Advancing beyond technicism when managing big data in companies' decision-making

Francesco Caputo, Barbara Keller, Michael Möhring, Luca Carrubbo and Rainer Schmidt

Abstract

Purpose – *In recognising the key role of business intelligence and big data analytics in influencing companies' decision-making processes, this paper aims to codify the main phases through which companies can approach, develop and manage big data analytics.*

Design/methodology/approach – *By adopting a research strategy based on case studies, this paper depicts the main phases and challenges that companies "live" through in approaching big data analytics as a way to support their decision-making processes. The analysis of case studies has been chosen as the main research method because it offers the possibility for different data sources to describe a phenomenon and subsequently to develop and test theories.*

Findings – *This paper provides a possible depiction of the main phases and challenges through which the approach(es) to big data analytics can emerge and evolve over time with reference to companies' decision-making processes.*

Research limitations/implications – *This paper recalls the attention of researchers in defining clear patterns through which technology-based approaches should be developed. In its depiction of the main phases of the development of big data analytics in companies' decision-making processes, this paper highlights the possible domains in which to define and renovate approaches to value. The proposed conceptual model derives from the adoption of an inductive approach. Despite its validity, it is discussed and questioned through multiple case studies. In addition, its generalisability requires further discussion and analysis in the light of alternative interpretative perspectives.*

Practical implications – *The reflections herein offer practitioners interested in company management the possibility to develop performance measurement tools that can evaluate how each phase can contribute to companies' value creation processes.*

Originality/value – *This paper contributes to the ongoing debate about the role of digital technologies in influencing managerial and social models. This paper provides a conceptual model that is able to support both researchers and practitioners in understanding through which phases big data analytics can be approached and managed to enhance value processes.*

Keywords *Big data, Big data analytics, Companies' decision-making, Smarter management*

Paper type *Technical paper*

Francesco Caputo is based at the Department of Economics, Management and Institutions, University of Naples Federico II, Naples, Italy.
Barbara Keller is based at the Duale Hochschule Baden-Württemberg Stuttgart, Stuttgart, Germany.
Michael Möhring is based at the Department of Informatics – HHZ Reutlingen University, Reutlingen, Germany.
Luca Carrubbo is based at the Department of Management and Innovation Systems, University of Salerno, Salerno, Italy.
Rainer Schmidt is based at the Department of Computer Science and Mathematics, University of Applied Sciences Munich, Munich, Germany.

Received 8 October 2022
Revised 26 January 2023
Accepted 25 February 2023

Corrigendum: It has come to the attention of the publisher that the article: Caputo, F., Keller, B., Möhring, M., Carrubbo, L. and Schmidt, R. (2023), "Advancing beyond technicism when managing big data in companies' decision-making", *Journal of Knowledge Management*, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/JKM-10-2022-0794 did not accurately display Möhring, M.'s affiliation.

Our guidelines state that affiliations should be supplied in full when the article is submitted.

The city corresponding to Reutlingen University has been amended from Munich to Reutlingen.

1. Preliminary reflections

In the past few decades, socio-economic configurations have profoundly changed because of the increasing use and accessibility of Information and Communication Technologies (ICT) in multiple domains of everyday life (Forester, 1987; Turban *et al.*, 1998; Drucker, 2011; Caputo *et al.*, 2019b). Consolidated views based on the representation of technologies for data management as "simple instruments" for supporting decision-making activities have progressively shown that they are incapable of explaining ongoing dynamics and trends (Caputo *et al.*, 2019c). Similarly, new interpretative approaches and managerial models are strongly required by researchers and practitioners interested in effectively understanding what the main implications, consequences and effects of the increasing use of ICT in business and social dynamics are (Castells, 1999; Markus and Topi, 2015).

[2015] Building upon this widely recognised need, in recent decades, a challenging debate has emerged around the topic of big data analytics as "a way of extracting value from these huge volumes of information, and it drives new market opportunities and maximizes customer retention" (Zakir *et al.*, 2015, p. 81). Several contributions have been provided with reference to the multiple advantages that it is possible to obtain for companies from a "new" approach in the collection, coding and management of data related to the multiple dimensions of shopping expeditions and evaluations (Griffin *et al.*, 2000; Mummalaneni, 2005; Demangeot and Broderick, 2006; Amendola *et al.*, 2018; Ardito *et al.*, 2018). Multiple stimuli for reflections have also been provided with reference to the ways in which people, processes and technologies can be combined to improve the quality of companies' and markets' approaches in data collection and use (Alter, 2006; Singh and Del Giudice, 2019).

As effectively summarised by Demchenko *et al.* (2012, p. 614), "Data Science is becoming a new technology driver and requires re-thinking a number of infrastructure, components, solutions and processes to address the following general challenges: Exponential growth of data volume produced by different research instruments and/or collected from sensors; Need to consolidate e-Infrastructure as [a] persistent research platform to ensure research continuity and oration, deliver/offer persistent services, with [an] adequate governance model." According to the authors' reflections, the challenging domain about big data should mainly refer to the infrastructure and processes required for ensuring the effective collection and organisation of a huge volume of data.

