

Understanding Big Data for Industrial Innovation and Design: The Missing Information Systems Perspective

Miguel Baptista Nunes[†]

School of Information Management, Sun Yat-sen University, Guangzhou 510006, China



Miguel Baptista Nunes, BSc, MSc, PhD, FHEA, FBCS, is a Professor and Dean of the School of Information Management at Sun Yat-sen University, Guangzhou, China. Prior to this appointment Miguel was at the Information School of the University of Sheffield in the UK for 23 years, with 19 years as faculty. Miguel's research interests are centered on information systems and information management, specifically on areas such as information systems modeling, information systems design and development, database design and development, project management, risk management, and information systems change management.

Citation: Miguel Baptista Nunes (2017). Understanding Big Data for Industrial Innovation and Design: The Missing Information Systems Perspective.

Vol. 2 No. 4, 2017

pp 1–6

DOI: 10.1515/jdis-2017-0017

Received: Mar. 3, 2017

Revised: May 4, 2017

Accepted: Jun. 10, 2017

Abstract: This paper identifies a need to complement the current rich technical and mathematical research agenda on big data with a more information systems and information science strand, which focuses on the business value of big data. An agenda of research for information systems would explore motives for using big data in real organizational contexts, and consider proposed benefits, such as increased effectiveness and efficiency, production of high-quality products/services, creation of added business value, and stimulation of innovation and design. Impacts of such research on the academic community, the industrial and business world, and policy-makers are discussed.

Keywords Big data; Business value; Information systems; Information science

Introduction

Big data is one of the most fashionable, yet most misused and misunderstood terms being circulated in policy making, academia, industry, business, and above all the media (Hartford, 2014). In China like almost everywhere else, the concept is of fundamental importance to national policies and has generated a huge hype,

[†] Corresponding author: Miguel Baptista Nunes (E-mail: miguelnunes@mail.sysu.edu.cn).



Perspective

with all types of big data applications being adopted at organizational, city, regional, and national levels. Universities, research groups, and individual academics in all disciplines also have readily seized this golden opportunity for funding.

Beyond the hype, big data is a term that was made globally well known by the 2011 McKinsey Global Institute report *Big data: The next frontier for innovation, competition, and productivity* (Manyika et al., 2011). The fundamental importance of this report consisted in the assertion that big data would change competition: by transforming processes, altering corporate ecosystems, and facilitating innovation. That is why this concept became rapidly fashionable in industrial, business, and policy making circles, where innovation, new designs, and global competition have become a key strategic aspect of enterprise as well as of national and regional economic development.

However, the definition and conceptualization of big data has changed considerably since the McKinsey report defined the term as follows:

“‘Big data’ refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze. This definition is intentionally subjective and incorporates a moving definition of how big a dataset needs to be in order to be considered big data—i.e., we don’t define big data in terms of being larger than a certain number of terabytes (thousands of gigabytes).”

This definition has now evolved and has been reinterpreted in different settings, contexts, and purposes. The most significant change is the acknowledgement that “When it comes to data, size isn’t everything” (Harford, 2014). This led IBM to reinterpret a 3V meta-data management model from the META Group[Ⓢ] and re-define big data using the now universally accepted 4Vs definition: volume, variety, velocity, and veracity. Later in 2015, IBM introduced an additional V: value, as in business value. This last “V” is probably the one that has been most important in encouraging the actual adoption and use of big data applications in real business practice, not just in research laboratories from large companies such as Google or Baidu or academic research groups.

Information Systems Perspective

“Like so many new technologies, Big Data will surely become a victim of Silicon Valley’s notorious hype cycle: after being feted on the cover of magazine and industry conferences, the trend will be dismissed.” (Mayer-Schönberger & Cukier, 2013, p. 7). This prediction is not only probable, but it is almost a certainty (as can be seen by recent examples of other technologies such as hypermedia or Web 2.0) and is already happening. A recent (12/19/2016) *Forbes* article advises: *Forget big data: What you need is deep data* (Chamorro-Premuzic, 2016). Other authors are



[Ⓢ] It was acquired by Gartner Inc. in 2004.

