Aspectual Analysis as an alternative way of understanding the definitions of Big Data

Sina Joneidy
Salford Business School

Maria Burke
Winchester Business School

Abstract

This paper considers the application of philosophy to the field of Big Data. In particular, the paper considers how "Dooyeweerd's aspects of everyday life" can contribute to the reconceptualization of Big Data. The paper reviews recent debates relating to Big Data as a concept by investigating the meaning of Big Data definitions gathered in the De Mauro study published in 2015. In doing so, Dooyeweerd's "philosophy of everyday life" can assist us, not only in finding the more precise meaning of definitions, but also in contributing to concepts that help our understanding of Big Data. In conclusion, this study shows a useful way of exploring the meaning of Big Data definitions towards affirming and enriching them.

Keywords: Big Data, definitions, diversity, everyday life, Dooyeweerd

Introduction

Big Data is seen as the core of a "new era". The application of Big Data in different fields of research and practice can be seen to be very promising. There is considerable literature regarding the potential benefits of Big Data (Manyika et al., 2011; Chen et al., 2012; Raghupathi and Raghupathi 2014).It has been said that examples demonstrating the applicability of Big Data only scratch the surface of how Big Data can effect business transformation. Many studies have tried to characterise and define the term Big Data, yet academic understanding of Big Data is fragmented, lacks clarity, and definitions are sometimes rather broad and vague. There have been attempts to address the diversity and ambiguity of the definitions of Big Data, but these attempts are not sufficient for today's – and tomorrow's – digital, seamless, connected society. There is therefore a need for a new approach which will help us to understand each definition, and provide a broad picture of the literature surrounding the definitions of Big Data.

The aim of this paper is to explore the meaning of Big Data definitions using "Dooyeweerd's modal theory" as a tool for analysis. Taking a conceptual approach, this paper demonstrates how "Dooyeweerd's suit of aspects" could help improve the conceptualisation of Big Data. This is achieved in the following ways. The first section allows for a discussion of the recent attempts in redefining Big Data. Second, the appropriateness of these attempts is considered. Third, Dooyeweerd's modal aspects are introduced as an alternative approach for analysing the concept. Fourth, the extant definitions of Big Data are aspectually analysed. Finally the results of the analysis and the possible contribution to the field are discussed.

Attempts to redefine Big Data

A convincing definition of a concept is an "enabler of its scientific development" (De Mauro et al., 2015). As Ronda-Pupo and Guerras-Martin (2012) suggest, the level of consensus shown by a scientific community regarding the definition of a concept can be used as a measure of progress of a discipline. Big Data however, has instead evolved so quickly and disorderly that such a universally accepted formal statement denoting its meaning does not currently exist. There have been many attempts to define what is meant by the term Big Data, However, none of the definitions to date have been fully satisfactory as scholars are still creating new definitions (De Mauro et al., 2015).

There is currently no single, accepted, unified definition of Big Data (Ward and Barker, 2013), although various stakeholders have provided diverse and often contradictory definitions. The lack of a consistent definition introduces ambiguity and hampers the discourse relating to Big Data. Ward and Barker (2013) aimed to gather the various definitions to create a concise definition of an otherwise ambiguous term. They argued that "Big" implies significance, complexity and challenge. The problem arises when the term "Big" also invites quantification and this then presents difficulty with definitions. From Ward and Barker's point of view most definitions have three factors in common: size, complexity and the use of technologies. According to Ward and Barker (2013) most definitions encompass at least one of these factors.

To give justice to the term, Ward and Barker believed there should be a combination of various technologies and the significant use of data sets. Therefore, as an outcome of the study, they defined Big Data as: "*Big data is a term that can be used to describe the storage and analysis of large and or complex data sets using a series of techniques including, but not limited to: NoSQL, MapReduce and machine."*

In taking the study of Ward and Barker (2013) further, De Mauro et al., (2015) examined a large number of abstracts of peer-reviewed conference and journal papers. They identified four top research themes in current literature, namely: Information, Technology, Methods and Impact. They reviewed a non-exhaustive list of previously proposed Big Data definitions (Table 1) and conceptually tie them to the aforementioned four themes of research. After analysing the commonalities between definitions they proposed a consensual definition of Big Data as *"Big Data represents the Information assets characterized by such a High Volume, Velocity and Variety to require specific Technology and Analytical Methods for its transformation into Value."*

To De Mauro et al., (2015), agreement among the definitions of Big Data comes from the acknowledgement of the centrality of some recurring attributes associated with Big Data. De Mauro et al., (2015) hope that a definition, carefully constructed, which takes into account the views of scholars and practitioners, would be less prone to attack from previous authors.