Despite the relevance of the aforementioned dimensions, it only represents a "small" part of the multiple reflections that seem to require the ongoing transitions towards a knowledge era based on technology infrastructure. Several relevant elements related to human approaches to big data, the consequences of big data analytics in companies' decision-making processes and the antecedents capable of addressing the ongoing digital transition (Caputo *et al.*, 2019a; Chinnaswamy *et al.*, 2018), among others, seem to be vastly underestimated. Accordingly, the paper proposes extending current perspectives in the study of big data analytics by focusing attention on the intriguing domain of big data analytics, specifically "the extraction of hidden sight about consumer behaviour from big data and the exploitation of that insight through advantageous interpretation" (Erevelles *et al.*, 2016, p. 897). Thanks to the adoption of a research strategy based on case studies, the paper aims to depict the main phases that companies face in the process of reshaping decision-making processes through big data analytics. The analysis of case studies has been chosen as the main research method because it offers the possibility for different data sources to describe a phenomenon and subsequently to develop and test theories.

The paper is structured as follows. In Section 2, the theoretical background will be presented by focusing attention on smart management and on the role of big data analytics in companies' decision-making processes as relevant domains with reference to which proposed reflections have been developed. In Section 3, the method and data collection of the proposed research will be reported, whilst in Section 4, the results of the proposed research will be summarised to enrich the current debate about the role of big data analytics in reshaping companies' decision-making processes. Finally, in Section 5, the study's preliminary conclusions, main limitations, implications and possible future directions will be presented.

2. Theoretical background

The way in which organisations apply data analysis has changed over time (Chen *et al.*, 2012). In recent years, different methods have been developed that depend on the different data sources and related data structures.

In general, different data sources with structured and/or unstructured data can be part of big data projects (Gandomi and Haider, 2015). In the past, enterprises were only able to analyse structured datasets like customer order data coming from, for example, CRM or ERP systems (Chen *et al.*, 2012). The data used for analyses mainly consisted of numbers or categorial variables, for example. The way of collecting, storing and analysing data was less complex in comparison to more recent data sources containing unstructured data (Buneman *et al.*, 1997; Blumberg and Atre, 2003; Baker and Thien, 2020; Del Giudice *et al.*, 2021). Today, however, up to 90% of the collected data is unstructured data like texts, images, audio and video (Harbart, 2021). The analysis of unstructured data is currently challenging organisations because of its unsuitability for use in conventional data models (Harbart, 2021). The use of unstructured data together with structured data is manifold. For instance, it can be used to improve the quality and the possibilities of prediction within big data analytics (Davenport *et al.*, 2021). Nevertheless, the more data types are included in analytical projects, the more different methods must be used. Today, more and more IoT-related data sources like connected home appliances (Bayer *et al.*, 2020) or services like Google Popular times (Möhring *et al.*, 2020) can be used to predict and better understand customer behaviour. These new data sources must be integrated within the analytical landscape to be used in related analysis. Another interesting use case that highlights the challenges of the benefits of big data analytics is product returns in e-commerce. This field is even more important because it meets both customer behaviour and the sustainability concept, as well as helping to easily understand the facets appearing in big data analysis. For instance, if an organisation wants to use online customer reviews (unstructured textual data) to predict the product returns probability (Schmidt and Möhring, 2013; Möhring *et al.*, 2013), past customer order data from the CRM and ERP system (structured data) as well as images (unstructured image data) from offered goods should also be integrated into the analysis to enhance the quality of the prediction. Therefore, they must apply different methods like text mining for textual data, image pattern recognition for images and traditional data mining techniques like regression or correlation analysis. In turn, this means that different results, various key figures and quality criteria must be aggregated and harmonised within one comprehensive result (Kaur *et al.*, 2019).

Furthermore, the data must be stored in different locations like relational databases for the order data and/or within NoSql databases (Stonebraker, 2010) like document-based databases for textual data. In sum, all these requirements will increase the complexity of big data analytics projects and generate challenges for organisations running an analytical project. In line with the identified methodological complexity and storing issues, the computational complexity also increases. The more variables are included in analytical approaches, the more steps for information processing and result calculation are necessary. Therefore, organisations that are considering applying big data analytics must explore the option of scalable public cloud computing services at major sites like Amazon AWS, Microsoft Azure and Google Cloud to capture the limitations of traditional non-scalable systems (Schmidt and Möhring, 2013).

2.1 Challenges and dynamics of smart management

Nowadays, the dynamics in decision-making in all contexts are increasingly guided and conditioned by the reception, filtering, processing and use of data (Raisinghani, 2000). The evolution of new technologies favours the development of virtuous processes [thanks to big data analytics techniques, data mining, machine learning, artificial intelligence (AI), etc.] that support decision-making processes (Nutt, 2008; Yang *et al.*, 2019). The growing uncertainty in all application areas accentuates the importance of the way in which decisions are made, especially if they involve significant consequences for the community. Decision making is a multidisciplinary topic that lends itself to different levels of analysis when we focus on the various elements (including technological ones) that condition or facilitate it (Papadakis *et al.*, 1998).

Decision-making processes are increasingly data-driven. Therefore, the decisions are more "informed" because the exchange of information is rapid (often in real-time); hence, it can be precise, punctual, efficient and valid. From the electronic medical record (now fully operational) to the development of information systems to new communication protocols, it is possible to record a continuous flow of data, information (and the contained in it) and the inputs to be filtered, processed, used and managed in a timely manner (Sharma *et al.*, 2014). The risk of "data-deluge" and the difficulty of having useful elements is very high, while the possibility of making quick, accurate, thoughtful decisions becomes more and more necessary, indeed fundamental (Sabherwal and King, 1995; Citroen, 2011). In this sense, the evolution of decision-support-systems (DSS) assumes increasing importance in many critical "moments", both for descriptive-analytics (e.g. diagnostics, evaluations and monitoring), as well as in the follow-up analytics in the operational phases and even for forecasting possible choices in the future and related reasons through predictive analytics and prescriptive analytics (Boonstra, 2003).

