Foresights for big data across industries

Fernando Almeida

Fernando Almeida is based at INESC TEC, University of Porto, Porto, Portugal.

Abstract

Purpose – *The purpose of this study is to explore the potential and growth of big data across several industries between 2016 and 2020. This study aims to analyze the behavior of interest in big data within the community and to identify areas with the greatest potential for future big data adoption.*

Design/methodology/approach – *This research uses Google Trends to characterize the community's interest in big data. Community interest is measured on a scale of 0–100 from weekly observations over the past five years. A total of 16 industries were considered to explore the relative interest in big data for each industry.*

Findings – *The findings revealed that big data has been of high interest to the community over the past five years, particularly in the manufacturing, computers and electronics industries. However, over the 2020s the interest in the theme decreased by more than 15%, especially in the areas where big data typically had the greatest potential interest. In contrast, areas with less potential interest in big data such as real estate, sport and travel have registered an average growth of less than 10%.*

Originality/value – *To the best of the author's knowledge, this study is original in complementing the traditional survey approaches launched among the business communities to discover the potential of big data in specific industries. The knowledge of big data growth potential is relevant for players in the field to identify saturation and emerging opportunities for big data adoption.*

Keywords *Big data, Internet of Things, Blockchain, Challenges, Opportunities*

Paper type *Research paper*

1. Introduction

Data analysis is a fundamental phase in a company's decision-making process. However, information is a key asset to ensure positive returns for the business. Therefore, the organization must have a high quality of data. According to Albright and Winston (2014), big data implies knowing the provenance of the data, checking the composition of the elements and assessing the consistency of the available information, among other procedures.

The evolution of information technology and the emergence of more powerful and distributed technological tools in the cloud have enabled companies of all sizes to access information quickly and dynamically (Cao, 2017). Studies performed by Atov *et al.* (2020) and Berman *et al.* (2018) also demonstrate that strategies based on accurate data are proven to be more effective than actions not grounded in any kind of analysis. However, the mere availability of data does not guarantee the assertiveness of the information and does not reveal the solution to the company's business problem. Therefore, in addition to ensuring the origin of the elements of the analysis, it is necessary to choose the appropriate methods to study and interpret the data. In this sense, Dunn and Davis (2017) emphasize the importance of respecting all stages of the analysis, from the information collected to the validation of the findings.

It becomes evident that more important than storing data is to know its usefulness and impact on the company. Furthermore, it is relevant to understand how big data can contribute to improving business. The generation and storage of data are a practice that

Received 26 February 2021
Revised 10 July 2021
4 September 2021
22 November 2021
Accepted 1 February 2022

happens since ancient times, but it was with the advancement of technology in handling information that the concept of big data emerged. Big data is defined by TechAmerica Foundation's Federal Big Data Commission (2012) as "a term that describes large volumes of high velocity, complex and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the information." According to Satish and Yusof (2017), the use of big data in organizations not only aims to know the consumer behavior but also to know the reasons that lead the customer to have a specific behavior. Issues related to big data are extremely timely and have been changing the way companies relate with consumers, enhancing fundamental efficiency gains in a highly competitive market.

Access to information is power and allows companies to explore new business opportunities. The availability of more information about customers enables the company to know their preferences regarding services and products. With this, Verdenhofs and Tambovceva (2019) note that companies can perform better market segmentation and consequently better differentiation they can make of the consumers to whom these products and services are intended. These benefits appear not to be focused on a specific industry. Several studies highlight the potential of big data in various industries such as healthcare, agriculture, retail and transportation. (Dash *et al.*, 2019; Dekimpe, 2020; Milne and Watling, 2019; Wolfert *et al.*, 2017). The potential of big data is essentially measured considering two approaches: questionnaires launched among the business communities to find out the big data practices they are implementing or intend to implement in the near future (Cabrera-Sánchez and Villarejo-Ramos, 2019) and demonstration of big data application scenarios in specific sectors (Harerimana *et al.*, 2018; Mani *et al.*, 2017). Despite the relevance of these approaches in the exploration of big data in a business context, they do not identify trends on a global scale, namely, regarding the relative importance of big data by different types of industries. This work aims to address these deficiencies by adopting Google Trends, which explores the most popular terms searched in Google in the past five years. This approach complements the work performed by Raguseo (2018) that looked to the potential of big data during the period from 2004 to the end of 2015. Additionally, this study allows us to know the relevance given by the community to big data and to segment it considering 16 types of industries (e.g. government, health, telecommunications, tourism, etc.). The following three research questions were established:

RQ1. What has been the evolution of big data in the community?

RQ2. In which industries have big data offered the greatest potential?

RQ3. In which industries have big data seen the most growth?

These three research lines allow us to quantify the growth of big data in the community and understand its relative potential in each industry because it is expected that the potential of big data can be asymmetric in some specific industries. Furthermore, and to trace future trends of big data growth, it is relevant to explore comparatively the growth of searches considering a lower granularity for each industry.

The rest of this manuscript is organized as follows: In the first phase, a theoretical framework about the concept of big data and its potential for companies is carried out. Next, the methodology of the study is presented, in which the Google Trends platform is described and the industries considered in this study are characterized. After that, the results are reported considering the three research questions previously established. After that, the results are discussed and contextualized. Finally, the conclusions of the study are presented. Also, in this section, the limitations of the study are stated and the theoretical and practical implications of the work are addressed.

2. Background

Big data is typically used to refer to the processing, storage and analysis of large amounts of data from diverse sources to extract knowledge that identifies patterns in them. Therefore, big data is not only limited to the mere existence of data but also includes the techniques and processes inherent in collecting and processing it (Almeida, 2018; Gandomi and Haider, 2015). These data are created in the most varied ways, namely, through users' interaction with services on the internet, and can be used for several purposes such as improving products or services, developing new products or improving data analysis techniques (Zhan *et al.*, 2018).

Three basic dimensions of big data (i.e. volume, velocity and variety) are commonly addressed in the literature (Favaretto *et al.*, 2020; Oussous *et al.*, 2018). Volume is associated with the large amount of data generated; velocity explores how quickly new data is created and moves over time; and variety includes the wide diversity of data types we can find, which include structured and unstructured data. However, several authors add other V's. Veracity is also a key element as it explores the relevance of the data collected to the purpose of the analysis (Reimer and Madigan, 2019). Value is equally relevant as analysis should be focused on the business direction (Raja and Prema, 2018). In this regard, the work developed by Hariri *et al.* (2019) summarizes the characteristics of big data into five components (i.e. volume, velocity, variety, value and veracity).

