International Journal of Information Management 38 (2018) 187–195

Contents lists available at ScienceDirect

International Journal of Information Management

journal homepage: www.elsevier.com/locate/ijinfomgt

Big data technologies: An empirical investigation on their adoption, benefits and risks for companies

Elisabetta Raguseo

Department of Management and Production Engineering, Politecnico di Torino, Corso Duca degli Abruzzi, 24, Turin, Italy

ARTICLE INFO

Keywords:
Big data
Big data technologies
Adoption
Benefits
Risks
Size
Industrial sector

ABSTRACT

Companies currently have to deal with profound changes in the way they manage their business, their customers and their business models, since they are overrun by a data-driven revolution in management. This revolution is due to the wide availability of big data and the fast evolution of big data technologies. Big data is recognized as one of the most important areas of future technology, and is fast gaining the attention of many industries, since it can provide high value to companies. This article investigates the adoption levels of big data technologies in companies, and the big data sources used by them. This article also points out the most frequently recognized strategic, transactional, transformational and informational benefits and risks related to the usage of big data technologies by companies. In order to achieve these aims, the paper looks at the differences that exist among companies of different sizes, by comparing medium-sized and large companies, and the differences among companies of different industrial sectors. It provides evidence that only in a few cases these differences are significant. This study could serve as a reference for managers who wish to initiate an evaluation cycle on the adoption and usage of big data technologies.

1. Introduction

Companies have to deal with profound changes in the way they manage their business, their customers and their business models, since they are overrun by a data-driven revolution in management (Tambe, 2014). This is taking place because new opportunities are emerging thanks to the fast evolution of big data technologies and to the enormous availability of data that firms can capture from many sources, such as social media activities, Radio Frequency Identification (RFID) tags, web information, mobile phone usage and consumer preferences expressed on the web (Davenport, 2014).

In view of the increasing importance of big data, companies can leverage on these data to exploit new opportunities and gain an in-depth understanding of the hidden values. Companies in many industries are increasingly interested in their potentialities, since they can generate high revenues. As stated by the International Data Corporation (IDC, 2016) "Big data and business analytics worldwide revenues will grow from nearly $122 B in 2015 to more than $187 B in 2019, an increase of more than 50% over the five-year forecast period".

Big data can be obtained from within and outside companies, in the form of structured and unstructured data, and it can generates difficulties for companies in capturing, aggregating, analysing and extracting value from them (Gandomi & Haider, 2015). As stated by McAfee, Brynjolfsson, Davenport, Patil and Barton (2012) "businesses are collecting more data than they know what to do with". However, they need to develop new skills and a new management style to turn all this information into a competitive advantage. In order to achieve this goal, and considering the managerial complexity involved in dealing with such an enormous quantity of data, as well as the difficulties that companies may have in managing them, managers should not underestimate the positive and negative issues that need to be taken into account in the exploitation of big data. For this reason, managers need to be aware not only of its positive effects, but also of its negative ones in order to avoid being unprepared when they decide to invest and to derive value from big data. In this direction, there is scant empirical evidence about the benefits and risks associated with the exploitation of big data for business goals.

Till now, literature has provided very little empirical evidence on these issues, and the key contribution of this paper is therefore to enrich the studies that investigate the issues related to big data by empirically studying, through a survey analysis on a sample of 200 companies, the benefits and risks of big data at a company level. This topic is acquiring importance in literature, since the benefits and risks need to be identified, managed and controlled if managers want to derive value from their investments in big data technologies. This paper also provides statistics about the type of big data sources used by companies, and the big data technologies that are adopted. In order to achieve these aims, the differences that emerge according to the size and to a company

E-mail address: elisabetta.raguseo@polito.it.

http://dx.doi.org/10.1016/j.ijinfomgt.2017.07.008
Received 3 January 2017; Received in revised form 22 June 2017; Accepted 26 July 2017

industrial sector have also been investigated.

Based on the increasing academic and managerial interest in big data, this paper first explains what big data is, and then discusses the main issues that need to be considered when managing big data. After this discussion, the research methodology is presented, and this is followed by the results of the study. Finally, the discussion and the conclusion sections are given. The paper concludes by highlighting the big data technologies that are adopted the most frequently by companies, the big data sources that are used the most by companies, and the benefits and risks of big data technologies that are recognised the most by companies, in an aggregated way, and by showing results according to the company size and industrial sector.

2. What big data is

How organizations capture, create and use data is changing the way these organizations work. Executives, academics and business analysts need to be aware of this change, which will transform how organizations are managed, and will also alter the economies and societies in which they operate. This revolution has introduced several changes in companies. For example, they now have more data to use than ever before. This data may be internal or external, structured or unstructured. By using internal and external data, companies are beginning to understand patterns of consumer activity that had once been impossible to perceive or act upon. Companies are also using new technological solutions to understand their own operations and behaviour at a much finer level of detail.

The term “big” began to be diffused in 2011 (Gandomi & Haider, 2015). Leading companies, such as IBM, have made huge investments in building a powerful platform for big data analytics in order to deliver new business insights with the goal, for example, of reducing storage and maintenance costs. Gartner provided the following definition for big data: “Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making”. In this definition, volume refers to the generation and collection of a huge amount of data where the data scale becomes increasingly high; velocity refers to the timeliness by which big data is produced, collected and analysed; variety indicates the different types of data that can be produced in a structured and unstructured way as audios, videos, webpages and texts. As mentioned by Davenport (2014), big data can be also classified as machine-generated, which refers to data that is created by a machine without human intervention, or as human-generated, which refers to data that humans, in interaction with computers, supply. The former refers to audio, music, image, speech and video data, to sensor data, such as RFID tags used to track locations, to Intelligent Lighting Control (ILC) sensors used to identify the location and conditions of goods on the supply chain for example, and to smart meter, medical device or Global Positioning System (GPS) data. The latter instead refers to Twitter tweets, social media posts, clickstream data or web contents.

