The influence of big data in the practices of strategic intelligence professionals in Brazil

[Priscila Nesello e Ana Cristina Fachinelli]

Abstract— **Strategic use of information is at the core of intelligence activity whose focus has been largely directed to the use of technology for data analysis. However, in practice the human factor still remains decisive for the correct and incorrect interpretation of data. In this article we present the outcome of a study on how the big data phenomenon is perceived and influences the activities of strategic intelligence (SI) professionals in Brazil. We used both the social constructionist research and the content analysis methods to analyzed in-depth interviews with strategic intelligence experts in Brazil. The results indicate that intelligence analysis is not as yet directly benefiting from the big data phenomenon. If on one side big data favours data collection, on the other it hinders the command of the subject. The study also showed that, for some interviewees big data is already integrated into the more elaborate professional practices and in the development of specific projects. However, for others, big date is not yet a reality hence the unperceived need of large volumes of data in analysis.**

Keywords— **information, big data, strategic intelligence, analysis, professional.**

I. Introduction

Half a century after computers became part of the routine of people and businesses, information rapidly began to accumulate. Modern information technology systems enabled the current deluge of data, named big data. However, analysis evolution started during the analogue era when data analysis was expensive and time consuming, then through digitalization, when analog information became meaningful to computers. Eventually, it reached the origin of what is now called datafication which consists of collecting everything existent and process it into quantifiable data [20]. The definition taken as the basis for our big data study - the data set whose size goes beyond the capacity to capture, store, manage and analyze - is typical of conventional database software tools [17].

Big data has the potential to be different from other times due to the possibility of analyzing data in its original form, unstructured; and analyze not only what happened in the past but indeed predict in great detail what will happen around the world [19]. This change not only involves the amount and frequency with which data arrives but also the

Priscila Nesello
Universidade de Caxias do Sul
Brasil

Ana Cristina Fachinelli
Universidade de Caxias do Sul
Brasil

diversity of sources and the availability of different types of data [2]. This compounds the volume, velocity and variety - considered the main differential factors of this phenomenon compared to other times [18].

The technologies that support big data can be analyzed from two points of view: the infrastructure technologies such as NoSQL[1] database which store and process petabytes of data, and the ones involved in analysis such as Hadoop and Map Reduce[2]. Nevertheless, in order to provide the company with the expected return, investment in technology needs to be aligned with the existing business model as well as a synchronized transformation process be developed between people, processes and technology. This is because it is necessary to explore in great detail scientific aspects such as testing, processing and storage of vast volumes of data. Even if all companies large or small have the same access to information, an analysis process is required to convert the available required information into intelligence. The analysis is supported by rating processes and data documents charts to produce maps, graphs and other communication tools [9]. The analysis also takes into consideration methodological and cognitive aspects to perform the interpretation of information [8]. The cognitive aspects are the ones that most seem to challenge the transformation of information into intelligence as they are less instrumentalized than the methodological ones and are associated with the psychological dimensions [12].

While analyzing the databases of indexers Scopus and Web of Science (WoS) to identify the number of types of publications produced in big data and analysis between 2010 and 2013, we could observe that the number of publications in these areas grew on average by more than 200% per year. However, no published papers with big data issues and analysis were identified related to strategic intelligence. Our target is to examine how the big data phenomenon affects the analysis process in the strategic intelligence activity.

I. Material and methods

This research was conducted as a qualitative study to examine how the big data phenomenon affects the analysis process in the activity of strategic intelligence. We adopted the social constructionist approach for its association with the postmodern era in qualitative research and its concern with the nature of knowledge and how it is created (Andrews, 2012). Semi-structured interviews were conducted with 18 leading individuals in the field of strategic intelligence in Brazil, using the content analysis technique for the data analysis. The ATLAS/ti software was used to encode the data and organize the batch of categories for analysis.

---

[1] Non-related database infrastructure technology [20].

[2] Processing Technologies which allow for the management of great volumes of data [20].