After identifying the characteristics of big data, it becomes relevant to understand the sources and techniques where big data can be used and provide value for the organization. Text analytics is adopted to extract information from textual data, such as interactions in social networks, emails, forums and response to questionnaires. According to Redondo and Sandoval (2016), this type of analysis involves statistical analysis, computational linguistics and machine learning. Audio analysis is a technique that analyzes and extracts information from unstructured audio data. The main areas of intervention of this technique are healthcare and call centers. Sivarajah *et al.* (2017) note that this technique can be used to improve the consumer experience, evaluate agent performance, support various types of diagnostics and aid in the treatments of certain medical conditions that affect patient communication patterns. Video analytics involves a variety of techniques to monitor, analyze and extract meaningful information from video streams. Subudhi *et al.* (2019) reveal that the main application of video analytics in recent years focuses on security systems and automated surveillance. Social media analytics refers to the analysis of data and unstructured under multiple social media access points. Esfahani *et al.* (2019) highlight that research on social media spans multiple disciplines (e.g. mathematics, computer science, sociology, psychology). Finally, predictive analytics includes a variety of techniques that predict future outcomes based on historical and current data. Predictive analytics can be applied in various areas such as predicting shopping habits or for supply chain demand (Boone *et al.*, 2019; Seyedan and Mafakheri, 2020).

Big data is emerging as a relevant topic from a scientific and business perspective. The results of the systematic review by Sardi *et al.* (2020) indicate a significant increase in the number of publications on big data and performance, but a shortage of studies of big data application in business, management and accounting. Taleb *et al.* (2021) look at big data as a holistic approach for managing, processing and analyzing data to deliver value in a sustained way and improve decision-making. Mikalef *et al.* (2019) add that big data enables to improve the efficiency and effectiveness of organizations and empowers the decision-making process based on evidence rather than intuition. Saritas *et al.* (2021) explore the role that big data can play in the development of mobile commerce-related businesses. In this sense, organizations are increasingly using key elements of big data in their information systems. From this reality, two aspects arise that must be considered: investment in specific training for employees and the introduction of a cultural change. Frisk and Bannister (2017) point out that this cultural change should not only involve employees but also the leaders of

organizations who often lack knowledge about the value that big data can bring to their organizations. Finally, implementing big data solutions into their systems and integrating them into the decision-making process is not enough to ensure the full utilization of these solutions. In line with this challenge, De Luca *et al.* (2020) emphasize the importance of having a continuous evaluation process of the applications' performance and a constant search for innovation possibilities.

A data-driven company tends to be forced to reconsider its structure and all the organization's processes considering the availability of new information. Thirathon *et al.* (2017) note that the benefits of implementing a data-driven culture are extensive, starting with informed decision-making. It puts an end to decisions based only on intuition and offers an immediate competitive advantage to businesses by providing greater accuracy and security to the actions taken by the company. Sen *et al.* (2016) emphasize that data driven is neither limited to large corporations nor is it just about using user information. Small- and medium-sized enterprises not only can but also must apply the culture of data-driven decision-making, which can come from statistical studies, market research and analysis and information from customer interactions. A data-driven company gains competitive advantages over other players in the market. Among the competitive advantages that a data-driven organization can offer are agility and efficiency in decision-making, more assertive forecasts, greater adaptability, quick reactions to change and the fostering of innovation (Fosso *et al.*, 2017; Gobble, 2013; McAfee and Brynjolfsson, 2012). Accordingly, data-driven organizations can perform better and achieve their financial and operational objectives.

Big data creates value in many different domains for businesses. The knowledge extracted from the data enables companies to offer innovative and personalized products through a more effective segmentation of markets (Yoseph *et al.*, 2020). Big data has also been used in human decision support which allows it to replace and/or assist decision-making through the application of automated algorithms (Power, 2014). In the manufacturing sector, the integration of information from the engineering and research departments significantly reduces the time to market for a new product and allows for an increase in its quality (Amado *et al.*, 2018). These examples show that in addition to the financial gains that companies can acquire by adopting big data, performance gains are also achieved.

Although the adoption of big data is more common in digital markets, this practice is not exclusive to them. In traditional markets, we increasingly realize the use of data analysis and customer profiling. According to Anshari *et al.* (2019), this approach is important for companies to make strategic decisions and apply better market strategies which will consequently attract more customers. In these situations, data can be collected not only from direct contact with the customer in the physical space, namely, through consumer surveys, but also by exploring the various forms of payment made by the consumer.

3. Methodology

This study adopts the quantitative methodology intending to quantitatively describe the interest in the theme of big data across industries. According to Swift and Piff (2014), this approach is appropriate when numerical data describe a given phenomenon, which facilitates the description of the change and transformations undergone by the objects under study. Furthermore, this approach enables the quantification of the evolution of big data and consequently a greater comparison between industries and a greater objectification of the information collected and its analysis. Finally, Apuke (2017) mentions the potential of this methodology in explaining the causal dependencies between social phenomena.

Google Trends was used as a source of information in this study. It is a tool provided by Google that allows tracking the evolution of the number of searches for a particular keyword

over time (Rogers, 2016). Furthermore, Google Trends has been successfully used to understand and predict the behavior of several areas, such as tourism flows (Siliverstovs and Wochner, 2018), financial markets (Preis *et al.*, 2013), health (Nuti *et al.*, 2014) or COVID-19 lockdown (Brodeur *et al.*, 2021). The base information provided by Google Trends includes a graph in which the volume of searches over time is represented. The numbers represent the search interest relative to the highest point on the graph for the specified region and time interval. A value of 100 is the peak popularity of the term; a value of 50 means that the term was half as popular; and a score of 0 means that there was not enough data for this term. Google Trend allows us to explore interest in a given phenomenon considering various degrees of granularity such as past hour, past day, past seven days, past month, past quarter and past year. This approach calibrates the Google Trends Time Series and adapts it to the specific granularity needs of the study (West, 2020). Therefore, it can be used to make comparisons between different search terms and evaluate the seasonality or progress of some topic over a period of time. Google Trends allows us to analyze the importance given by the community to the big data phenomenon and identify trends within a market segment or industry. The adoption of Google Trends has been followed in several investigations particularly for forecast and consequence analysis of COVID-19 (Kurian *et al.*, 2020; Mangono *et al.*, 2021; Timoneda and Wibbels, 2021).

