Working together
www.rcis.ro

Revista de cercetare și intervenție socială

ISSN: 1583-3410 (print), ISSN: 1584-5397 (electronic)
Selected by coverage in Social Sciences Citation Index, ISI databases

Big Questions on Big Data

Dumitru OPREA

Revista de cercetare și intervenție socială, 2016, vol. 55, pp. 112-126

The online version of this article can be found at:
www.rcis.ro, *www.doaj.org* and *www.scopus.com*

Published by:
Expert Projects Publishing House

expert projects
publishing

On behalf of:
„Alexandru Ioan Cuza" University,
Department of Sociology and Social Work
and
Holt Romania Foundation
REVISTA DE CERCETARE SI INTERVENTIE SOCIALA
is indexed by ISI Thomson Reuters - Social Sciences Citation Index
(Sociology and Social Work Domains)

Big Questions on Big Data

Dumitru OPREA[1]

Abstract

The present paper aims to explore the Big Data concept in order to answer the question whether this is a necessity or a hidden agenda to promote disrupting new technologies. It was ascertained that, unlike the *traditional data*, *the new data deluge* would have special characteristics; therefore they would be completely new. The performance of the current technologies could be enhanced by using the concepts of *bionic data* and *abionic data*. And this way, many of the *abionic data* would become *bionic data*, through the use of IoT, Big Data Analytics and Big Data Technologies. The analysis of Big Data definitions delivers three types of statements: the new avalanche of data has become uncontrollable; the current technologies, including the software for relational databases management, are unable to cope with the new requirements; new technologies are required to access and process the new types of semi-structured and unstructured data. In our opinion, the characteristics of traditional data cover also those of the Big Data. The characteristics described by words beginning with the letter V do not refer to data, but to Big Data Analytics and Big Data Technologies. Both are extensions of previous data analysis and existing technologies. Therefore, the concept of Big Data promotes the disruption process, embodied in new types of data, analysis models and technologies. The major changes in our society in the last decades are so dramatic that the previous global, national or organizational information systems are easily categorized as traditional systems. Are they becoming the target of the disruption phenomenon due to new technologies?

Keywords: Big Data, disruption process, bionic data, abionic data, V's Big Data characteristics.

[1] Alexandru Ioan Cuza University of Iasi, Faculty of Economics and Business Administration, Department of Business Information Systems, Iasi, ROMANIA. E-mail: doprea@uaic.ro

Introduction

When searching on the Internet, the term “big data” shows up in the results twice as much as the term “globalization”. Beer (2016) stated that “big data is a concept that has achieved a profile and vitality that very few concepts attain”. One could run from the big data, as well as from globalization, however one cannot escape them. The big data topic, regardless of industry and profession, is pursued by everybody; as also Moorthy *et al.* (2015) observed “different stakeholders such as consumers, companies and businesses are likely to exploit the potential of Big Data”. Etzion and Aragon-Correa (2016) also note: “Big Data, The Internet of Things, and the Deep Web have captured the imagination not only of Silicon Valley but also of other industries, governments, and non-profits all developing novel offerings that employ vast troves of data”. Kitchin (2014) has an interesting point of view, “Big Data, coupled with new data analytics, challenges established epistemologies across the sciences, social sciences and humanities, and assesses the extent to which they are engendering paradigm shifts across multiple disciplines”. A similar opinion we found in (Taylor, 2014). Also, Beer (2016) notes: “Big Data is a concept that has achieved a profile and vitality that very few concepts attain”.

We believe that it is natural to use the big data concept within the vocabulary of a world that’s information flat. Data is the foundation of all information, which is one of the most widely utilized concepts. The word “information” is used almost four times more frequent than „big data”. This is expected, since Substance, Energy and Information are all aspects of the same Fundamental Existence (Stancovici, 1975) and they are part of the ecosystem flow.

Even if, conceptually, Big Data has surfaced quite recently, many success stories have been circulated about them, many contradictory opinions have been exchanged and some authors (Beer, 2016) have attempted to write the story of Big Data by asking the question “How should we do the history of Big Data?”, “treating Big Data as both a material phenomenon and also a concept”. The answers are different, based on the industry of those formulating them; so results “there are multiple forms of Big Data” (Kitchin & McArdle, 2016) and “Big Data is a hugely diverse term, and it can be applied in many other ways” (Harper, 2016). They may be either in business, or culture, politics, sports, sociology, research, teaching staff, experts, etc. For instance, the granularity of data has various meanings from one field to another, but it follows the general rule of the system granularity. Even the bit, that could take the values *0* or *1*, representing a certain state at a certain time, is „complemented” by the qubit that can be both *0* and *1* at the same time, according to the principles of quantum computers. These complement the traditional computers, however they do not replace them. Therefore, the interdependence, complementarity phenomenon takes a more defined shape. From this perspective and given the role played by new technologies and new types of semi-structured or non-structured data, Big Data can be conceptually viewed as complementary to the traditional data.

