

# Assessing the Impact of Big Data on Large Organizations' Strategies\*

Riccardo Rialti\*\*, Monica Faraoni\*\*\*,  
Lamberto Zollo\*\*\*\*, Irene Pieraccini\*\*\*\*\*

## Abstract

*Big data and big data analytics' (BDA) importance for business management is growing at an unmatched pace. Such datasets, and the techniques to analyse them, are indeed able to inundate organizations with huge quantities of precious information. Building on this premise, the purpose of this research is to develop a unitary holistic framework on the strategic impact of big data on large organizations. In order to do so, a cognitive map was developed from a coding analysis of 523 job offers on the European labour market coming from 124 large organizations. The results show the effects of big data and BDA on marketing strategies, strategic business intelligence and, finally, operation and supply chain management.*

**Keywords:** Big Data; Big Data Era; Big Data Marketing; Management Information Systems; Coding Analysis; NVivo; Strategic Management; Global Competition

## 1. The Reshaping Role of Big Data in Modern Organization

Big data is the name given to unstructured datasets characterized by dimensions and complexity that cannot be explored using any traditional statistical data analysis software (Erevelles et al., 2016; Manyika et al., 2011). Their importance has been highlighted by studies on business management and management information systems due to the extremely valuable information that these datasets may hold (McAfee et al., 2012). Scholars have assessed how the informative content of big data allows organizations to predict with a certain degree of reliability future sales trends, the forecast consumption of raw materials, and workforce performance (Chen et al., 2012). Yet, in relation to their complexity, the implementation of big data analytics (BDA) has been considered among the biggest challenges an organization could face (Manyika et al., 2011).

---

\* The Authors: Rialti, R. §§ 2, 3, Faraoni, M. § 4, Zollo, L. § 1, Pieraccini, I. § 5

\*\* Research Fellow in Management, University of Florence ([riccardo.rialti@unifi.it](mailto:riccardo.rialti@unifi.it))

\*\*\* Associate Professor of Management, University of Florence ([monica.faraoni@unifi.it](mailto:monica.faraoni@unifi.it))

\*\*\*\* Assistant Professor of Management, University of Florence ([lamberto.zollo@unifi.it](mailto:lamberto.zollo@unifi.it))

\*\*\*\*\* Research Fellow in Management, University of Florence ([irene.pieraccini@unifi.it](mailto:irene.pieraccini@unifi.it))

Despite this growing attention, however, scant attention has been paid to holistic investigation of the strategic use of big data and BDA in modern organizations (Davenport et al., 2015; Erevelles et al., 2016; Rialti et al., 2019a; Rialti et al., 2019b). In detail, while several papers have observed how BDA capabilities (personal, managerial and infrastructural capabilities) could affect organizational performance, literature is still missing a comprehensive view of the impact that more skilled personnel could have on the overall organization. Therefore, starting from this gap, the purpose of this research is to explore how the presence of personnel skilled in BDA within modern large organization could reshape the organization and its functioning, and then affect performance. In particular, the objective of the research is to develop a framework on the impact of big data and BDA skilled personnel moving from qualitative data. Large organizations have been chosen as the context of this research as they are the organizational typology most affected by big data and the ones that can obtain the greatest benefits from these datasets. Large organizations are in fact the organizations that can usually more easily implement BDA processes like (1) data acquisition, (2) data cleansing, (3) data integration, (4) data modelling, and (5) data interpretation (Cillo et al., 2019; Labrinidis & Jagadish, 2012). Furthermore, the implementation of big data analytical processes requires a considerable investment. Therefore, only large organizations can afford to take advantage of these resources frequently without having significant problems covering investments for BDA processes. Finally, large organizations are the organizational typology that can obtain the greatest benefits from the implementation of processes for big data analysis (Motamarri et al., 2017). On the one hand, they are usually the only type of organization that can collect quantities of information sufficiently large to be classified as big data. Moreover, the greatest benefits associated are usually linked to efficiency gains deriving from the reduction of production wastage, and these kinds of benefits are usually easier to see in organizations that produce many types of product and operate in different markets (De Mauro et al., 2016, 2017; Camussone & Biffi, 2018).

After this brief introduction, the present research is structured as follows. The next section will review the existing literature on big data in management science. In particular, the focus will be on defining the expression 'big data', reviewing the existing studies in terms of their strategic potential, and exploring the professional figures involved. The third section describes the methodology that was used. Finally, the third and fourth sections present the discussion of results, conclusions and suggestions for future research.

## **2. The Increasing Attention to Big Data in Management Science**

The advent of big data generated considerable interest among scholars dealing with economic and managerial disciplines. Many authors focused on identifying the potential of big data availability for business management (Chen et al., 2012; McAfee et al., 2012; Provost & Fawcett, 2013). Nevertheless, the research conducted so far is uneven and too anchored to study of the impact of this phenomenon on specific areas. In this sense, the need to provide unitary frameworks capable of holistically representing its strategic potential emerges.

Starting from this premise, in the following subsections, we deal first with identifying the characteristics; subsequently, the professional figures necessary for organizations to be able to derive benefits from those, will be explored.

## 2.1 A Taxonomy of the Expression ‘Big Data’

As previously suggested, it is possible to define as big data any set of unstructured data that is too large and complex to be analysed using traditional information technology tools (Manyika et al., 2011). Therefore, any dataset that cannot be analysed by means of rigid computer programs such as Excel, SPSS or Stata can potentially be identified as “big data”. To analyse big data, an advanced knowledge of non-relational databases (e.g. NoSQL) and computer programming languages (e.g. Python, Java, C++) is necessary.