The importance of big data is demonstrated by the fact that data are produced extensively every day in many forms and from many different sources. For example, more than 98,000 tweets are written every sixty seconds, 695,000 status updates are posted on Facebook, 11 million instant messages are written, 685,445 Google searches are lunched, more than 169 million emails are sent, more than 1820 TB of data are created, and there are 217 new mobile web users. Furthermore, as shown by Google trends (2016), interest in big data has also increased over time (Fig. 1).

3. Managing big data technologies in companies

Davenport (2014) highlighted the importance of big data technologies, such as Hadoop or Natural Languages Processes, to analyse a huge amount of data for cost reduction purposes, to take faster and better decisions and to improve the products and services offered. For example, Wal-Mart’s sematic analysis search engine, Polaris, a platform that was in-house designed, relies on text analysis and machine learning to produce relevant search results. Adding a semantic search has improved the possibility of online shoppers completing a purchase in Wal-Mart by 10% to 15%.

When new technological solutions are adopted by companies, as big data technologies are, these companies can exploit the benefits, but they also need to be aware of the existence of the possible risks they can incur. On the one hand, investments in new technologies necessitate changes over time in organizations, in order to reap the full benefits. In fact, companies need to sustain a range of management activities in order to ensure that they obtain benefits from a technological investment (Farbey, Land, & Targett, 1999). According to Gregor, Martin, Fernandez, Stern, & Vitale (2006), the benefits associated with technological investments can be classified in four typologies: strategic, informational, transactional and transformational. The strategic benefits are those that can alter the way companies compete or the nature of their products. The informational benefits are those that provide information and communication that can be used to improve decision making in a company. The transactional benefits refer to investments that support operational management and which are able to cut the costs sustained by companies. Finally, transformational benefits refer to the results of changes that a firm has to make to the structure and to the capacity of implementing a technological investment.

Organizations, on the other hand, need to enhance their ability to manage the risks associated with a new technological investment, if they want to have successful implementations (Willcocks & Graeser, 2001). Managers feel the need to minimize the systemic risks that arise from the use of Information Technology (IT). Accordingly, the assessment of the risks that managers might incur is very important, and it allows managers to identify any outcomes that could mine the achievement of the expected benefits. For example, there may be risks related to a security breach of information systems or risks related to the costs of the breakdown of systems for the company (Dhillon & Backhouse, 1996).

Considering all these reasons, the following sections show the research methodology that has been used to investigate the specific benefits and risks of big data technologies and their adoption in companies.

4. Research methodology

4.1. Data collection

A questionnaire was distributed to medium and large-sized French companies in order to evaluate the benefits and risks that these companies have to face when adopting and implementing big data solutions. In order to decide whether a company belonged to the medium or large sized categories, we adopted the European definition that considers the number of employees. A medium-sized firm is a company with a number of employees of between 50 and 249, while a large company has more than 249 employees.

As our study has been conducted at the firm level, we followed the lines of previous studies (Ji-fan Ren, Fosso Wamba, Akter, Dubey, & Childe, 2016) and targeted the Chief Information Officer (CIO). We implemented a random sampling method to select 1,962 medium and large French companies to interview from a population of 19,875 companies registered in the Bureau Van Dijk database. The questionnaire was divided into two sections. In the first section, questions were included about the big data sources used by the companies and the big data technologies. In the second part, questions were asked about the benefits and risks of the usage of big data.

First, a pilot study was conducted on a subsample of companies in order to test the comprehensibility of the questions included in the questionnaire, as well as to identify any possible criticalities and to

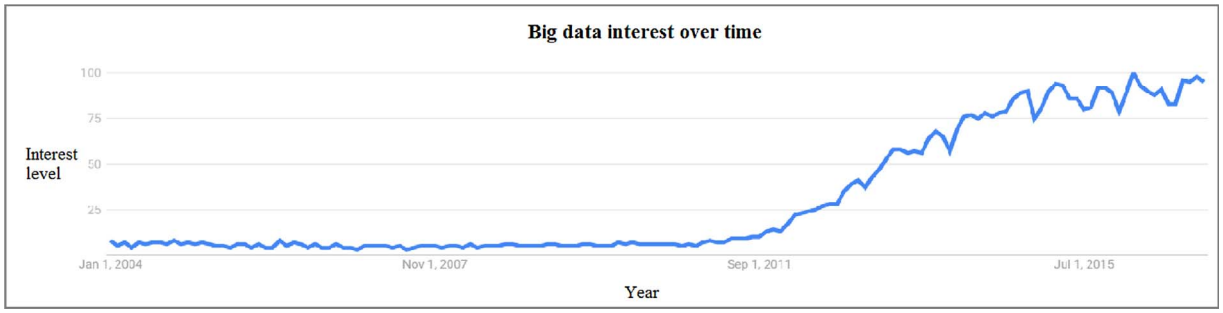

Fig. 1. Big data interest over time (Google trends, 2016).

establish the expected response rate. All the responses were positive, the questions appeared clear and the final questionnaire was therefore left unchanged. The response rate of the pilot test was equal to 21.13%, with 30 companies responding.