It has explored the searches on big data on Google over the past five years (i.e. 2016–2020). This analysis was conducted over the course of February 2021. A total of 260 samples (i.e. one per week) over this period were considered, which allowed us to perform a seasonality analysis by trimester. In the specific analysis of the relative importance of big data for each industry, 16 types of industries were analyzed as described in Table 1. Furthermore, we have explored the growth trend by industry considering the past available year (i.e. 2020). Accordingly, the level of granularity was reduced to a time horizon of one year and the growth trend by quarter was explored.

Table 1 Description of the industries involved in the study

Industry	*Description*
Arts and entertainment	It comprises artistic, musical, and literary production and presentation activities
Automobiles and vehicles	All forms of trade and repair of motor vehicles and motorcycles are included in this section
Beauty and fitness	It comprises activities related to physical maintenance and wellness. It also includes beauty salons and gyms
Computers and electronics	It comprises the manufacture of computers and their peripheral equipment, communications equipment and similar electronic products, as well as the development of software
Education	It comprises the activities of public, private and cooperative education of all kinds in any field
Finance	Financial activities include banking, insurance, pension funds and activities auxiliary to financial intermediation, insurance and pension funds
Food and beverage	Businesses in the catering and similar fields are found here, which includes companies in the area of services in which food and drink are consumed
Games	It includes companies that trade computer games and access those games via streaming
Government	It comprises the activities of the public administration, national and local
Health	Included are all companies that provide health and social support services, whether they are public or private entities
Leisure	It comprises leisure-related activities that are not included in the sports category
Manufacturing	It includes a comprehensive set of companies that engage in the transformation of raw materials or parts into finished goods through the use of machinery and human labor
Real estate	In this section are real estate activities such as buying, selling, renting, administration and real estate mediation
Sport	It comprises both professional and amateur sports activities
Telecommunications	Sector of activity within electrical engineering, whose purpose is to control, implement and manage the various communications systems (internet, telephony, satellites, etc.)
Travel	It comprises the activities of tourism agents and tourism-related businesses

Stata v.16 was used in the exploration and analysis of the results. Several statistical techniques were used which included both descriptive analyses of the sample data and analysis of variance (ANOVA) which was employed to compare the distribution of samples across industries. According to Judd *et al.* (2017), the ANOVA is also a way to summarize a linear regression model by decomposing the sum of squares for each source of variation in the model and, using the *F*-test, test the hypothesis that any source of variation in the model is equal to zero.

4. Results

4.1 Big data evolution

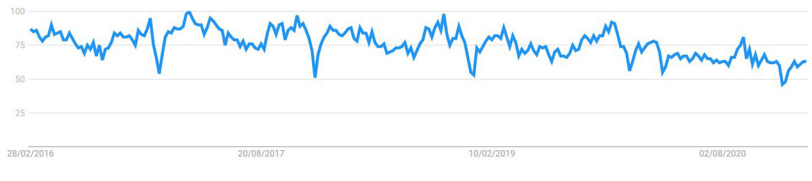
Figure 1 presents the evolution of the community's worldwide interest in big data. This information was collected from Google Trends on 21 February 2021. The results show an interesting phenomenon as big data has received strong interest from the community between 2016 and 2019. The average weekly estimates indicate an average level of interest of 76.08. However, in 2020 there was a reduction of 17.41% of interest in this topic. The average was 79.21 between 2016 and the end of 2019, whereas in the year 2020 the average was only 65.42.

Figure 2 explores the seasonality of interest in big data considering quarterly analysis over five years (i.e. from 2016 to 2020). There is a gradual decline in interest in big data, most significantly in the third quarter. The first quarter of each year is the quarter with the highest interest in big data (mean equal to 52.82), except for 2019, which was surpassed by the fourth quarter with a mean equal to 51.38, whereas the first quarter only got a mean of 48.46.

4.2 Comparative analysis by industry

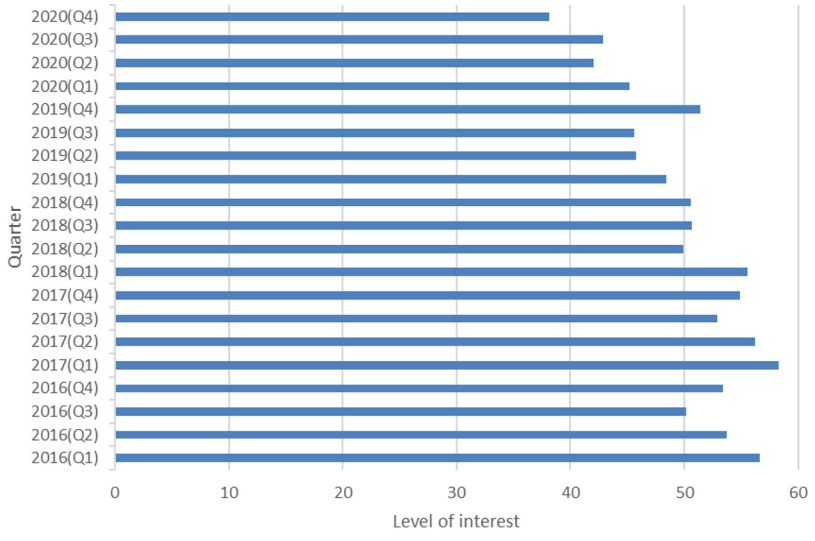
Table 2 explores the interest of big data across industries. For this purpose, we have initially done a descriptive statistical analysis like it is performed in Raguseo (2018). It is calculated as the overall ranking, mean, median and standard deviation. We also performed a hypothesis test considering the difference between two groups (i.e. all industries and a specific industry). A significance level of 5% (α = 0.05) was adopted. The results obtained indicate that the industries where big data has attracted the most interest are manufacturing and computers and electronics. On the opposite side, the industries that have captured the least interest from the community are real estate and beauty and fitness. It was also possible to identify that interest in big data, in general, is higher (mean equal to 76.08) than the interest in big data in a specific industry as evidenced by significant differences across industries (mean equal to 47.88). Real estate is the sector of activity in which there was a higher value of standard deviation (19.199), indicating a very heterogeneous behavior throughout the considered period. The findings also indicate that the interest in big data crosses multiple areas.