And if Big Data is perceived as a "universal good", it should be supported, according to (Beer, 2016), by "an analytic framework for structuring the analysis of the work done by the concept Big Data". It can also justify the current paper title, "Big Questions on Big Data", because Beer presented 27 'analytical questions and issues'. If the author's table would have been referred to Big Data Analytics and not only to Big Data, traditional data existing in the systems, but not yet explored would have also been included.

If we extend the discussion to the existing, but not yet explored or identified data, it would be time to launch a challenge. We are referring to the "biotic" and "abiotic" concepts used especially in biology and ecology to separate the "living" from the "non-living", knowing that "an ecosystem contains *living and non-living things*". The goal of the specialized researchers is to identify *abiotic factors*. Likewise, with the new technologies and the possibility to identify several types of data in the world of Internet of Things (IoT), by using ultra-miniature sensors like *smart dust*, we could be *bringing to life* or activating certain long-time data within the *non-living* systems. The performance of the current technologies could be enhanced by using the concepts of *bionic data* and *abionic data*. And this way, many of the *abionic data* would become *bionic data*, through the use of IoT, Big Data Analytics and Big Data Technologies.

The new technologies as triggers of disruption

Disruption represents a drastic change within a certain period of time. In the Romanian language, there is also the concept of dysthanasia, which means a certain and difficult death, with prolonged coma and pain. We can state that the old, traditional systems, due to the disruption process, through dysthanasia, disappear and new systems emerge from new technologies. Therefore, the concept of sustainable development could be replaced by continuous and sustainable disruption. An interesting question about relation between big data and disruption could be finding in Wessel (2016). The traditional businesses world, also called the offline world, undergoes a major transformation enabled by new technologies and caused by digitalization and the creation of the digital world. New business models are shaped around the concepts of Social Media, Cloud computing, Big Data, Mobile Solutions, Video and unified communication, Blockchain ledger, Internet of Things and Internet of Everything. A similar point of view we found at Ksherti (2014): "Massive amount of data generated by Social Media, cellphones, and other digital communication tools, [...], are a true form of big Data".

This significant global digital disruption introduces new concepts correlated to sharing, such as: sharing economy, sharing city, sharing region, sharing country, sharing global etc. New related words are also surfacing: co-interests, co-creation, co-dependence, co-habitation, co-vision, co-mission, co-strategy, co-innovation, co-design, co-manufacturing, co-operation, co-experience, co-activities, co-beneficiary, etc. With the deployment of digitalization and sharing process, the data world was powerfully impacted by a real flood of data, widely known as Big Data. Is this a true disruption phenomenon? Is this a natural, spontaneous phenomena or a controlled one?

Big Data, necessity or marketing strategy

In a 2010 posting on my personal blog (Oprea, 2010), I asserted that, from the very beginning, we were somewhat holding back our need for more data. Is this a real need or an induced one? Is there a wish to create a certain behavior towards data? Is it intended to prove the inability of organizations faced with information overload? The answers to these questions seem to be affirmative, since more often than not these statements are common sense. It is stated that data cannot be controlled anymore; that the volume of data is too high when compared to those stored in the current database management systems; that the data analysis must be handled by increasingly performant statistical packages which need simultaneous processing on dozens, hundreds or thousands of servers. Also, Henke *et al.* (2016) notice that “for all the enormous promise, most companies have so far struggled to realize anything more than modest gains from their investments in big data, advanced analytics, and machine learning. Many organizations remain preoccupied with classic large-scale IT-infrastructure programs and have not yet mastered the task of creating clean, powerful, linked data assets; building the capabilities they need to extract insight from them; and creating the management capacity required to lead and direct all this”.

When analyzing the above statements, three hypotheses emerge: (1) The induction of the idea that data abundance is so high that they have to become more necessary, but also more difficult to control; (2) The new technologies, including the software for relational databases management, are outdated and unable to cope with the new demands; (3) It is necessary to transition to higher performance new technologies in order to access the new types of data.