Volume, velocity and variety were the first three characteristics identified by scholars (McAfee et al., 2012). These are also known as the 3Vs. In addition to these historical features, four other modern features have recently been identified: veracity, variability, visualization and value. The four additional characteristics, together with the three historical ones, have been called the 7Vs that distinguish big data (Mishra, 2015; Zaïane, 2015).

In relation to the historical characteristics, McAfee et al. (2012) state that the difference from traditional datasets is the volume. In fact, only datasets exceeding one petabyte ( $10^6$  gigabytes) in terms of size can be classified as big data (McAfee et al., 2012). The choice of this unit as a reference is consistent with the technological evolution. For example, Facebook every day accumulates more than 100 petabytes of data deriving from the actions of online users (Kaplan & Haenlein, 2010). However, a dataset cannot be classified as big data only according to its volume: huge structured datasets should be considered only as large datasets, which are datasets of unusual dimension (Erevelles et al., 2016). Moving on from this, the importance of the other historical characteristics of big data can be understood. Another aspect is the difference from traditional large datasets because their size can increase rapidly (velocity) and constantly (IBM, 2012). Finally, for a dataset to be classified as big data, it must be characterized by variety; therefore it must be unstructured – that is, not initially organized when it originated – and made of not necessarily homogeneous sub-components (Davenport et al., 2015). Tweets, posts and reactions on Facebook, feedback on Amazon or eBay, videos uploaded on YouTube, photos shared on Instagram, text messages and voice messages have been identified over the last decade as components of most big data (George et al., 2014; McAfee et al., 2012).

□ *Regarding the modern characteristics, IBM's report (2012) titled What is Big Data? has identified veracity as the fourth characteristic: it has been stated that only authentic data originating from reliable sources can be considered big data (Mishra, 2015). Therefore, it is possible to state that they are not wholly reliable if it is not possible to identify the process that generated them (Mishra et al., 2017). Regarding the characteristic of variability, on the one hand, it has emerged how*

*the sub-components of big data can derive from numerous inhomogeneous sources (McAfee et al., 2012). On the other hand, since the process of generating is constant due to the speed (velocity) that distinguishes them, the content can change quickly, following consumer preference changes, for example (Demchenko et al., 2014). In summary, it is possible to explain the fifth feature by describing them as dynamic datasets or as a continuous flow of information (Davenport et al., 2015; Brondoni & Zaninotto, 2018).*

Big data must also be characterized by potential visualization through computerized analysis technologies (Zaiane, 2015). The value of the information contained is the last modern feature (Everelles et al., 2016; Rialti et al., 2016).

In light of the stated characteristics, therefore, it is possible to classify as big data any dynamic and unstructured dataset that cannot be analysed through the use of traditional data analysis software programs and that is characterized by high information potential due to its origin from reliable sources.

Over the last decade, the literature dealing with managerial science has explored the importance of this phenomenon. A first stream of research focused on the methods by which big data can be treated and analysed. The topics of greatest interest have concerned the storage of big data, the advanced calculation methodologies to process them and, finally, the development of information technology architectures composed of multiple connected hardware. The solutions proposed for their analysis have been the use of techniques based on machine learning or systems based on cloud computing. It was observed how architectures centred on data lakes (which are huge cloud-based databases often shared between several organizations) can be a viable and technologically suitable solution. Meanwhile, a second stream of research dealt with investigation into their informative potential and how organizations must necessarily update themselves in order to be able to take advantage (Demchenko et al., 2014). In particular, the authors working in this field have highlighted how organizations must develop the technical skills needed to analyse big data, or try to increase the organizational BDA capabilities (Wamba et al., 2017). In addition, the literature has explored how organizations must also prepare ad hoc information analysis processes in order to analyse it (Labrinidis & Jagadish, 2012).

## **2.2 Professional Figures in the Big Data Era**

In order fully to take advantage of the information potential of big data, in addition to technologies and processes, the need for organizations to endow themselves with new professionals (Saccardi, 2003; Mishra, 2015) capable of analysing such data has recently emerged. Davenport and Patil (2012, p. 70) have identified the data scientist as “the sexiest job of the 21st century”. Given the various needs present in the organisation, the data scientist is only one of the many professional figures required. In this sense, the relevant literature has shown that the presence of data analysts (Davenport et al., 2015), data architects (Gardiner et al., 2017), data engineers (Provost & Fawcett, 2013) and data managers (Miller, 2014)

is also usually necessary for organizations that try to use big data to obtain competitive advantages.

The data analyst is technician who deals with the structuring and analysis of data; his/her skills are related to the programming capability needed to analyse and organize unstructured computer data (De Mauro et al., 2016). Differently, the data scientist is responsible for extracting the most significant content from the data and identifying the most relevant patterns from them. Since there is a link between the analysts and the upper hierarchical levels, the data scientist must also be able to render the data readable and must be equipped with management, marketing and finance skills (Davenport & Patil, 2012; De Castillo et al., 2015). The data architect, instead, is a figure who is concerned with organizing and maintaining the hardware structures for the collection, storage and analysis of big data (Gardiner et al., 2017). Similarly, the data engineer is a type of electronic-computer engineer who takes care of the software infrastructure (Provost & Fawcett, 2013). Finally, the data manager is designated to interpret the data and extract strategic decisions. Therefore, data managers are the people who actually make strategic data-driven decisions – that is, decisions based on both big data analysis results and managerial insight (Ciappei & Cinque, 2015) – in order to turn big data information into competitive advantages (Provost & Fawcett, 2013).