Figure 1 Interest in big data over time (beginning of 2016 to the end of 2020)



Source: Google Trends

Figure 2 Seasonality analysis from first quarter of 2016 to fourth quarter of 2020



Source: Developed by the authors

Table 2 Interest in big data by industry for January 2016 to December 2020

Industry	*Ranking*	*Mean*	*Median*	*SD*	*Sig.*
Arts and entertainment	4	57.81	57	14.966	$<1.10^{-3}$
Automobiles and vehicles	9	42.53	40	17.475	$<1.10^{-3}$
Beauty and fitness	15	31.43	30	15.414	$<1.10^{-3}$
Computers and electronics	2	68.30	67	11.605	$<1.10^{-3}$
Education	6	54.28	55	10.525	$<1.10^{-3}$
Finance	10	42.10	41	13.317	$<1.10^{-3}$
Food and drinks	7	53.19	52	15.648	$<1.10^{-3}$
Games	13	37.02	35	17.287	$<1.10^{-3}$
Government	5	57.75	57	16.841	$<1.10^{-3}$
Health	3	58.52	57	16.135	$<1.10^{-3}$
Leisure	12	38.96	37	15.033	$<1.10^{-3}$
Manufacturing	1	72.96	73	10.864	$<1.10^{-3}$
Real estate	16	25.19	24	19.199	$<1.10^{-3}$
Sport	11	39.10	37	16.777	$<1.10^{-3}$
Telecommunications	8	50.31	48	14.336	$<1.10^{-3}$
Travel	14	36.66	35	16.827	$<1.10^{-3}$
All industries	n/a	76.08	76	10.007	n/a

4.3 Growth rates

Table 3 shows that in a relatively generalized way the interest in the application of big data to specific areas has decreased. Significant declines are registered in areas such as arts and entertainment (−9.33%), education (−9.66%) or games (−7,67%). The decline in these industries is greater than the average drop in interest in big data for all industries (−6.33%). However, and despite a significant decline in interest for big data in most industries, the sport experienced a growth of 58% in the third quarter of 2020 and travel had a growth of 15.96% in the same period.

Table 3 Growth of interest in big data in 2020

Industry	First quarter		Second quarter		Third quarter		Fourth quarter	
	Mean	Growth (%)	Mean	Growth (%)	Mean	Growth (%)	Mean	Growth (%)
Arts and entertainment	58.38	–0.82	48.23	–17.39	46.31	–3.98	39.31	–15.12
Automobiles and vehicles	35.85	–4.97	33.00	–7.95	39.38	19.33	24.00	–39.06
Beauty and fitness	31.15	7.90	24.77	–20.45	31.92	28.74	22.54	–29.82
Computers and electronics	60.15	–9.55	58.15	–3.33	57.08	–1.84	52.77	–7.59
Education	51.54	–14.50	44.15	–14.36	48.46	9.76	39.00	–19.52
Finance	39.38	–16.71	37.15	–5.66	37.85	1.88	30.85	–18.49
Food and drinks	48.77	–15.80	47.46	–2.69	41.46	–12.64	44.69	7.72
Games	34.08	–18.15	28.85	–15.49	33.15	15.94	28.85	–12.97
Government	51.15	–23.83	55.62	8.82	48.15	–13.39	43.00	–10.70
Health	61.54	1.39	62.38	2.89	49.62	–19.85	47.77	–3.87
Leisure	36.92	–0.22	34.23	–7.56	36.54	6.46	29.15	–19.95
Manufacturing	70.15	–5.96	63.62	–9.35	61.54	–3.27	60.92	–1.01
Real estate	14.85	–46.38	19.46	3.12	19.31	–0.77	17.08	–12.18
Sport	28.85	–26.02	20.69	–28.28	32.69	58.00	32.85	0.49
Telecommunications	45.15	–19.59	42.00	–6.98	45.31	7.88	39.15	–13.71
Travel	30.38	–2.72	28.38	–6.58	32.85	15.96	35.15	7.00
All industries	70.15	–12.35	66.38	–5.33	66.46	0.12	61.31	–7.76

5. Discussion

A good starting point to discuss the relevance of the findings presented in this study was developed by Raguseo (2018), which explored the potential of big data and characterized the evolution of interest in the topic using Google Trends. Raguseo (2018) identified that from 2004 until the end of 2015, the interest in big data had an exponential growth from the fourth quarter of 2011 until the end of 2014. From this date, the growth was around only 14%, but interest in the community remained high (mean equal to 73) at the end of 2015. The work done in this study addresses and complements Raguseo's (2018) view and allows us to characterize the evolution of big data in the community as established in *RQ1*. The findings revealed that from the beginning of 2016 until the end of 2019, big data continued to have high community interest (mean equal to 79.21). However, from the beginning of 2020, there was a downward trend at an annual percentage rate of 17.41%, which addresses *RQ1* and indicates that the attention given by the community to big data has not experienced a homogeneous growth. The current literature does not directly address this phenomenon which seems contradictory when we look at the data from the Digital 2020 report discussed by Kemp (2020) which indicates that nearly 60% of the online population has access to the internet (e.g. 4.5 billion people use the internet and 3.8 billion of social media users). However, it is possible to identify reasons that indirectly allow us to understand this phenomenon and the findings of this study. Khvoynitskaya (2020) reports that in the next five years, the big data analytics field will grow significantly in areas such as machine learning, artificial intelligence and cybersecurity. A similar study in the same field, but considering a time horizon of 10 years, also highlights the growth of big data because of the growth of quantum data and supported by technological growth on the internet of Things (IoT) that will significantly increase the volume of data available for analysis (Insight Slice, 2020). COVID-19 is also an element that can explain this phenomenon as organizations needed to allocate resources to address the immediate challenges of digitization and remote working as indicated in the studies by Almeida *et al.* (2020) and Phillips (2020). Therefore, organizations have struggled to allocate resources and time to devote to strategic business growth areas. The conclusion we can draw is that the decrease in the interest of the big data community is not exactly because of a lack of interest in the phenomenon but because of the growth of interest in other areas complementary to big data such as IoT, Industry 4.0, machine learning and deep learning, among others.