The 18 interviewees were selected through the following agents: Brazilian Association of Competitive Intelligence Analysts (ABRAIC), IBM Brazil and Brazilian Society of Knowledge Management (SBGC). Respondents work as professionals, professors or researchers in the field of Strategic Intelligence and were selected via the snowball sampling technique [11], chosen because it is in accordance with the social approach adopted in our method. This is used in social research where the original participants (seeds) introduce new participants (children or fruits), which in turn introduce new participants and so on, until the final objective or saturation point is reached (depletion of accessible participants).

According to the content analysis method, the categories of analysis were defined a priori based on theoretical assumptions about the topic of the study. For strategic intelligence, categories were taken as required attributes to the analyst: subject knowledge; understanding of research methods; imagination and scientific rigor; knowledge of collection methods; self-awareness potential and other cognitive influences; receptiveness to divergent opinions or alternative models; and self-confidence to concede and learn from analytical errors [3]. The adopted dimensions for the analysis of these categories were the three Vs - volume, variety and velocity of big data [2].

II. Results and Discussion

One of the biggest challenges in the coming years will be the lack of professionals with adequate analytical skills to the big data context. The United States alone will face up to the year 2018, a shortage of 140.000 to 190.000 people with analytical skills, which represents a gap of 50% to 60% on the demand [17]. Another recent study, portrays the gap in analytical prowess as a major improvement opportunity for companies of all sectors [14]. This study, conducted with marketing executives from around the world, reveals that this issue is among the most important challenges and concerns of the digital marketing. This is because market research has relied upon the capacity of field-work staff, while studies on the skills of analysts are limited.

This limitation has impacted the growth of the analytic function in marketing, skilled resources attraction and also how the various analytical specificities will cohabit within the organizational environment [14]. This scenario indicates research space for market scientists with a focus on understanding how analysts can work more effectively to improve their impact on the marketing function.

In our study, the characteristics of big data analysis professionals explain the shortage scenario of analytical and managerial talent, essential to transform large volumes of data into intelligence. This is because these qualities were not recognized in most of the professionals interviewed despite their importance being mentioned by some of them. They complement the attributes used as categories for strategic intelligence and share the development of cross-functional skills in three areas: anthropology, sociology and psychology, as well as information technology (IT), with emphasis on mathematics. These different skills are integrated to achieve the understanding of how people behave, relate and how the human psyche works, since this professional is handling data generated by people. These characteristics in the big data context designate the "data scientist" [5]. In the view of these authors, this is the professional capable of finding answers to organizational problems in the midst of unstructured information. This definition can be associated with the professional of strategic intelligence in a modern approach, related to the big data context. The power to synthesize topics and communicate this knowledge within the organization was also mentioned as important for the professional. The command of technological tools becomes essential for a robust management of unstructured data. Other features related to the SI literature were also cited as "curious", "proactive" and "inquisitive" [10].

In general, professionals who do not possess these characteristics well developed see big data as a threat in the analysis of information for strategic intelligence. Instead of developing these scarce attributes, they take the opposite approach by ignoring the potential of big data in generating unprecedented intelligence products. The consequence is a bias towards specialization instead of multidisciplinarity; the restriction of sources for information collection and a decrease in the volume of analyzed data instead of the analysis of all data available; the definition of restrictive intelligence processes instead of the curiosity to make discoveries through diversified methods of analysis. This behaviour obliges the analyst to simplify the processing of information to increase the effectiveness of the analysis and reduce the risk of failures in the process [16]. Specifically, mistakes in the analysis can occur at different levels and stem from: false assumptions, lack of time, excessive consensus among analysts, dysfunctions within organizations, motivated interpretations and problems with information gathering [13;16]. In our study, we identified that respondents do perceive that these elements are in the root of errors of analysis. Through the adopted method, we organized the responses into categories and dimensions as follows.