### **2.3 The Fragmented Literature on Big Data in Organizations and Competitive Strategies**

As mentioned, during the last decade the literature that deals with business management has begun to explore critically the impact of big data on organizations and on competitive strategies (Davenport & Patil, 2012). Specifically, scholars have focused on identifying the impact of big data on traditional business intelligence processes (Chen et al., 2012), on marketing strategies (Erevelles et al., 2016), on operations management (LaValle et al., 2011) and on investment management (Brown et al., 2011).

Regarding their impact on business intelligence, Chen et al. (2012) have recently identified big data as the trigger for the transition from Business Intelligence 2.0 – a business intelligence paradigm based on the use of semi-structured datasets and typical Web 2.0 technologies (Doan et al., 2011) – to Business Intelligence 3.0, which is instead based on the analysis in real time of unstructured data coming from users' activities and machinery sensors (Chen et al., 2012). With this new technological paradigm, the possibility emerged to base the analyses directly on the single customer, product and process; to have data concerning the various locations of the individuals in whom the organization is interested; and, finally, to carry out context-relevant analyses on every aspect concerning business performance (Chen et al., 2012). Moreover, thanks to the use of cloud-based analysis and computation systems aimed at analysing big data, analyses can be performed in real time (Demirkan & Delen, 2013). To be able to take advantage of all the benefits associated with Business Intelligence 3.0, the need to adapt internal and external information systems in a 3.0 perspective has emerged among the organizations (Goes et al., 2014).



As a result of the evolution of business intelligence systems, the other potentialities of these resources have recently been identified. In fact, thanks to big data and their analysis systems, it is possible to carry out surveys directly on individuals and now marketers can carry out sales campaigns aimed at the individual customer (Erevelles et al., 2016). Moreover, thanks to the speed with which the data reach the organization and the speed of calculation of the 3.0 information systems, customized marketing campaigns can be performed almost in real time (Provost & Fawcett, 2013). Therefore, big data allow managers to offer their customers combinations of products and services more pertinent to their requests, thus improving the organization-customer relationship (Mayer-Schönberger & Cukier, 2013; Barile & Polese 2018). Thanks to the information potential of big data, moreover, marketing campaign planning managers can today try to extrapolate forecasts about the future behaviours of consumers observing a high number of previous behaviours (Erevelles et al., 2016; Zirpoli and Cabigiosu, 2018). In terms of the impact on the management of organizations' operations, on the other hand, the possibility of making a triple check on organizations' operations has emerged, thanks to the machineries' sensors and tracking technologies. Specifically, 3.0 information systems capable of analysing big data have made it possible to increase the efficiency of procurement and distribution processes (Waller & Fawcett, 2013). In addition, thanks to the sensors placed on the organization's machinery, which in turn are connected to the internet (Rieple & Pisano, 2015; Caputo et al., 2016), managers are able to monitor each stage of the production process constantly through consoles that are typical of Business Intelligence 3.0 systems, and act accordingly (Waller & Fawcett, 2013). Finally, it has been observed that thanks to big data – in particular, those deriving from customers' credit cards or concerning the credit history of subjects that have relations with the organization – it is possible for the organization to plan more favourable extension policies with customers and suppliers, thus improving the management of working capital (Kitchin, 2017).

Despite the high number of studies concerning the definition of big data, their impact on competitive business strategies and the professional figures required, the absence of frameworks able to represent the challenges related to the availability of big data holistically is evident, and the same can be said for big data analytics capabilities in large organizations (Erevelles et al., 2016; Trento et al., 2018). This research seeks to contribute to the literature on the relevance of big data in business management, with particular attention to large organizations, which are the type of organizations mostly impacted by the advent of these new elements and the ones that can benefit from these data the most.

Therefore, the research question underlying this work is the following:

□ *RQ: Is it possible to create a framework capable of portraying professional figures necessary in the era of big data, disciplines of business management impacted by big data and the strategic potential of this type of data?*

### 3. Methodology

In order to answer the aforementioned research question, a coding analysis was performed on a dataset composed of job offers, a procedure that is consistent with previous studies (De Mauro et al., 2016; Gardiner et al., 2017). Indeed, job offers have extraordinary information potential. In particular, since they contain the description of the offered position and the goals that the candidate will have to achieve, if well explored they may be able to reveal precious information about organizational strategies. Starting from these considerations, this research focuses on identifying the latent strategic function contained in the description of each offered position. In fact, following this procedure, it was possible to obtain the strategic functions of each figure. Then, it was possible to create a conceptual map indirectly inferring the strategic impact on the organization from each single strategic function identified in more than one job offer related to the same type of professional figure required.