The findings also revealed significant differences in interest in the big data phenomenon when looking at specific industries (significance level below 1×10^{-3}). The three sectors of activity that have aroused the most interest are manufacturing, computers and electronics and health. Furthermore, these data allow us to conclude that big data has raised interest in the community in general and that it often does not translate into an immediate application of big data technologies in the business environment. Indeed, as shown in the study by Mikalef *et al.* (2020), the business adoption of big data requires tangible (e.g. basic resources, data, technology), intangible (e.g. data-driven culture, organizational learning) and human skills (e.g. technical skills, managerial skills). Other key factors in its adoption also emerge. According to Lozada *et al.* (2019), collaboration is a key factor if the knowledge gained from data analysis is to be exploited commercially. Akhtar *et al.* (2019) add that big data requires an increasingly skilled workforce. Finding professionals in the market with interdisciplinary knowledge in areas such as mathematics, statistics and programming is a difficult task. As shown by the US Bureau of Labor Statistics (BLS) study, data science jobs are expected to grow by 12% in the next decade (BLS, 2020). Therefore, the adoption of big data poses challenges and opportunities for companies in various sectors of activity, although from *RQ2*'s perspective the manufacturing and computers and electronics industries stand out as having the greatest potential.

Big data has attracted great interest in the manufacturing sector mainly in studies that seek to relate its adoption with organizational performance, thereby highlighting its strategic role (O'Donovan *et al.*, 2015). The second line of research is also emerging that aims to explore the application of predictive and prescriptive models (Dubey *et al.*, 2019; Poornima and Pushpalatha, 2020). Through this data analysis, it becomes possible to improve manufacturing methods and product quality, also improving the final delivery to the customer. Finally, the third line of research is also identified that explores the role of big data in building the factories of the future (Gao *et al.*, 2020). At this level, data can be used by smart devices to proactively improve manufacturing processes and respond to new demands. Cloud computing, artificial intelligence and IoT are typically used in conjunction with big data, which suggest that the potential of big data is significantly greater when applied in conjunction with other technologies.

The average community interest in the first quarter of 2020 for the computers and electronics sector was equal to 60.15, which is 108% higher than the sports industry or 98% higher than the travel industry. The big data ecosystem is supported by a diverse set of technology platforms that can be organized into several groups according to the challenges they address. In a first group technologies for distributed storage and processing can be found. Research in this domain focuses on horizontal scalability and efficient large-scale distributed processing over a variety of different infrastructures (Avci *et al.*, 2020). In a second group, we find research lines that seek to explore technologies for building dynamic low-latency applications that use semistructured data and web applications with personalization through real-time analysis and updates (Talia, 2019). Finally, a third group explores the use of technology to implement real-time feed ingestion and analytics systems, real-time trading system and real-time processing of information from machines and sensors (Jabbar *et al.*, 2020).

The findings also reveal that overall interest in big data has decreased 6.33% over the year 2020. This decline is particularly relevant in areas where big data has traditionally held strong community interest, such as education, health and arts and entertainment. However, this data should be interpreted with caution, as it may not mean that big data has become less relevant to these industries. In recent years, there have been emerging technologies such as IoT or blockchain that make big data even more attractive to businesses. For example, as Taylor *et al.* (2020) point out, one of the main potentialities of blockchain is its application in the digital security field, because each record is individually encrypted which does not allow this data to be altered or deleted. Therefore, one of the possibilities that

blockchain allows, besides financial transactions, is the secure storage of strategic documents for a company. The two technologies (i.e. big data and blockchain) can be used together to provide more transparent and trustworthy information, as the system rejects input that is suspicious or cannot be verified (Deepa *et al.*, 2021). With this, the patterns of customer behaviors identified by the company become more genuine. In contrast, the exploration of *RQ3* allowed the identification of emerging areas of big data application such as real estate, sport and travel. At the real estate level, big data can be used to obtain information about customer preferences and interconnect with IoT and sensor networks to obtain information about the urban environment (Oluwunmi *et al.*, 2019); in sports, big data can be used to measure and improve the performance of athletes and obtain more complete and reliable information about opponents (Morgulev *et al.*, 2018); and in the travel industry, big data can offer benefits for conducting personalized marketing campaigns, more efficient operations and better customer satisfaction (Yallop and Seraphin, 2020).

Finally, the pandemic put strategic business decisions on hold because of the need to respond to the immediate challenges posed by COVID-19. In the year 2020, companies that had not anticipated the digitalization movement are more vulnerable. Among them are those that still lack technological structure and have systems without connectivity, documents and data only physical and little investment in technologies and digitalized processes (Almeida *et al.*, 2020). This pandemic has come to change the way business is done and generate a breakthrough in innovation and digital transformation. As stated by Tavoletti *et al.* (2021), the need for digitization of services and business models, which was already moving at a fast pace, grows as a large part of the world's population needs to stay at home. However, new habits of consumption and interaction with companies are here to stay and are increasingly supported on digital channels. In areas where big data typically raised less interest such as leisure and real estate now have new growth potential with online shows or remote visits to enterprises. Also, in traditional areas of strong interest for big data such as health, new specific markets are emerging such as telemedicine or tools that encourage a healthy lifestyle.

6. Conclusions

Big data has captured strong community interest with an average of 79.21 by the end of 2019. However, from the beginning of 2020 onward, there was a reduction of 17.41%. This drop was more accentuated in industries that have traditionally captured significant interest from the community, such as health, education and arts and entertainment. On the contrary, other areas that until the beginning of 2020 had recorded little interest such as real estate, sports and travel are emerging areas of application of big data.

Several industries have emerged that stand out with greater potential for big data application such as manufacturing and computers and electronics. These areas have sparked a high level of interest from the scientific community, which has developed studies to explore the role of big data in organizational strategy and company performance. Many emerging studies also look at the technological ecosystem of big data, namely, exploring architectures for the distributed storage and processing of large volumes of data and analyzing the data to draw useful insights for business.

The decrease in interest in big data over the year 2020, essentially because of the pressure on organizations to meet the direct and immediate challenges posed by COVID-19, does not mean that big data has lost interest for companies. On the contrary, its role may be even more relevant when combined with other emerging technologies such as IoT and blockchain. IoT emerges to facilitate the connectivity process, making devices even more connected and producers of information, which will increase the volume of data that needs to be analyzed to have business value. On the other hand, blockchain will increase the security of the data storage process and, together with big data, will offer more transparent and reliable information about the business. Furthermore, the acceleration of digitization

processes driven by COVID-19 will cause the volume and variety of data to grow significantly. This offers new potential for organizations to adopt big data in areas where digitization processes were lagging.