now proposing “Dirty Data” (e.g. Waterman and Hendler, 2013) and many other derivatives of the term “big data.” Having been a professional and a researcher in the field of information systems for more than 30 years now, I have witnessed the emergence, evolution, and disappearance of such fads and trends on other occasions (e.g. MRP, MRP II, and ERP[®]) as well as the merger of technologies (SCM, CRM[®], data warehousing, and RDBMS[®] all into ERP). These examples show that, although disappearing and merging technologies and concepts served a purpose in organizational evolution, they also influenced business models, organizational structures, information architectures as well as the manufacturing and design of product and services. These influences, associated with an adequate understanding and the grasping of opportunities afforded by the technologies, are indeed more lasting and important than the technologies themselves in enabling innovation, new designs, and development.

Big data is indeed the start of a global transformation in business, government, and society. But from an information systems and social science perspective, it is fundamental to understand the transformation beyond the technical aspects of data science, data analytics, and data processing technologies. Specifically, an information systems perspective must focus on basic questions such as:

- What are the needs in industrial environments for big data?
- Why are industrial organizations using big data?
- What can be done with big data that was not previously possible?
- What changes are occurring in organizational structures, cultures, technological infrastructures, and business models?
- What are the changes in working practice, use of technology, and efficiency?

These are the transformations that will persist long after the disappearance of big data as a trend or it is normalized, so that people no longer talk about it as being anything special or radical. These questions and their impact are something that computer and data scientists are not equipped to deal with and not even interested in doing so. Moreover, without addressing this type of question in depth, the widespread use and adoption of big data applications is unlikely ever to go beyond the pages of policies, academic papers, and speculative industrial articles. The real world of government, industry, and business is a pragmatic one, driven by business value, efficiency, and competition. If the added value of big data applications and services cannot be clearly established beyond the realm of speculation and theory, then its survival is doomed.

[®] MRP refers to “material requirement planning,” MRP II refers to “manufacturing resource planning”, and ERP refers to “enterprise resource planning.”

[®] SCM refers to “supply chain management,” and CRM refers to “customer relationship management.”

[®] RDBMS refers to “relational database management system.”



The Need for New Agendas in Information Systems

Some attempt to address the set of basic questions posed above is crucial if regional and national policies aiming at promoting big data are to succeed. Such questions would bring clarity to a discussion, where there has been much confusion, misunderstanding, and opportunistic use of the term “big data” as a buzz word rather than a scientific one. Lazer et al. (2014) proposed in a very highly cited *Science* article in 2014 that there is a “big data hubris.” This refers to the implicit assumption that big data is a substitute for, rather than a supplement to, traditional data collection and analysis. They stress that foundational issues of measurement, construct validity, reliability, and dependencies among data should not be ignored. This is certainly true in academic and research institute environments, but in industrial and business environments this hubris is even more significant and damaging. There is no uniform understanding of what big data means in industry and business, nor how it could or should be used. More importantly, it is not clear what specific needs for innovation and design big data would serve. So the term is thrown into policy, and from policy into practice with no clearly defined purpose, identified need or sometimes understanding on how big data could even be used. Almost by an act of magic, data-driven statistics and data mining are expected to resolve industry problems by themselves.

In order to mitigate the effects of this “Industrial Big Data Hubris” it is necessary to clearly define the concept of big data in terms of its business value and the information that contributes to this value. This is of fundamental importance since there is a clear difference between data and information. Data comprises facts and figures which have been collected from a variety of sources, both from within the organization and from outside it. Data is the record of an event or a fact. Data is not information until it has been arranged in a manner that allows a particular individual to comprehend and extract meaning. Consequently, information is data endowed with relevance and purpose (Drucker, 1995). Therefore, and this may choke many applied mathematicians, data scientists, and data miners, processed data may still just be data if it does not serve a specific organizational need. In other words, information is data processed for a purpose that is meaningful to users when performing their tasks in a particular organizational environment (e.g. business, industry, or government). This process of meaning attribution is a uniquely human act (Checkland, 1993). It depends on individual and group perceptions, objectives, and motives. If big data developments aim to have an impact on the real world of practice, then it must be recognized that “organizations are complex and paradoxical phenomena that can be understood in many different ways” (Morgan, 1997). Such recognition will enable researchers, data scientists, and developers to look beyond



the hard data and into the complex, interconnected, and constantly evolving issues that pervade every human activity system. Organizations are not laboratorial environments where experimental artificial intelligence and neural networks engage in simplified tasks, but they are complex human activity systems where subjective concerns with mission, efficiency, and business value are at the forefront. In particular, business value needs to be understood and measurable. In an acclaimed article on big data in *Harvard Business Review*, McAfee and Brynjolfsson (2012) stated that: “*You can’t manage what you don’t measure.*” I want to add to this statement that: *You can’t manage what you do not understand.*