In general, information sharing with shared databases, data-storage, data extraction and data processing favours the design of a more functional, versatile, scalable, context-friendly service provision, where the smart management can make a difference thus deserves to be further explored. For this, it becomes important to study the main characteristics of data that can be acquired. Here, the so-called "10V[s] of big-data" (Volume, Velocity, Variety, Veracity, Value, Validity, Variability, Venue, Vocabulary and Vagueness) are often taken into consideration to understand how new knowledge is generated and, consequently, how much decision-making processes are affected; particularly with reference to the possible advantages of meta dating, data modelling, architecture and data integration (Manogaran *et al.*, 2022).

Similarly, the most frequently used methods are studied to improve decision-making from data management's point of view. Typical topics of interest here are cloud computing for information sharing, artificial intelligence for the data interpretation available and the generation of new ones like data mining and machine learning. The aim here is to better understand how the information flow works, what criticalities it presents, how it feeds the activation and management of known protocols, how it integrates the various data-sources and how it supports the management of queries (Hicks *et al.*, 2006).

All this effectively integrates decision-making techniques (cost benefits, grid analysis, paired comparison, compensatory strategies, etc.), with particular reference to conditions of uncertainty because of, for example, systematic errors, cognitive biases, risk situations, external distortions, information asymmetries, misalignments, internal friction, misunderstandings, technical or administrative misunderstandings, legal aspects, technological crashes or even weak signals escaping, somatic markers and negative contingencies.

These issues are so fundamental and interesting that in the period between 2021 and 2027, European investments will be geared towards building a smarter Europe through innovation, digitalisation, economic transformation and support for small- and medium-sized enterprises. EIT Digital has launched the 2022 call to promote entrepreneurship and education for the construction of a strong digital Europe and contribute to the development of digital technology, digital industry, digital cities, digital wellbeing and digital finance.

Since 2014, the European Commission has spoken out in favour of a thriving data-driven economy (European Commission, 2014); in 2015, it discussed a strategy for the digital single market in Europe (European Commission, 2015). In 2018, the International Data Corporation estimated an increase of 16 trillion gigabytes of data, with an annual growth rate of 236% in terms of data generation to date; they linked this to the fact that decisions based on knowledge generated by big data can lead to increased productivity and competitiveness and GDP (equal to 1.9% by 2020) (Reinsel *et al.*, 2018).

Today, the evolving trend of Big Data Analyses is an integral part of a new digital market. According to the European Commission, it guarantees the development of innovative and competitive business models. However, while having to comply with the EU data protection framework, big data can involve significant risks and challenges, especially in fundamental rights like privacy and data protection. More recently, the European Parliament discussed the role of the data-based economy in the strategy for the digital union against the background of all stakeholders and their daily life situations, such as consumers (ease of use, efficiency and savings), businesses (industry 4.0) and public administration (e-government), housing (smart cities), science, medicine (Mhealth), disaster response capacity and the fight against crime, etc.

2.2 Big data in companies' decision-making processes

Scientists and researchers have long since faced the challenges of data management, focusing their attention on possible ways to collect data both directly and indirectly (Sapsford and Jupp, 1996; Hajian and Domingo-Ferrer, 2012). Several experiments have been conducted aiming to define the processes and protocols that enhance the effectiveness of data collection as a relevant way to extend consolidated knowledge about the reasons, antecedents and motivations behind actors' behaviours and decisions in multiple domains (Grant and Mayer, 2009; Guiot and Roux, 2010; Daunt and Harris, 2012; Rahrovani and Pinsonneault, 2020). Along this line, studies focusing on companies' decision-making have also been developed and multiple approaches for collecting and analysing data have been investigated (Goulding, 1999; Rokka and Uusitalo, 2008; Paço and Lavrador, 2017).

Nowadays, all these approaches and contributions seem to be outmoded against the background of the disruptive role of big data analytics in the data and knowledge management processes (Pauleen and Wang, 2017). Today, big data infrastructure supports the handling of data operations by facilitating the source's integration and collaboration in real time with high standards for control and data safety (Sagiroglu and Sinanc, 2013).

Demchenko *et al.* (2014, p. 105) reports "the Big Data definition as having the following 5V properties: Volume, Velocity, Variety that constitute native/original Big Data properties, and Value and Veracity as acquired as a result of data['s] initial classification and processing in the context of a specific process or model." These properties effectively summarise the relevant contributions that big data can provide the management of a high volume of data in real time without "damaging" the granularity of information to ensure a realistic representation of the phenomenon (Polyakova *et al.*, 2019).

According to Erevelles *et al.* (2016), the properties of big data seem to provide a valuable solution for organisations striving to find an answer to environmental and social changes through predictive approaches about market trends. More comprehensively, big data offers organisations the opportunities to increase:

- their *dynamic capabilities* – their "ability to respond to change incorporates skills and knowledge embedded within the organization to alter existing resources and create new value" (Erevelles *et al.*, 2016, pp. 898–899); and
- *adaptive capabilities* – as capabilities that do not derive "from a specific change in organizational structure but from the overall ability to capture consumer activities and extract hidden in-sights" (Erevelles *et al.*, 2016, p. 899).