6.1 Economic contributions

This study highlights the role that big data can play for organizations to become competitive within the context in which they operate. Within a digital context, where data-driven industries operate, the market power of organizations is more dynamic and difficult to measure. Business competitiveness is largely driven by the ability of organizations to collect and use information gathered internally and externally. In the digital economy, especially in the context of rapidly expanding dynamic markets and the growing advance of the data market, the possibility of greater detail about target audience profiles has emerged, allowing organizations to anticipate their needs and offer products/services that are increasingly segmented and oriented to customer needs. Furthermore, big data enables organizations in various industries to gather rapid feedback from consumers, thus allowing them to adjust business strategies almost in real time, attracting even more users as a result. Accordingly, the presence of network effects and strong economies of scale can prevent the entry of new competitors, especially in markets with lower growth rates and in which the adoption of big data is already more popular. In this sense, the analysis performed in this study enables organizations to anticipate a strategy in the adoption of big data at a time when they had to respond to immediate challenges of digitization of their activities because of the pandemic and in which the adoption of emerging technologies (e.g. IoT, artificial intelligence, blockchain) will cause the adoption of big data to become even more important for the competitiveness of these companies.

6.2 Social contributions

The technological revolution brought the awareness that, more than the concern with the product, it is necessary to worry about how to do business. In this sense, technology emerges as a vehicle to promote innovation and facilitate previously complex and challenging processes. The revolution in data analysis should not look exclusively at the economic perspective and potential in each industry but should also have as its main objective to make data act for the good of humanity, in an increasingly precise way in initiatives with social impact and also in actions of nonprofit organizations. There are many benefits that big data can bring from data collection, processing and analysis in the social dimension, such as inspiring solutions to improve people's quality of life, increase financial inclusion, make streets safer, reduce environmental impacts and create a network to fight corruption. The challenge for organizations is not to look exclusively at the individual potential of big data in each of the industries, but to realize points of synergy between them to build solutions for social inclusion that can meet the challenges of multidimensional sustainability and corporate social responsibility of organizations.

6.3 Limitations and future research directions

This study presents some limitations that it is relevant to highlight. The information collected comes only from Google Trends and, therefore, in future studies it is important to consider other sources of complementary information, such as surveys to better understand the attractiveness of big data in specific industries. Another limitation of the study is that the reasons for the growth of big data in industries that traditionally showed less potential were not explored. In this sense, and as future work, it would be relevant to conduct studies with these industries to understand this recent phenomenon that was identified in this study. Similarly, another interesting future work would be to explore how emerging technologies such as IoT and blockchain will influence scientific and technical developments in the big data field.

References

Akhtar, P., Frynas, J.G., Mellahi, K. and Ullah, S. (2019), "Big data-savvy teams' skills, big data-driven actions and business performance", *British Journal of Management*, Vol. 30 No. 2, pp. 252-271.

Albright, S.C. and Winston, W.L. (2014), *Business Analytics: Data Analysis*, Cengage Learning, Boston, MA.

Almeida, F. (2018), "Big data: concept, potentialities and vulnerabilities", *Emerging Science Journal*, Vol. 2 No. 1, pp. 1-10.

Almeida, F., Santos, J.D. and Monteiro, J.A. (2020), "The challenges and opportunities in the digitalization of companies in a post COVID-19 world", *IEEE Engineering Management Review*, Vol. 48 No. 3, pp. 97-103.

Amado, A., Cortez, P., Rita, P. and Moro, S. (2018), "Research trends on big data in marketing: a text mining and topic modeling based literature analysis", *European Research on Management and Business Economics*, Vol. 24 No. 1, pp. 1-7.

Anshari, M., Almunawar, M.N., Lim, S.A. and Al-Mudimigh, A. (2019), "Customer relationship management and big data enabled: personalization & customization of services", *Applied Computing and Informatics*, Vol. 15 No. 2, pp. 94-101.

Apuke, O.D. (2017), "Quantitative research methods: a synopsis approach", *Kuwait Chapter of Arabian Journal of Business and Management Review*, Vol. 6 No. 11, pp. 40-47.

Atov, I., Chen, K.C., Kamal, A.E. and Louta, M. (2020), "Data science and artificial intelligence", *IEEE Communications Magazine*, Vol. 58 No. 10, pp. 10-11.

Avci, C., Tekinerdogan, B. and Athanasiadis, J.N. (2020), "Software architectures for big data: a systematic literature review", *Big Data Analytics*, Vol. 5 No. 1, pp. 1-53.

Berman, F., Rutenbar, R., Hailpern, B., Christensen, H., Davidson, S., Estrin, D., Franklin, M., Martonosi, M., Raghavan, P., Stodden, V. and Szalay, A. (2018), "Realizing the potential of data science", *Communications of the ACM*, Vol. 61 No. 4, pp. 65-72.

BLS (2020), "Computer and information technology occupations", available at: www.bls.gov/ooh/computer-and-information-technology/home.htm (accessed 24 February 2021).

Boone, T., Ganeshan, R., Jain, A. and Sanders, N.R. (2019), "Forecasting sales in the supply chain: consumer analytics in the big data era", *International Journal of Forecasting*, Vol. 35 No. 1, pp. 170-180.

Brodeur, A., Clark, A.E., Fleche, S. and Powdthavee, N. (2021), "COVID-19, lockdowns and well-being: evidence from Google Trends", *Journal of Public Economics*, Vol. 193, pp. 1-8.

Cabrera-Sánchez, J.P. and Villarejo-Ramos, A.F. (2019), "Factors affecting the adoption of big data analytics in companies", *Revista de Administração de Empresas*, Vol. 59 No. 6, pp. 413-427.

Cao, L. (2017), "Data science: a comprehensive overview", *ACM Computing Surveys*, Vol. 50 No. 3, pp. 1-42.

Dash, S., Shakyawar, S.K., Sharma, M. and Kaushik, S. (2019), "Big data in healthcare: management, analysis and future prospects", *Journal of Big Data*, Vol. 6 No. 1, pp. 1-25.

De Luca, L.M., Herhausen, D., Troilo, G. and Rossi, A. (2020), "How and when do big data investments pay off? The role of marketing affordances and service innovation", *Journal of the Academy of Marketing Science*, Vol. 49 No. 4, pp. 790-810.

Deepa, N., Quoc-Viet, P., Dinh, C.N., Bhattacharya, S., Prabadevi, B., Gadekallu, T.R., Maddikunta, P.K., Fang, F. and Pathirana, P.N. (2021), "A survey on blockchain for big data: approaches, opportunities, and future directions", available at: https://arxiv.org/pdf/2009.00858.pdf (accessed 25 February 2021).

Dekimpe, M.G. (2020), "Retailing and retailing research in the age of big data analytics", *International Journal of Research in Marketing*, Vol. 37 No. 1, pp. 3-14.