Table 1. Existing definitions of Big Data, adapted from De Mauro et al. (2015)

No.	Authors	Definitions
1	Beyer and Laney (2012)	High volume, velocity and variety information assets that demand cost-effective, innovative forms of information

		processing for enhanced insight and decision making.
2	Dijcks (2012)	The four characteristics defining big data are Volume,Velocity, Variety and Value.
3	Intel, I. T. Center.(2012)	Complex, unstructured, or large amounts of data.
4	Suthaharan (2014).	Can be defined using three data characteristics: Cardinality, Continuity and Complexity.
5	Schroeck et al.,(2012)	Big data is a combination of Volume, Variety, Velocity and Veracity that creates an opportunity for organizations to gain competitive advantage in today's digitized marketplace.
6	NIST (2014)	Extensive datasets, primarily in the characteristics of volume, velocity and/or variety that require a scalable architecture for efficient storage, manipulation, and analysis.
7	Ward and Barker (2013)	The storage and analysis of large and or complex data sets using a series of techniques including, but not limited to: NoSQL, MapReduce and machine learning.
8	Microsoft (2013)	The process of applying serious computing power, the latest in machine learning and artificial intelligence, to massive and often highly complex sets of information.
9	Dumbill (2013)	Data that exceeds the processing capacity of conventional database systems.
10	Fisher et al.,(2012)	Data that cannot be handled and processed in a straightforward manner.
11	Shneiderman (2008)	A dataset that is too big to fit on a screen.
12	Manyika et al., (2011)	Datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze.
13	Chen et al., (2012)	The data sets and analytical techniques in applications that are so large and complex that they require advanced and unique data storage, management, analysis, and visualization technologies.

14	Boyd and Crawford(2012)	A cultural, technological, and scholarly phenomenon that rests on the interplay of Technology, Analysis and Mythology.
15	Mayer-Schönberger and Cukier (2013).	Phenomenon that brings three key shifts in the way we analyse information that transform how we understand and organize society: 1. More data, 2. Messier (incomplete) data, 3. Correlation overtakes causality.

It is clear that none of the presented definitions make it easy to understand the ambiguity associated to the term. Even the definition of De Mauro et al., (2015) is more of finding the relationship between the four Vs., by rearranging and shuffling the four Vs., than actually reducing the ambiguity of Big Data concept. A "good" definition of Big Data should represent the everyday experience of all the stakeholders, which is what may be necessary in a real-life application. However, before generating and introducing another definition to the field, one needs to understand the insight within the existing definitions and the reasons why they are ambiguous.

An alternative approach

This section outlines the need for an alternative approach to defining Big Data.

- First, some of the definitions are complicated and difficult to understand. They have a tendency to be rather broad and vague, and sometimes even poetic in nature. For example Shneiderman (2008) defined Big Data as "A dataset that is too big to fit on a screen."

- Second, there are overlaps between some of these definitions as they are sharing the same core understanding, have common elements with each other and yet treated as an exclusive definition from a different perspective. For example, the overlap between definitions provided from Oracle and Intel perspectives.

- Third, looking at the list of definitions presented in the second part of this paper indicates that these definitions are the product of recent years. Considering the issues and challenges in the field, in the years ahead, there is a chance of adding more definitions to the field. We need to remember that the potential of technologies, people, and organisations cannot be limited to these definitions. So as both attempts show there is, and will be, a diversity of definitions. This being so then it would be useful to recommend an approach for accommodating the diversity in the field both for now and the future.

With all the points mentioned above it is tempting to think that the two attempts in redefining Big Data provided by Ward and Baker (2013) and De Mauro et al., (2015) are not probably sufficient and perhaps a new approach is required for analysing the definitions with the hope of opening up new avenues for defining the concept. Therefore a new solution needs to be explored which is applicable in many organisational contexts.

For tackling these problems, unlike the two attempts already discussed, this paper takes an approach which is pre-theoretical - one that allows us to consider the definitions from everyday life perspectives. This requires an approach that enables us to identify distinctly what is/are important in each definition, especially where this is multidimensional, which does not presuppose a certain context, and can view definitions as constituted in a coherence of diverse human experiences, across time and context. One approach that facilitates all these is based on modal aspects of the Dutch philosopher, Herman Dooyeweerd.

