were analyzed in IBM SPSS and SmartPLS. In this research, the G*Power 3.1.9 software was used to calculate the test power (1-b err prob) is equal 98.78%.

The research approach was quantitative, using an online questionnaire with closed-ended questions. Two management academics were consulted to refine the questionnaire and validate the adaptation of the research instrument to the Brazilian context. Based on their feedback, some items were reformulated. Through operationalization, the research variables were reduced to constructs that can be observed and measured. The method of data collection was convenience sampling, a non-probabilistic technique, a fact that limits the generalization of research results. To measure each item of the constructs, the Likert scale was used with end points anchored in "totally disagree" (1) and "totally agree" (5) for all 70 statements. For aspects of the characterization of the demographic profile and organizations, specific objective questions were elaborated.

5 ANALYSIS OF RESULTS

5.1 Profile of respondents

The sample had a majority of male individuals (n=104/73.8%). With regard to age, the largest number of respondents in the survey is concentrated up to 35 years (n=83/58.9%), and respondents were characterized by a sample of managers (45.4%) and professionals (54.6%).

With regard to education, one third of the respondents have completed graduate studies (n=47/33.3%) and are working in middle and senior management positions (n=116/82.2%). The average length of time working in the organization is $\bar{x}$=66.61 months, or approximately 5 years and 6 months. Further observing the length of employment of these professionals in the company, it appears that, by separating into groups that are "less" and "more" time than the average, those who stay less are identified as middle management professionals, such as: Analysts, Experts, Developers, Designers, Statistics, Data Analysts and BI Developers (n=50/35.7%). Conversely, when analyzing the longest stay in the company, senior management is identified, such as: Managers, Customers Success Managers, Coordinators, Supervisors, Consultants and Technology Leaders (n =23/16.4%).

When performing an intergroup analysis of length of stay in organizations, respondents who spend "less" time had an average age $\bar{x}$=30 years, and those who spend "more" time in the organization had an average age $\bar{x}$=40 years ($t_{(132)}$=6.278; p<.001). This indicates that professionals aged 30 years or less are seeking employment within the field in other companies, and those aged 40 years and over remain in organizations longer. In the intergroup analysis of the variable "sex" in relation to "age", the data indicated that the respondents do not have differences in means. Male professionals represented in the sample were an average age of approximately $\bar{x}$=33 years and the female group $\bar{x}$=36 years. Despite the small difference, the T test indicated equal variances ($t_{(133)}$=1.393; p=.166).

5.2 Characterization of organizations

Most of the organizations participating in this research were national (n=82/58.2%) compared to international ones (n=57/40.4%). The service sector was the one with the highest participation in this research with 70.9% (n=100). Organizations with a "large" size also indicated most of the sample with 65.2% (n=92). When analyzing the person or persons responsible for the investment of technological systems that involve Big Data, the responsibility in general rests with the "Board", with 66% (n=93) and the item "other" was indicated by 4.3% (n=6) representing management outside the country or executive councils. It was also possible to verify that no respondents identified theirselves with the item "low management" - thus, for the purposes of this research, the terms "middle management" and "top management" were used. The average age of organizations is $\bar{x}$=572.67 months or approximately 47 years and 7 months, which indicates that most organizations have been established for almost half a century in the country, having gone through different economic scenarios, as well as technological transformations. It was also possible to verify that 24.8% (n=35) of the organizations are national, in the service sector and large.

The order of chronological occurrence in each stage of the Big Data Analytical capabilities with the respective analysis of variances (ANOVA) of the groups for each of the indicated variables is presented below.

As a result, the data indicated that the "Diagnostic" capacity was the one that presented the highest averages in practically all the variables analyzed, but always maintaining the same result of "there is no related effect in the group". Therefore, only the results that "related effects in the groups" of the analyzed variables were presented.

One-way ANOVA showed that there is an effect of the group on the Predictive ability to analyze Big Data [F(3,133)=3.451; p=.019]. This difference is observed in the average of large companies ($\bar{x}$=3,62). Indicating that they are the ones that most invest in predictive solutions. Similarly, one-way ANOVA also showed that there is an effect of the group on the Predictive ability to analyze Big Data

[$F_{(1,136)}$=10.213; p=.002]. This difference is observed in the difference in the mean of respondents, and the group of professionals ($\bar{x}$=3.70) indicates greater experience and long-term vision than the managers ($\bar{x}$=3.02) predictive solutions.