Dubey, R., Gunasekaran, A., Childe, S.J., Blome, C. and Papadopoulos, T. (2019), "Big data and predictive analytics and manufacturing performance: integrating institutional theory, resource-based view and big data culture", *British Journal of Management*, Vol. 30 No. 2, pp. 341-361.

Dunn, P.F. and Davis, M.P. (2017), *Measurement and Data Analysis for Engineering and Science*, CRC Press, Boca Raton, FL.

Esfahani, H.J., Tavasoli, K. and Jabbarzadeh, A. (2019), "Big data and social media: a scientometrics analysis", *International Journal of Data and Network Science*, Vol. 3 No. 3, pp. 145-164.

Favaretto, M., De Clercq, E., Schneble, C.O. and Elger, B.S. (2020), "What is your definition of big data? Researchers' understanding of the phenomenon of the decade", *PloS One*, Vol. 15 No. 2, p. e0228987.

Fosso, W.S., Gunasekaran, A., Akter, S., Ren, S.J., Dubey, R. and Childe, S.J. (2017), "Big data analytics and firm performance: effect of dynamic capabilities", *Journal of Business Research*, Vol. 70, pp. 356-365.

Frisk, J.E. and Bannister, F. (2017), "Improving the use of analytics and big data by changing the decision-making culture: a design approach", *Management Decision*, Vol. 55 No. 10, pp. 2074-2088.

Gandomi, A. and Haider, M. (2015), "Beyond the hype: big data concepts, methods, and analytics", *International Journal of Information Management*, Vol. 35 No. 2, pp. 137-144.

Gao, R., Wang, L., Helu, M. and Teti, R. (2020), "Big data analytics for smart factories of the future", *CIRP Annals*, Vol. 69 No. 2, pp. 668-692.

Gobble, M. (2013), "Resources: big data: the next big thing in innovation", *Research-Technology Management*, Vol. 56 No. 1, pp. 64-66.

Harerimana, G., Jang, B., Kim, J.W. and Park, H.K. (2018), "Health big data analytics: a technology survey", *IEEE Access*, Vol. 6, pp. 65661-65678.

Hariri, R.H., Fredericks, E.M. and Bowers, K.M. (2019), "Uncertainty in big data analytics: survey, opportunities, and challenges", *Journal of Big Data*, Vol. 6 No. 1, pp. 1-16.

Insight Slice (2020), "Big data analytics market – global market share, trends, analysis and forecasts, 2020-2030", available at: www.insightslice.com/big-data-analytics-market (accessed 24 February 2021).

Jabbar, A., Akhtar, P. and Dani, S. (2020), "Real-time big data processing for instantaneous marketing decisions: a problematization approach", *Industrial Marketing Management*, Vol. 90, pp. 558-569.

Judd, C.M., McClelland, G.H. and Ryan, C.S. (2017), *Data Analysis: A Model Comparison Approach to Regression, ANOVA, and Beyond*, Routledge, Abingdon.

Kemp, S. (2020), "Digital 2020: global digital overview", available at: https://datareportal.com/reports/digital-2020-global-digital-overview (accessed 24 February 2021).

Khvoynitskaya, S. (2020), "The future of big data: 5 predictions from experts for 2020-2025", available at: www.itransition.com/blog/the-future-of-big-data (accessed 24 February 2021).

Kurian, S.J., Bhatti, A., Alvi, M., Ting, H., Storlie, C., Wilson, P.M., Shah, N.D., Liu, H. and Bydon, M. (2020), "Correlations between COVID-19 cases and Google Trends data in the United States: a state-by-state analysis", *Mayo Clinic Proceedings*, Vol. 95 No. 11, pp. 2370-2381.

Lozada, N., Arias-Pérez, J. and Perdomo-Charry, G. (2019), "Big data analytics capability and co-innovation: an empirical study", *Heliyon*, Vol. 5 No. 10, pp. 1-7.

McAfee, A. and Brynjolfsson, E. (2012), "Big data: the management revolution", *Harvard Business Review*, Vol. 90 No. 10, pp. 60-68.

Mangono, T., Smittenaar, P., Caplan, Y., Huang, V.S., Sutermaster, S., Kemp, H. and Sgaier, S.K. (2021), "Information-seeking patterns during the COVID-19 pandemic across the United States: longitudinal analysis of Google Trends data", *Journal of Medical Internet Research*, Vol. 23 No. 5, pp. 1-16.

Mani, V., Delgado, C., Hazen, B.T. and Patel, P. (2017), "Mitigating supply chain risk via sustainability using big data analytics: evidence from the manufacturing supply chain", *Sustainability*, Vol. 9 No. 4, pp. 1-21.

Mikalef, P., Boura, M., Lekabos, G. and Krogstie, J. (2019), "Big data analytics and firm performance: findings from a mixed-method approach", *Journal of Business Research*, Vol. 98, pp. 261-276.

Mikalef, P., Krogstie, J., Pappas, I. and Pavlou, P. (2020), "Exploring the relationship between big data analytics capability and competitive performance: the mediating roles of dynamic and operational capabilities", *Information & Management*, Vol. 57 No. 2, pp. 1-15.

Milne, D. and Watling, D. (2019), "Big data and understanding change in the context of planning transport systems", *Journal of Transport Geography*, Vol. 76, pp. 235-244.

Morgulev, E., Azar, O.H. and Lidor, R. (2018), "Sports analytics and the big-data era", *International Journal of Data Science and Analytics*, Vol. 5 No. 4, pp. 213-222.

Nuti, S.V., Wayda, B., Ranasinghe, I., Wang, S., Dreyer, R.P., Chen, S.I. and Murugiah, K. (2014), "The use of Google Trends in health care research: a systematic review", *PLoS One*, Vol. 9 No. 10, p. e109583.

O'Donovan, P., Leahy, K., Bruton, K. and O'Sullivan, D.T. (2015), "Big data in manufacturing: a systematic mapping study", *Journal of Big Data*, Vol. 2 No. 1, pp. 1-22.

Oluwunmi, A.O., Role, B.A., Akinwale, O.M., Oladayo, O.P. and Afolabi, T.O. (2019), "Big data and real estate: a review of literature", *Journal of Physics: Conference Series*, Vol. 1378 No. 3, pp. 1-7.

Oussous, A., Benjelloun, F.Z., Lahcen, A. and Belfkih, S. (2018), "Big data technologies: a survey", *Journal of King Saud University – Computer and Information Sciences*, Vol. 30 No. 4, pp. 431-448.