After the pilot test, another 170 companies took part and compiled in the survey. The data gathering process involved three steps. In the first step, the company was contacted in order to inform them about the aim of the research study and to ask permission to contact of the Chief Information Officer (CIO). In the second step, the CIO was contacted and asked about his/her willingness to participate in the survey. When the CIO was not available at the time agreed upon in the first call, a second appointment was fixed. Therefore, the questionnaire was filled in, either in the second or third step, according to the availability of the CIO. When the CIO was definitely not willing to reply to the questionnaire, another person qualified about investments and the adoption of big data analytics solutions was identified and was invited to fill in the questionnaire. Overall, 200 questionnaires were filled in. The final sample was composed of 86.5% medium sized companies, and 13.5% large sized companies, thus assuring a 95% confidence level and a 6.9% confidence interval (Table 1).

4.2. Measures

The usage of big data sources and technologies was operationalized through dummy variables. The big data sources were operationalized using 13 dummy variables (see Table 3), as by asking the company about the availability and use of clickstream data, RFID system data and/or smart meter data. Instead, eight dummies were included for the usage of big data technologies, such as Hadoop and MapReduce (see Table 3).

The taxonomy used to investigate the benefits and risks related to big data was based on previous research (Gregor et al., 2006; Ji-fan Ren et al., 2016; Love, Irani, Standing, Lin, & Burn, 2005; Wu & Wang, 2005), and readapted to the big data context. Table 2 shows all the measurement items, which were based on a seven-point Likert scale, with responses ranging from “completely disagree” (-3) to “completely agree” (+3). Respondents were asked to evaluate the extent to which the big data technology solutions adopted by their companies had led to benefits (transactional, strategic, transformational or informational) and risks.

Table 1
Sample characteristics.

	No. = 200	Percentage (%)
Role of the respondent		
General Director	53	26.50%
Chief Executive Officer	8	4.00%
Chief Information Officer	127	63.50%
Other person qualified about big data analytics investments	12	6.00%
Firm size according to the number of employees		
Medium (50–249 employees)	173	86.50%
Large (more than 249 employees)	27	13.50%
Industrial sectors		
Manufacturing	57	28.50%
Services	85	42.50%
Retail	41	20.50%
Construction	17	8.50%

4.3. Data analysis

A four-steps procedure was followed to present the results. First, rankings were computed for each big data source and each big data technology as well as for the overall number of big data sources used by companies and for the overall number of the big data technologies adopted by the companies. The ranking was also computed for the benefits and risks that respondents had experienced in relation to big data technologies. The ranking was based on an ordering of the variables according to their mean values. Second, the rankings were computed according to the size of the company and the industrial sector the companies belonged to in order to establish the existence of differences. Third, the Kruskal-Wallis test was conducted for each big data source, big data technology, benefit and risk in order to evaluate any significant statistical differences, according to the size of the company and industrial sector, in the means of the variables considered in this study. A Chi-square test was conducted, and information about the significance level (p-value) was provided when the difference was statistically different by less than 5%. Fourth, when the difference in industrial sector was significant, the Scheffé test was conducted, as a post hoc analysis, in order to identify the couples of industrial sectors that had statistically significant different values.

5. Findings

5.1. Big data sources and technologies adopted

Table 3 provides information about the usage of the big data sources in an aggregated way and about the existence of any differences according to the size of the company and to the industrial sector the companies belong to. Taking into account that the 38% of the sample adopted big data technologies to capture, process and analyse big data sources, Table 3 shows the big data sources that are used by the companies.

Considering all the companies in the sample, the source used most frequently by the surveyed companies was the online portal content, and this was followed by the Point Of Sales (POS) data and smart meter data. Smart meters offer households and businesses the opportunity of understanding and reducing the usage of energy throughout the different parts of a day in much greater detail than was previously possible, since, in the past, meter readings were taken once a quarter, or even annually. The statistics also show that very few companies have such data sources as weblog posts or microblog posts available that they can analyse in order to extract information with the aim of obtaining valuable insights. The data shown in Table 3 also shows that around the

Table 2
Benefits and risks of big data technologies.