There has been a precedent, known in the academic literature as Moore’s law, which had a similar goal: creating certain behaviors for the users of equipment incorporating integrated microprocessors. The law was originally launched in 1965, while in 1968 Robert Noyce and Gordon Moore pioneered the semiconductors industry, founding Integrated Electronics Corporation, which later became the famous Intel Corporation.

Many opined that, in fact, it was not a practical law, but a theoretical concept driven by the desire to control the industry progress. Regardless, it has been „honored" for almost 50 years. The law says that, every 18 months, the microprocessor performance doubles, while the costs, respectively the prices, remain constant or are further reduced. However, there have been exceptions to the doubling of the performance rule. For example, in the case of bitcoins mining, more powerful microprocessors were needed, forming Application-Specific Integrated Circuits (ASICs). Despite of what the law previously stated, the timing of the doubling in performance was also cut back significantly to accommodate the high market demanded for such a fast progress.

Moore's law is somewhat applicable to the Big Data. Back in 2003-2004, the early adaptor was Google. Later on, Apache Software Foundation was started, which further developed Apache Hadoop. They focused on the storage, processing and analysis of high data volumes, which became Big Data. IBM, Dell, Amazon, Cloudera, eBay, Yahoo, Twitter, Facebook etc. entered the stage, embracing Hadoop's framework. We could call them all „New Moore", since, similar to Moore, they create new behaviors by bombarding with information about the spectacular evolution of the data flood. These data cannot be managed with relational databases and are categorized as NoSQL (referred to as „noSQL", „not only SQL" or „NewSQL"); relevant examples would be BigTable, Name_ value (Key_Value), Document, Graphical.

To verify the three hypotheses, we analyzed a multitude of definitions of the Big Data concept and are presenting below the most relevant ones. According to Deloitte's vision, "Big Data refers to the set of problems – and subsequent technologies developed to solve them - that are hard or expensive to solve in traditional relational databases" (Delloite, 2013). IBM considers that ½Big Data has been used to convey all sorts of concepts, including: huge quantities of data, social media analytics, next generation data management capabilities, real-time data, and much more. Whatever the label, organizations are starting to understand and explore how to process and analyze a vast array of information in new ways. In doing so, a small, but growing group of pioneers is achieving breakthrough business outcomes.½(Schroeck *et al.*, 2012).

According to ISO/IEC JTC 1, "Big Data is a data set(s) with characteristics (e.g. *volume, velocity, variety, variability, veracity*, etc.) that for a particular problem domain at a given point in time cannot be efficiently processed using current/existing/established/traditional technologies and techniques in order to extract *value*" (ISO/IEC JTC 1, 2014).

Financial Times says that „Big Data is a vague term for a massive phenomenon that has rapidly become an obsession with entrepreneurs, scientists, governments and the media" (Harford, 2014). Press (2014), a Forbes collaborator, presents 12 definitions provided by various personalities from different fields of study, as well as a link to the views of over 40 opinion leaders. Next we highlight five of these opinions which are relevant for our goal: (1) “Big Data is data that requires novel processing techniques to handle. Typically, big data requires massive parallelism in some fashion (storage and/or compute) to deal with volume and processing variety” (Brad Peters, Chief Product Officer at Birst); (2) “Big Data is data that can’t be processed using standard databases because it is too big, too fast-moving, or too complex for traditional data processing tools” (AnaLee Saxenian, UC Berkely School of Information); (3) “Big Data is datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze”. (McKinsey researchers); (4) “The problems with definition of big data (as other experts I have consulted) predict a relatively short life span for this unfortunate term” (Tom Davenport, Harvard Business School); (5) *Large* and *traditional* are relative and ambiguous (and potentially self-serving for IT vendors selling either „more resources” of the „traditional” variety or new, non-„traditional” technologies) (NASA researchers). These opinions are summarized in *Table 1*.

Table 1. Summary of opinions on Big Data definitions

The definition author	Common ideas in the Big Data definitions				
	One or several datasets with special characteristics	No. of Big Data characteristics	Inability of current, soft and RDBMS technologies	New data flows	New technologies
ISO	✓	5V	✓		
Deloitte	✓	-	✓	✓	✓
IBM		4V		✓	✓
Gartner 2012		4V		✓	✓
Brad Peters				✓	✓
AnaLee Saxenian			✓	✓	✓
McKensey researchers	✓		✓		
Tom Davenport	He does not agree with Big Data				
NASA researchers	Ambiguity		✓ Growth of sales for traditional resources	Ambiguity	✓ Growth of new technology sales

Analyzing the viewpoints of specialized organizations and experts, we observed the following definition highlights: (1) Big data means datasets with special characteristics, described by a different number of Vs (derived from the words that define the characteristics); (2) Big Data solves the problem of the „traditional” technologies inability, especially the capability of the relational database management systems to deal with new sources, to cope with the new types of data and other progress elements in the industry; (3) Big Data includes the new flows of data, especially the semi-structured and unstructured ones, as well as the new technologies, therefore visibly improving the decision making process in organizations.

