Computers & Industrial Engineering 101 (2016) 528–543

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

Big data applications in operations/supply-chain management: A literature review

Richard Addo-Tenkorang *, Petri T. Helo

University of Vaasa, Faculty of Technology, Department of Production, Industrial Management Unit, PL 700, 65101 Vaasa, Finland

ARTICLE INFO

Article history:
Available online 28 September 2016

Keywords:
Big data – applications and analysis
Internet of Things (IoT)
Cloud computing
Master database management
Operations/supply-chain management

ABSTRACT

Purpose: Big data is increasingly becoming a major organizational enterprise force to reckon with in this global era for all sizes of industries. It is a trending new enterprise system or platform which seemingly offers more features for acquiring, storing and analysing voluminous generated data from various sources to obtain value-additions. However, current research reveals that there is limited agreement regarding the performance of "big data." Therefore, this paper attempts to thoroughly investigate "big data," its application and analysis in operations or supply-chain management, as well as the trends and perspectives in this research area. This paper is organized in the form of a literature review, discussing the main issues of "big data" and its extension into "big data II"/IoT–value-adding perspectives by proposing a value-adding framework.
Methodology/research approach: The research approach employed is a comprehensive literature review. About 100 or more peer-reviewed journal articles/conference proceedings as well as industrial white papers are reviewed. Harzing Publish or Perish software was employed to investigate and critically analyse the trends and perspectives of "big data" applications between 2010 and 2015.
Findings/results: The four main attributes or factors identified with "big data" include – big data development sources (Variety – V_1), big data acquisition (Velocity – V_2), big data storage (Volume – V_3), and finally big data analysis (Veracity – V_4). However, the study of "big data" has evolved and expanded a lot based on its application and implementation processes in specific industries in order to create value (Value-adding – V_5) – "Big Data cloud computing perspective/Internet of Things (IoT)". Hence, the four Vs of "big data" is now expanded into five Vs.
Originality/value of research: This paper presents original literature review research discussing "big data" issues, trends and perspectives in operations/supply-chain management in order to propose *"Big data II" (IoT – Value-adding) framework*. This proposed framework is supposed or assumed to be an extension of "big data" in a value-adding perspective, thus proposing that "big data" be explored thoroughly in order to enable industrial managers and businesses executives to make pre-informed strategic operational and management decisions for increased return-on-investment (ROI). It could also empower organizations with a value-adding stream of information to have a competitive edge over their competitors.

1. Introduction

Data has increased on a large scale in various fields over the last two decades; hence, the term "big data" has been coined. It has been largely anticipated that the amount of data will continue to increase greatly in the coming years in this digital era, where huge amounts of data are constantly being generated from several sources. A report from International Data Corporation (IDC), Gantz and Reinsel (2011) indicates that the overall created and copied data volume in the world was 1.8ZB (≈ 1021B), which increased by nearly nine times within a five year period. The world generated over 1ZB of data in 2010, and by 2014 7ZB per year (Richard, Matthew, & Carl, 2011). A great amount of this data increase is the result of diverse devices employed at the periphery of industrial enterprise supply chain (SC) networks including embedded sensors, smartphones, computer systems and computerized devices. All of this data creates new opportunities to extract more value. Therefore, "big data" could be defined as a fast-growing amount of data from various sources that increasingly poses a challenge to industrial organizations and also presents them with a complex range of valuable-use, storage and analysis issues. Current research on "big data" reveals that there is limited

* Corresponding author.
E-mail addresses: ratenko@uva.fi (R. Addo-Tenkorang), petri.helo@uva.fi (P.T. Helo).

http://dx.doi.org/10.1016/j.cie.2016.09.023
0360-8352/

agreement regarding the most valuable use or performance of "big data" in industrial operations and neither is there even a single agreed definition of it. Attempts to compare traditional datasets with big data typically include masses of unstructured data that needs more real-time analysis. Furthermore, big data is also believed to enable new opportunities to discover new value-adding and make strategic decisions that would help industrial enterprises to gain a better understanding of the hidden values of "big data" and also rise to new challenges, e.g. how to effectively organize, manage or store and analyse such datasets.

Industrial enterprises as well as governmental and parastatal institutions such as ministries *(e.g., of finance, health care, telecommunications, immigration, education, agriculture, etc.)* have recently become keenly interested in the high value-adding potential of big data. Hence, many government agencies have initiated major plans to accelerate big data research and applications (Report to the President Big Data and Privacy: A Technological Perspective, 2014). Big data research has also gained a lot popularity in the academic world, which has motivated publications from a lot of publishing houses as well as congress and conference proceeding themes in addition to the media and industrial white papers. Therefore, the impact, performance and challenges of big data have been discussed widely across all the various sectors. However, big data is still seen as an abstract concept; thus, differentiating between itself and "huge data" or "massively big data" is still rather fuzzy. Although the essence of big data value-adding has been generally acknowledged, industrial enterprise managers still have different opinions on its definition mainly relative to the nature of their operations. Big data could generally be referred to as the datasets that could not be perceived, acquired, managed or stored and finally analysed by legacy IT and software/hardware systems within a reasonable time frame. Thus, various stakeholders have their own definition of big data which are most likely relative to the nature of their organizational operations.