ID	Item description	Construct	Reference
TAB1	Saving on supply chain management	Transactional benefit	Ji-fan Ren et al. (2016)
TAB2	Reducing operating costs	Transactional benefit	Ji-fan Ren et al. (2016)
TAB3	Reducing communication costs	Transactional benefit	Ji-fan Ren et al. (2016)
TAB4	Avoiding the need to increase the workforce	Transactional benefit	Ji-fan Ren et al. (2016)
TAB5	Increasing return on financial assets	Transactional benefit	Ji-fan Ren et al. (2016)
TAB6	Enhancing employee productivity	Transactional benefit	Ji-fan Ren et al. (2016)
SB1	Creating a competitive advantage	Strategic benefit	Ji-fan Ren et al. (2016)
SB2	Aligning IT with a business strategy	Strategic benefit	Ji-fan Ren et al. (2016)
SB3	Establishing useful links with other organizations	Strategic benefit	Ji-fan Ren et al. (2016)
SB4	Enabling a quicker response to change	Strategic benefit	Ji-fan Ren et al. (2016)
SB5	Improving customer relations	Strategic benefit	Ji-fan Ren et al. (2016)
SB6	Providing better products or services	Strategic benefit	Ji-fan Ren et al. (2016)
TFB1	Achieving an improved skill level for the employees	Transformational benefit	Ji-fan Ren et al. (2016)
TFB2	Developing new business opportunities	Transformational benefit	Ji-fan Ren et al. (2016)
TFB3	Expanding capabilities	Transformational benefit	Ji-fan Ren et al. (2016)
TFB4	Improving business models	Transformational benefit	Ji-fan Ren et al. (2016)
IB1	Enabling faster access to data	Informational benefit	Gregor et al. (2006)
IB2	Enabling easier access to data	Informational benefit	Gregor et al. (2006)
IB3	Improving management data	Informational benefit	Gregor et al. (2006)
IB4	Improving data accuracy	Informational benefit	Gregor et al. (2006)
IB5	Providing data in more useable formats	Informational benefit	Gregor et al. (2006)
R1	Reluctance of employees to adapt to changes	Risk	Love et al. (2005)
R2	Lack of Information System infrastructure support	Risk	Love et al. (2005)
R3	Technical uncertainty	Risk	Love et al. (2005)
R4	Minimal IT expertise	Risk	Love et al. (2005)
R5	Uncertainty about how to measure potential benefits	Risk	Love et al. (2005)
R6	Uncertainty about how to measure the involved costs	Risk	Love et al. (2005)
R7	Capital outlay with no guarantee of likely returns	Risk	Love et al. (2005)
R8	Security issues	Risk	Love et al. (2005)
R9	Reassignment of personnel trained on big data analytics solutions	Risk	Love et al. (2005)
R10	Privacy issues	Risk	Wu and Wang (2005)

20% of the companies have GPS data and social media posts about customers' preferences. The big data sources are extremely important for marketing purposes, since they include customers' posts on social media that enable the companies to know their customers' preferences in real time considering their profile features and locations; this information enables the companies to know, for example, when a customer has entered a store or a website. The data in Table 3 highlights that the mean number of the big data sources used by companies is equal to 3.250, and that this number is not statistically different when considering the size of the companies or the industrial sector the companies belong to.

Table 3 highlights that there are significant differences in the usage of some big data sources according to the size and industrial sector of the company. A significant difference has emerged in the usage of RFID data (Chi-squared = 4.269, p-value < 5%) and GPS data (Chi-squared = 7.082, p-value < 1%), according to the size of the company, both used more by large companies than medium sized companies. Moreover, there are statistically significant differences in the usage of POS data and other transactional data sources (Chi-squared = 10.472, p-value < 5%), according to the industrial sector. In order the establish the couples related to these statistical significant differences, a Scheffé test was conducted. The results show that companies in the manufacturing sector and those in the service sector use fewer POS data or other transactional data sources than companies in the retail sector, as expected given the peculiarities of these sectors.

Table 4 provides information about the adoption of big data technologies in an aggregated way and shows the existence of any differences according to the size of the company and to the industrial sector the companies belong to. Table 4 shows that the majority of companies adopt visual analytics software or other software to display their analytical results in visual or graphic formats (40.79% of the companies). These forms of software allow companies to turn large sets of data into valuable insights in order to support operational processes and the decision-making of managers. They also allow trends that necessitate scrutiny to be tracked in order to take the correct strategic decisions. The second most frequently adopted technology for big data is related to scripting languages (38.16% of the companies) and to in-memory analytics software (30.26% of the companies). MapReduce and Hadoop software were chosen by 27.63% of the companies, while machine learning was chosen by 22.40%. The least frequently adopted technologies were Natural Language Processing (NLP) software, social media analytics software and predictive analytics software. This could be explained by the fact that the considered companies were less likely to rely on social media data because they have probably not understood the high value that social media content can provide, in terms of customers' needs and preferences. The data in Table 4 also shows that the mean number of the big data technologies used by companies is equal to 2.474, and that this number does not differ statistically according of the size of the companies or the industrial sector the companies belong to.

Table 4 highlights that there are significant differences in the adoption of big data technologies according to the company's size, but not according to the company's industrial sector. The results reveal that there are only significant differences for the adoption of machine learning software or other software used to rapidly find the model that best fits a data set (Chi-squared = 6.183, p-value < 5%), and for the adoption of visual analytics software or other software used for the display of analytical results in visual or graphic formats (Chi-squared = 3.899, p-value < 5%), thus indicating that the adoption levels of these two technologies are higher for large companies than for medium sized ones.

5.2. The benefits of big data

Various kinds of benefits can be attained from an IT investment. In the case where firms do not achieve the desired goal, they need to reconsider their investments and the way they leverage on them. Table 5 shows the types of benefits that companies can achieve by adopting big

Table 3
Big data sources.