Besides that, one-way ANOVA showed that there is an effect of the group on the Prescriptive capacity of Big Data analysis [$F_{(1,133)}$=6.397; p=.013]. This difference is observed in the average of International companies ($\bar{x}$=3,67). In the case of the highest average, indicating that international companies are the ones that seek the greatest prescriptive solutions. One-way ANOVA also showed that there is an effect of the group on the Prescriptive ability to analyze Big Data [$F_{(3,133)}$=2.911; p=.037]. This difference is observed in the average of Public Services ($\bar{x}$=1,75) and Services ($\bar{x}$=3,49). Indicating that Public Services and the sector that most needs to develop prescriptive solutions and Services in general and the one that is most active. One-way ANOVA showed that there is an effect of the group on the Prescriptive ability to analyze Big Data [$F_{(3,132)}$=2,679; p<.05].This difference is observed in the average of large companies ($\bar{x}$=3.58). Indicating that they are the ones that most invest in predictive solutions.

5.3 Analysis of the structural model

The normality of the data was verified by the Kolmogorov-Smirnov test (K-S) and the respective p-value of each variable. This procedure was necessary to limit the possibility of using some statistical analysis techniques that have the normal distribution of data as a characteristic. All the Z values and individual p-values of the K-S test for the indicators were "very significant" with p<.001. Regarding the predictor variables related to the latent variables "Process-oriented dynamic Capabilities", "Product Innovation Performance", "Business Value" and "Satisfaction" it was possible to accommodate multicollinearity in the model - all values of the Variance Inflation Factors (VIFs) were below 5, with the lowest being SBDS03=1.168, and the greatest, SBDS01=4.583.

After the first interaction, the results of the factorial loads obtained by the variables were presented, it was found that it was only necessary to exclude two variables that had factor loads below .5 (SUBD02=0.419 and SBA01=0.406). In the sequence, convergent validity was verified, which demonstrates the extent to which the latent variable correlates with the items chosen to measure that variable and the discriminant validity involves the correlation between the constructs of the theoretical model.

The analysis of the measurement model must precede the analysis of the relationships between the constructs or latent variables. The next step was to examine the measurement model, which involved: Cronbach's Alpha (CA), Composite Reliability (CR), Average Variance Extracted (AVE), determination coefficients (R^2), predictive relevance (Q^2), effect size (f^2) and the Goodness of Fit (GoF) (Hair et al., 2014) according to Table 1:

Constructs	**CA**	**CR**	**AVE**	**R^2**	**f^2**	**Q^2**	**nº of items**
Strategic Business Alignment	.851	.910	.771	.660	1.938	.482	3
Advanced Analytics	.874	.914	.728	.708	2.430	.474	4
Updating IT	.883	.919	.740	.677	2.098	.462	4
Big Data Analytics Capability	.929	.938	.522				15
Financial Performance	.927	.945	.776	.859	6.109	.621	5
Market Performance	.913	.939	.795	.851	5.696	.634	4
Big Data Operations	.831	.889	.670	.735	2.776	.473	4
Product Innovation Performance	.946	.953	.630	.571	.274	.329	12
Process-oriented capabilities	.864	.908	.713	.360	.563	.235	4
Satisfaction	.836	.895	.692	.683	.081	.453	4
Strategic Uses of Big Data	.858	.914	.779	.691	2.232	.506	3

Business Value	.938	.951	.764	.547	1.208	.392	6

Table 1 – Convergent validity

Source: authors' elaboration

Cronbach's Alphas coefficients ranged between .831 and .946. These results indicated a high internal consistency of the scales used. The composite reliability obtained in this study varied between .889 and .953 indicating that the results were satisfactory in this regard (Hair et al., 2010).

For this model, the AVEs ranged between .522 and .795. All latent variables showed an average extracted variance greater than 50%, which reaches the criteria of Chin (1998) for the indication of the existence of convergent validity.

