A Preliminary Systematic Performance on Critical Success Factors Categories for Big Data Analytics

Zaher Ali Al-Sai, Rosni Abdullah, Mohd Heikal Husin, Sharifah Mashita Syed-Mohamad

Abstract: Big Data could be used in any industry to make effective data-driven decisions. The successful implementation of Big Data projects requires a combination of innovative technological, organizational, and processing approaches. Over the last decade, the research on Critical Success Factors (CSFs) within Big Data has developed rapidly but the number of available publications is still at a low level. Developing an understanding of the Critical Success Factors (CSFs) and their categories are essential to support management in making effective data-driven decisions which could increase their returns on investments. There is limited research conducted on the Critical Success Factors (CSFs) of Big Data Analytics (BDA) development and implementation. This paper aims to provide more understanding about the available Critical Success Factors (CSFs) categories for Big Data Analytics implementation and answer the research question (RQ) “What are the existing categories of Critical Success Factors for Big Data Analytics?”. Based on a preliminary Systematic Literature Review (SLR) for the available publications related to Big Data CSFs and their categories in the last twelve years (2007-2019), this paper identifies five categories for Big Data Analytics Critical Success Factors (CSFs), namely, Organization, People, Technology, Data Management, and Governance categories.

Keywords: Big Data, Big Data Analytics, Critical Success Factors (CSFs), Big Data Critical Success Factors.

I. INTRODUCTION

The term Big Data (BD) was coined to describe the increased collection of various data from several sources such as dealing with customers, vendors, and even among the internal employees (Koronios et al., 2014). As this phenomenon is occurring in various organizations at the moment, the concept has attracted the attention of managers and executives toward implementing and investing in Big Data projects. Due to the large volume of data generated, the current infrastructure in several areas such as technology capabilities, organizational structure, and processing capacity often fails to deal with the requirements in capturing, storing, and even processing Big Data efficiently (Eybers & Hattingh 2017; Alexandros Labrinidis & H. V. Jagadish 2012). Big Data projects often differ from other technology projects, where the implementation of such projects sometimes requires new technical and organizational approaches (Cato, Golzer, & Demmelhuber, 2016). This often requires organizations to be ready for additional requirements in those areas that should be aligned with the increased complexity of Big Data characteristics (namely Volume, Velocity, and Variety) to ensure the achievement of a successful implementation, and the ability to gain a suitable value from the Big Data investments (Eybers & Hattingh 2017; Alexandros Labrinidis & H. V. Jagadish 2012).

Despite the obvious benefits of harvesting, BD projects often represent one of the largest and riskiest investments that require a sufficient level of preparation from the organization side (Evers 2014). Organizations are currently at a crossroad, where they are either able to handle these challenges and reap the benefits from Big Data or struggle badly with Big Data challenges (Eybers & Hattingh 2017). As such, organizations can decide to be more prepared before investing in a BD technology to ensure that they would not lose their competitive advantage over their competitors (Eybers & Hattingh 2017). Many organizations found out later on that their adoption of Big Data was not as successful as they expected. These results could stem from several issues that occur for most organizations; from the ineligible capabilities of Big Data and analytics infrastructures to process the large, varied and constant flow of data during the implementation phase. Therefore, a combination of such issues related to low-quality data and insufficient organizational requirements from Big Data could affect an organization’s overall performance, business outcomes, and decision-making (Adrian et al. 2017). Big Data Analytics (BDA) implementation is enveloped by several factors that organizations should examine to guarantee a successful implementation (Halaweh & Massry, 2015). The concept and the approach of Critical Success Factors (CSFs) are very powerful and applicable to address many of the challenges in the information technology fields (Caralli 2004). The growing interest in Big Data requires a focused investigation on the critical success factors as well as their effect on the implementation process (Mikalefet al., 2016).

