



Contents lists available at ScienceDirect

## Journal of Business Research



## Factors influencing big data decision-making quality☆

Marijn Janssen\*, Haiko van der Voort, Agung Wahyudi

Delft University of Technology, Faculty of Technology, Policy and Management, Jaffalaan 5, 2628 BX Delft, The Netherlands

## ARTICLE INFO

Available online xxx

## Keywords:

Big data  
Big data analytics  
Big data chain  
E-government  
Governance  
Decision-making  
Decision-making quality

## ABSTRACT

Organizations are looking for ways to harness the power of big data (BD) to improve their decision making. Despite its significance the effects of BD on decision-making quality has been given scant attention in the literature. In this paper factors influencing decision-making based on BD are identified using a case study. BD is collected from different sources that have various data qualities and are processed by various organizational entities resulting in the creation of a big data chain. The veracity (manipulation, noise), variety (heterogeneity of data) and velocity (constantly changing data sources) amplified by the size of big data calls for relational and contractual governance mechanisms to ensure BD quality and being able to contextualize data. The case study reveals that taking advantage of big data is an evolutionary process in which the gradually understanding of the potential of big data and the routinization of processes plays a crucial role.

© 2016 Elsevier Inc. All rights reserved.

## 1. Introduction

Big data (BD) has rapidly moved to being a mainstream activity of organizations. Tapping into large-scale, fast-moving, complex streams of datasets has the potential to fundamentally transform the way organizations make their decisions. Big data refers to datasets that are both big and high in variety and velocity, which makes them difficult to handle using traditional tools and techniques (Elgendy & Elragal, 2014). The ability to take advantage of all available information has become a critical ability for organizational success (Olszak, 2016). The creation of value from data requires combining large datasets originating from different and heterogeneous data sources (Janssen, Estevez, & Janowski, 2014). Big data is closely related to Big Data Analytics (BDA) which are needed to create value of the data (Elgendy & Elragal, 2014; Holsapple, Lee-Post, & Pakath, 2014).

Data is often generated by other organizations, by users on social media or provided by devices of the Internet of Things (IoT). In practice there is often a whole chain of activities in which various actors plays a role (Anderson, 2015). The variety of data sources, the need to combine various sources, and the use of BDA often requires the collaboration between organizations and departments to create a flow of activities. Organizational silos are locking the use of big data for decision-making (Economist Intelligence Unit, 2012). Data collection, processing and use is not done within a single department or organization. Instead data is collected by many parties and organizations might use

collaborations and partnerships for acquiring the resources and capabilities for analyzing BD. All these actors and steps increase the difficulty of using BD for decision-making.

The value of BD often originates from the ability to take better decisions (Economist Intelligence Unit, 2012). The quality is not solely dependent on the data, but also on the process in which the data is collected and the way data is processed. BD and BDA often requires bringing together multiple actors from different disciplines and diverse practices to examine the underexplored relationships between types of data (Janssen & Kuk, 2016). Each activity may be carried out by different actors with different capabilities and skills.

The involvement of a variety of organizations results in a flow or chain of activities which can be labeled as the 'big data chain'. A BD chains begins with collecting the data from the sources and ends when data-based decisions are taken. A big data chain consists of subsequent activities that can be distinguished analytically. The term 'chain' refers to the analytical view taken on the collaboration (Stank, Keller, & Daugherty, 2001). In reality, there are many data sources, variations in flows and decisions. In such a chain there are many efforts to increase the quantity and quality of published data over time. These include removing noise, converting selected datasets into machine readable and linked data and adding meta-data (Kitchin, 2014). These activities can affect how BD can be used for decision-making. The chain perspective is hardly taken as an analytical view of looking at BD. A search using the keywords "big data" and "chain" results only in a few hits. Brown, Chui, and Manyika (2011) are the only one who mention the chain in relationship to big data to express the need to cooperate with supply chain partners and the role other parties can play in generating data.

The goal of the research is to identify factors influencing BD decision-making quality. Often it is assumed that BD results in better decisions,

☆ The author is grateful for proofreading the paper by Karin Boersma and the comments by anonymous reviewers on prior drafts of this article.

\* Corresponding author.

E-mail address: m.f.w.h.a.janssen@tudelft.nl (M. Janssen).

but it is unclear which factors influence the decision-making quality and how decision-making quality can be improved by organizations. As BD and BDA become more common, understanding the big data chain and factors influencing decision-making quality chain becomes of paramount importance for organizations.

This structure of this paper is as follows. First, factors influencing BD a decision-making quality as found in the literature. Thereafter the research method in which a BD chain is analyzed in a large, administrative organization is presented. Based on the case study factors influencing decision-making quality are derived. Finally, business implications are discussed and conclusions are drawn.

## 2. Big data for decision-making

Several steps for the BD process starting with data capturing and resulting in decision-making can be found in the literature. For example, Bizer, Boncz, Brodie, and Erling (2012) identify six steps; data capturing, data storage, data searching, data sharing, data analysis, and data visualization. In contrast Chen and Liu (2014) only use three steps; data handling, data processing, and data moving. Marx (2013) proposes five steps; problem definition, data searching, data transformation, data entity resolution, answer the query/solve the problem. Whereas others used other names for denoting similar steps. For example Zhou, Chawla, Jin, and Williams (2014) use data collection, data storage, data management, data manipulation, data cleansing, and data transformation. Although steps are identified little attention is paid to who executes these steps and the effects of one step on the other steps.