#### 3.1 Data Collection Procedure and Creation of the Dataset

The final dataset used in this research was built through an iterative process divided into several phases. Firstly, 14,931 job vacancies for positions in Europe, containing the expression 'big data' in the job description were extracted using a web scraper software. As a preliminary step, we did a web scraping (De Mauro et al., 2016) from the LinkedIn site in order to create the initial unstructured dataset to be analysed and transformed into a usable database and from which to conceptualize the themes. The queries made to the web crawler using the Portia web scraper software, freely downloadable, were structured as follows: *Big\_Data U EU\_Country*. In order to collect data from all 28 member countries of the European Community, 28 queries were carried out, from time to time replacing *EU\_Country* with the country name. Job offers were collected from an online platform that aggregates those already existing on multiple sites or on social media. Only those offers that were present for less than a month were selected, thus reducing the dataset to 10,430 units. In fact, it was necessary to consider only recent offers in order to exclude expired offers, offers incorrectly posted by organizations or any duplicates. The offers included in the dataset correspond to those posted for less than one month as of 23 March 2017, the day the data collection ended. Furthermore, in order to have a homogeneous dataset in terms of language, non-English offers (3,544) and further duplicates (490) were eliminated. Although only job offers in English were selected, the dataset was composed of well distributed job offers in all countries. In fact, from the selection process described above, a dataset consisting of 467 job offers from all 28 countries resulted. Specifically, 71 are for positions in the United Kingdom, 70 in Germany, 45 in France and 28 Italy. The rest are distributed in the other 23 countries. Approximately 65 per cent of the job positions of our dataset were offered by multinational organizations. About 45 per cent came from organizations that operate directly in the information technology sector. Finally, since the goal was to identify the impact on large organizations, the fact that the job offer was posted by large organizations with more than 500 employees was considered as an inclusion parameter.

The final dataset consisted of 592 pages of text (Microsoft Word, single line spacing, Times New Roman pt. 10) and 185,945 words.

### 3.2 Preliminary Analysis

The collected data were analysed using NVivo 11 Plus software. NVivo 11 Plus is a Qualitative Data Analysis (QDA) software deemed particularly suitable for the identification of common patterns and common words within small and large volumes of qualitative data. In our case, we selected NVivo 11 Plus as it allows to organize non-numerical unstructured data, obtain information about linkages between words, and to develop models deriving from the interpretation of data. In order to have more information on the organization of the subsequent coding phase, an analysis of the 100 most common words in the dataset was performed (see Table 1 and Figure 1).

**Table 1:** *The 100 Most Frequent Words in the Dataset*

Word	Length	Count	Weighted Percentage
data	4	3504	2.81
experience	10	1605	1.28
business	8	1293	1.04
team	4	1023	0.82
big	3	936	0.75
work	4	867	0.69
skills	6	846	0.68
analytics	9	706	0.57
solutions	9	687	0.55
development	11	636	0.51
management	10	581	0.47
knowledge	9	573	0.46
working	7	551	0.44
new	3	497	0.40
technical	9	496	0.40
software	8	495	0.40
technology	10	488	0.39
services	8	458	0.37
ability	7	448	0.36
technologies	12	424	0.34
science	7	415	0.33
role	4	396	0.32
customer	8	393	0.31
strong	6	391	0.31
company	7	382	0.31
tools	5	379	0.30
design	6	373	0.30
support	7	371	0.30
learning	8	369	0.30
analysis	8	360	0.29
industry	8	333	0.27
projects	8	328	0.26



customers	9	316	0.25
systems	7	314	0.25
environment	11	312	0.25
product	7	310	0.25
marketing	9	308	0.25
teams	5	305	0.24
within	6	298	0.24
world	5	298	0.24
global	6	295	0.24
help	4	289	0.23
level	5	289	0.23
project	7	287	0.23
clients	7	286	0.23
machine	7	285	0.23
looking	7	284	0.23
sales	5	283	0.23
digital	7	282	0.23
job	3	280	0.22
client	6	279	0.22
cloud	5	279	0.22
one	3	279	0.22
across	6	274	0.22
using	5	273	0.22
jobs	6	271	0.22
develop	7	267	0.21
del	3	266	0.21
description	11	263	0.21
products	8	260	0.21
information	11	253	0.20
years	5	252	0.20
including	9	250	0.20
engineering	11	246	0.20
analytical	10	244	0.20
SQL	3	244	0.20
scientist	9	243	0.19
informatica	11	242	0.19
leading	7	242	0.19
requirements	12	240	0.19
communication	13	239	0.19
excellent	9	237	0.19
use	3	236	0.19
platform	8	233	0.19
solution	8	232	0.19
key	3	229	0.18
Etc.	3	227	0.18
degree	6	226	0.18
high	4	225	0.18
based	5	224	0.18
understanding	13	224	0.18
hadoop	6	222	0.18
insights	8	222	0.18
provide	7	220	0.18

best	4	219	0.18
opportunity	11	218	0.17
required	8	216	0.17
opportunities	13	213	0.17
good	4	212	0.17
part	4	209	0.17
python	6	207	0.17
well	4	207	0.17
intelligence	12	205	0.16
responsibilities	16	204	0.16
like	4	203	0.16
performance	11	203	0.16
value	5	203	0.16
research	8	201	0.16
build	5	200	0.16
lead	4	199	0.16

Source: Our processing obtained through NVivo11 Plus

**Figure 1: Word Cloud of the 100 Most Common Words**



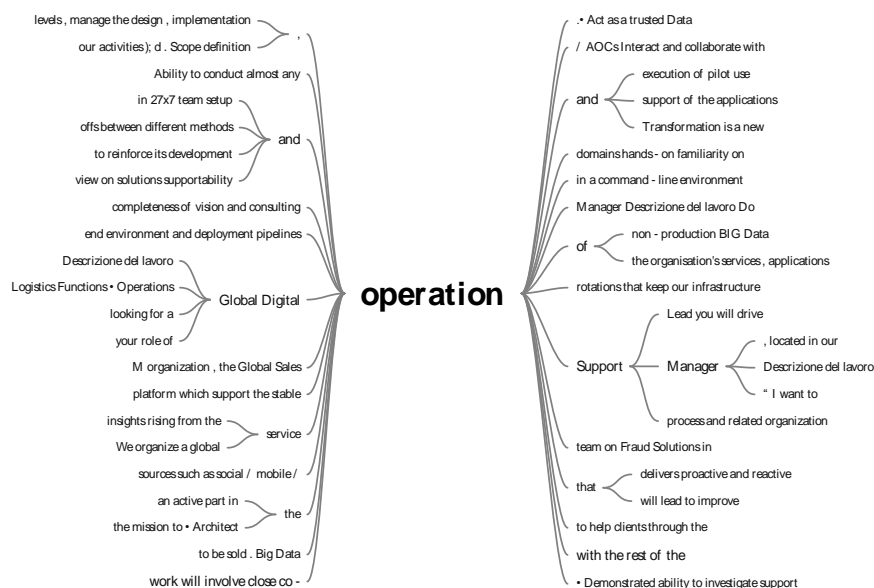
Source: Our processing obtained through NVivo11 Plus