Critical Success Factors are key areas where everything will be done in the right way through the business process in order to achieve the organization’s goals (Kaur & Singh, 2016). The concept of Critical Success Factors (CSFs) has...
A Preliminary Systematic Performance on Critical Success Factors Categories for Big Data Analytics

II. MATERIALS AND METHODS

This research aims to summarize the existing literature relevant to Critical Success Factors for Big Data Analytics to identify their categories in the available literature. As such, a systematic literature review (SLR) is a suitable, reliable and accurate technique and method can be used to identify and evaluate all existing research related to a particular research question, topic domain, or phenomenon of interest (Kitchenham 2004). While the individual studies that contribute to a systematic review are called primary studies, the systematic review considers as a form of secondary study can be used in this research (Kitchenham 2004; Kitchenham 2007; Saltz & Shamshurin 2016a).

In this research, the SLR method was used to evaluate all the available research related to the research question (RQ) “What are the existing categories of Critical Success Factors for Big Data Analytics”. The selection of primary studies is governed by the inclusion and exclusion criteria in the context of Big Data Critical Success Factors and their categories. This research identified several keywords were: “Big Data”, “Big Data analytics”, “Big Data projects”, “success”, “critical success factors”, “issues”, “challenges”, “barriers”, “requirements”, “capabilities”, and “successful implementation”. This research included publications that were in “English” and were published between the years of 2007 and 2019.

A total of seven search engines were used as main electronic databases to identify the available literature relevant to the research question, namely Research Gate, Google, Google Scholar, EBSCOhost, Science Direct, Scopus, and IEEE. Based on the criteria of this SLR paper, 16 relevant articles were identified as possible sources for this paper. The relevant literature and their sources (journals, conferences, thesis, and dissertation) which are recognized in this research paper are shown in Table 1.

Table 1 Selected sources and studies for the systematic literature review

<table>
<thead>
<tr>
<th>Source</th>
<th>Number of selected studies</th>
<th>Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal of International Technology and Information Management</td>
<td>1</td>
<td>(Halaweh &amp; Massry 2015)</td>
</tr>
<tr>
<td>Journal of The Association for Information Science and Technology</td>
<td>1</td>
<td>(Yeoh &amp; Popovic 2015)</td>
</tr>
<tr>
<td>Journal of Technological Forecasting and Social Change – Elsevier</td>
<td>1</td>
<td>(Wang et al. 2016)</td>
</tr>
<tr>
<td>Journal of Information Development - Sage</td>
<td>1</td>
<td>(Kim &amp; Park 2017)</td>
</tr>
<tr>
<td>Review of Business Management</td>
<td>1</td>
<td>(Félix et al. 2018)</td>
</tr>
<tr>
<td>Development Informatics - Working Paper</td>
<td>1</td>
<td>(Gómez &amp; Heeks 2016)</td>
</tr>
<tr>
<td>Twenty-First Americas Conference on Information Systems</td>
<td>1</td>
<td>(Gao et al. 2015a)</td>
</tr>
<tr>
<td>2016 IEEE International Conference on Big Data</td>
<td>1</td>
<td>(Saltz &amp; Shamshurin 2016a)</td>
</tr>
</tbody>
</table>
As shown in Table 1, eight conferences papers, six journals, one MSc thesis, and one dissertation were included in the Systematic Literature Review.

III. RESULTS AND DISCUSSION

Previous studies by Halaweh & Massry (2015); Yeoh & Popovic (2015); Wang et al. (2016); Kim & Park (2017); Félix et al. (2018); Cato et al. (2016); Gómez & Heeks (2016); Koronios et al. (2014); Gao et al. (2015); Saltz & Shamshurin (2016); Chen et al. (2016); Eybers & Hattingh (2017); Adrian et al. (2017); Schüll & Maslan (2018); Evers (2014); Nieder (2016), have been selected and analyzed to identify the categories related to the research question of this paper. The criteria used to classify the categories of CSFs for BDA are based on the most common categories identified in the systematic review process and its relevance to the context of this study. As a result, the categories of critical success factors for BDA have been identified and classified into five main categories, namely Organization, People, Technology, Data Management, and Governance. The five main categories for Big Data Analytics CSFs and their CSFs from the existing literature are shown in Table 2.