III. Dimensions related to analytical activity which is being affected by the big data phenomenon

Interviewees were questioned in relation to the effects of volume, velocity and the variety (3 V's) in each of the Attributes of the SI analyst used as categories in this study. In Table 1 the effects of 3 V's according to the respondents are shown:

TABLE I. EFFECT OF 3 V'S IN THE ANALYSIS.

Categories	1	2	3	4	5	6	7
Volume	D/C	C/D	Un/C	C/Un	C/Un	Un	Un/C
Velocity	D	NC	Un	NC	D/Un	Un	Un
Variety	D	C/D/Un	NC	D/C	D/Un	Un	Un

Key: (D) Detrimental; (C) Contribute; (Un) Unaffected (NC) Not cited or irrelevant.

Having identified the perception of respondents about the effects of big data 3 V's on the SI process of analysis, this study will proceed to the individual analysis of the proposed categories in relation to the dimensions: volume, variety and velocity.

A. *Subject Command*

The category subject command (1) is said by respondents to be impaired by volume, velocity and variety. According to respondents, big data affect the subject command on three aspects: quality of information (volume), absence of tools to deal with large volumes of data (variety) and obsolescence of knowledge (velocity).

There are a paradox between experience and skill. Senior analysts are characterized by top command of their fields, but not a lot of familiarity with big data and its complexity. They run the risk of becoming quickly obsolete due to the high velocity of information output and update. On the other hand, junior analysts lack command in specific areas, but generally have a multidisciplinary approach and familiarity with big data. They are the professionals more likely to succeed in the current environment. However, to derive value from large volumes of data the professional must be able to analyze this volume of information. The smaller his experience, the more subject to the cognitive influences of the process he will be. This is because, unlike the already well developed methodological aspects, the cognitive aspects of the review process are neither epistemological nor pragmatically consolidated [7].

Heuer highlights the importance of mental models for the cognitive aspects of the review process [12]. The author expresses that professionals build their own version of reality on the basis of the information provided by the senses. This is mediated by complex mental processes that determine which information is perceived, how it is organized, and what meaning it is given. What and how easily people perceive information and how it is processed after receiving it are all elements strongly influenced by past experience, education, cultural values, requirements and organizational standards, as well as the specificities of the information received. According to respondents, the volume contributes to subject command, as during analysis both content and theme command bring multiple perspectives to the process. But in order for this to occur, more time for information analysis should be spent.

B. *Comprehension of the investigation methods*

The category comprehension of research methods (2) in the view of the respondents can be either aided or upset by big data. Also, in some quotes, big data does not affect this category. According to the responses, the deleterious effects of big data in this category are related to the lack of some characteristics by the professional to help conduct the investigation process with large volumes of data. These features may involve the inability to apply mathematics and statistics to the process. The lack of information technology knowledge was also mentioned alongside of the application of traditional methods of analysis. These methods are eventually structured for the treatment of limited data, represented by the samples, and not by all the available data from big data.

In this scenario, we found that some professionals opt to remain with the same research methods used before the big data phenomenon, and if asked for more refined intelligence products, they commonly outsource it to specialized companies or consultants. This interpretation is in conformity with the work done by Leeflang, Verhoef, Dahlström and Freundt [14]. The researchers identified that the shortage of skilled professionals capable of working with large volumes of data increases the need for outsourcing, a practice which poses risks because it shares among other professionals and the market the products of intelligence.

According to respondents, volume and variety can also contribute to the improvement and development of research methods. Furthermore, robust statistical data based on masses of data reduces sampling error and leads to refined views. For the interviewees, data variety contributes to the development of a predictive model, thus complementing the dimension volume. The predictive power of big data is being exploited and proves promising [15]. While descriptive statistics became popular reporting past events, predictive analysis uses past information to forecast future results with some degree of probability.

C. *Imagination and scientific rigour to propose and test hypotheses*

According to most respondents, imagination and scientific rigor (3) are not affected by volume and velocity, though in some cases the volume may contribute to the category. The analysis of the responses indicated that imagination to reach the hypotheses has not been explored by the analysts in the strategic intelligence activity, and for this reason it is not affected. This occurs to avoid errors and because of time restrictions for the delivery of products.