The R^2 value measures the predictive accuracy of the model, representing the combined effects of endogenous variables on exogenous variables. In the present study, the R^2 values demonstrated that the model has accuracy and predictive relevance in all constructs.

Cohen's f^2 is used to estimate the effect size in correlated samples (repeated measures, longitudinal data, grouped data) for two continuous variables. It is evaluated how much each construct is "useful" to adjust the model. Values of .02, .15 and .35 are considered small, medium and large, respectively (Hair et al., 2014). The f^2 values are obtained by reading the communalities. What is perceived in this research is that the vast majority were considered medium and large, with the exception of the construct "Satisfaction" with .081.

The Stone-Geisser's Q^2 test assesses how close the model is to what was expected of it (or the quality of the model prediction or the accuracy of the adjusted model). According to the criterion used to assess the accuracy of the adjusted model, Q^2, all constructs were considered to be adequately accurate, as they presented values greater than zero (Akter et al., 2017).

The GoF is the global adjustment measure, that is, the square root of the multiplication of both the averages of the AVE and the R^2 of the endogenous variables. The value found for the GoF is .691 (69.1%) and was considered large enough for the validity of the model in the PLS (Wetzels, Odekerken-Schröder & Van Oppen, 2009). The discriminant validity assessed items that reflect the factor or that are correlated with other factors. In this research, the average variances extracted were greater than or equal to the square of the correlation between the factors, as shown in Table 2, with the definitive Fornell-Larcker criterion with all the factor loads of each indicator with values above .5. It was not necessary to exclude variables to adjust the model.

Latent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1)Strategic Business Alignment	**.878**									
(2)Advanced Analytics	.615	**.853**								
(3)Updating IT	.617	.516	**.860**							
(4)Financial Performance	.625	.471	.422	**.881**						
(5)Market Performance	.664	.578	.456	.778	**.891**					
(6)Big Data Operations	.560	.668	.628	.353	.446	**.818**				
(7)Process-oriented capabilities	.601	.597	.527	.575	.641	.402	**.845**			
(8)Satisfaction	.720	.609	.543	.682	.637	.566	.530	**.832**		
(9)Strategic Uses of Big Data	.640	.587	.613	.533	.578	.635	.455	.670	**.883**	
(10)Business Value	.777	.577	.647	.713	.706	.577	.594	.811	.700	**.874**

Table 2 - Discriminant validity

Source: authors' elaboration.

Note: the highlighted diagonal shows the square roots of the AVE.

The path model shown in Figure 2 in the structural equations consists of two elements: the internal model in which the relationships between the constructs are displayed, and the measurement model that displays the relationships between the constructs and the indicators. Thus, the measures allow for the comparison of the structural models with the theoretically established reality, that is, how well the theory is adjusted to the data (Hair et al., 2014).