The most frequent words that emerged from this preliminary analysis were ‘data’, ‘experience’, ‘business’, ‘team’, ‘big’, ‘work’, ‘skills’, ‘analytics’, ‘solutions’ and ‘development’. These terms were the starting points of the coding analysis, together with the professional figures required in the big data era and the impacted strategic areas already identified by the literature.

### 3.3 Coding Analysis and Framework Design

The dataset was analysed according to the key words. The first step was to identify the key words related to the main professions regarding big data, thus identifying the first level of the framework. Then, based on the main tasks, the strategic impacts of them were recognized. The coding analysis of the dataset was carried out through the word tree function of NVivo 11 Plus, which creates word trees through the analysis of the co-occurrences of words in the text (an example is shown in Figure 2). In the specific case of this research, each word tree has a theme as its starting point from which the branches representing the characteristics of the theme start. The use of co-occurrences as the initial nodes in word tree is a particularly useful method in the research as it allows to understand the linkages existing between two words through the considered text. Then, moving from several existing links it is possible to create more complex trees and models based on the existing associations. The five main professions represent the themes selected for the first level of the framework: data analyst, data scientist, data engineer, data architect and data manager. The branches that originate from here show the main tasks that these professions perform. Regarding the second level, those aspects of organizations' management most affected by the availability of big data, according to the prevailing literature, have been selected as themes: strategic marketing, human resource management, operations management, corporate finance and information management. The branches that originate from there contain information about the strategic impact of these data on large organizations (Gunasekaran et al., 2017; Wamba et al., 2015; Büchi et al., 2018).

**Figure 2:** Example of a Word Tree Obtained Using 'Operation' as a Keyword for the Query



Source: Our processing obtained through NVivo11 Plus

Therefore, it is possible to describe the whole process, from the identification of the keywords to the conceptualization phase. We followed the original four subsequent phases characterizing the so-called documentary coding method (Wrightson, 1976):

- 1) The text encoding – that is, the identification of the main variables (nodes of the map and word trees starting points), in turn categorized in this specific case as main themes, characteristics and managerial implications. It is worth noting that in this study we already had an absolute starting point represented by the professional figures (which are common to both literature and job offers).
- 2) The realization of the ‘dictionary’ – that is, the list of all the present concepts and the relative verbalizations to unify those with the same meaning (merging). In doing this, the authors manually observed all the posts and identified any similar concepts in the various posts expressed differently. This operation was very useful in order to avoid having double second level nodes or replicas of concepts on more levels or in multiple positions of the framework.
- 3) The elaboration of the so-called ‘relationship card’ – that is, a table that indicates all the causal relationships and connotations between the identified themes, characteristics and managerial implications. In order to do this, of course, first inspiration was taken from the word trees (which represent relationship cards themselves); then the authors re-interpreted all the relationships of each word tree, identified the branches that could serve as a link between multiple separate word trees, and developed an initial draft of the framework.
- 4) The conceptualization of the framework.

After writing down each single word tree, and after their qualitative analysis, a preliminary map was elaborated through the systematization of the main positions emerging in a single cognitive framework.

#### **4. Discussion of Results and Managerial Implications**

The results obtained from our analysis show the effects of big data on organizations and the challenges and opportunities connected to big data and BDA. The two main contributions of this research to literature are the identification of professionals able to seize the emerging opportunities and the way in which organizations can benefit from big data.

In respect of the first contribution, it has emerged how managers characterized by skills in big data are fundamental. These managers are extremely important in the marketing sector, as well as in the corporate finance, human resource management, information management and operation management sectors. These managers should indeed orchestrate all the other players that will have to deal with these data. Another professional figure of extreme importance is the data analyst, who plays a

crucial role in the analysis of data. We observed that data analysts must take care to help managers to develop marketing and operations strategies. The data scientist, on the other hand, is fundamental to the use of big data in financial strategies. Finally, data architects and engineers have roles that are mainly related to the use of big data for operations management and for the modernization and design of information systems.

Regarding the opportunities connected to the availability of big data, numerous implications concerning strategic marketing have emerged. In fact, thanks to them, it is possible both to analyse the markets in an aggregated manner while being able to track the behaviour of the individual consumer. Thus, it emerges that this phenomenon can play a crucial role in terms of consumer intelligence. Moreover, consumers relationship management (CRM) systems, customer communication and online experience can be improved by the newly available consumer data. Thus, big data can be used for more effective marketing campaigns (Faraoni et al., 2019). Specifically, they can help marketers in consumer segmentation and in measuring advertising effectiveness.