In other words, since one is provided with modern flows of data and the latest equipment, one must use them, but Poon (2016) note that "If social scientist are going to have a critical perspective on how big data analytics and the Internet of Things are about to reformat economic life, they must retain two points from this story. First, that *technical accuracy is not the foundation of profitability in data-driven industries*. [...] The second takeaway message is that *financial activity does not follow a singular logic*".

As Shaw (2015) stated "The study of Big Data could lead to a more comprehensive understanding of social reality. But achieving that understanding will require developing a sense of the complex materiality of our Big Data-producing information systems, and empathy for the people who fund, design, build, use, and exploit them. Without that sense and empathy, when we are asked what we have learned from Big Data, we may be left pointing mutely at our data centers".

There are many opinions about the relatively general character of the Big Data concept, which some of the authors (Guszcza *et al*., 2013) refer to as *vagueness*. Hence, certain clarifications would be appropriate: we believe that as the traditional systems based on Enterprise Data Warehouse (EDW) platforms became Big Data, the transition to more advanced *data volume* performance stages could happen in a similar manner. This proposal is presented in *Figure 1*.

Daniel Gutierrez's statement (2014) is also relevant to this point of view: "The definition [of Big Data - author's note] has changed where big is more influenced by data variety and velocity. When it comes to data volume, big is relative". Similar position we find in (Dalton, 2015).

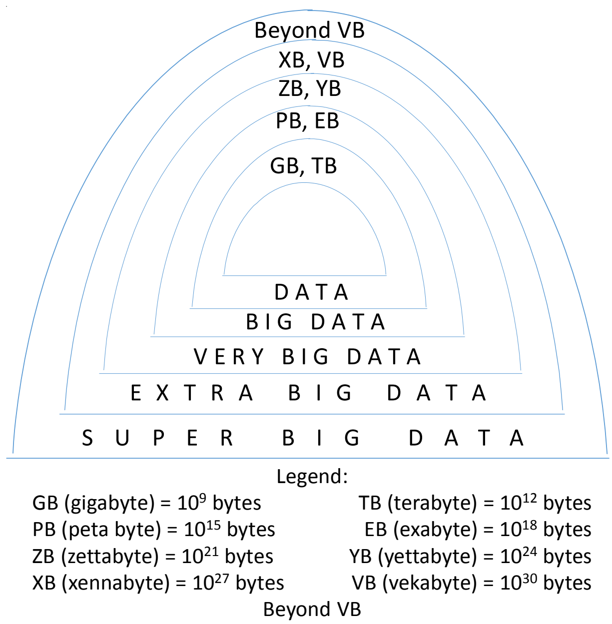

Figure 1. The evolution of the Big Data concept based on the data volume measurement units

Big Data characteristics

The related academic literature deals with data, information, knowledge and wisdom in a pyramidal structure, with the data positioned at the base of the pyramid.

These concepts have undergone a complex processing and were included in multiple standards, the most significant being ISO 8000, which presents the data quality characteristics and ISO 9000, which defines quality. The scope of this paper does not cover the various standards in the data field, but the potential changes of data characteristics when they appear from a multitude of new sources, under multiple forms and at incredibly high speeds. The information quality improvement is the result of the mankind's progress, causing technical changes in the multitude of operations they are submitted to; meantime, the data, as a "disruption of an ongoing process" (Benson, 2005), remain very well defined, including their qualities and characteristics.

The data concept does not seem to be the focus of the debate, but the Big Data concept, which was defined in the first part of this analysis. Most of the references are made to the helpless *traditional technologies* and to the *new technologies* capability to collect data and render them under various forms to interested users. This would be also the reason why Big Data is defined through specific characteristics, without reference to the traditional characteristics of data or data quality.