Dooyeweerd's Philosophy

In order to fully understand how Big Data impacts on our lives we need a philosophy that acknowledges the possibility of a genuine point of contact between technology and human beings. Being mostly of the "life world", with the human being in the social context, Big Data requires a philosophy that affords dignity to everyday life and to what it means to be fully and socially human. Thus materialist and rationalist philosophies are unlikely to be helpful (Eriksson, 2001). To deal with the definitions that are mostly of human origin, a philosophy is required to transcend and yet uphold the perspectives of the Big Data stakeholders.

Herman Dooyeweerd (1894-1977) who was a Dutch philosopher introduced the concept of "philosophy of everyday life". His philosophy was a reaction against the Neo-Kantian trend in continental thought prevalent at that time. One of his significant domains of thought is the modal theory. For the purpose of this study we found the modal theory was useful in meeting the research aim.

Modal Theory

The Modal Theory emerged from Dooyeweerd's comprehensive studies of theoretical thought and its relation to human reality. Dooyeweerd maintained that our thought is based upon, and bound to, our experience and that this experience exhibited a number of distinct modalities (or aspects) of organization or laws (Dooyeweerd, 1955). Accordingly a modality emerges out of human interaction with reality, which includes both perceptions and conceptions (Eriksson, 2001), and it is a particular type of knowledge that has its own unique and distinct characteristics. Dooyeweerd proposed 15 modalities which he termed "Aspects of everyday life" and these are shown in Table 2. The first column is aspects and the second column shows their kernel meaning.

Early aspects anticipate the later aspects (for example, the lingual anticipates the social) and later aspects give more meaning to earlier ones. Each aspect is a sphere of meaning that is centred on a kernel meaning. Dooyeweerd believed that kernel meaning of aspects cannot be defined by theoretical thought, but can be grasped by intuition. The aspects cannot be directly observed, but they are expressed in things, events, situations, and so on as ways these can be meaningful. All things in real-life involve functionality in a variety of aspects, usually all the aspects. By this we do not mean that aspects are different parts of things in reality, but rather that they are different ways in which it occurs meaningfully. To Dooyeweerd "each aspect plays a different but necessary part in making life richly good" (Basden, 2008). Therefore, all things within our experience make sense by reference to one or more of the aspects.

Table 2. Dooyeweerd's aspects	
Aspects	(Meaning)
Quantitative	(Discrete amount)
Spatial	(Continuous extension)
Kinematic	(Flowing movement)
Physical	(Fields, Energy, mass)
Biotic/organic	(Life, organism)
Sensitive/psychic	(Sensing, feeling, emotion)
Analytical	(Distinction, concepts, Abstraction, logic)
Formative	(Deliberate shaping, Technology, skill, history)
Lingual	(Symbolic signification)
Social	(Relationships, roles)
Economic	(Frugality, resources; Management)
Aesthetic	(Harmony, delight)
Juridical	('Due', appropriateness; Rights, responsibilities)
Ethical	(Attitude, Self-giving love)
Pistic/Faith	(Faith, commitment, belief; Vision of who we are)

Big Data such as patients' health records, IMF datasets, etc. are the product of everyday human experiences with the system and so can be thought about in terms of aspects. The present study uses the modal theory as a tool for finding and understanding the everyday life meaning of each definition of Big Data.

The next section presents an analysis of each of the definitions (as shown in the earlier Table 1) using Dooyeweerd's aspects.

Aspectual Analysis of the definitions

This section reviews ways in which Big Data researchers and commentators have conceptualized and defined Big Data. Here we look at the words which directly explain Big Data. The present investigation uses the modal theory of aspects as a tool of investigation. We represent our analysis in Table 3. The second column indicates the authors, and the third column shows our analysis of the definitions.