Big data sources	Size							Industrial sector									
	All the sample		Medium		Large		Significant differences?	Manufacturing		Construction		Retail		Services		Significant differences?	
	Ranking	Mean	Ranking	Mean	Ranking	Mean	Chi-squared	Ranking	Mean	Ranking	Mean	Ranking	Mean	Ranking	Mean	Chi-squared	Scheffé test
Online portal content	1	0.474	1	0.484	2	0.417	0.183	1	0.462	1	0.429	2	0.400	1	0.515	0.511	–
POS data or other transactional data sources	2	0.355	2	0.328	1	0.500	1.286	4	0.231	2	0.286	1	0.800	2	0.333	10.472*	Manufacturing-Retail, Retail-Service
Smart meter data	3	0.329	3	0.313	2	0.417	0.490	2	0.385	2	0.286	2	0.400	3	0.273	1.111	–
Video sources	4	0.276	4	0.266	3	0.333	0.229	2	0.385	2	0.286	3	0.300	5	0.182	1.588	–
Image sources	5	0.263	5	0.250	3	0.333	0.357	3	0.346	2	0.286	3	0.300	5	0.182	2.110	–
Other natural language text sources	6	0.237	6	0.203	2	0.417	2.516	6	0.154	1	0.429	4	0.200	3	0.273	2.689	–
Social media posts	7	0.224	5	0.250	6	0.083	1.595	6	0.154	4	0.000	2	0.400	3	0.273	4.929	–
RFID system data	8	0.197	9	0.156	2	0.417	4.269*	8	0.077	2	0.286	3	0.300	4	0.242	3.764	–
Audio sources	9	0.184	8	0.172	4	0.250	0.405	5	0.192	3	0.143	4	0.200	5	0.182	0.107	–
Clickstream data	9	0.184	7	0.188	5	0.167	0.029	5	0.192	3	0.143	4	0.200	5	0.182	0.107	–
GPS data	10	0.158	10	0.109	2	0.417	7.082**	7	0.115	2	0.286	5	0.100	5	0.182	1.586	–
Weblog posts	11	0.079	11	0.094	7	0.000	1.205	9	0.038	4	0.000	5	0.100	6	0.121	2.031	–
Microblog (eg. Tweets) posts	12	0.066	12	0.078	7	0.000	0.990	9	0.038	4	0.000	6	0.000	6	0.121	3.121	–
Total	n.a.	3.250	n.a.	3.250	n.a.	3.250	0.010	n.a.	3.000	n.a.	3.286	n.a.	3.600	n.a.	3.333	0.797	–

Note: ***p-value $< 0.1\%$; ** $p < 1\%$; * $p < 5\%$; "n.a." stands for "not applicable".

data technologies, and the constructs for each type of benefit. Four different kinds of benefits can be achieved by companies: transactional, strategic, transformational and informational.

According to the data shown in Table 5, the transactional benefit most frequently recognized by the surveyed companies is the enhancement of productivity growth. This means that big data technology companies are able to achieve goals in terms of efficiency. One key motivation, which was ranked second, behind the adoption of big data technologies to achieve transformational benefits, is the reduction of the operating costs, while the enhancement of the returns on financial assets was ranked third, and saving in supply chain management activities was ranked fourth. Instead, the adoption of big data technologies to reduce communication costs and avoid an increase in the workforce was not recognized as a transactional benefit. This means that these technologies do not have any particular impact on the entity of the workforce or on the intra-firm and inter-firm communication costs.

Considering the strategic benefits of big data technologies, the most frequently recognized is related to the ability of proving better products and services. In fact, many products have undergone changes in their features thanks to big data. For example, the Babolat company has put sensors and connectivity devices into their tennis racket handles, in order to allow users to track and analyse the ball speed, spin, and impact location in real time so as to improve the players' game (Porter & Heppelmann, 2015). A second strategic benefit is the ability to align IT with the business strategy, which demonstrates that big data represents one way of improving the internal processes of a company. The third benefit most frequently recognized by the surveyed companies concerns the ability of enabling a quick response to changes that occur in the business environment where companies act.

The most frequently recognized transformational benefit of big data concerns the ability of enabling the expansion of a company's capabilities, because when a company invests in new technologies, it needs to develop new skills and hire new people who are able to handle the new technology. Big data capabilities enable companies to collect and integrate structured and unstructured data from different sources over a short timeframe. This benefit is followed by the development of new business opportunities, since there are companies that, thanks to big data solutions, are able to operate in new value chains that they were not previously in. Let us consider, for example, the data produced by Fitbit products. They enable the health of people to be improved by tracking their activity, exercise, food, weight and sleep, and these data at the same time can be sold to insurance companies in order to allow them to understand the profiles of different people and provide different insurance packages according to their profile. In this way, thanks to big data, new business opportunities can arise between two different companies that operate in different industrial sectors. Accurate weather data is beneficial for many organizations, since some companies can use weather information to improve business activities as their supply chain planning or advertising. The third ranked transformational benefit is related to employees, since companies have declared that, thanks to big data, there has been an improvement in the skill level of their employees.

Finally, the informational benefits most frequently recognized are those related to data management, to data accuracy and to enabling easier access to data. Thanks to big data technologies, companies obtain a higher quality of data and have less trouble in accessing them.

The data in Table 5 also shows that the companies in the sample are homogenous in their judgement of the transactional, strategic, transformational and the informational benefits of big data technologies. This result was confirmed by the Kruskal-Wallis tests, which did not reveal any significant difference, according to the size or industrial sector of the companies, in the benefits of big data technologies.

Table 4
Big data technologies.