Recognising the disruptive role of big data in reinventing firms' market approaches, it is possible to underline its contribution in supporting enterprises in innovating their relationships with the market by focusing on the "implementation of creative ideas" (Gumusluoglu and Ilsev, 2009, p. 61). From this perspective, big data analytics can be seen as a valuable approach that supports firms to enforce their relationship by focusing on the definition of the innovation management path based on their "ability to effectively acquire and exploit new information" (Chaston *et al.*, 2001, p. 147). Data acquisition and exploitation became the bridge with the capacity to link innovation management, information management and market analysis under the common umbrella of big data analytics; this offers the opportunity to understand current interest in developing an effective model for information management, allowing firms to better understand (and predict) market trends and expectations based on big data analytics (Erevelles *et al.*, 2016).

In a nutshell, big data can be considered a disruptive innovation (Caputo *et al.*, 2017) that is potentially able to reinvent firms' approach to market analysis. Accordingly, Davenport *et al.* (2012, p. 43) stated that big data supports firms "to understand their business environments at a more granular level, [...] creating new products and services, and [...] responding more quickly to change as it occurs." As a result, a new challenge emerges concerning how to decode the pattern for companies' decision-making processes through big data analytics.

3. Method and data collection

With the aim to enrich current debate about the role of big data in companies' decision-making, a case study approach was set as the research strategy (Kohlbacher, 2016). The reasons why this approach was chosen are multi-faceted. On the one hand, the approach follows the recommendations of Yin (2003), who described the importance of case study research when a contemporary phenomenon is investigated in its real-world setting, and the boundaries between the phenomena itself and the related context are blurred. As a matter of fact, this method allows for a variety of research methods (Yin, 2003; Kohlbacher, 2016). Case studies allow researchers to combine different data sources (such as interviews, texts and observations), as well as using qualitative and quantitative data analysis. Therefore, they can be used to describe a phenomenon and Subsequently to develop and test theories (Darke *et al.*, 1998).

A widespread procedure is to use case studies in qualitative inquiries (Stake, 2000; Kohlbacher, 2016). This is especially relevant in contexts where the "why" and the "how" of a phenomenon are the focus of an investigation. Consequently, a case study research strategy with a qualitative inquiry thus seems to be an appropriate approach for an investigation and the provision of new insights. It is therefore unsurprising that case studies are an appropriate and popular way of investigating the implementation and use of information systems within organisations. This is particularly true in information systems research and related scientific areas, in which it is quite important to examine and understand the context of the phenomenon, because often researchers are unclear about how a phenomenon arises or how individuals' experiences and doings are critical to its actions and effects. Furthermore, numerous research approaches demand that with regards to the research question the number and topic of the cases must be determined at the outset. Whilst a single case study is applied to gain deep and rich insights, multi-case studies have the advantage of allowing replications (literal, theoretical) and comparisons between cases (Darke *et al.*, 1998).

Here, a topic highly related to information systems research is investigated. Besides managerial and human factors, the research question also aims to understand the technical issues and their related problems. Following the recommendations given in the literature, as described previously, a multiple case study research strategy was chosen as an appropriate approach in line with our research question. As the research focuses on different aspects, a single case study approach did not seem to be appropriate to best gain the desired insights about the subject. Therefore, multiple cases were investigated by collecting different data from different sources and conducting a qualitative analysis (Yin, 1994, 2012).

Consequently, three different cases were examined. The investigated cases were a manufacturing enterprise, an enterprise from the IT sector and a supplier for IT solutions. It is assumed that all the branches are equally affected by the challenges of implementing big data analytics. In addition, the cases highlight and clarify that all sectors are affected by the challenges of Big Data Analysis. The IT sector is no exception. The investigated enterprises have different sizes and turnovers. This circumstance is useful in terms of the generalisability of the findings. More details about the companies' characteristics are reported in Table 1.

In all cases, the process to implement the possibility of big data analytics was accompanied and supported by at least one of the researchers. As a result, a minimum of one person was involved as an "action researcher" within the organisations (Walsham, 1995). Subsequently, both the data and the contextual insights gathered are very rich and useful. Every case was comprehensively investigated and hence a strong understanding of the phenomenon was achieved (Darke *et al.*, 1998). Furthermore, the action researchers accompanied different big data analytics projects within the companies chosen as cases. This allowed them to prove and control the generalisability of the insights and findings in different settings (Darke *et al.*, 1998). As recommended in the literature, different data sources such as observations, interviews and questionnaires were picked-up and combined (Darke *et al.*, 1998). An overview about the data sources used in this investigation is provided in Table 1.

For the data analysis, the Grounded Theory approach was conducted (Strauss and Corbin, 1994). This approach is very common and widespread in Information Systems research (Aarnikoivu *et al.*, 2019). In the first step, the open coding process was conducted. The data was investigated, and the relevant aspects were tagged with abstract labels. This step is followed by the so-called axial coding process. As the second step of the procedure, the axial coding process examines the relationships between the labels and tries to build networks containing relevant aspects. Hence, the identified labels were aggregated and networks were built. In the third step, selective coding was applied, meaning that the networks were subsumed into categories. In each step, all the team members did the coding process alone and the results were discussed afterwards.

4. Results

The data analysis revealed that in all cases along the project's timeline specific patterns occurred at special points in time. The findings are summarised in Table 2 and explained in more detail subsequently.