Table 3. Aspectual analysis of the definitions of Big Data

No.	Authors	Aspectual Analysis of the definitions in Table 1
1	Beyer and Laney (2012)	Volume is functioning in Quantitative and Spatial aspects. Velocity is functioning in Quantitative and Kinematic aspects. Variety is Analytical aspect. Cost-effective is functioning in Economic aspect. An innovative form of processing information is functioning in Formative aspect.
2	Dijcks (2012)	Volume is functioning in Quantitative and Spatial aspects. Velocity is functioning in Quantitative and Kinematic aspects. Variety is Analytical aspect. Value is referring to the facilitating conditions. It shows Oracle belief that existing technical infrastructures are available to support storage and analysis of Big data. That is giving due to the infrastructures which were supporting data management. This is functioning in Juridical aspect.
3	Intel, I. T. Center.(2012)	Being Complex is functioning in Analytic aspect. Unstructured is mainly Formative aspect. But here by Complex and unstructured the author means large amounts of data which is a functioning in Quantitative and spatial aspects.
4	Suthaharan (2014)	Cardinality is a functioning in Quantitative aspect. Continuity is a function in Spatial aspect. Complexity on its own is a functioning in Analytical aspect.
5	Schroeck et al., (2012)	Volume is functioning in Quantitative and Spatial aspects. Velocity is functioning in Quantitative and Kinematic aspects. Variety is Analytical aspect. Veracity refers to trust which is mainly a functioning in Pistic aspect.
6	NIST (2014)	Emphasize on volume, velocity and variety. Volume is functioning in Quantitative and Spatial aspects. Velocity is

		functioning in Quantitative and Kinematic aspects. Variety is Analytical aspect. Here the author refers to the efficiency in storing, manipulating and analysis of Big Data which is mainly functioning in Economic aspect.
7	Ward and Barker (2013)	Here Ward and Baker (2013) emphasize in on the large and/or complex data sets. Large is a functioning in both Quantitative and spatial aspects. If complex here means large then it is function in the same aspects, but if not then it is Analytical aspect.
8	Microsoft (2013)	Massive is mainly both Quantitative and Spatial aspects, highly complex is mainly Analytical aspect.
9	Dumbill (2013)	Emphasize is on exceeding the processing capacity of conventional database systems. This exceeding is referring to a kind of mastery and power which is a functioning in Formative aspect.
10	Fisher et al.,(2012)	This is a vague definition. The author is telling us what the Big Data is not. Not handled and processed in a straightforward manner is an emphasis on complexity in handling which is both Analytical and Formative aspect.
11	Shneiderman (2008)	Dataset that is too big to fit on a screen is a function in Economic aspect.
12	Manyika et al., (2011)	Emphasis is on the size which is beyond the ability of typical software tools. Similar to Dumbill (2013) this is referring to the mastery power of Big Data which is mainly Formative aspect.
13	Chen et al.,(2012)	Similar to Microsoft (2013), again Large here is a functioning in Quantitative, Spatial aspects and complex is a functioning in Analytic aspect.
14	Boyd and Crawford (2012)	Both cultural and technological are functioning in Formative aspect. Being scholarly is functioning in Analytical and Formative aspect.

15	Mayer-Schönberger and Cukier (2013)	Key shift and organising society is a functioning in Formative aspect.

Summary of analysis

Table 4 indicates all the results together. The first column shows the numbers associated with the definitions, the second column shows their related main aspects and the third column is the secondary aspects. Main aspects were derived via aspectual analysis, an understanding of their kernel meaning. Notice that some definitions are the manifestation of three aspects, which is mainly because the three aspects were considered as equally important to the desired definition. For most definitions there was a chance of finding other aspects. These aspects are deduced from their associated definitions. In this initial study we only focus on the main aspects.

Table 4. Summary of the aspectual analysis

Definitions No.	Main aspects	Secondary aspects in their order
1	Quantitative, Spatial , Kinematic, Analytical	Economic, Formative
2	Quantitative, Spatial , Kinematic, Analytical, Juridical	None
3	Quantitative, Spatial	None
4	Quantitative, Spatial, Analytical	None
5	Quantitative, Spatial , Kinematic, Analytical, Pistic	None
6	Quantitative, Spatial, Kinematic,	Economic
7	Quantitative, Spatial, Analytical	None
8	Quantitative, Spatial, Analytical	None
9	Formative	None
10	Analytical, Formative	None

11	Economic	None
12	Formative	None
13	Quantitative, Spatial, Analytical	None
14	Analytical, Formative	None
15	Formative	None

To have a better view of comparison between the different aspects, Table 5 below presents the results in another format. The first column shows fifteen aspects in their usual order and the second column shows the number of times one aspect has been the main sphere of meaning for different definitions of Big Data.

Table 5. The frequency of the aspects in the definitions of Big Data

Aspects	Frequency of aspects
Quantitative	9
Spatial	9
Kinematic	4
Physical	0
Biotic/organic	0
Sensitive/psychic	0
Analytical	9
Formative	5
Lingual	0
Social	0

Economic	1
Aesthetic	0
Juridical	1
Ethical	0
Pistic/Faith	1

The Table 5 illustrates that Quantitative, Spatial and Analytical aspects have appeared more as the main aspects in different definitions compared with other aspects. Formative aspect has been the main component of conceptualising Big Data five times. Kinematic repeated four times as the main aspect. From Social and Normative level of real-life only one aspect has received attention one time, which is Economic aspect. Interestingly, from Societal and Normative level two aspects (i.e. Juridical and Pistic) have received attention once for each.