With regard to the impact of big data on corporate finance, it emerged that, in addition to being generically useful for risk management and the prevention of online fraud, they play a fundamental role in customer credit scoring (Santoro et al., 2018). In human resources management, big data analysis allows the identification of online talents and online head hunting. As for the impact on information systems, big data are revolutionizing the classic concept (Wang & Cotton, 2018). It is in fact necessary that organizations adapt their information systems and, to do so, the process of collecting and storing information is central, since in the current economic system the value of information, also with a view to reselling it to third parties, is a source of potential income (Rialti et al., 2018). Finally, big data impact both operations management and supply chain management planning (Gunasekaran et al., 2017). As in any research related to a specific geographical area, results are anyway to be considered as context dependant. In fact, findings derive from how European companies' employees deal with the analysis of big data.

This notwithstanding, from the analysis, it emerges that big data are fundamental for marketing intelligence, revenue and financial intelligence, operations intelligence and information intelligence, that together compose big data-driven business intelligence. Therefore, thanks to big data and the 3.0 information systems that analyse them, managers can now make sales forecasts and monitor production constantly and precisely. Given the possibility of predicting the behaviour of the individual customer, moreover, these forecasts are characterized by extreme reliability (Erevelles et al., 2016; Rialti et al., 2017; Fornari et al., 2018).

In terms of suggestions for practitioners, big data-driven decision making is therefore emerging as an innovative approach allowing managers to take big data-based decisions in combination with their intuition. From this, in respect of the potential managerial implications, it is possible to state that managers should continue to encourage investment in BDA. Such investment, in fact, can provide organizations with innovative technological tools capable of generating significant competitive advantages. In any case, it emerges that these investments must also be linked to investments aimed at changing the organization's culture in order to accept big data inside the decision-making processes. This is necessary because

otherwise some internal professional figures (in particular, those not data skilled) could resist the implementation of BDA.

## 5. Conclusions, Limitations and Suggestions for Future Research

Based on our analysis, it was found that big data are fundamental, especially for business intelligence improvement and, subsequently, strategic marketing campaigns. Our results are hereby consistent with the existing literature; they help to shed light on why big data and BDA matter. It is clear, in fact, that they must be managerial persons and manage the strategy-making aspect (Fawcett & Provost, 2013; Brondoni, 2015) using data from the business intelligence systems. Furthermore, we have seen that organizations mainly seek to use them to predict customer behaviour.

The results obtained are not fully generalizable, because the job offers explored are located in Europe in a particular period of time, and only job offers from large organizations have been collected. Moreover, in terms of the limitations connected to the method, the coding process performed was supervised in nature: in fact, the data were only partially analysed by automatic software based on well-identified principles. Given such methodological limitations, possible future research should use larger datasets with unsupervised analysis methodologies based on machine learning, such as latent Dirichlet allocation (Blei et al., 2003).

As for the implications for future research, we suggest that future researchers who want to address issues related to big data explore how managers can encourage organizational BDA capabilities. Furthermore, the potential benefits that can be obtained from the entry of big data into the organization are still to be explored in detail. For example, it may be necessary to explore how they change the micro-mechanisms of the internal operation of large organizations. It would be useful to analyse whether big data can help make large organizations more agile and dynamic, or if they represent a burden. Finally, it is suggested that scholars who want to address the topic from a more academic perspective also review the main theories used in the big data literature.

## Bibliography

- Barile, S., & Polese, F. (2018). Grocery Retailing in the I4.0 Era. *Symphonya. Emerging Issues in Management (symphonya.unimib.it)*, (2), 38-51.  
<http://dx.doi.org/10.4468/2018.2.4barile.polese.sarno>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Brondoni, S. M. (2015). Global Networks, Outside-In Capabilities and Smart Innovation. *Symphonya. Emerging Issues in Management (symphonya.unimib.it)*, (1), 6-21.  
<http://dx.doi.org/10.4468/2015.1.02brondoni>
- Brondoni, S. M., & Zaninotto, E. (2018). Ouverture de 'The 4th Industrial Revolution. Business Model Innovation & Global Competition'. *Symphonya. Emerging Issues in Management (symphonya.unimib.it)*, (2), 1-7.  
<http://dx.doi.org/10.4468/2018.2.01ouverture>