The category imagination and scientific rigour to reach hypotheses can also be benefitted by the volume of big data, in particular when the professional does not start from prior assumptions in order to test them later, according to the traditional method. In this view, the big data phenomenon offers a new angle in relation to the scientific method, because actually we express more aspects of reality in the form of data [20]. So, though the entire process of collection and analysis is based on theories, with datafication, the comprehension of a phenomenon will be increasingly driven by data rather than hypotheses. The authors mention that in the future, the big data correlations will be routinely used to disagree with our causal intuition, showing that normally there is little or no statistical connection between the effect and what is supposed to be the cause. However, at present, the most advanced use of datafication is in the world of business.

D. *Understanding the collection methods*

Regarding the understanding of collection methods (4) volume contributes, but variety upsets. Besides, citations were identified in which the volume does not affect, and others in which the variety contributes to the category.

Big data contributes with volume and variety when the adequate technological tools are used for the capture and storage of large volumes of data, from a variety of sources. This category is also influenced by technological developments for data collection. In the old days data collection was performed manually, which was time consuming and led to errors. Today this collection can already be performed automatically, resorting to social networking sources or mobile device resources. This allows

for the use of more data or all data available, which contributes to the generation of better intelligence products.

For some interviewees the volume does not affect this category. This occurs when methods that facilitate both the organization and data analysis are adopted. According to the interviewees, it is the method and not the data that will affect the product of intelligence. Structured and robust methods of analysis will protect the analyst from the possible effects of big data. However, the mechanization of analysis eliminates the effect of important characteristics such as imagination, creativity and the quest for answers to big questions in the midst of large volumes of unstructured information.

E. *Potentialities and other cognitive influences*

In our study we found that this category (5) is hampered by the volume and variety even if in some cases volume contributes while in others it does not. This category was cited by respondents as one of the main challenges for strategic intelligence professionals. This is due to the limits of human cognition with regards to the prioritization of what will be considered for analysis.

The informational complexity of the world exhausts the human cognitive abilities. This fact forces us to create diversions and simplifying cognitive shortcuts that automatically act in the subconscious, ensuring our capacity to deal with complex environments. However, these mechanisms are limited and may lead us to process certain pieces of information inaccurately [4]. Also mentioned was the fact that the professional is not at ease to execute the construction of hypotheses in the analysis, due to the large volumes of information that rapidly change. As a result, he specializes without getting the benefits derived from big data. This thinking goes against the multidisciplinarity, one of the main characteristics of the data scientist [5]. The authors state that unlike the specialization, the analyst must meet a very peculiar set of characteristics, among them curiosity and associative reasoning. For this, the areas of recruitment and training have successfully sought individuals with an academic background or experience in physical or social sciences, though with a strong focus on data and computing [5].

Data volume contributes to the awareness of the potentialities because it is possible to use a larger volume of data with less collection effort. Some respondents, however, tend to value experience more than the volume of information. The idea about how data can be used in this process is a change more important and subtle than the technological. In order for this to occur, it is required among various aspects that the analyst accepts the real confusion of data instead of privileging the accuracy: the increase in the amount of data opens the doors to the inaccuracy [20].

F. *Open-mindedness to contrary opinions or alternative models*

The activity open-mindedness to contrary opinions or alternative models (6) is not affected by volume, velocity and variety. In the view of the interviewees, this category takes into consideration behavioural and environmental questions in which the professional operates, considering

market and technological aspects. In this sense, these would be the aspects that affect the category and not big data. Another point to be emphasized is the importance of free access to information and the collaboration in organizations.