Dooyeweerd's aspects were used as a tool to analyse the definitions of Big Data. In the next part of the paper, the value of the Dooyeweerd's approach is discussed, based on these results.

Discussion

Investigation of meaningfulness based on what is more important in each definition yielded in the aspectual picture which provides various ways of discerning characteristics of Big Data as a whole, some of which support what we might already suspect, while some disclose new and surprising things.

So what?

The most important factor that has resulted from this analysis is that eight of Dooyeweerd's fifteen aspects are represented. The range of aspects represented in the collection means that the current definitions do not offer a wide range of exemplars. Few aspects predominate, and the others feel like outliers. Though certain aspects - the Quantitative, Spatial, and Analytical - occur more frequently, there is no more or less attention to other aspects. These most occurred aspects are not isolated from the rest.

This shows that, using Dooyeweerd we can identify gaps in conceptualising Big Data. Showing gaps allows us to provide a way of reducing the ambiguity in the definition of Big Data, one factor which was the aim of research for Ward and Barker (2013) and De Mauro et al., (2015). Ward and Barker (2013) intention was to give justice to the term. But there is an alternative way of doing that. If things are centred on meaningfulness, then this gives a philosophical basis for considering each of the definitions to understand which aspects were and were not taken into account. Dooyeweerd's aspects are all equally important in our everyday life. To give justice to the term Big Data, the existing definitions of Big Data would benefit from the Modal Theory.

Dooyeweerd argued that there is no incompatibility between the aspects and all work in harmony with the others, as instruments of an orchestra do when playing a symphony. This leads to the Shalom Principle: that if we function well in every aspect then things will go well, but if we function poorly in any aspect, then our success will be jeopardized (Basden, 2008). For example the definition "complex unstructured large amounts of information" is too general, and would benefit from other aspects. It may be helpful, to take into account equally other aspects such as Pistic, Ethical, Juridical, Aesthetic, Social, Lingual. It is Dooyeweerd's contention that we function in all aspects: that all of these aspects work in harmony in a thing.

Lingual aspect is concerned with the meaning of the data. This would seem important in Big Data; surely if we do not know what the source of data means, we will not deal with it correctly. Social aspect is about the status or affiliations of the sources. For example, the head of a political party might lead to biased data. Looking at Economic aspect, waste and superfluity is a problem. Big Data is massive, but if there is a way of reducing waste without losing data, that is surely important. Looking at Aesthetic aspect, one can think of the harmony of the data and the data sources and question how they all fit in a large picture. Both rightness of the Big Data source and appropriateness of the analysis techniques are highlighting the importance of Juridical aspect. Thinking about the user of the findings of Big Data analytics is Ethical. So much data is collected and mined by Facebook, Google, mobile phone companies that raise the issue of users' awareness. It is not clear where in the four or five Vs. we would be able to fit the Ethical aspect. Finally, trustworthiness of the data and the processing algorithms or techniques is centred on Pistic aspect.

These all go beyond the Vs in the existing definitions of Big Data and need our attention in conceptualising the concept. From Dooyeweerdian point of view each aspect brings a blessing to temporal reality, which cannot be obtained from other aspects, even in combination. None of the aspects could be alleviated against others.

We also observed that while De Mauro et al., (2015) see the consensus between definitions coming from centrality of some attributes (i.e. information, method, technology and impact), the Dooyeweerdian approach enables us to see the fifteen aspects of everyday life as a common ground for all definitions and upcoming ones. This way Dooyeweerd helped us to shed a new light on the definitions of Big Data by extracting the essence of what Big Data means to academics and practitioners.

Conclusion

The aim of this paper was to explore the meaning of Big Data definitions using "Dooyeweerd's modal theory" as a tool for analysis. This has been achieved through the aspectual analysis of the extant definitions in the literature.

The paper discussed the possibility of applying Dooyeweerd's aspects to the definitions of Big Data by seeking to understand in which sense the definitions are meaningful. In the first place, the vagueness of the concept inspired the scholars to make an attempt in redefining the concept of Big Data. To some extent Ward and Barker (2013) and De Mauro et al., (2015) have addressed the ambiguity of the definitions. But in doing so, their approach seems not to be sufficient. This initial study has provided a way of showing gaps in the definitions of Big Data.