- Brown, B., Chui, M., & Manyika, J. (2011). Are You Ready for the Era of 'Big Data'? *McKinsey Quarterly*, 4(1), 24-35.
- Büchi, G., Cugno, M., & Castagnoli, R. (2018). Economies of Scale and Network Economies in Industry 4.0. *Symphonia. Emerging Issues in Management (symphonia.unimib.it)*, (2), 66-76.  
<http://dx.doi.org/10.4468/2018.2.06buchi.cugno.castagnoli>
- Camussone, P., & Biffi, A. (2018). Information Technology and Work: Threats and Opportunities. *Symphonia. Emerging Issues in Management (symphonia.unimib.it)*, (2), 125-144.  
<http://dx.doi.org/10.4468/2018.2.10camussone.biffi>
- Caputo, A., Marzi, G., & Pellegrini, M. M. (2016). The Internet of Things in Manufacturing Innovation Processes: Development and Application of a Conceptual Framework. *Business Process Management Journal*, 22(2), 383-402.  
<http://dx.doi.org/10.1108/BPMJ-05-2015-0072>
- Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, 36(4), 1165-1188.  
<http://dx.doi.org/10.2307/41703503>
- Ciappei, C., & Cinque, M. (2015). *Strategic Data-Based Wisdom: Applying Aristotle's Theory of Strategic Data-Based Wisdom in the Big Data Era*, in Girard, J. (ed.), *Strategic Data-Based Wisdom in the Big Data Era* (pp. 172-195). New York, NY: IGI Global.
- Cillo, V., Rialti, R., Del Giudice, M., & Usai, A. (2019). Niche Tourism Destinations' Online Reputation Management and Competitiveness in Big Data Era: Evidence from Three Italian Cases. *Current Issues in Tourism*, in press.  
<http://dx.doi.org/10.1080/13683500.2019.1608918>
- Davenport, T. H., & Patil, D. J. (2012). Data Scientist: The Sexiest Job of the 21st Century. *Harvard Business Review*, 90, 70-76.
- Davenport, T. H., Barth, P., & Bean, R. (2015). How 'Big Data' Is Different. *MIT Sloan Management Review*, 54(1), 343-351.
- De Mauro, A., Greco, M., Grimaldi, M., & Nobili, G. (2016). Beyond Data Scientists: A Review of Big Data Skills and Job Families. In *11th International Forum on Knowledge Assets Dynamics*.
- De Mauro, A., Greco, M., Grimaldi, M., & Ritala, P. (2017). Human Resources for Big Data Professions: A Systematic Classification of Job Roles and Required Skill Sets. *Information Processing & Management*, in press.  
<http://dx.doi.org/10.1016/j.ipm.2017.05.004>
- Del Castillo, J., Paton, J., & Barroeta, B. (2015). Smart Specialisation for Economic Change: The Case of Spain. *Symphonia. Emerging Issues in Management (symphonia.unimib.it)*, (1), 30-43.  
<http://dx.doi.org/10.4468/2015.1.04delcastillo.paton.barroeta>
- Demchenko, Y., De Laat, C., & Membrey, P. (2014). Defining Architecture Components of the Big Data Ecosystem. In *2014 International Conference on Collaboration Technologies and Systems*, pp. 104-112. IEEE.  
<http://dx.doi.org/10.1109/CTS.2014.6867550>
- Demirkan, H., & Delen, D. (2013). Leveraging the Capabilities of Service-Oriented Decision Support Systems: Putting Analytics and Big Data in Cloud. *Decision Support Systems*, 55(1), 412-421.  
<http://dx.doi.org/10.1016/j.dss.2012.05.048>
- Doan, A., Ramakrishnan, R., & Halevy, A. Y. (2011). Crowdsourcing Systems on the Worldwide Web. *Communications of the ACM*, 54(4), 86-96.
- Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data Consumer Analytics and the Transformation of Marketing. *Journal of Business Research*, 69(2), 897-904.  
<http://dx.doi.org/10.1016/j.jbusres.2015.07.001>
- Faraoni, M., & Petretto, L., (2009), Market-Driven Management and Global Supply Chain. *Symphonia. Emerging Issues in Management (symphonia.unimib.it)*, (2), 58-74.

- <http://dx.doi.org/10.4468/2009.2.06faraoni.petretto>  
Faraoni, M., Rialti, R., Zollo, L., & Pellicelli, A. C. (2019). Exploring E-Loyalty Antecedents in B2C E-Commerce: Empirical Results from Italian Grocery Retailers. *British Food Journal*, 121(2), 574-589.
- <http://dx.doi.org/10.1108/BFJ-04-2018-0216>  
Fornari, E., Grandi, S., & Fornari, D. (2018). Retailing 4.0: The New Era of E-commerce in Fast Moving Consumer Goods. *Symphonia. Emerging Issues in Management (symphonia.unimib.it)*, (2), 77-90.
- <http://dx.doi.org/10.4468/2018.2.07fornari.grandi.fornari>  
Gardiner, A., Aasheim, C., Rutner, P., & Williams, S. (2017). Skill Requirements in Big Data: A Content Analysis of Job Advertisements. *Journal of Computer Information Systems*, 43(2), 1-11.
- <http://dx.doi.org/10.1080/08874417.2017.1289354>  
George, G., Haas, M. R., & Pentland, A. (2014). Big Data and Management. *Academy of Management Journal*, 57 (2), 321-326.
- <http://dx.doi.org/10.5465/amj.2014.4006>  
Goes, P. B. (2014). Editor's Comments: Big Data and IS Research. *MIS Quarterly*, 38(3), 3-7.
- Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S. F., Childe, S. J., Hazen, B., & Akter, S. (2017). Big Data and Predictive Analytics for Supply Chain and Organizational Performance. *Journal of Business Research*, 70, 308-317.
- <http://dx.doi.org/10.1016/j.jbusres.2016.08.004>  
IBM (2012). *What is big data?* IBM Corporate Website. Retrieved from <http://www-01.ibm.com/software/data/bigdata/>
- Kaplan, A. M., & Haenlein, M. (2010). Users of the World, Unite! The Challenges and Opportunities of Social Media. *Business Horizons*, 53(1), 59-68.
- <http://dx.doi.org/10.1016/j.bushor.2009.09.003>  
Kitchin, R. (2017). Big Data – Hype or Revolution in Sloan L., & Quan-Haase, A. (Eds.), *The Sage Handbook of Social Media Research Methods*. Thousand Oaks, CA: Sage.
- Labrinidis, A., & Jagadish, H. V. (2012). Challenges and Opportunities with Big Data. *Proceedings of the VLDB Endowment*, 5 (12), 2032-2033.
- <http://dx.doi.org/10.14778/2367502.2367572>  
LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big Data, Analytics and the Path from Insights to Value. *MIT Sloan Management Review*, 52(2), 21-36.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. Report of McKinsey Global Institute.
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. London: Houghton Mifflin Harcourt.
- McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big Data. The Management Revolution. *Harvard Business Review*, 90 (10), 61-67.
- Mishra, A. S. (2015). Information Professionals and Big Data. *International Journal of Advanced Research in Computer Science and Software Engineering*, 5 (9), 123-129.
- Mishra, D., Luo, Z., Jiang, S., Papadopoulos, T., & Dubey, R. (2017). A Bibliographic Study on Big Data: Concepts, Trends and Challenges. *Business Process Management Journal*, 23(3), 555-573.
- <http://dx.doi.org/10.1108/BPMJ-10-2015-0149>  
Motamarri, S., Akter, S., & Yanamandram, V. (2017). Does Big Data Analytics Influence Frontline Employees in Services Marketing? *Business Process Management Journal*, 23(3), 623-644.
- <http://dx.doi.org/10.1108/BPMJ-12-2015-0182>  
Provost, F., & Fawcett, T. (2013). Data Science and Its Relationship to Big Data and Data-Driven Decision Making. *Big Data*, 1(1), 51-59.
- <http://dx.doi.org/10.1089/big.2013.1508>