There is therefore the need to establish an agenda of research for information systems that complement the current calls for strictly technical and mathematical proposals. Such an agenda would aim to understand perceptions of the nature and value of big data. It would explore motives for using big data in real organizational contexts, and consider proposed benefits, such as increased effectiveness and efficiency, production of high-quality products/services, creation of added business value, and stimulation of innovation and design. However, the world trend in funding of big data, both at national and international levels, has been devoted to technical and mathematical research that focuses on the concept and its theoretical implementation. The vast majority of these projects are highly theoretical, based on algorithm development and the technology to support it. Data-driven analytics, data mining, and all sorts of applied mathematical propositions have been made in academic and technical journals and conferences. Nonetheless, the reality is that there is little permeation of these theoretical insights into the real world of daily practice in industry and business. If the investment by national and regional government and the significant academic effort and development are to bear fruit in practice, then a totally different type of study must now be undertaken. Studies focusing on social aspects of the implementation of big data would help address the changes in information needs, information behaviors and information architectures that are emerging due to the fast development in smart, cloud, and big data technologies. Information management schools like mine are ideally placed to undertake this type of study.

Such an agenda would have a target audience in the academic community, the industrial and business world, and among policy-makers. Academics would be better informed about the real world applications of their data analytics and data mining algorithms. Business leaders and chief information officers (CIOs) would benefit from a clarification of uses and purposes, as well as a better understanding of models of adoption. Finally, policy-makers and government could use the reports to fine tune national and regional policies and plans.



Conclusions

This paper identifies a need to complement the current rich technical and mathematical research agenda on big data with a more information systems and information science strand, which focuses on the business value of big data, and explores aspects of the way in which it is perceived and used. This would require a shift in the understanding of data as raw material for business, government, and society leading to it being regarded as a vital economic input that could help create new forms of business and social value. Consequently, if used effectively, data can become a fountain of innovation, new designs, and new services (Mayer-Schönberger & Cukier, 2013, p. 5). Such studies would help policy-makers make better policies, scientists to produce better science, and industry leaders to be better competitors. Finally, the findings of this type of research will inform universities and colleges so that they can improve curriculums, syllabuses, and courses, making them better at developing talent, and to produce graduates who are more useful, efficient, and productive members to the workforce of the future.

References

- Chamorro-Premuzic, T. (2016). Forget big data: What you need is deep data. *Forbes*, December 19, 2016. Retrieved on March 12, 2017, from <https://www.forbes.com/sites/tomaspremuzic/2016/12/19/forget-big-data-what-you-need-is-deep-data/#7ffa841b5cec>.
- Checkland, P. (1993). *Systems thinking, systems practice*. Chichester, UK: Wiley.
- Drucker, P.E. (1995). The post capitalistic executive. In P.E. Drucker (Ed.), *Management in a Time of Great Change*. New York: Penguin.
- Hartford, T. (2014). Big data: A big mistake? *Significance*, 11 (5), 14–19. Retrieved on March 12, 2017, from <http://onlinelibrary.wiley.com/doi/10.1111/j.1740-9713.2014.00778.x/full>.
- Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). [The parable of Google flu: Traps in big data analysis](#). *Science*, 343(6176), 1203–1205.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. (2011). *Big data: The next frontier for innovation, competition, and productivity*. San Francisco, CA: McKinsey Global Institute, McKinsey & Company.
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Boston: Eamon Dolan Books.
- McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*, 90 (10), 60–66, 68, 128.
- Morgan, G. (1997). *Images of organization*. Thousand Oaks, CA: Sage.
- Waterman, K.K., & Hendler, J. (2013). Getting the dirt on big data. *Big Data*, 1(3), 137–140.