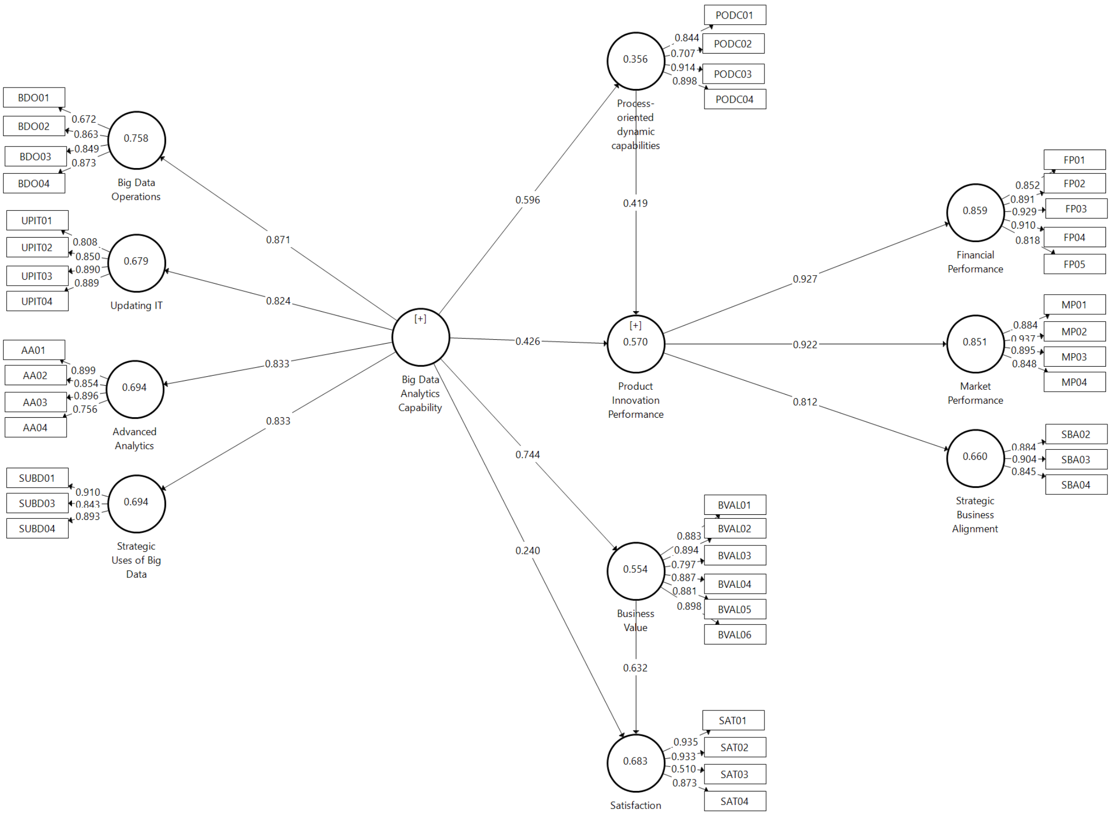

Figure 2: Proposed research model
Source: output from SmartPLS.

In the practical application of the structural equation modeling for the proposed research model - Figure 2, the approach suggests the evaluation of the measurement models individually, seeking to verify their validity and consistency, and then carrying out the structural model approach (Chin, 1998).

6 DISCUSSION

The **H$_{1a}$** hypothesis ((β=.429; p<.001) described in the causal relationship of the path "Big Data Analytics Capability" and "Product Innovation Performance" revealed a significant relationship because the companies understood that product innovation is necessary. This realization happened since they are in a constantly changing environment, and only by adding new technologies and tools - as is the case of Big Data and its analysis capacity - is it possible to improve Product Innovation Performance (Hooi, Abu & Raim, 2018). By using Big Data within the product innovation process, companies allow for greater accuracy in their investments in research and development, because there will be greater prior intelligence in knowing what will actually return to the organization. This intelligence will come from the larger amount and variety of information that the organization will have, allowing for more refinement in its investment strategy.

The **H_{1b}** hypothesis (β=.600; p<.001), described as the causal relationship in "Big Data Analytics Capability" and "Process-oriented dynamic capabilities" was significant because organizations note that Digital Transformation will be a crucial competitive differentiator (Wixom et al., 2013) allowing the companies to stand out from the competition. It was noted in the survey that the responding managers associated the Big Data Analytics Capability with the dynamic capabilities, as the organization is facing its competition., That is, the better the Big Data analytics applied, the better its differential is to the rest of the market (Wamba et al., 2017).
In the case of hypothesis **H_{1c}** (β=.239; p<.05), "Big Data Analytics Capability" and "Satisfaction", there was a positive relationship because managers note that the use of Big Data tools is linked to the way in which organizations' analysts and managers view their role, leading to their satisfaction with Big Data. A positive relationship was found in the research, that is, the way in which Big Data's Analytics Capability are used, as Christensen (2016) states that there is an emphasis in the way companies manage their digital environment. IT managers and professionals can observe the integration of digital technology in all areas of a business. It is clear that in all sectors analyzed in this research there is a presence of companies that seek innovative and modern service, and this favors companies that create conditions for innovation to continue in the process of digital transformation with a speed greater than their competitors.
Regarding the **H_{1d}** hypothesis (β=.740; p<.001), described as "Big Data Analytics Capability" and "Business Value", it was significant because the survey respondents note that there are significant gains to the company when the Big Data tools are applied correctly. That is, the perception of managers and analysts is that the organization is ahead of its competition, enabling a better understanding of its customers and their relationship with them (Wamba et al., 2017).
The H_2 hypothesis (β=.416; p<.001) described in the causal relationship between "Process-oriented dynamic capabilities" and "Product Innovation Performance", was positive. This is because regarding Product Innovation, it is necessary for the organization to be ahead of the market, directing its resources for process improvement to be constant, maintaining a competitive differential, thus distancing the organization from the competition. Furthermore, the company will have better knowledge of its customers, so it will be able to focus on perfecting the main characteristics of its products and business strategies that will be more prominent with its target audience, consequently its performance in product innovation will also improve (Hooi, Abu & Raim, 2018).