Therefore, it is stated that the Big Data does not have a single, widely recognized definition. Most of the definitions are based on the 3Vs: "Volume, Velocity and Variety" (European Commission, 2016), referencing an opinion published on a Gartner blog from 2001 (Laney, 2001) and ignoring a 2012 a post on the same blog which added another V, standing for "Value".

The 2016 European Commission's position did not even took into account the ISO opinion released on "Big Data - Preliminary Report 2014, ISO/IEC JTC 1 - Information Technology". It mentioned the Big Data characteristics as the following 5Vs: Volume, Variety, Velocity, Variability, and Veracity, however Value is left out of this list of Vs.

At any rate, we may conclude that a real competition of searching for V-words has been launched, despite not having much in common with the Big Data characteristics and not being clearly defined. Therefore, new words were introduced: *visualization*, *viscosity*, *virality*, *validity*, *venue*, *vocabulary*, *vagueness*, *volatility*, *virtuosity* and many more. Seeing this list, we are tempted to find non-existing variants like *variform* and *vastitude*. But we need to ask ourselves the following questions: How long will the competitions for as many Vs as possible last? How many of the characteristics are specific to Big Data? How many of them refer to Big Data Analytics?

Looks like few trends are emerging and some are contradictory: (1) establishing standards for the specific Big Data characteristics, including their number; (2) abandoning the requirement that a characteristic should be defined by a word beginning with the letter V; (3) changing the word-characteristic to another beginning letter (for instance, the letter P), would be, in our opinion, entirely uninspired; (4) separating the Big Data characteristics from those of the Big Data Analytics; (5) abandoning the Big Data concept and, implicitly, the Big Data characteristics debated here.

The description of the Big Data Vs frenzy reminded us of a similar case in marketing, where to the 4Ps originally introduced by McCarty to define the marketing mix concept (Product, Price, Promotion, Placement), over time, other characteristics were added, reaching over 30Ps. The example is quite relevant. Still, the topic is unfairly discussed without taking in consideration the data characteristics or properties, "preparing big social data for analysis and conducting actual analytics involves a plethora of decisions, some of which are already embedded in previously collected data and built tools” (Diesner, 2015). A matrix approach would have been more appropriate, since it identifies the required data quality characteristics (Colorado STEM Pipeline, 2013) based on the behavior of Big Data characteristics (ISO/IEC JTC1, 2014), as presented in *Table 2*.

Table 2. Assessing Big Data characteristics based on data quality characteristics

Data Quality Metrics / Big Data Characteristics	Accuracy	Validity	Reliability	Timelines	Relevance	Completeness
Volume	No / Yes	No / Yes	No / Yes	No / Yes	No / Yes	No / Yes
Variety	No / Yes	No / Yes	No / Yes	No / Yes	No / Yes	No / Yes
Velocity	No / Yes	No / Yes	No / Yes	No / Yes	No / Yes	No / Yes
Variability	No / Yes	No / Yes	No / Yes	No / Yes	No / Yes	No / Yes
Veracity	No / Yes	No / Yes	No / Yes	No / Yes	No / Yes	No / Yes

The elements presented in the table above allow for a complete assessment of the data quality characteristics and for flexibility to adjust them to new requirements, specific to Big Data Analytics, based on data sources, types, purchase, processing, use, legality and economic feasibility procedures and as presented in *Figure 2*. This way, *Table 1* could be compatible with Google Cloud Big Table, which represents the new noSQL database service for scalable bulk data.

Hence, while focusing on Big Data characteristics, we reached the procedures specific to the Big Data Analytics, which actually means a consolidated need to transition to new technologies such as: Apache Hadoop, Apache Spark, NoSQL, Machine learning and Data Mining, Statistical and Quantitative Analysis, Data Visualization etc.

All of these things to happen, the labor market, in turn, as is stated also in (Bughin, 2016), need to promote new types of profession for coming years, like: Big Data Engineer, Machine Learning Scientist, Data Visualization Developer, Analytics Manager, Machine Learning Engineer, Data Scientist, Data Visualizer, Data Architect, Data Analyst, Research Analyst, Business Analyst, etc.