Phillips, S. (2020), "Working through the pandemic: accelerating the transition to remote working", *Business Information Review*, Vol. 37 No. 3, pp. 129-134.

Poornima, S. and Pushpalatha, M. (2020), "A survey on various applications of prescriptive analytics", *International Journal of Intelligent Networks*, Vol. 1, pp. 76-84.

Power, D.J. (2014), "Using 'big data' for analytics and decision support", *Journal of Decision Systems*, Vol. 23 No. 2, pp. 222-228.

Preis, T., Moat, H.S. and Stanley, H.E. (2013), "Quantifying trading behavior in financial markets using Google Trends", *Scientific Reports*, Vol. 3 No. 1, pp. 1-6.

Raguseo, E. (2018), "Big data technologies: an empirical investigation on their adoption, benefits and risks for companies", *International Journal of Information Management*, Vol. 38 No. 1, pp. 187-195.

Raja, A. and Prema, S. (2018), "A review on big data mining", *International Journal of Applied Engineering Research*, Vol. 13 No. 22, pp. 16021-16026.

Redondo, T. and Sandoval, A.M. (2016), "Text analytics: the convergence of big data and artificial intelligence", *International Journal of Interactive Multimedia and Artificial Intelligence*, Vol. 3 No. 6, pp. 57-64.

Reimer, A.P. and Madigan, E.A. (2019), "Veracity in big data: how good is good enough", *Health Informatics Journal*, Vol. 25 No. 4, pp. 1290-1298.

Rogers, S. (2016), "What is Google Trends data and what does it mean?", available at: https://medium.com/google-news-lab/what-is-google-trends-data-and-what-does-it-mean-b48f07342ee8 (accessed 5 February 2021).

Sardi, A., Sorano, E., Cantino, V. and Garengo, P. (2020), "Big data and performance measurement research: trends, evolution and future opportunities", *Measuring Business Excellence*.

Saritas, O., Bakhtin, P., Kuzminov, I. and Khabirova, E. (2021), "Big data augmentated business trend identification: the case of mobile commerce", *Scientometrics*, Vol. 126 No. 2, pp. 1553-1579.

Satish, L. and Yusof, N. (2017), "A review: big data analytics for enhanced customer experiences with crowd sourcing", *Procedia Computer Science*, Vol. 116, pp. 274-283.

Sen, D., Ozturk, M. and Vayvay, O. (2016), "An overview of big data for growth in SMEs", *Procedia – Social and Behavioral Sciences*, Vol. 235, pp. 159-167.

Seyedan, M. and Mafakheri, F. (2020), "Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities", *Journal of Big Data*, Vol. 7 No. 1, pp. 1-22.

Siliverstovs, B. and Wochner, D.S. (2018), "Google Trends and reality: do the proportions match?: appraising the informational value of online search behavior: evidence from Swiss tourism regions", *Journal of Economic Behavior & Organization*, Vol. 145, pp. 1-23.

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

Subudhi, B.N., Rout, D.K. and Ghosh, A. (2019), "Big data analytics for video surveillance", *Multimedia Tools and Applications*, Vol. 78 No. 18, pp. 26129-26162.

Swift, L. and Piff, S. (2014), *Quantitative Methods: For Business, Management and Finance*, Red Globe Press, New York, NY.

Taleb, I., Serhani, M.A., Bouhaddioui, C. and Dssouli, R. (2021), "Big data quality framework: a holistic approach to continuous quality management", *Journal of Big Data*, Vol. 8 No. 1, pp. 1-41.

Talia, D. (2019), "A view of programming scalable data analysis: from clouds to exascale", *Journal of Cloud Computing*, Vol. 8 No. 4, pp. 1-16.

Tavoletti, E., Kazemargi, N., Cerruti, C., Grieco, C. and Appolloni, A. (2021), "Business model innovation and digital transformation in global management consulting firms", *European Journal of Innovation Management*, doi: 10.1108/EJIM-11-2020-0443.

Taylor, P.J., Dargahi, T., Dehghantanha, A., Parizi, R.M. and Choo, K.K. (2020), "A systematic literature review of blockchain cyber security", *Digital Communications and Networks*, Vol. 6 No. 2, pp. 147-156.

TechAmerica Foundation's Federal Big Data Commission (2012), "Demystifying big data: a practical guide to transforming the business of government", available at: https://bigdatawg.nist.gov/_uploadfiles/M0068_v1_3903747095.pdf (accessed 21 February 2021).

Thirathon, U., Wieder, B., Matolcsy, Z. and Ossimitz, M.L. (2017), "Big data, analytic culture and analytic-based decision making evidence from Australia", *Procedia Computer Science*, Vol. 121, pp. 775-783.

Timoneda, J. and Wibbels, E. (2021), "Spikes and variance: using Google Trends to detect and forecast protests", *Political Analysis*, Vol. 30 No. 1, pp. 1-18.

Verdenhofs, A. and Tambovceva, T. (2019), "Evolution of customer segmentation in the era of big data", *Marketing and Management of Innovations*, Vol. 1, pp. 238-243.

West, R. (2020), "Calibration of Google Trends time series", *Proceedings of the 29th ACM International Conference on Information & Knowledge Management (CIKM'20), Dublin, Ireland*, pp. 2257-2260.

Wolfert, S., Ge, L., Verdouw, C. and Bogaardt, M.J. (2017), "Big data in smart farming – a review", *Agricultural Systems*, Vol. 153, pp. 69-80.

Yallop, A. and Seraphin, H. (2020), "Big data and analytics in tourism and hospitality: opportunities and risks", *Journal of Tourism Futures*, Vol. 6 No. 3, pp. 257-262.

Yoseph, F., Heikkilä, M., Brezulianu, A., Geman, O. and Rostam, N.A. (2020), "The impact of big data market segmentation using data mining and clustering techniques", *Journal of Intelligent & Fuzzy Systems*, Vol. 38 No. 5, pp. 6159-6173.

Zhan, Y., Tan, K.H. and Tse, Y.K. (2018), "Unlocking the power of big data in new product development", *Annals of Operations Research*, Vol. 270 Nos 1/2, pp. 577-595.

Further reading

Walimbe, P. (2016), "To overcome communication challenges in distributed/virtual scrum teams", available at: http://digitalcommons.harrisburgu.edu/pmgt_dandt/4 (accessed 15 December 2020).

Corresponding author

Fernando Almeida can be contacted at: almd@fe.up.pt