Currently, information is a commodity relatively cheap and easily obtainable. Companies and government agencies aggregate data from multiple private and/or public sources [1]. Private data are pieces of information that are available to an individual organization, such as employee data, customer data and machine data (for example, transactions and customer behaviour of users). Public data are pieces of information that are available to the public by means of a payment or for free, such as credit ratings and social media contents (e.g. LinkedIn, Facebook and Twitter). Professionals can also use public information, but not necessarily published. In other words, the information that professionals look for is readily available and identified through legal means of open sources, such as public documents, interviews, announcements and internal documents. One way to understand intelligence is to view it as a process of raw material input and of finished materials output [9].

G. *Open-mindedness to contrary opinions or alternative models*

The gategory self-confidence to admit and learn from analytical errors (7) is not affected by volume, velocity and variety. However, in some cases the volume can contribute to this activity. In the view of interviewees, big data does not affect this category, but the absence of a clear guideline from the decision maker as to what should be searched. Another point mentioned refers to the lack of a new cognitive pattern and evolution of the decision-making process. This makes the products generated by the analysis - knowledge - not intelligently effective for organizations.

On the future of the decision-making process, the big data phenomenon is intimately linked to the emergence of the data science discipline [6]. The term refers to the use of data applied to scientific method and to businesses. So, benefitting from big data means investing in teams with a set of specific skills, and to surround them with the organizational will to understand and use the data to gain an advantage. Though with no evidence of further studies, that the decisions taken based on data will compete or excel human judgment [20]. According to the authors, big data in the analysis process announces the end not only of casualty, but also of the influence of specialists in many areas. According to this view the specialist would give way to the statistician or data analyst who will let data manifest itself.

Big data also provides a better basis for analysis. This reflects greater self-confidence on the part of the professional. However, Big Data phenomenon the context is presented overloaded with data and information. Our study identified that even being an object of intelligence, the growing availability of information in all areas and its rapid production growth are not a direct contribution to the analysis process.

IV. **Conclusion**

In this study, we present how the big data phenomenon is perceived by Brazilian strategic intelligence professionals

in their analytical activities. This was based on 18 in-depth interviews conducted with experts in the field. The results show that the challenges presented to handle large volumes of data are not only technological, but also cognitive and can be observed in the analytical activity. Other results indicate that the big data phenomenon is not yet a reality for professionals in traditional SI organizations. According to the results, the effects of big data are less influential over the cognitive dimension of the analyst, particularly in the skills to keep their minds open to contrary opinions and alternative models and in the self-confidence to accept and learn from analytical errors.

On the other hand, we identified volume contributes to the understanding of big data collection methods, although according to some respondents it troubles the command of the subject. To better deal with the amount of information associated with the big data phenomenon, we identified that many professionals resort to adjusting the focus and adopt structured and robust analytical methods supported by information technology. The research also identified that even in the big data context, the intelligence products produced by professional analysts compete strongly with the analysis based on intuition. Notwithstanding, this proves that even without an artificial technological limit, our mindset is not geared to work with big data [20]. Hence, endowed with the same mentality and methods, professionals of intelligence have not yet realized the full potential advertised by the phenomenon.

If we consider that the big data phenomenon is still not a reality for SI professionals of traditional organizations, it is possible to suppose that this deficiency tends to increase with the same velocity as the actual complexity. Thus, making SI professionals dependent on the few companies that have data scientists capable of dealing with big data. The fact is that fundamental aspects that were before pursued by analysts such as the access to rare and still undivulged information, allied to information confidentiality and experience in the focus of observation, no longer represent the only ways to generate intelligence products. This is because in the digital environment with big data, the access to primary information is already shared among competitors, and the knowledge command of the analyst that comes from experience, if not implemented or updated in a short time, tends to become obsolete. Currently, with analysis support and appropriate technological tools, professionals have the opportunity to identify unknown patterns amid large volumes of data. So, it is not just about identifying rare information prior to competition, but indeed, from the information available, to try to identify relevant patterns until now unperceived by others. It is a change of paradigm. So, besides the attributes required for the analyst, here referred to as categories, the ability to handle the set of technological tools and multidisciplinary training emerge as increasingly key issues for the scope of SI activity in relation to Big Data. It is a fundamental complement of analytical activities, and not a replacement of the existing ones.