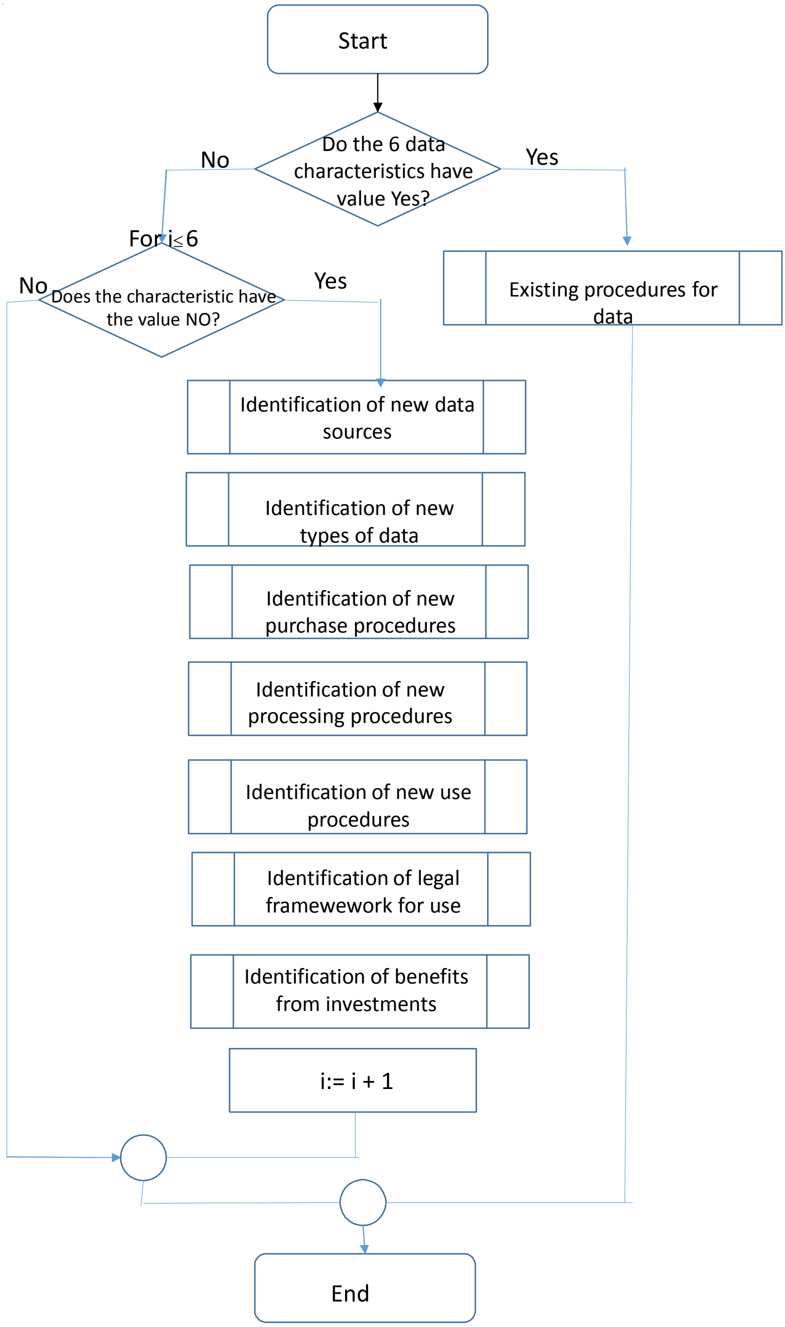

Figure 2. The extension of existing ½traditional½data procedures to Big Data procedures

The financial and accounting treatment of the Big Data concept

Towards the end of the XXth century, it was common knowledge that the masters of the global information fortunes will rule the world. Let's assume that the planet's wealth will be approximately 1 billion of billions of dollars. Which percentage of this wealth would belong to data, information and knowledge? Nobody can ever estimate. It is also well known that these fortunes, as *information wealth, are intangible assets*. However, nobody argues that, regardless of their preciousness, they are the worst managed resources, if one can even say that they are managed. More often than not, it is commonly accepted that if the financial resources of an organization would be managed the same way as its data, that organization would go bankrupt in few seconds. This is the reason for Taylor *et al*. (2014) to note: *"the value of Big Data to the discipline may lie partly in creating a stimulus for new ways of thinking, but specifically in challenging economists to imaginatively apply an economic perspective to the evolving digital landscape"*.