This research has implications in line with the interests of those who are concerned about the concept, challenges and improvement of research in the field of Big Data. Regarding the discussion section, this study contributes to the attempts in defining and conceptualising the concept of Big Data.

First, this study has concentrated on the definitions of Big Data in the literature. This paper has not aimed to criticise the definitions of Big Data, but, by finding the spheres of meaning in the definitions, it has shown the gaps in the concept of Big Data. We hope this opens up new avenues for thinking for those interested in redefining the concept of Big Data. Dooyeweerd's philosophy helped us to have a pre-theoretical view and consider the definitions from the everyday life experience. The idea that early aspects anticipate the later ones will help academics to ponder about a new definition of Big Data and recommend one, which includes multiple aspects.

Second, if we see each aspect as a distinct category but interrelated to other aspects, then we are able to categories about 16 definitions into 8 aspects. The advantage of this is that, compared to 4 aspects considered by De Mauro et al., (2015), future research would be able to examine current and upcoming definitions on the basis 15 aspects.

This preliminary review provides a basis from which to investigate and target the gaps in how the Big Data is defined. Future research should i) draw attention to these definitions and provide a way of enriching them, ii) look at the ways of accommodating diversity of, and addressing the overlaps in, the definitions of Big Data.

Acknowledgement:

The authors would like to acknowledge thanking Professor Andrew Basden from the University of Salford for providing insight and expertise that greatly assisted the research.

References

Abawajy, J. (2015). Comprehensive analysis of big data variety landscape. *International Journal of Parallel, Emergent and Distributed Systems*, *30*(1), 5-14.

Basden, A. (2008). Philosophical Frameworks for Understanding Information Systems. Herschey, PA, USA: IGI Global.

Beyer, M. A., & Laney, D. (2012). The importance of 'big data': a definition. *Stamford, CT: Gartner*.

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*, *15*(5), 662-679.

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.

De Mauro, A., Greco, M., & Grimaldi, M. (2015). What is big data? A consensual definition and a review of key research topics. In *AIP Conference Proceedings* (Vol. 1644, pp. 97-104).

Dijcks, J. P. (2012). Oracle: Big data for the enterprise. *Oracle White Paper*.

Dooyeweerd H. (1955). A New Critique of Theoretical Thought, Vol. I-IV, Paideia Press (1975 edition), Jordan Station, Ontario.

Dumbill, E. (2013). Making sense of big data. *Big Data*, *1*(1), 1-2.

Fisher, D., DeLine, R., Czerwinski, M., & Drucker, S. (2012). Interactions with big data analytics. *interactions*, *19*(3), 50-59.

Intel, I. T. Center.(2012) Big data analytics: Intel's it manager survey on how organizations are using big data.

Larson, E. B. (2013). Building trust in the power of "big data" research to serve the public good. *JAMA*, *309*(23), 2443-2444.

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.

Mayer-Schönberger, V., & Cukier, K. (2013). Big data: A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt.

Microsoft, (2013), available at https://www.microsoft.com/en-us/news/features/2013/feb13/02-11bigdata.aspx.

Nafus, D., & Sherman, J. (2014). Big Data, Big Questions, This One Does Not Go Up To 11: The Quantified Self Movement as an Alternative Big Data Practice. *International Journal of Communication*, *8*, 11.

NIST (2014) Big Data Public Working Group, *Big Data Interoperability Framework: Definitions (draft)*.

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health Information Science and Systems*, *2*(1), 3.

Ronda-Pupo, G. A., & Guerras-Martin, L. Á. (2012). Dynamics of the evolution of the strategy concept 1962–2008: a co-word analysis. *Strategic Management Journal*, 33(2), 162-188.

Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., & Tufano, P. (2012). Analytics: The real-world use of big data. *IBM Global Business Services, Somers*.

Shneiderman, B. (2008, June). Extreme visualization: squeezing a billion records into a million pixels. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data* (pp. 3-12). ACM.

Suthaharan, S. (2014). Big data classification: Problems and challenges in network intrusion prediction with machine learning. *ACM SIGMETRICS Performance Evaluation Review*, *41*(4), 70-73.

Ward, J. S., & Barker, A. (2013). Undefined by data: a survey of big data definitions. *arXiv preprint arXiv:1309.5821*.