Hypothesis **H_3** (β=.634; p<.001), described in the causal relationship of path "Business Value" and "Satisfaction" was also positive because managers and direct users of Big Data note that, when attributes are ascribed to the tool that is directly related to the company's competence in having a competitive advantage, is adaptable, and establishes a better relationship with its customers, it is seen that there is greater Satisfaction with Big Data. Therefore, it is possible to affirm that when perceived Business Value in using Big Data, companies will achieve a higher level of Satisfaction with it, avoiding failed Big Data implementation projects, without any delivery of Value (Golder et al., 2012).
When observing the constructs that comprise the operational definitions of Product Innovation Performance, it appears that both the Financial Performance (R^2=0.859) and the Market Performance (R^2=0.851) presented high determination coefficients (R^2). Thus, it is possible to verify that companies see relevance when they place Product Innovation Performance linked to Financial Performance. In other words, there is a greater return on investment, as well as greater sales growth, customer retention and profitability, when using the Big Data Analytics Capability directed to Product Innovation.

In the case of the operational definition of business strategy alignment (R^2=0.660), it was possible to identify, despite a value considered high for the coefficient of determination, the construct will probably need greater articulation by companies in relation to the use of Big Data. As much as organizations see advantages in using Big Data, there is little strategic alignment with the other areas that will use this tool.

As seen in the survey, approximately 89.4% of respondents stated that those responsible for investments in Big Data systems are Senior Management (23.4%) and the Board (66%), as well as the decisions of when and how the project will be implemented. Big Data has become top-down, with little end-user involvement, thereby increasing the chances of noise between the defined strategies and those who should execute them, so there is no clarity in the quantified objectives, or the expected performance when using tools Big Data analytics.

Mediation analysis was used to verify the relationship between three variables. In this research it was found that the two mediations proposed in the model were partial, that is, both the direct and the indirect relationship had similar significant results.

7 IMPLICATIONS AND CONCLUSIONS

This paper both theoretical and practical implications to identify the consequent factors of Big Data Analytics Capability in companies that seek digital transformation, identifying the main models of its use (descriptive, diagnostic, predictive and prescriptive). In this research, companies located in the state of São Paulo - Brazil were selected for how they are contributing to Digital Transformation.

Another factor was observing how they are influenced by technology in strategic market decisions, through the relationships between Product Innovation Performance directly related, or through the Process-oriented dynamic capabilities construct. Product Innovation Performance was built from the perspective of three different performance-related constructs, namely: Financial Performance, Market Performance and Strategic Business Alignment. The work also sought to define the best way to evaluate the Big Data Analytics Capability and its relationship with the constructs described above when intermediated by the Product Innovation Performance. In addition, other key factors for Digital Transformation were observed: Satisfaction and Business Value.

Satisfaction was the construct that had the most explanatory power in the model with R^2=68.3%, followed by Product Innovation Performance with R^2=57.1%, Business Value with R^2=54.7% and finally, Process-oriented dynamic capabilities R^2=36.0%. This indicates that Process-oriented dynamic capabilities should provide the company's ability to change (improve, adapt, adjust, reconfigure, update, renew, among others) into a better business process than the competition. Therefore, the competence of managers is essential for such changes. In fact, dynamic capabilities can be understood as an organization's ability to purposely create, extend or modify its resource base (Helfat et al., 2007). This is implied when the company has the capacity to recognize opportunities for change in internal and external environments, configuring organizational processes and deploying resources efficiently and immediately being able to capitalize on them (Eisenhardt & Martin, 2000; Kim et al., 2011; Wamba et al., 2017).