Big data technologies	All the sample		Size					Industrial sector								Significant differences?	
			Medium		Large		Significant differences?	Manufacturing		Construction		Retail		Services			
	Ranking	Mean	Ranking	Mean	Ranking	Mean	Chi-squared	Ranking	Mean	Ranking	Mean	Ranking	Mean	Ranking	Mean	Chi-squared	Scheffé test
Visual analytics software or other software used to display analytical results in visual or graphic formats	1	0.408	2	0.359	1	0.667	3.899*	2	0.308	1	0.571	1	0.500	1	0.424	2.214	–
Scripting languages or other programming languages that work well with big data (e.g., Python, Pig, Hive)	2	0.382	1	0.375	3	0.417	0.073	1	0.346	2	0.429	2	0.300	1	0.424	0.731	–
In-memory analytics software or other processing big data memorised in the computer for greater speed	3	0.303	3	0.313	4	0.250	0.185	2	0.308	4	0.000	2	0.300	2	0.364	3.575	–
MapReduce and Hadoop software or other software used to process big data across multiple parallel servers	4	0.276	4	0.281	4	0.250	0.589	1	0.346	3	0.143	3	0.200	3	0.273	1.531	–
Machine learning software or other software used to rapidly find the model that best fits a certain data set	5	0.224	6	0.172	2	0.500	6.183*	3	0.269	4	0.000	2	0.300	5	0.212	2.653	–
NLP or other software used to analyse texts − information extraction, text summarization, question answering, sentiment analysis	6	0.184	5	0.188	5	0.167	0.029	5	0.154	3	0.143	4	0.100	4	0.242	1.436	–
Social media analytics software (content-based analytics, structure-based analytics)	7	0.171	6	0.172	5	0.167	0.002	4	0.192	4	0.000	4	0.100	5	0.212	2.264	–
Predictive analytics software used to extract information from data and predict trends and behaviour patterns	8	0.158	7	0.156	5	0.167	0.008	6	0.038	3	0.143	3	0.200	4	0.242	4.646	–
Total	n.a.	2.474	n.a.	2.406	n.a.	2.833	0.589	n.a.	2.308	n.a.	1.714	n.a.	2.400	n.a.	2.788	1.304	–

Note: Base: all the companies that use at least one technology for big data; ***p-value $< 0.1\%$; ** $p < 1\%$; * $p < 5\%$; "n.a." stands for "not applicable".

Table 5
Benefits of big data.

ID	Benefits	All the sample		Size					Industrial sector								Significant differences?	
				Medium		Large		Significant differences?	Manufacturing		Construction		Retail		Services			
		Ranking	Mean	Ranking	Mean	Ranking	Mean	Chi-squared	Ranking	Mean	Ranking	Mean	Ranking	Mean	Ranking	Mean	Chi-squared	Scheffé test
	Transactional benefits																	
TAB6	Enhancing employee productivity	1	1.079	1	1.000	1	1.500	0.838	1	1.077	1	1.429	1	0.900	1	1.061	0.783	–
TAB2	Reducing operating costs	2	0.658	2	0.688	3	0.500	0.336	2	0.385	4	0.286	3	0.400	2	1.030	1.374	–
TAB5	Increasing returns on financial assets	3	0.500	4	0.391	2	1.083	3.540	2	0.385	3	1.000	2	0.800	4	0.394	4.068	–
TAB1	Savings in supply chain management	4	0.487	3	0.563	5	0.083	2.555	3	0.346	2	1.143	4	−0.400	3	0.727	4.811	–
TAB3	Reducing communication costs	5	0.013	5	−0.016	4	0.167	0.039	4	−0.077	5	−0.286	3	0.400	5	0.030	1.374	–
TAB4	Avoiding the need to increase the workforce	6	−0.579	6	−0.531	6	−0.833	0.391	5	−0.846	6	−0.571	5	−0.500	6	−0.394	0.847	–
	Strategic benefits																	
SB6	Providing better products or services	1	1.092	1	1.063	1	1.250	0.295	1	1.077	3	0.714	1	1.200	1	1.152	0.200	–
SB2	Aligning IT with a business strategy	2	0.987	2	0.938	1	1.250	0.529	2	0.962	1	1.714	2	0.800	3	0.909	2.863	–
SB4	Enabling a quicker response to change	3	0.855	3	0.813	2	1.083	0.225	3	0.885	5	0.429	4	0.500	2	1.030	0.967	–
SB5	Improving customer relations	4	0.461	4	0.406	4	0.750	0.126	4	0.692	6	−0.143	5	0.300	5	0.455	1.943	–
SB3	Establishing useful links with other organizations	5	0.382	5	0.375	5	0.417	0.075	5	0.615	4	0.571	3	0.700	6	0.061	3.594	–
SB1	Creating a competitive advantages	6	0.368	6	0.250	3	1.000	2.916	6	−0.115	2	1.286	6	0.000	4	0.667	1.542	–
	Transformational benefits																	
TFB3	Expanding capabilities	1	0.908	1	0.938	3	0.750	0.599	1	0.962	2	0.571	1	1.400	1	0.788	2.083	–
TFB2	Developing new business opportunities	2	0.592	2	0.531	1	0.917	0.343	2	0.385	2	0.571	2	0.900	2	0.667	1.072	–
TFB1	Achieving an improved skill level for employees	3	0.513	3	0.453	2	0.833	0.295	3	0.308	1	0.714	3	0.600	3	0.606	0.568	–
TFB4	Improving business models	4	0.158	4	0.250	4	−0.333	3.087	4	0.077	3	0.143	4	0.300	4	0.182	0.254	–
	Informational benefits																	
IB3	Improving management data	1	1.487	1	1.453	1	1.667	0.031	1	1.462	4	0.857	1	2.000	1	1.485	1.780	–
IB4	Improving data accuracy	2	1.211	2	1.313	3	0.667	2.976	2	1.269	2	1.286	2	1.500	3	1.061	2.021	–
IB5	Providing data in more useable formats	3	1.145	3	1.047	1	1.667	2.407	3	1.231	1	1.714	4	0.700	3	1.091	2.029	–
IB2	Enabling easier access to data	4	0.974	5	0.953	2	1.083	0.039	5	0.923	4	0.857	4	0.700	2	1.121	0.691	–
IB1	Enabling faster access to data	5	0.908	4	1.016	4	0.333	1.817	4	1.000	3	1.143	3	1.400	4	0.636	2.665	–

Note: ***p-value < 0.1%; ** p < 1%; * p < 5%.