Phase (a): Nearly all enterprises have recognised that the customer data they own is a hidden gem. Hence, it is not surprising that companies want to exploit this potential. Consequently, organisations have recognised the need for big data analytics to realise the benefits provided by the data. Often, the top management takes the initiative to create plans for big data analytics projects. They set ambitious goals and objectives that frequently consist of a mix of dreams, wishes and reality. In many cases, the intended big

Table 1 Overview of case studies and data gathering process

	Enterprise no. 1 (case 1)	*Enterprise no. 2 (case 2)*	*Enterprise no. 3 (case 3)*
Sector	Manufacturing	IT	IT solution supplier
Company size	Large	Medium	Small
No. employees	>550	>200	63
Turnover	~200 Mio€	~200 Mio €	~5 Mio €
Observations by accompanying/supportive researcher	x	x	x
Cross-divisional e-mail traffic	x	x	x
Interviews and expert talks	x	x	x
Surveys	x		

Source: Authors' elaboration

Table 2 Main results about companies' approach to big data

Phase (a.): Before/at the beginning of the project	*Phase (b.): During the project*	*Phase (c.): At the end/ finalization of the project*
Need for big data analytics	Staff with adequate knowledge is missing or cannot be found	Not all requirements/ automation tasks can be fulfilled
Mix-up of dreams, wishes and reality	Missing openness/restricted mindset	Predictions by the algorithms are not always better than the human ones
Budget and available staff	Data sources (e.g. databases) do not fit	Usability issues
Implementation/time horizon	Identification of the best Big Data algorithm(s)	Time, costs and effort was underestimated (run of time and budget)
Trust in databases	Must re-design the project and re-start	
	IT infrastructure is old and not flexible	
	Data protection rules	

Source: Authors' elaboration

data analytics projects are not realisable for several reasons. Firstly, the company lacks concrete processes, possibilities and outcomes along with the initial vague and imaginative assumptions. Hence, big data analytics projects begin similarly and specific requirements are often not respected because of the company's inexperience with such projects. Subsequently, wrong estimations in terms of budget and staffing, as well as time and scope occur. In addition, some of the most prominent aspects in big data analytics projects are also neglected. Furthermore, the availability of data is a crucial factor that is often misjudged. Organisations trust in their databases. However, it is not uncommon for data to be unusable because of poor data management and questionable data quality. There are also often assumptions about data sources that do not, in fact, exist in the reality of the company. In one of the cases in this study, an expert in case (1) stated that the management proclaimed that all the needed data is stored and available in their proAlpha ERP system. However, it turned out that this was a false estimation from the management. Even if the data is available, wrong judgement can be taken as case (3) revealed. The responsible persons in case (3) assumed that they have high quality data about their customers and their behaviour. Although data about the customers was available, it did not meet the requirements. Relevant aspects of customers' behaviour were missing and, therefore, the potential for the analysis was quite restricted.

Phase (b): Once a project is started, challenges because of human factors, as well as technical issues arise. On the human side, the challenges are two fold. On the one hand, it might be that the assigned employees did not have the relevant knowledge for conducting the project or cannot be identified. During the project, the management of case (1) discovered that their internal staff were not able to implement the AI models into their systems. Therefore, they had to find an external service provider who was able to cope with this challenge. On the other hand, missing openness and/or a restricted mindset are a critical human factor too. This often results in staff hiding their knowledge to avoid changes that could lead to more work or that has a negative impact on their job position.

Besides challenges occurring because of human factors, we also observed technical aspects that were crucial for the continuation of big data analytics projects. On the technical side, it might be the case that the database does not even contain the expected data or that the data did not fit the requirements, as described previously. In many cases, the missing data cannot be procured because the IT infrastructure is too old and inflexible. There are missing interfaces, hence new analytical systems can connect to it and collect the data needed for analysis (in cases 1, 2 and 3). Modern standard application interfaces like REST-APIs (Masse, 2011) were not provided, which hindered the seamless collection of data. Furthermore, the implementation of modern big data analytic and data visualisation tools into old systems might be difficult.

Both human and technical factors might stop or delay the project. In all cases (1)–(3), it was hard to find the correct experts with business and domain specific know-how. In cases (1) and (3), often the most suitable employees for the task were also not known by the management. Sometimes, a step back to the first phase was needed to re-define the responsibilities and even the technical possibilities. In cases (1) and (3), the project had to be restarted (a). In case (1), adjustments during the project were done. Hence, the aim of the project must be reviewed and re-defined. Another aspect that sometimes occurs is that the best algorithm cannot be found. In all cases (1)–(3), there was no generally available algorithm or approach fitting the project's goal that would deliver a result within the expected quality range from the very beginning. Furthermore, the available IT infrastructure resources (e.g. CPU, RAM, disk) for the analysis hindered the evaluation of different algorithms. For example, (sample) data was split and patterns were reconstructed to evaluate the algorithms. Different algorithms were combined in all cases to accommodate issues such as linear and non-linear behaviour (e.g. linear regression and neuronal networks) and selected based on different rules (rule-based algorithm selection and combination), as well as patterns that could only be identified during the actual data analysis. For instance, after starting the project the responsible persons in case (1) found out that their systems could not be used to run analytical services. They did not anticipate in advance that the necessary infrastructure capabilities (e.g. CPU/RAM) would be missing.

Phase (c): In the final project phase, further patterns were identified within the selected cases. Regarding the definition and targets of the big data projects at the beginning of the project, not all requirements and automation tasks could be fulfilled. This is often a consequence of the fact that the challenges from the two preceding project phases could not be sufficiently taken into account. In cases (1) and (3), only a minor set of requirements could be fulfilled because of the issues in the prior project phases. It was only in case (2) that important requirements during the project could be delivered. Sometimes, the prediction of human experts with years of experience is faster and more accurate compared to the developed systems. This might mean that not all the necessary data is available and the data behaviour patterns may not be recognised by the system. This was particularly true of case (1), where the system was not accurate compared to experienced experts. The developed big data systems are very complex. Therefore, their usability and user friendliness are severely limited. Experts must configure the systems in advance by entering specific parameters. Consequently, the staff must be trained to use the system and to interpret the results with regards to the specific business demands. In all cases (1)–(3), the effort needed to complete the project, in terms of, for instance, time, costs and budget, was underestimated. In cases (1) and (3), the project ran out of time and budget and had to be adjusted. Again, this might be a consequence of the identified patterns in the first two phases of the projects (a) and (b).