It is important to highlight that, with the massive amount of data collected by companies, it is increasingly necessary to have strategies aimed at their use and management. From the use and management of these data, which are accessible in digital media, along with previous organizational knowledge, different inputs could be synthesized that will value the way the company does business. These final inputs can be used to generate new ways of looking at the market, as well as new ideas and value proposals.

A relevant aspect to be highlighted is the impact that Big Data has been generating in organizations. The literature review revealed that the transformation that occurred in companies is not limited to the area of technologies and systems, but also in the way professionals are being inserted in this context. As the survey revealed, one third of respondents have completed postgraduate studies, that is, professionals with higher levels of graduation are increasingly required, with 76.5% (n=108) having at least one degree.

When checking each stage of Big Data Analytics Capability in the companies surveyed, most of them use predominantly Diagnostic analytics, with no statistical differences in the means in the analysis of variances (ANOVA). This demonstrates that they are working on their analytical models based on past events, to help them understand the historical data series. Market studies claim this type of analysis is sufficient, however, there is also a need for Predictive and Prescriptive analysis, with data from the past, to predict ways to improve future results. The use of predictive analysis was evident for "big companies" and "professional's" respondents (characterized by working with predictions, simulations and modeling), and the use of prescriptive analysis for "international" companies, from the "service sector" and "big companies".

For this, greater investment in analytical tools is necessary, as well as people trained in these models in analysis (descriptive, diagnostic, predictive and prescriptive), in addition to a change of mindset, which will dictate the bias of the company's strategic decision making. It will also be necessary to invest in the training of these professionals, so that they can properly handle the Big Data tools and increase their influence in the organizational decision-making process.

The limitations of this study are mainly related to its external validity. The respondents to this survey came from companies located in the state of São Paulo. Although many of these companies are identified as reference in the use of Big Data for future research, the extension of companies in other cities and states could be considered. Consequently, these results can be generalized only for companies located in the state of São Paulo.

Referências

Akter, S., Wamba, S. F. & Dewan, S. (2017). Why PLS-SEM is suitable for complex modeling? An empirical illustration in Big Data Analytics Quality. *Production Planning and Control*, 28 (11-12), 1011-1021.

Ananyin, V. I., Zimin, K. V., Lugachev, M. I., Gimranov, R. D., & Skripkin, K. G. (2018). Digital organization: Transformation into the new reality.

Andreeva, T., Chaika, V. (2006) Dynamic Capabilities: What they need to be Dynamic? St. Petersburg University: Institute of Management.

Baines, T. S., Lightfoot, H. W., Benedettini, O. & Kay, J. M. (2009). The servitization of manufacturing: A review of literature and reflection on future challenges. *Journal of Manufacturing Technology Management*, 20(5), 547-567.

Baron, R. M., & Kenny, D. A. (1986). "The moderator-mediator variable distinction in social psychological research: conceptual, strategic and statistical considerations". Review of. *Journal of personality and social psychology* 51 (6):1173–82.

Berry, L. L., Shankar, V., Parish, J. T., Cadwallader, S. & Dotzel, T. (2006). Creating new markets through service innovation. *MIT Sloan Management Review*, 47(2), 56-63.

Bhatt, G. D. & Grover, V. (2005). Types of Information Technology Capabilities and Their Role in Competitive Advantage: An Empirical Study. *Journal of Management Information Systems*, 22, 253-277.

Bhimani, A. & Willcocks, L. (2014). Digitisation, 'Big Data' and the Transformation of Accounting Information, *Accounting and Business Research*, 44(4): 1–22.

Chen, H., Chiang, R. H. L. & Storey, V. C. (2014) Business Intelligence and Analytics: from Big Data to Big Impact. *MIS Quarterly*, 36(4), 1165-1188.