Table 6
Risks of big data.

ID	Risks	Size							Industrial sector								Significant differences?	
		All the sample		Medium		Large		Significant differences?	Manufacturing		Construction		Retail		Services			
		Ranking	Mean	Ranking	Mean	Ranking	Mean	Chi-squared	Ranking	Mean	Ranking	Mean	Ranking	Mean	Ranking	Mean	Chi-squared	Scheffé test
R10	Privacy issues	1	0.974	1	1.016	1	0.750	0.689	1	1.269	1	1.143	1	1.200	1	0.636	3.200	–
R8	Security issues	2	0.816	2	0.859	3	0.583	0.606	2	1.000	2	1.000	2	1.100	2	0.545	2.554	–
R2	Lack of Information System and infrastructure support	3	0.553	3	0.531	2	0.667	0.036	3	0.846	4	0.571	3	0.800	5	0.242	2.086	–
R7	Capital outlay with no guarantee of likely returns	4	0.461	5	0.438	3	0.583	0.001	4	0.808	3	0.857	8	−0.400	3	0.364	10.404*	Manufacturing-Retail
R4	Minimal IT expertise	4	0.461	4	0.500	5	0.250	1.041	4	0.808	4	0.571	4	0.600	6	0.121	3.408	–
R3	Technical uncertainty	5	0.368	6	0.375	4	0.333	0.043	5	0.654	4	0.571	4	0.600	7	0.030	2.264	–
R1	Reluctance of employees to adapt to changes	6	0.184	7	0.297	9	−0.417	3.609	6	0.500	6	−0.571	5	0.400	7	0.030	3.559	–
R5	Uncertainty about how to measure potential benefits	7	0.145	8	0.188	7	−0.083	0.672	7	0.462	6	−0.571	9	−0.600	4	0.273	4.845	–
R6	Uncertainty about how to measure the involved costs	8	−0.237	9	−0.234	8	−0.250	0.090	8	−0.154	7	−0.714	7	−0.300	8	−0.182	0.992	–
R9	Reassignment of personnel trained on big data analytics solutions	9	−0.263	10	−0.313	6	0.000	0.602	9	−0.462	5	0.143	6	−0.100	9	−0.242	1.355	–

Note: ***p-value < 0.1%; ** p < 1%; * p < 5%.

5.3. The risks of big data

Risks are elements that should be considered when companies decide to invest in big data solutions. As shown in Table 6, even though big data can be of great value, companies have to take into account privacy and security issues before using big data technologies. These are the two risks most frequently recognized by the surveyed companies. This happens because different data sources can be found in distributed places and they therefore determine security vulnerabilities. Big data technologies enable data from multiple locations to be gathered, processed and analysed, and challenges to be determined in terms of satisfying different regulations around the world in the countries where the big data originated from. In fact, similar data can be treated in different ways, according to where they are generated or gathered.

In terms of privacy risks, for example, in the case of medical data, it is sometimes not clear who is the owner of the data, but using the data without the right legal foundation or consent of the patient may cause serious problems (Ernst & Young, 2014). The third most frequently recognized risk is related to the lack of Information System structure support. Instead, other risks, such as those related to the uncertainty about how to measure the potential benefits, the involved costs, and the reassignment of a company's personnel trained on big data analytics solutions have not been considered priority risks for the surveyed companies.

No particular differences emerged in the judgement of the risks outlined in Table 6 by companies according to their size or industrial sector, with the only exception being that of "Capital outlay with no guarantee of likely returns", which was judged riskier by companies in the manufacturing sector than those in the retail sector (Chi-squared = 10.404, *p*-value < 5%).

6. Discussion

The benefits and the risks of big data technologies should not be underestimated by companies that decide to make investments in big data technologies. The risks and benefits of the use of IT in organizations have received a great deal of attention from academics and practitioners, but there is still scant empirical evidence on the risks and benefits of big data technologies chosen by companies and on evaluating whether differences exist according to the size and the industrial sector of the companies. Previous studies investigated two main aspects related to big data. First, they looked at the financial impacts of big data in companies (Ji-fan Ren et al., 2016; Tambe, 2014). Second, prior work was conducted to provide insight into big data challenges and opportunities in a qualitative and aggregated way (Chen, Chiang, & Storey, 2012; Gandomi & Haider, 2015), thus demonstrating that there is still a lack of empirical evidence on these issues.

Considering this research gap, the present research offers an opportunity to empirically understand more about the adoption of big data technologies in companies, and about their benefits and risks. This has been achieved by investigating whether any differences exist according to the size and the industrial sector.

The results shed light on the adoption level of big data technologies and demonstrate that the most diffused technologies are visual analytics technologies, scripting languages, and in-memory analytics software. The big data sources that companies use the most are online portal contents, POS data and smart meter data. This result highlights that companies are still more likely to use big data that are proprietary than to buy other data on customers from third parties, such as those that can be produced by the social media.

This research has also shed light on the most frequent benefits of big data technologies. These technologies, in terms of transactional benefits, enable employee productivity to be enhanced to a greater extent, better products and services to be provided as strategic benefits, company's capabilities to be expanded, in terms of transformational benefits, and management data to be improved, in terms of informational

benefits. These finding allow researchers and managers to be aware, at the same time, of the most potential benefits that big data technologies can offer their companies, and to act accordingly in order to maximize the returns of their investments and extract high value from big data.