References

[1] Akerkar, R. (2014). Big data computing. Boca Raton: Taylor & Francis Group, LLC.

[2] Betser, J. & Belanger, D. (2013). Architecting the enterprise via big data analytics. In: J. Liebowitz, (Ed.), Big data and business analytics. Boca Raton: Taylor & Francis Group, LLC.

[3] Bruce, J. B. & George, R. Z. (2008). Analyzing intelligence: origins, obstacles, and innovations. Washington, D. C.: Georgetown University Press.

[4] Cepik, M. A. C. & Ambros, C. C. (2012). Explicando falhas de inteligência governamental. Varia Historia, 28, 79-99.

[5] Davenport, T. H. & Patil, D. J. (2012). Data scientist: the sexiest job of the 21st century. Harvard Business Review, 90, 46-53.

[6] Dumbill, E. (2014). What is big data? An introduction to the big data landscape. O'Reilly Media, Inc. Retrieved from: http://strata.oreilly.com/2012/01/what-is-big-data.html. Date: 2014/03/05.

[7] Fachinelli, A. C., Rech, J. & Bertolini, A. L. (2011). La dinámica cognitiva de los processos de análisis en la actividad de inteligencia estratégica. In: A. M. Sánchez & M. C. Estrada (Eds.), Administración de conocimiento y desarrollo basado en conocimiento. Cidade do México: Cengage Learning.

[8] Fachineli, A. C. & Alberdi, A. M. (2014). Structural Integrity of Strategic Intelligence: an Assessment in a Cooperative Corporation. Brazilian Business Rewiew, 11, 76-99.

[9] Fleisher, C. S. & Bensoussan, B. (2007). Business and competitive analysis: effective application of new and classic methods. New Jersey: Pearson Education, Inc.

[10] Fuld, L. M. (1995). The new competitor intelligence: the complete resource for finding, analyzing, and using information about your competitors. John Wiley & Sons, Inc.

[11] Goodman L. (1961) Snowball sampling. Annals of Mathematical Statistics.

[12] Heuer, R. J. (1999). Psychology of intelligence analysis. Center for the study of intelligence.

[13] Jordán, J. (2011). Introducción al análisis de inteligência. Retrieved from: http://www.seguridadinternacional.es/?q=node/2. Date: 2014/03/05.

[14] Leeflang, P. S. H., Verhoef, C. V., Dahlström, P. & Freundt, T. Challenges and solutions for marketing in a digital era. European Management Journal, 32, 1-12.

[15] Lohr, S. (2012). The age of big data. The New York Times. Retrieved from: http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html?pagewanted=all&_r=0. Date: 2014/03/05.

[16] Lowenthal, M. (2006). Intelligence: from secrets to policy. 3.ed. Washington-DC: CQ Press.

[17] Manyika, J., Et Al. (2011). Big data: the next frontier for innovation, competition, and productivity. McKinsey Global Institute.

[18] Mcafee, A. & Brynjolfsson, E. (2012). Big data: a revolução da gestão. Harvard Business Review Brasil, 90, 39-45.

[19] Mineli, M., Chambers, M. & Dhiraj A. (2013). Big data, big analytics: emerging business intelligence and analytic trends for today's businesses. New Jersey: John Wiley & Sons, Inc.

[20] Mayer-Schönberger, V. & Cukier, K. (2013). Big Data: A Revolution That Will Transform How We Live, Work and Think. John Murray.

About Author (s):

Priscila Nesello
Universidade de Caxias do Sul
Rua Francisco Getúlio Vargas, 1130
Caxias do Sul – RS - Brazill
CEP 95070-560

Corresponding author

Ana Cristina Fachinelli
Universidade de Caxias do Sul
Rua Francisco Getúlio Vargas, 1130
Caxias do Sul – RS - Brazill
CEP 95070-560