Chen, M., Mao, S., & Liu, Y. (2014). "Big data: A survey." Mobile Networks and Applications, Vol. 19, No. 2, pp. 171-209.

Chin, W. W. (1998). The partial Least Squares Approach for Structural Equation Modeling. In: Marcoulides G. A. (ed.). *Modern Methods of Business Research*. New Jersey: Lawrence Erlbaum Associates, Publishers, p. 295-336.

Christensen, J. (2016). Digital Economics: The Digital Transformation of Global Business. Copenhagen: Books on Demand.

Cohen, S. L., & Tripsas, M. (2018). Managing Technological Transitions by Building Bridges. *Academy of Management Journal*, 61(6), 2319-2342.

Day, G., & Schoemaker, P. J. H. (2016). Adapting to fast-changing Markets and Technologies. *California Management Review*, 58(4), v. 58, 59-77.

Eisenhardt, K. M., & Martin, J. (2000). Dynamic capabilities: What are they? *Strategic Management Journal*, 21(10/11), 1105–1121.

Eggers, J. P., & Park, K. F. (2018). Incumbent adaptation to technological change: The past, present, and future of research on heterogeneous incumbent response. *Academy of Management Annals*, 12(1), 357-389.

Ekbia, H., Mattioli, M., Kouper, I., Arave, G., Ghazinejad, A., Bowman, T., Suri, V.R., Tsou, A., Weingart, S. & Sugimoto, C.R. (2015). Big Data, Bigger Dilemmas: A critical review, *Journal of the American Society for Information Science & Technology*, 1-39.

Gim, J., Lee, S. & Joo, W., (2018) "A Study of Prescriptive Analysis Framework for Human are Services Based on CKAN Cloud", vol. 2018, *Journal of Sensors*, 1-10.

Goes, P. B. (2014). Big Data and IS Research. *MIS Quarterly*, 38(3), 3-8.

Golder, P. N., Debanjan, M., & Moorman, C. (2012). What is quality? an integrative framework of processes and states. *Journal of marketing,* 76(4):1-23.

Gunday, G., Ulusoy, G., Kilic, K., & Alpkan, L. (2011). Effects of innovation types on firm performance. *International Journal of production economics*, 133(2), 662-676.

Hair, J. F., Black, W. C., Babin, B. J. & Anderson, R. E. (2010). Multivariate Data Analysis: A Global Perspective, New Jersey, Pearson Prentice Hall.

Hair, J. F., Hult, G.; Tomas, M.; Ringle, C., Sarstedt, M. (2014). A primer on partial least squares structural equation modeling (PLS-SEM). *Sage Publications*.

Halevi, G., Moed, H. (2012). The evolution of big data as a research and scientific topic: Overview of the literature, Res. Trends 3–6.

Hannachi, Y. (2015). Development and Validation of a Measure for Product Innovation Performance: The PIP Scale. *Journal of Business Studies Quarterly*, 6(3), 23-35.

Helfat, C. E., Finkelstein, S., Mitchell, W., Peteraf, M., Singh, H., Teece, D., & Winter, S. (2007*) Dynamic capabilities: Understanding strategic change in organizations*. Oxford: Blackwell.

Hess, T., Matt, C., Benlian, A., & Wiesböck, F. (2016). Options for formulating a digital transformation strategy. *MIS Quarterly Executive*, 15(2), 123-139.

Hilbert, M. & Lopez, P. (2011). The World's Technological Capacity to Store, Communicate, and Compute Information, *Science*, 332 (6025), 60-65.

Hooi, T. K., Abu, N. H. B., & Raim, M. J. I. A. (2018). Relationship of Big Data Analytics Capability and Product Innovation Performance using SmartPLS 3.2.6: Hierarchical Component Modelling in PLS-SEM. *International Journal of Supply Chain Management*, 7(1), 51-64.

Im, S., Bayus, B. L., & Mason, C. H. (2003). An empirical study of innate consumer innovativeness, personal characteristics, and new-product adoption behavior. *Journal of the Academy of Marketing Science*, 31(1), 61-73.

Kacfah Emani, C., Cullot, N., Nicolle, C. (2015). "Understandable Big Data: A survey, Computer Science Review".