BD is commonly characterized with three or more Vs: Volume, Velocity, Variety (McAfee & Brynjolfsson, 2012). Gandomi and Haider (2015) add three other Vs to this list; Value, Variability and Veracity. These Vs makes the datasets difficult to handle in traditional ways (Elgendy & Elragal, 2014). BD often originates from many sources which are often beyond the control of a single actor like social media and devices. Veracity refers to that data may be incomplete, out-of-date and contains noise (Gandomi & Haider, 2015). BD sources have a variety of data quality. Data quality is a multidimensional concept describing properties of the information such as accuracy, timeliness, completeness, consistency, relevance and fitness for use (Miller, 1996; Strong, Lee, & Wang, 1997). Data quality can be viewed as set of dimensions describing the quality of the information produced by the information system (DeLone & McLean, 1992). Past research about the use of data shows that data quality influences decision-making quality (Keller & Staelin, 1987; O'Reilly, 1982). As such, BD quality might influence decision-making quality.

An effective BD chain requires to build capabilities and capacity for data management and BDA (Chen & Hsieh, 2014). BDA capabilities include descriptive, exploratory, inferential, predictive, causal and mechanistic techniques (Schutt & O'Neil, 2013). For that reason, various methods are employed such as natural language processing, text mining, linguistic computation, machine learning, search and sort algorithms, syntax and lexical analysis, and so on. BD is often related to predictive analytics which comprises a variety of techniques that predict future outcomes to uncover patterns and find relationships in data (Gandomi & Haider, 2015). Past research in data processing shows that the organizational capability to process information impacts its performance (Galbraith, 1973; Premkumar, Ramamurthy, & Saunders, 2005). Activities for processing BD and BDA capabilities likely influence decision-making quality.

The capabilities of each organizational entity involved in the BD chain influence the outcomes. Capabilities include skills and processes that transform inputs into outputs of greater value (Wade & Hulland, 2004). The ability of organizational entities and firms to collect, prepare and analyze BD might be different. Firms may possess a capability that is idiosyncratic to the firm or difficult to imitate due to path dependencies (Dierickx & Cool, 1989). For example, they might have experts with specialized knowledge in BDA that are rare on the market (Davenport

& Patil, 2012). Firms might also have developed a customized infrastructure enabling them to take advantage of BD.

Raghunathan (1999) defines decision-making quality as the accuracy and correctness of decisions. Decision quality may improve or degrade when information quality and processing improves (Raghunathan, 1999). As data becomes larger, more complex, and more inexplicable, the limited mental capacities of humans pose difficulties in deciphering and interpreting an unknown environment (Sammur & Sartawi, 2012). In BD there might be little understanding of what the data actually means and in which context data is collected. Lack of knowledge about BD sources influences decision-making quality.

Decision-makers should be able to interpret the outcomes of BDA and should not be manipulated by fancy graphics (Huff, 1993). Raghunathan (1999) found that decision quality improves if the decision-maker has knowledge about the relationships among problem variables. In contrast, the decision quality of a decision-maker may degrade if the decision-maker does not understand the relationships. Interactions with those who collected and processed data produce better decisions than those without (Burlinson, Levine, & Samter, 1984). This may equally apply for BD, suggesting that interactions with others persons involved in the BD chain involved result in higher decision-making quality.

The previous overview shows that the quality of the decision depends on the quality of the inputs and on the quality of the process that transforms the inputs into outputs. Factors affecting the decision-making quality of BD include the characteristics and quality of the BD sources, the quality of the BDA process, the BDA capacity and capabilities of persons involved in collecting and processing BD, and the availability of an BD infrastructure. In addition, research in data processing shows that the ability of decision-makers to understand the data and collaborating with others in the BD chain results in better decision quality.

## 3. Research approach

Literature shows that there is a wide range of factors influencing the BD decision-making. Particular in situations in which several actors are involved and it is hard to oversee all steps of the BD chain, the quality of decisions might be compromised. A deep understanding of the context is necessary to understand factors affecting it (Davenport, 2010; Goodhue, Wybo, & Kirsch, 1992). In a similar vein, a deep understanding of the BD chain is required to understand the factors influencing decision-making quality. Therefore an indepth case study within a large information-processing organization was conducted.

A qualitative approach based on a case study research (Yin, 1989) was adopted to gain a deep understanding of the factors influencing decision-making quality. The case study research methodology is particularly well-suited for investigating organizational issues (Benbasat, Goldstein, & Mead, 1987). A single case study can contribute to scientific development through a deep understanding of the context and by capturing experiences (Flyvbjerg & Budzier, 2011). Deep understanding is necessary to identify a broad range of factors influencing decision-making quality, whereas understanding experiences results in the identification of mechanisms for improving the decision-making quality. As decision-making quality is dependent on the decision-maker, the data collecting and processing, all these aspects were taken into account when analyzing the case study.

The number of cases that could reveal factors influencing BD for decision-making quality and using BDA for decision-making were found to be limited. This was further complicated as some of the cases considered did not want to disclose their practices. The Dutch Tax organization was selected as this organization was willingly to share their practices and much information was available publicly. Furthermore, this large information processing organization is considered as a frontrunner in the use of BD and BDA within the Dutch government.

Download English Version:

<https://daneshyari.com/en/article/5109826>

Download Persian Version:

<https://daneshyari.com/article/5109826>

[Daneshyari.com](https://daneshyari.com)