Table 2 The main categories for Big Data Analytics CSFs and their CSFs in literature

<table>
<thead>
<tr>
<th>CSFs category</th>
<th>CSFs</th>
<th>Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 Governance</td>
<td>Data Standardization, Legal, Data and Information Quality, System Quality, Data Availability and Quality</td>
<td>(Saltz &amp; Shamshurin 2016a; Cato et al. 2016; Kim &amp; Park 2017)</td>
</tr>
</tbody>
</table>

This section describes the five categories for Big Data Analytics CSFs that identified during a preliminary analysis of the selected studies. A Top-Down technique was applied to define the categories for the Big Data CSFs. The five categories related to the research question for this paper are:

**Organization**

The Organization category considered the most prevalent critical success factors (CSFs) within the existing literature (Eybers & Hattingh 2017).
The category refers to the vision, mission, and strategy of the organization. For successful implementation for the BDA investments, the organization should be clear on what they want to achieve with the alignment to the overall objectives of the organization (Eybers & Hattingh 2017). The characteristics of the organization affect the successful implementation of Big Data systems within the organization including the subcategory related to close collaboration between IT and Business as well as organizational structure (Cato, Golzer, & Demmelhuber, 2016).

People

The category of People is a key element to determine the success of Big Data projects. According to Hayen et al. (2007) findings, the availability of adequate resources, appropriate people on the project team and user participation have positive effects on the project's success. This category is often referred to as Human capabilities, team skills, and statistical and analytical skills required for team members working on Big Data projects (Eybers & Hattingh 2017).

Technology

The Technology category is related to data collection, storage, processing, analytics platforms, and applications which would act as a facilitator of Big Data projects (Kim & Park 2017). The category focused on technological tools investments with particular analytical capabilities and data visualization tools such as real-time availability, system performance, data quality, and integration with the existing tools. At the same time, the system infrastructure and applications refer to factors like system flexibility, scalability, and security. The Data sources refer to any factor related to data and will influence the trustworthiness of data, such as data quality, metadata management, and integrity (Eybers & Hattingh 2017).

Data Management

This category is related to the administrative process that comprises acquiring, processing, validating, storing, and protecting any required data to guarantee the accessibility, reliability, and timeliness of the data (Galetto 2018). Data management is a common term that overlay a wide range of data applications. This may refer to the basic concepts of data management or to specific technologies like Big Data. Some applications of Big Data management include data design, data storage, and data security (TechTerms 2012).

Governance

The category of Governance refers to the processes, policies, social activities, and practices undertaken by either institutions or other relevant actors. The category is commonly related to the act of “governing” where multiple actors are involving in the Big Data projects. The required Data often has to cross several organizational, departments, and even public-private borders. Governance is a key category because it contributes to the effectiveness and efficiency of processes that overlay through those different borders (Veeneman et al. 2018).

IV. CONCLUSION

During the last decade, the Critical Success Factors (CSFs) for Big Data Analytics (BDA) have developed rapidly but with limited research publications. Understanding the Critical Success Factors (CSFs) is essential to support the data-driven decisions making and to increase the returns on investment (ROI). Through the development of a comprehensive Information System (IS) success model for the Big Data context, it needs to be investigated which factors and their categories that could affect the successful implementation of BDA. Based on a preliminary systematic literature review of the existing literature related to the Critical Success Factors categories for Big Data Analytics in the last twelve years (2007-2019), sixteen (16) previous studies have been selected and preliminarily analyzed to identify the most common categories for successful Big Data Analytics implementation. As a result, the most common categories have been classified into five main categories, namely Organization, People, Technology, Data Management, and Governance. These categories can be used by the organizations to handle the issues related to identifying the key aspects in Big Data issues, highlighting the critical factors of Big Data Analytics implementation, and creating a platform for a successful implementation overall Big Data initiatives. Finally, this paper provides a grounded referenced with critical insights for the managers to formulate and execute the Big Data strategies and decision-making. Future work may study the key elements of the CSFs categories of Big Data Analytics, the importance of these CSFs, and the priority for each factor.

ACKNOWLEDGMENT

The authors would like to express appreciation for the support of the sponsors [Grant Number = 1001/PKOMP/PKOMP/8014076].

REFERENCES

1. Adrian, C. et al., 2017. Factors Influencing to the Implementation Success of Big Data Analytics: A Systematic Literature Review. IEEE.