We mainly mentioned data because they represent the foundation of an organization's information wealth. Despite the multitude of national and international regulations, the data do not benefit yet from the appropriate financial and accounting treatment. If land, equipment, buildings, animals have clearly defined accounts within their respective account charts, data and people employed by organizations do not have a similar treatment. There are no dedicated accounts for these items. In other words, the accounting need to find the answer for the following question (Liu & Vasarhelyi, 2014): *"Have new data analysis techniques affected accounting practice?"*

The data, through quality standards, have well defined characteristics, to increase the efficiency of the data collection and meet the expectations of the potential beneficiaries. We believe that Big Data was intentionally introduced as a separate concept in order to convey the message that it is about something else than the data already known. Therefore, characteristics specific to Big Data were introduced without enough justification. The high volume, the high velocity, the big variety, etc. suggest that, in comparison to what already exists, the values of certain measurements of data characteristics would lead one to the conclude that there would be new characteristics for Big Data. This is totally false. The speed of the trip from Iasi to Bucharest by bike or by plane does not lead to the replacement of Velocity as characteristic to Big Velocity, even if the performances of the same characteristic are so visibly different: 20 km/h by bike versus 500-700 km/h by plane.

Therefore, from a financial and accounting perspective, both the data and the Big Data, are also intangible assets; according to the International Accounting Standards 38 (IAS 38) *Intangible Assets* represent "an identifiable monetary asset without physical substance and possessed for use in the production process or the supply of goods or services, for rent to third parties or for administrative goals".

An asset is categorized as intangible only if the entity that has the property rights proves that: (1) the asset is identifiable; (2) future economic profits are obtained through that respective intangible asset, for the entity that has the property right; (3) the cost of intangible asset can be evaluated in a credible manner. Identifiable intangible assets are classified in four categories, which may represent ownership rights obtained legally, by means of various relationships, grouped intangible assets and intellectual property rights.

The legal and financial-accounting treatment of intangible assets is very subjective and it is not the focus of this paper. This also explains the inability to accurately evaluate the information assets in organizations.

Conclusion

If initially, the transition to Big Data seems a voluntary action based on the previously presented findings, in reality, things could not be more different. The more Big Data characteristics, meaning more Vs, the more complicated the procedures. Acquiring new technologies and new types of data means new tangible, intangible and financial investments, leading to new internal and external costs; in these cases the treatment should be done differently. All issues derived from regulatory requirements must be solved in a commonly agreed manner by all the Big Data market participants. It was also the reason for the European Union, similarly to the American trend, to propose setting up functional structures coordinated at each state level by Chief Data Officers, therefore capitalizing on the Big Data advantages (European Commission, 2016). More specifically, it is about a strategy of complex analysis of the huge amount of data which was not addressed by the Big Data phenomena, but in a natural way, through the performance of new technologies.

And finally, instead of focusing exclusively on the potential consequences of Big Data phenomenon, we can gain additional insight from examining its social and political, but also its technical and epistemic roots (Rieder & Simon, 2016). Gary King (2016), professor at Harvard University, stated that “Big Data is not About the Data!”, substantially and indirectly contributing to the confirmation of our research hypotheses. It is not the data that create problems, but their analysis and to do so the performance of new technologies is needed. With all these ideas in mind and paraphrasing Gary King, we may say that Big Data is not About the Data! It is About the Big Data Analytics and the Big Data Technologies!

References

Beer, D. (2016). *How should we do the history of Big Data?*, Big Data & Society, January–June 2016: 1–10, http://bds.sagepub.com/content/3/1/20539517166461-35

Benson, P.R. (2005). ISO 8000 Data Quality - The fundamentals Part 1. In Stanford Encyclopedia of Philosophy, Stanford.

Bughin, J. (2016). *Big data: Getting a better read on performance*, McKinsey Quarterly, http://www.mckinsey.com/industries/high-tech/our-insights/big-data-getting-a-better-read-on-performance

Dalton, C.M., & Thatcher, J. (2015). Inflated granularity: Spatial Big Data and geodemographics. *Big Data & Society*, *July-December* 2015, 1-15.

Deloitte Touche Tohmatsu (2013). *Big Data Challenges and Success Factors*, http://www2.deloitte.com/content/dam/Deloitte/it/Documents/deloitte-analytics/bigdata_challenges_success_factors.pdf

Diesner, J. (2015). *Small decisions with big impact on data analytics*. *Big Data & Society*, *July-December 2015*, 1-6.

Etzion, D., & Aragon-Correa, J.A. (2016). Big Data, Management, and Sustainability, Strategic Opportunities Ahead. *Organization & Environment*, *29*(2) 147-155.