Kallinikos, J., Aaltonen, A. & Marton, A. (2013). The Ambivalent Ontology of Digital Artifacts, *MIS Quarterly,* 37(2): 357–370.

Kim, G., Shin, B., Kim, K. K., & Lee, H. G. (2011). IT Capabilities, Process-Oriented Dynamic Capabilities, and *Firm Financial Performance. Journal of the Association for Information Systems,* 12(7), 487-517.

Kraska, T. (2013). Finding the needle in the Big Data systems haystack. *IEEE Internet Computing*, 17(1), 84-86.

Kwon, O., Lee, N. & Shinb, B. (2014) Data quality management, data usage experience and acquisition intention of big data analytics, *International Journal of Information Management*, 34(3), 387-394.

Lavalle, S., Lesser, E., Shockley, R., Hopkins M. S., & Kruschwitz, N. (2011). "Big Data, Analytics and the Path from Insights to Value". Review of. *MIT Sloan Management Review* 52 (2):21-32.

Leeflang, P. S. H., Verhoef, P. C., Dahlström, P. & Freundt, T. (2014). Challenges and solutions for marketing in a digital era, *European management journal*, 32(1), 1-12.

Mcafee, A. & Brynjolfsson, E. (2012). Big data: the management revolution. *Harvard business review*, 60-6, 68, 128.

Miles, I. (2008). Patterns of innovation in service industries. *IBM Systems Journal*, 47(1), 115-128.

Nambisan, S., Lyytinen, K., Majchrzak, A., & Song, M. (2017). Digital innovation management: Reinventing innovation management research in a digital world. *MIS Quarterly*, 41(1), 223-238.

Nwankpa, J.; Roumani, J. (2016). IT Capability and Digital Transformation: A Firm Performance Perspective. *Thirty Seventh International Conference on Information Systems*, Dublin.

Ostrom, A. L., Parasuraman, A., Bowen, D. E., Patrício, L., Voss, C. A. & Lemon, K. (2015). Service research priorities in a rapidly changing context. *Journal of Service Research*, 18(2), 127-159.

Rothenberg, S. (2007). Sustainability through servicizing. *MIT Sloan Management Review*, 48(2), 83-89.

Santhanam, R. & Hartono, E. (2003). Issues in Linking Information Technology Capability to Firm Performance. *MIS Quarterly*, 27(1), 125-153.

Schallmo, D., Williams, C., Boardman, L. (2017). Digital Transformation of Business Models-Best Practice, Enablers and Roadmap. *International Journal of Innovation Management*, 21(8), 1740014(1-17).

Song, S., Jeong, D.-H., Kim, J., Hwang, M., Gim, J., & Jung, H. (2014). Research Advising System Based on Prescriptive Analytics. *Future Information Technology*, 569–574.

Subramaniam, M., Iyer, B., & Venkatraman, V. (2019). Competing in digital ecosystems. *Business Horizons*, 62(1), 83-94.

Tippins, M. J., Sohi, R. S. (2003). IT competency and firm performance: is organizational learning a missing link? *Strategic Management Journal,* 24(8), 745-761.

Waldrop, M. (2008). Wikiomics. *Nature*, 455(7290), 22-25.

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

Wang, N., Liang, H., Zhong, W., Xue, Y. & Xiao, J. (2012). Resource Structuring or Capability Building? An Empirical Study of the Business Value of Information Technology. *Journal of Management Information Systems*, 29(2), 325-367.

Warner, K. S., & Wäger, M. (2019). Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal. *Long Range Planning*, *52(3)*, 326-349

Wetzels, Odekerken-Schroder, G. & Van Oppen, C. (2009). Using PLS Path Modeling for Assessing Hierarchical Construct Models: Guidelines and Empirical Illustration. *MIS Quarterly*, v. 33, p. 177-195.

Wixom, B. Yen, B. & Relich, M. (2013). "Maximizing value from business analytics". *MIS Quarterly Executive,* 12:111-23.

Xu, Z., Frankwick, G. L., & Ramirez, E. (2016). Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective. *Journal of Business Research*, 69(5), 1562-1566.