Technological advances in storage and computations have also determined many risks for companies. Managers need to recognize that it is not possible to associate only positive outcomes to big data, and that they also need to handle the major risks they determine, such as security and privacy issues. In such a way, they will be ready to take part in the big data revolution and take care of the risks of big data investments in order to minimize the risks associated with them.

Looking beyond this research, future studies could enrich the findings of this paper by conducting case studies among companies that use big data technologies in order to enter into the black box of the dynamics of their adoption. Future studies could also investigate the trends of the adoption of big data technologies in other settings, in order to establish whether there are differences among countries and, where present, to identify the factors that determine them. Future research could also look at the impacts of big data solutions on several financial performances, such as the productivity of companies, by integrating balance sheet data with data gathered though surveys, in order to verify the extent to which the adoption of big data technologies has objectively impacted the financial performance of firms. In this context, future research could look at this phenomenon from an objective perspective. In this direction, future research could add further knowledge for company managers and could help them to avoid underestimating any aspects that may be related to the benefits and risks of investments related to big data solutions.

7. Conclusion

This research has identified the benefits and risks of big data technologies, and of the adoption of these technologies in firms. Given the scant empirical evidence on these issues in literature, and also the lack of evidence about the empirical investigation of the existence of differences according to the size and industrial sector of companies, this paper has added knowledge to the recent literature on big data. The findings of this study could also be used by managers as an instrument to help them understand the most important transactional, transformational, strategic and informational benefits and risks of big data technologies. This could help them to avoid a common form of management behaviour, based on ignoring or ineffectively evaluating the effect of big data investments on their operations and strategies. The study has identified the following nine key findings:

- online portal content is the source of big data most frequently used by companies;
- visual analytics software is the big data technology most frequently adopted by organizations;
- an improvement in data management is the benefit that was ranked highest by all the industrial sectors and for the two considered company dimensions;
- considering the risks associated with the adoption of big data technologies, organizations in different industrial sectors significantly differ only for one risk;
- privacy is the most frequently mentioned risk associated with the adoption of big data technologies;
- only investments on two big data sources were influenced by the organizational size of the companies, namely RFID system data and GPS data, and only one big data source was considered to be influenced by the industrial sector of the examined companies, namely, POS data or other transactional data sources;
- only investments in two big data technologies were influenced by the organizational size of the companies, namely the adoption of machine learning software, or other software, to rapidly find the model that best fits a given data set, and the adoption of visual analytics software, or other software, to display analytical results in visual or graphic formats;
- organizations in different industrial sectors do not differ statistically in the adoption of big data technologies;
- the benefits achieved by organizations, of different size and in different industrial sectors, from the adoption of big data technologies do not differ statistically.

Acknowledgements

The author acknowledges the support of the European Community through a Marie Curie Intra European Fellowship, call “FP7-PEOPLE-2013-IEF”, Project No: 622531.

References

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.
Davenport, T. (2014). *Big data at work: Dispelling the myths, uncovering the opportunities.* Harvard Business Review Press.
Dhillon, G., & Backhouse, J. (1996). Risks in the use of information technology within organizations. *International Journal of Information Management, 16*(1), 65–74.
Ernst, & Young (2014). *Big data: changing the way businesses compete and operate*.
Farbey, B., Land, F., & Targett, D. (1999). The moving staircase-problems of appraisal and evaluation in a turbulent environment. *Information Technology & People, 12*(3), 238–252.
Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137–144.
Google trends (2016). *Big data interest over time*.
Gregor, S., Martin, M., Fernandez, W., Stern, S., & Vitale, M. (2006). The transformational dimension in the realization of business value from information technology. *The Journal of Strategic Information Systems, 15*(3), 249–270.
IDC (2016). *Worldwide big data and business analytics revenues forecast to reach $187 billion in according to IDC*.
Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., & Childe, S. J. (2016). Modelling quality dynamics, business value and firm performance in a big data analytics environment. *International Journal of Production Research,* 1–16.
Love, P. E., Irani, Z., Standing, C., Lin, C., & Burn, J. M. (2005). The enigma of evaluation: Benefits, costs and risks of IT in Australian small–medium-sized enterprises. *Information & Management, 42*(7), 947–964.
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.
Porter, M. E., & Heppelmann, J. E. (2015). How smart, connected products are transforming companies. *Harvard Business Review, 93*(10), 53–71.
Tambe, P. (2014). Big data investment, skills, and firm value. *Management Science, 60*(6), 1452–1469.
Willcocks, L., & Graeser, V. (2001). *Delivering IT and E-business value.* Boston, MA: Butterworth–Heinemann.
Wu, J.-H., & Wang, S.-C. (2005). What drives mobile commerce?: An empirical evaluation of the revised technology acceptance model. *Information & Management, 42*(5), 719–729.

Elisabetta Raguseo is Assistant Professor at Politecnico di Torino (Italy) since January 2017. Previously, she was a Marie Curie research fellow at the business school Grenoble Ecole de Management (France) from 2014 to 2016. Her research and teaching expertise is in digital innovation, big data, impacts of digital innovation on the value chain configurations, business economics, and smart working. Dr. Raguseo is author of several research papers published in international journals including International Journal of Electronic Commerce, Information & Management, International Journal of Information Management, and New Technology, Work and Employment.