According to Min, Shiwen, and Yunhao (2014), big data typically comprises masses of unstructured data that needs more real-time analysis. Manyika et al. (2011) defined big data as the next frontier for innovation, competition, and productivity (Intel IT Centre – Peer Research, 2012). Richard et al. (2011) in their IDC whitepaper stated that big data technology could be described as a new generation of technologies and architectures, designed so that enterprise organizations could economically extract value from very large volumes of a wide variety of data by enabling high-velocity capture, discovery, storage and analysis. This definition is largely agreed to by many researchers and enterprise industrial R&D managers (Seref and Duygu, 2013; Min et al., 2014; Manyika et al., 2011; Janusz, 2013), although it is also obvious others have different views. The IDC, one of the most influential leaders in big data and its research fields, defines big data in two of its reports (Gantz & Reinsel, 2011; Richard et al., 2011), and outlines some attributes of big data as the four Vs, that is, big data development sources (Variety – V_1), big data acquisition (Velocity – V_2), big data storage (Volume – V_3), big data analysis (Veracity – V_4), and finally modulating towards big data value adding or implementation benefits to industry (Value-adding – V_5). Hence, the five Vs of big data can be seen in Fig. 1, which illustrates the big data S-cure. This implies that big data is thus the data of which the data volume, acquisition speed or data representation limits the capacity of using classical database management methods to conduct effective analysis (Mayer-Schönberger & Cukier, 2013) and therefore that efficient methods or technologies need to be developed and used to analyse and process big data.

The S-curve model in Fig. 1 illustrates the need to scrutinize the attributes of big data for a more optimum value-adding approach by squeezing a huge data volume (V_3) for more enhanced data flow velocity (V_2) into a lesser time, and also stepping up the data variety (V_1) into a more enhanced date veracity (V_4) in a lesser time. According to Hauang, Zhong, and Tsui (2015), an automatic data collection approach with high level systems/technologies and networking sensors such as RFID brings new challenges. They further elaborate that these challenges could be summarized from the horizontal and vertical dimensions. The horizontal dimensions indicate the dynamics of big data, which means the interaction and interlinking features of data among the manufacturing, logistics, and retailing phases. The vertical dimension describes the characteristics of big data, which are highlighted as the "5Vs" – volume, velocity, variety, verification, and value (Hauang et al., 2015). In recent times there have been extensive discussions in both enterprise industrial organizations and academia about a consensus definition of big data (Team, 2011; Grobelnik, 2012). Furthermore, it has been identified that effective and efficient value-adding (V_5) acquisition, storage, development and analysis of big data for enterprise industrial SCM is important in this era and can never be over-emphasized.

The remaining sections of this research paper investigate, discuss and elaborate on the following: first, big data in the perspective of operations/SCM; second, big data applications in operations/SCM; third, various analysis tools for big data in operations/SCM; fourth, the trends and perspective of big data, followed by big data extension; fifth comprises the operational/managerial implications of big data application and analysis in SCM, and finally the paper ends with the conclusion and recommendations of the research.

2. Big data in operations/supply chain management perspective

As already stated in the previous sections of this research paper, the actual definition or an agreed definition of big data has not yet been settled on. However, the aligning definition or what big data really is; is relative and based on the operations of various enterprises industrial organizations. Although there is as yet no agreed definition of big data, many enterprise industrial SCM stakeholders and experts predict that big data will have a positive impact on their operations and activities, enabling them to make more strategic data-oriented and informed decisions. Furthermore, one of the reports from International Data Corporation (IDC), Gantz and Reinsel (2011) predicted that the return-on-investment (ROI) for the big data market would reach $16.1 billion in 2014, thus representing a growth about six-times faster than Information Technology (IT) businesses overall. Therefore, it has become imperative that more effort is put into arriving at a common consensus definition for big data in an operations or supply-chain management perspective to obtain more informed and data-oriented strategic decision making.

Identifying a clear understanding and a common definition of big data in operations or supply-chain management has been long overdue as it is imperative for enterprise SCM (eSCM) stakeholders to work together collaboratively with consistency in definition and terminology. This will enhance efficiency and effectiveness in their Information and Communication Technology (ICT) processes and applications to obtain sustainable competitive advantage. According to Milan (2015), big data provides ample opportunities in SCM as an invaluable instrument for spending analysis in terms of supply-chain risks or measuring supplier performance for senior stakeholders with an accuracy never seen before. Furthermore, Milan stated that big data comes with huge possibilities as well as the ability to drill down and identify credible areas for optimization. Big data has been making huge strides in enterprise industrial circles recently as a prospective and feasible solution to almost every organizational operations challenge facing industrial decision makers today. The research question *(RQ)* here is how