European Commision (2016). *Big data and B2B digital platforms: the next frontier for Europe's industry and enterprises*, Recommendations of the Strategic Policy Forum on Digital Entrepreneurship, Ref. Ares(2016)1913442 - 22/04/2016

Guszcza, J., Lucker, J., Lewis, H., Steier, D., & Gopalkrishnan, V. (2013). Too big to ignore. When does big data provide big value?, *Delloite Review*, *12*, 2013, http://d27n205l7rookf.cloudfront.net/wp-content/uploads/2013/01/DR12_Too_Big_to_Ignore1.pdf

Gutierrez, D. (2015). *Big Data Industry Predictions for 2016, inside BIG DATA*, http://insidebigdata.com/2015/12/08/big-data-industry-predictions-2016/

Harford, T. (2014). Big data: are we making a big mistake?, *Financial Times Magazine*, http://www.ft.com/cms/s/2/21a6e7d8-b479-11e3-a09a-00144feabdc0.html

Harper, T. (2016). The big data public and its problems: Big data and the structural transformation of the public sphere. *New Media & Society*, 1–16, http://nms.sagepub.com/content/early/2016/04/12/1461444816642167.full.pdf+html.

Hence, N., Libarikian, A., & Wiseman, B. (2016). *Straight talk about big data*, McKinsey&Company, http://www.mckinsey.com/business-functions/business-technology/our-insights/straight-talk-about-big-data.

ISO/IEC JTC 1 (2014). *Information technology, Big Data*, Preliminary Report, ISO 2015, Published in Switzerland, Geneva.

King, G., 2016, *Preface: Big Data is Not About the Data!* , in Computational Social Science: Discovery and Prediction, edited by R. Michael Alvarez. Cambridge: Cambridge University Press at http://j.mp/1PP466V

Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, *April-June 2014, 1-12*.

Kitchin, R., & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society*, *April- June 2016*, 1-10.

Kshetri, N. (2014). The emerging role of Big Data in key development issues: Opportunities, challenges, and concerns. *Big Data & Society*, *July-December 2014*, 1-20.

Laney, D. (2001). *3D Data management: Controlling Data, Volume, Velocity and Variety, Application Delivery Strategies*, META Group, February 2001, https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf

Liu,Q., & Vasarhelyi, M.A. (2014). Big Questions in AIS Research: Measurement, Information Processing, Data Analysis, and Reporting. *Journal of Information Systems*, *28*(1), 1-17.

Moorthy, K., Lahiri, R., Biswas, N., Sanyal, D., Ranjan, J. Nanath, K., & Ghosh, P. (2015). Big Data: Prospects and Challenges. *VIKALPA: The Journal for Decision Makers*, *40*(1) 74-96.

Oprea, D. (2010). *Big Data – necesitate sau strategie de marketing?* http://blog.dumitruoprea.ro/2010/08/17/big-data-necesitate-sau-strategie-de-marketing/

Poon, M. (2016). Corporate Capitalism and the Growing Power of Big Data. Review Essay. *Science Technology Human Values*, http://sth.sagepub.com/content/early/2016/05/31/0162243916650491.refs

Press, G. (2014). *12 Big Data Definitions. What's yours?*, http://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-yours/#7182df221a97

Rieder, G., & Simon, J. (2014). Datatrust: Or, the political quest for numerical evidence and the epistemologies of Big Data. *Big Data & Society*, *January–June 2016*, 1–6, DOI: 10.1177/2053951716649398

Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., & Tufano, P. (2012). *Analytics: The real-world use of big data. How innovative enterprises extract value from uncertain data*, Executive Report, IBM Global Business Services Business Analytics and Optimization in collaboration with with Saï d Business School at the University of Oxford, https://www.ibm.com/smarterplanet/global/files/se__sv_-se__intelligence__Analytics_-_The_real-world_use_of_big_data.pdf

Shaw, R. (2015). Big Data and reality. *Big Data & Society*, *July–December 2015*, 1-4.

Stancovici, V. (1975). *Filosofia informaiei*, Bucuresti: Editura Politica.

Taylor, L., Schroeder, E., & Meyer, E. (2014). Emerging practices and perspectives on Big Data analysis in economics: Bigger and better or more of the same?. *Big Data & Society*, *July–December 2014*, 1-10.

Wessel, M. (2016). *How Big Data Is Changing Disruptive Innovation*, Harvard Business Review, https://hbr.org/2016/01/how-big-data-is-changing-disruptive-innovation

*** *Colorado STEM Pipeline, Quality Data System Strategic Plan* (Draft), October 2013, http://www.coloradostateplan.com/STEM/STEMDataStrategyDraft121313-1.pdf

*** *IAS 38 - Intangible Assets*, http://www.iasplus.com/en/standards/ias/ias38