5. Conclusions, implications, limitations and future research

In the past few years, big data and big data analytics tools have been presented as the new "miracle" for efficiency, survival and increased performance for any type of organised entities (Schmarzo, 2013). These approaches attracted the interest of multiple researchers and the investment of multiple companies interested in the possibility of obtaining multiple advantages by simply buying new instruments, software and digital devices. Despite the summarised scenarios, the proposed research shows different scenarios in which collected and analysed data demonstrate that the predictions made by the algorithms do not naturally offer value in isolation. Sometimes, human predictions are even better because they can involve more variable factors and are more intuitive.

In such a perspective, the research offers several practical implications because it underlines how automation may not even be possible, and several manual steps are needed as the usability of the tool decreases. Sometimes, users cannot work with the system because it is hard to handle or because they are not able to interpret the output of the system and relate it to adequate strategical or operational measures. In addition, because of delays and re-definitions the project may run out of time and budget. Thus, the expenses overcome the estimated benefit. Sometimes, projects must even be abandoned. Furthermore, issues related, e.g. to the technical foundation of the enterprises, used algorithms and data quality hinder a good implementation and positive value of the system.

In the same perspective, the research also underlines several theoretical implications by ascertaining that to run a big data analytics project successfully it is important to focus on the challenges and anticipate consequences. Therefore, current interpretative paths and managerial models require radical rethinking to better catch and depict the interconnections that could be possible between humans and technology.

Despite the conceptual and empirical advancements in the knowledge offered by the reflections herein, several limitations can be identified with reference to the proposed research approach because the results offered by the analyses of the case studies are subjective and related to the background in which they have been approached and analysed. In such a vein, the next steps for the research are required to test to what extent the proposed results and observations can be generalised to different cognitive and geographical domains.

References

Aarnikoivu, M., Nokkala, T., Siekkinen, T., Kuoppala, K. and Pekkola, E. (2019), "Working outside academia? Perceptions of early-career, fixed-term researchers on changing careers", *European Journal of Higher Education*, Vol. 9, pp. 172-189.

Alter, S. (2006), *The Work System Method: Connecting People, Processes, and IT for Business Results*, Work System Method.

Amendola, C., Calabrese, M. and Caputo, F. (2018), "Fashion companies and customer satisfaction: a relation mediated by information and communication technologies", *Journal of Retailing and Consumer Services*, Vol. 43, pp. 251-257.

Ardito, L., Scuotto, V., Del Giudice, M. and Petruzzelli, A.M. (2018), "A bibliometric analysis of research on big data analytics for business and management", *Management Decision*, Vol. 57 No. 8, pp. 1993-2009.

Baker, O. and Thien, C.N. (2020), "A new approach to use big data tools to substitute unstructured data warehouse", *2020 IEEE Conference on Big Data and Analytics (ICBDA)*, IEEE, pp. 26-31.

Bayer, S., Gimpel, H. and Rau, D. (2020), "IoT-commerce-opportunities for customers through an affordance lens", *Electronic Markets*, Vol. 31 No. 1, pp. 27-50.

Blumberg, R. and Atre, S. (2003), "The problem with unstructured data", *Dm Review*, Vol. 13, pp. 42-49.

Boonstra, A. (2003), "Structure and analysis of IS decision-making processes", *European Journal of Information Systems*, Vol. 12 No. 3, pp. 195-209.

Buneman, P., Davidson, S., Fernandez, M. and Suciu, D. (1997), "Adding structure to unstructured data", International Conference on Database Theory, Springer, pp. 336-350.

Caputo, F., Cillo, V., Candelo, E. and Liu, Y. (2019a), "Innovating through digital revolution: the role of soft skills and big data in increasing firm performance", *Management Decision*, Vol. 57 No. 8, pp. 2032-2051.

Caputo, F., Evangelista, F., Perko, I. and Russo, G. (2017), "The role of big data in value co-creation for the knowledge economy", in Vrontis, S., Weber, T., Tsoukatos, E. (Eds), *Global and National Business Theories and Practice: bridging the past with the Future*, EuroMed Press, pp. 269-280.

Caputo, F., Garcia-Perez, A., Cillo, V. and Giacosa, E. (2019b), "A knowledge-based view of people and technology: directions for a value co-creation-based learning organisation", *Journal of Knowledge Management*, Vol. 3 No. 7, pp. 1314-1334.

Caputo, F., Walletzky, L. and Štepánek, P. (2019c), "Towards a systems thinking based view for the governance of a smart city's ecosystem", *Kybernetes*, Vol. 48 No. 1, pp. 108-123.

Castells, M. (1999), *The Social Implications of Information and Communication Technologies*, *UNESCO's World Social Science Report*.

Chaston, I., Badger, B. and Sadler-Smith, E. (2001), "Organizational learning: an empirical assessment of process in small UK manufacturing firms", *Journal of Small Business Management*, Vol. 39 No. 2, pp. 139-151.

Chen, H., Chiang, R.H. and Storey, V.C. (2012), "Business intelligence and analytics: from big data to big impact", *MIS Quarterly*, pp. 1165-1188.