- Rialti R., Ciappei, C., Zollo, L., & Boccardi, A. (2016). Big Data Oriented Business Models: The 7Vs of Value Creation. XXVIII Sinergie Annual Conference: *Management in a Digital World. Decisions, Production, Communication*.  
<http://dx.doi.org/10.7433/SRECP.EA.2016.18>
- Rialti, R., Marzi, G., Ciappei, C., & Busso, D. (2019a). Big Data and Dynamic Capabilities: A Bibliometric Analysis and Systematic Literature Review. *Management Decision*, in press.  
<http://dx.doi.org/10.1108/MD-07-2018-0821>
- Rialti, R., Marzi, G., Silic, M., & Ciappei, C. (2018). Ambidextrous Organization and Agility in Big Data Era: The Role of Business Process Management Systems. *Business Process Management Journal*, 24(5), 1091-1109.  
<http://dx.doi.org/10.1108/BPMJ-07-2017-0210>
- Rialti, R., Zollo, L., Ferraris, A., & Alon, I. (2019b). Big Data Analytics Capabilities and Performance: Evidence from a Moderated Multi-Mediation Model. *Technological Forecasting and Social Change*, in press.  
<http://dx.doi.org/10.1016/j.techfore.2019.119781>
- Rialti, R., Zollo, L., Pellegrini, M. M., & Ciappei, C. (2017). Exploring the Antecedents of Brand Loyalty and Electronic Word of Mouth in Social-Media-Based Brand Communities: Do Gender Differences Matter? *Journal of Global Marketing*, 30(3), 147-160.  
<http://dx.doi.org/10.1080/08911762.2017.1306899>
- Rieple, A., & Pisano, P. (2015). Business Models in a New Digital Culture: The Open Long Tail Model. *Symphonia. Emerging Issues in Management (symphonia.unimib.it)*, (2), 75-88.  
<http://dx.doi.org/10.4468/2015.2.06rieple.pisano>
- Saccardi, A. (2003). Data Mining and Marketing Intelligence. *Symphonia. Emerging Issues in Management (symphonia.unimib.it)*, (2), 51-57.  
<http://dx.doi.org/10.4468/2003.2.05saccardi>
- Santoro, G., Vrontis, D., Thrassou, A., & Dezi, L. (2018). The Internet of Things: Building a knowledge Management System for Open Innovation and Knowledge Management Capacity. *Technological Forecasting and Social Change*, (136), 347-354.  
<http://dx.doi.org/10.1016/j.techfore.2017.02.034>
- Trento, S., Bannò, M., & D'Allura, G. M. (2018). The Impact of the 4<sup>th</sup> Industrial Revolution on the High-Tech Industry. *Symphonia. Emerging Issues in Management (symphonia.unimib.it)*, (2), 145-157.  
<http://dx.doi.org/10.4468/2018.2.11trento.banno.dallura>
- Waller, M. A., & Fawcett, S. E. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution that Will Transform Supply Chain Design and Management. *Journal of Business Logistics*, 34(2), 77-84.  
<http://dx.doi.org/10.1111/jbl.12010>
- Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'Big Data' Can Make a Big Impact: Findings from a Systematic Review and a Longitudinal Case Study. *International Journal of Production Economics*, 165, 234-246.  
<http://dx.doi.org/10.1016/j.ijpe.2014.12.031>
- Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big Data Analytics and Firm Performance: Effects of Dynamic Capabilities. *Journal of Business Research*, 70(1), 356-365.  
<http://dx.doi.org/10.1016/j.jbusres.2016.08.009>
- Wang, L., & Cotton, R. (2018). Beyond Moneyball to Social Capital Inside and Out: The Value of Differentiated Workforce Experience Ties to Performance. *Human Resource Management*, 57(3), 761-780.  
<http://dx.doi.org/10.1002/hrm.21856>

- Wrightson, M. T. (1976). The Documentary Coding Method. In R. M. Axelrod (Ed.), *The Structure of Decision: Cognitive Maps of Political Elites*. Princeton, NJ: Princeton University Press.
- Zaiane, O. R. (2015). Rich data: Risks, Issues, Controversies & Hype. In *Proceedings of 2<sup>nd</sup> Annual International Symposium on Information Management and Big Data* (pp. 21-24).
- Zirpoli, F., & Cabigiosu, A. (2018). Digitalization in the Italian Auto Industry. *Symphonia. Emerging Issues in Management (symphonia.unimib.it)*, (2), 158-169.  
<http://dx.doi.org/10.4468/2018.2.12cabigiosu.zirpoli>