Chinnaswamy, A., Papa, A., Dezi, L. and Mattiacci, A. (2018), "Big data visualisation, geographic information systems and decision making in healthcare management", *Management Decision*, Vol. 57 No. 8, pp. 1937-1959.

Citroen, C.L. (2011), "The role of information in strategic decision-making", *International Journal of Information Management*, Vol. 31 No. 6, pp. 493-501.

Darke, P., Shanks, G. and Broadbent, M. (1998), "Successfully completing case study research: combining rigour, relevance and pragmatism", *Information Systems Journal*, Vol. 8 No. 4, pp. 273-289.

Daunt, K.L. and Harris, L.C. (2012), "Motives of dysfunctional customer behavior: an empirical study", *Journal of Services Marketing*, Vol. 26 No. 4, pp. 293-308.

Davenport, T.H., Barth, P. and Bean, R. (2012), "How big data is different", *MIT Sloan Management Review*, Vol. 54 No. 1, pp. 43-46.

Davenport, T., Guszcza, J., Smith, T. and Stiller, B. (2021), *Analytics and AI-Driven Enterprises Thrive in the Age of With*, Deloitte Insights.

Del Giudice, M., Scuotto, V., Papa, A., Tarba, S.Y., Bresciani, S. and Warkentin, M. (2021), "A self-tuning model for smart manufacturing SMEs: effects on digital innovation", *Journal of Product Innovation Management*, Vol. 38 No. 1, pp. 68-89.

Demangeot, C. and Broderick, A.J. (2006), "Exploring the experiential intensity of online shopping environments", *Qualitative Market Research: An International Journal*, Vol. 9 No. 4, pp. 325-351.

Demchenko, Y., De Laat, C. and Membrey, P. (2014), "Defining architecture components of the big data ecosystem", *2014 International Conference on Collaboration Technologies and Systems (CTS)*, *IEEE*, pp. 104-112.

Demchenko, Y., Zhao, Z., Grosso, P., Wibisono, A. and De Laat, C. (2012), "Addressing big data challenges for scientific data infrastructure", *4th IEEE International Conference on Cloud Computing Technology and Science Proceedings*, *IEEE*, pp. 614-617.

Drucker, P.F. (2011), *Technology, Management, and Society*, Harvard Business Press.

Erevelles, S., Fukawa, N. and Swayne, L. (2016), "Big data consumer analytics and the transformation of marketing", *Journal of Business Research*, Vol. 69 No. 2, pp. 897-904.

European Commission (2014), "Communication from the commission to the European parliament, the council, the european economic and social committee and the committee of the regions", available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52014DC0442andfrom=EN

European Commission (2015), "Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of the regions", A Digital Single Market Strategy for Europe, available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52015DC0192andfrom=EN

Forester, T. (1987), *High-Tech Society: The Story of the Information Technology Revolution*, MIT Press.

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.

Goulding, C. (1999), "Consumer research, interpretive paradigms and methodological ambiguities", *European Journal of Marketing*, Vol. 33 Nos 9/10, pp. 859-873.

Grant, A.M. and Mayer, D.M. (2009), "Good soldiers and good actors: prosocial and impression management motives as interactive predictors of affiliative citizenship behaviors", *Journal of Applied Psychology*, Vol. 94 No. 4, pp. 900-920.

Griffin, M., Babin, B.J. and Modianos, D. (2000), "Shopping values of Russian consumers: the impact of habituation in a developing economy", *Journal of Retailing*, Vol. 76 No. 1, pp. 33-52.

Guiot, D. and Roux, D. (2010), "A second-hand shoppers' motivation scale: antecedents, consequences, and implications for retailers", *Journal of Retailing*, Vol. 86 No. 4, pp. 355-371.

Gumusluoglu, L. and Ilsev, A. (2009), "Transformational leadership, creativity, and organizational innovation", *Journal of Business Research*, Vol. 62 No. 4, pp. 461-473.

Hajian, S. and Domingo-Ferrer, J. (2012), "A methodology for direct and indirect discrimination prevention in data mining", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 25 No. 7, pp. 1445-1459.

Harbart, T. (2021), "Tapping the power of unstructured data", MIT Sloan Management School, available at: https://mitsloan.mit.edu/ideas-made-to-matter/tapping-power-unstructured-data

Hicks, B.J., Culley, S.J. and McMahon, C.A. (2006), "A study of issues relating to information management across engineering SMEs", *International Journal of Information Management*, Vol. 26 No. 4, pp. 267-289.

Kaur, S., Gupta, S., Singh, S.K. and Perano, M. (2019), "Organizational ambidexterity through global strategic partnerships: a cognitive computing perspective", *Technological Forecasting and Social Change*, Vol. 145, pp. 43-54.

Kohlbacher, F. (2016), "The use of qualitative content analysis in case study research", *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, Vol. 7 No. 1, pp. 1-30.

Manogaran, G., Thota, C. and Lopez, D. (2022), "Human-computer interaction with big data analytics", Research Anthology on Big Data Analytics, Architectures, and Applications, IGI Global, pp. 1578-1596.

Markus, M.L. and Topi, H. (2015), *Big Data, Big Decisions for Science, Society, and Business*, National Science Foundation.

Masse, M. (2011), *REST API Design Rulebook: designing Consistent RESTful Web Service Interfaces*, O'Reilly Media.

Möhring, M., Keller, B., Schmidt, R. and Dacko, S. (2020), "Google popular times: towards a better understanding of tourist customer patronage behavior", *Tourism Review*, Vol. 76 No. 3, pp. 553-593.

Möhring, M., Walsh, G., Schmidt, R., Koot, C. and Härting, R.C. (2013), "Returns management in eCommerce", *HMD*, Vol. 50 No. 5, pp. 66-75.

Mummalaneni, V. (2005), "An empirical investigation of web site characteristics, consumer emotional states and on-line shopping behaviors", *Journal of Business Research*, Vol. 58 No. 4, pp. 526-532.

Nutt, P.C. (2008), "Investigating the success of decision making processes", *Journal of Management Studies*, Vol. 45 No. 2, pp. 425-455.

Paço, A. and Lavrador, T. (2017), "Environmental knowledge and attitudes and behaviours towards energy consumption", *Journal of Environmental Management*, Vol. 197, pp. 384-392.

Papadakis, V.M., Lioukas, S. and Chambers, D. (1998), "Strategic decision-making processes: the role of management and context", *Strategic Management Journal*, Vol. 19 No. 2, pp. 115-147.

Pauleen, D.J. and Wang, W.Y. (2017), "Does big data mean big knowledge? KM perspectives on big data and analytics", *Journal of Knowledge Management*, Vol. 21 No. 1, pp. 1-6.

Polyakova, A., Loginov, M., Serebrennikova, A. and Thalassinos, E. (2019), "Design of a socio-economic processes monitoring system based on network analysis and big data", *International Journal of Economics and Business Administration*, Vol. 7 No. 1, pp. 30-139.

Rahrovani, Y. and Pinsonneault, A. (2020), "Innovative IT use and innovating with IT: a study of the motivational antecedents of two different types of innovative behaviors", *Journal of the Association for Information Systems*, Vol. 21 No. 4, pp. 5-14.

Raisinghani, M.S. (2000), "Knowledge management: a cognitive perspective on business and education", *American Business Review*, Vol. 18 No. 2, pp. 105-131.

Reinsel, D., Gantz, J. and Rydning, J. (2018), *The Digitization of the World. From Edge to Core*, An IDC White Paper, Seagate.

Rokka, J. and Uusitalo, L. (2008), "Preference for green packaging in consumer product choices–do consumer's care?", *International Journal of Consumer Studies*, Vol. 32 No. 5, pp. 516-525.

Sabherwal, R. and King, W.R. (1995), "An empirical taxonomy of the decision-making processes concerning strategic applications of information systems", *Journal of Management Information Systems*, Vol. 11 No. 4, pp. 177-214.

Sagiroglu, S. and Sinanc, D. (2013), "Big data: a review", *2013 International Conference on Collaboration Technologies and Systems (CTS)*, *IEEE*, pp. 42-47.

Sapsford, R. and Jupp, V. (Eds) (1996), *Data Collection and Analysis*, Sage.

Schmarzo, B. (2013), *Big Data: Understanding How Data Powers Big Business*, John Wiley and Sons.

Schmidt, R. and Möhring, M. (2013), "Strategic alignment of cloud-based architectures for big data", *2013 17th IEEE International Enterprise Distributed Object Computing Conference*, *IEEE*, pp. 136-143.

Sharma, R., Mithas, S. and Kankanhalli, A. (2014), "Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations", *European Journal of Information Systems*, Vol. 23 No. 4, pp. 433-441.

Singh, S.K. and Del Giudice, M. (2019), "Big data analytics, dynamic capabilities and firm performance", *Management Decision*, Vol. 57 No. 8, pp. 1729-1733.

Stake, R.E. (2000), "Case studies", in Denzin, N.K and Lincoln, Y.S (Eds), *Handbook of Qualitative Research*, Sage, pp. 435-453.

Stonebraker, M. (2010), "SQL databases v. NoSQL databases", *Communications of the ACM*, Vol. 53 No. 4, pp. 10-11.

Strauss, A. and Corbin, J. (1994), "Grounded theory methodology: an overview", in Denzin, N.K. and Lincoln, Y.S. (Eds), *Handbook of Qualitative Research*, Sage, pp. 273-285.

Turban, E., McLean, E. and Wetherbe, J. (1998), *Information Technology for Management Making Connections for Strategic Advantage*, John Wiley and Sons, Inc.

Walsham, G. (1995), "Interpretive case studies in IS research: nature and method", *European Journal of Information Systems*, Vol. 4 No. 2, pp. 74-81.

Yang, Q., Steinfeld, A. and Zimmerman, J. (2019), "Unremarkable AI: fitting intelligent decision support into critical, clinical decision-making processes", *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1-11.

Yin, R.K. (1994), "Designing single-and multiple-case. Improving educational management: through research and consultancy", in Bennett, N., Glatter, R. and Levacic, R. (Eds), *Improving Educational Management: Through Research and Consultancy*, Sage, pp. 135-155.

Yin, R.K. (2003), *Case Study Research, Design and Methods*, 3rd ed., Sage, Vol. 5.

Yin, R.K. (2012), "Case study methods", in Cooper, H., Camic, P.M., Long, D.L., Panter, A.T., Rindskopf, D. and Sher, K.J. (Eds), *APA Handbook of Research Methods in Psychology*, Vol. *2. Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological*, American Psychological Association, pp. 141-155.

Zakir, J., Seymour, T. and Berg, K. (2015), "Big data analytics", *Issues in Information Systems*, Vol. 16 No. 2, pp. 81-90.

Corresponding author

Francesco Caputo can be contacted at: francesco.caputo2@unina.it