International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-9 Issue-1, October, 2019

A Preliminary Systematic Performanceon Critical Success Factors Categories for Big Data Analytics

ZaherAliAl-Sai, Rosni Abdullah, MohdHeikalHusin, SharifahMashita 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 understandingof the Critical Success Factors (CSFs) and their categoriesare essential to support management in making effective data-driven decisions which could increase their returns on investments.There islimited research conducted on the Critical Success Factors (CSFs) of Big DataAnalytics (BDA) development and implementation.This paper aims to provide more understanding about the availableCritical Success Factors (CSFs) categoriesfor 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 identifiesfive categoriesfor Big Data AnalyticsCritical Success Factors(CSFs), namelyOrganization, 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 managersand executives toward implementingand 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 Dataefficiently(Eybers & Hattingh 2017; Alexandros Labrinidis & H. V. Jagadish 2012).

Revised Manuscript Received on October 30, 2019.
* Correspondence Author
ZaherAliAl-Sai*, School of Computer Sciences, Universiti Sains Malaysia
Rosni Abdullah, School of Computer Sciences, Universiti Sains Malaysia
MohdHeikalHusin, School of Computer Sciences, Universiti Sains Malaysia
SharifahMashita Syed-Mohamad, School of Computer Sciences, Universiti Sains Malaysia

Big Data projects oftendiffer from other technology projects, where the implementation of such projectssometimes requiresnew 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 toensure 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 DataAnalytics (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

been proven useful as it could beapplied to address many challenges within the Information Technology (IT)area(Caralli 2004). Critical Success Factorswas originally developed to align technology planning with the strategic planning of an organization by identifying the essential elements of successfor a project(Eybers & Hattingh 2017; Caralli 2004). Most organizations have adoptedthe CSF assessment method to identify the business value of their Big Data Analytics projects and to assess the risk before investing in the technology (Eybers & Hattingh 2017). By developing an understanding ofthe Critical Factors related to the successful management and implementation of Big Data Analytics, organizations will be able to identify their readiness, management, implementation, and to gain more Returns On their Investment (ROI). According to Lavalle et al. (2011); Eybers & Hattingh (2017), the organizational challenges are most critical determinates influencing the adoption of Big Data (Eybers & Hattingh 2017).

This paperattempts to answer the research question (RQ) "What are the existing categories of Critical Success Factors for Big Data Analytics" bysummarizes themost common categories of the Big DataCritical Success Factors through a preliminary analysis of the Systematic Literature Review (SLR) that could assist organizations in developing their plans for Big Data Analytics implementation. The following section presents the paper's research methodology. This is then followed by the results' discussion which focuses on the identificationof the Critical Success Factors categories for Big Data and discusses the results based on the categories ofTechnical, Organizational, People, Data Management, and Governance. The final section provides a conclusion for this paper. The findings from this paper can be used as an initial guide for future research in identifying the key elements of the Big DataAnalytics (BDA) CSFs categories, the importance of these CSFs, and the priority for each factor for an effective Big Data adoption.

II.MATERIALS AND METHODS

This research aims to summarize the existing literature relevantto Critical Success Factors for Big Data Analytics to identify their categoriesin the available literature. As such, a systematic literature review (SLR) is a suitable, reliable and accurate technique and methodcan be used to 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(Kitchen ham 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 andstudies for the systematic literature review

Source	Number of selected studies	Literature
Journal of International Technology and Information Management	1	(Halaweh & Massry 2015)
Journal of The Association for Information Science and Technology	1	(Yeoh & Popovic 2015)
Journal of Technological Forecasting and Social Change – Elsevier	1	(Wang et al. 2016)
Journal of Information Development- Sage	1	(Kim & Park 2017)
Review of Business Management	1	(Félix et al. 2018)
Development Informatics - Working Paper	1	(Gómez & Heeks 2016)
Proceedings - Pacific Asia Conference on Information Systems, PACIS 2014	1	(Koronios et al. 2014)
Twenty-First Americas Conference on Information Systems	1	(Gao et al. 2015a)
2016 IEEE International Conference on Big Data	1	(Saltz & Shamshurin 2016a)
Proceedings - 2015 11th International Conference on Innovations in Information Technology, IIT 2015	1	(Cato, Golzer, & Demmelhuber, 2016)

Proceedings of The Annual Hawaii International Conference on System Sciences 2016	1	(Chen et al. 2016)
2017 1st-Africa Week Conference (1st-Africa)	1	(Eybers & Hattingh 2017)
2017 International Conference on Research and Innovation in Information Systems (ICRIIS)	1	(Adrian et al. 2017)
International Conference on Enterprise Information Systems (ICEIS 2018)	1	(Schüll & Maslan 2018b)
Master Thesis	1	(Evers 2014)
MSc Dissertation	1	(Nieder 2016)

As shown in Table 1, eight conferences papers, six journals, one MSc thesis, and one dissertationwere included in theSystematic Literature Review.

III.RESULTS AND DISCUSSION

Previous studies byHalaweh & 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 categoriesrelated to the research question of this paper. The criteria used to classify the categories of CSFs for BDAare based on the mostcommon categories identified in the systematic review processand its relevance to the context of this study. As a result, thecategories of critical success factors forBDA have been identified and classifiedinto five main categories,namely Organization, People, Technology, Data Management, and Governance. Thefive maincategories forBig Data Analytics CSFs and their CSFsfrom the existingliterature are shown in Table 2.

Table. 2 The main categories for Big DataAnalyticsCSFsand their CSFs in literature

	CSFs category	CSFs	Literature
1	**Organization**	Organization Capability, Objectives and Values, Strategic Alignment, Top Management Support And Involvement, Support, Compatibility with Organizational Processes, Structure andFunctions For Implementation, Communication, Expectations Management, Organizational Change, Strategic, vision, mission, Identifiable Business Value, Performance Environment and Market Pressure, Analytical Culture, Process, Organizational Analytic Capability, Dynamic Capability, Other Resources (Partnership)	(Cato et al. 2016; Evers 2014; Eybers & Hattingh 2017; Kim & Park 2017; Schüll & Maslan 2018a; Nieder 2016; Saltz & Shamshurin 2016a; Gómez & Heeks 2016; Chen et al. 2016; Koronios et al. 2014; Félix et al. 2018; Adrian et al. 2017; Schüll & Maslan 2018b)
2	**People**	Skills And Knowledge Human Capability Teams Overcoming the Talent Gap	(Cato et al. 2016; Koronios et al. 2014; Evers 2014; Gao et al. 2015b; Saltz & Shamshurin 2016b; Eybers & Hattingh 2017; Wang et al. 2016; Nieder 2016; Kim & Park 2017; Chen et al. 2016; Gómez & Heeks 2016)
3	**Technology**	Information Technology (IT), IT Infrastructure and application, Technology Capability, Tools, Analytics Capability	(Cato et al. 2016; Gao et al. 2015a; Koronios et al. 2014; Evers 2014; Saltz & Shamshurin 2016b; Eybers & Hattingh 2017; Kim & Park 2017; Khan et al. 2017; Chen et al. 2016; Schüll & Maslan 2018b; Gómez & Heeks 2016; Wang et al. 2016)
4	**Data Management**	Management, Management Systems And Structure, Information Sharing, Cultural Change For Data Orientation And Experimentation, Privacy Management, Privacy And Security, Information Security For Big Data, Security And Integrity,Data	(Cato et al. 2016; Adrian et al. 2017; Kim & Park 2017; Saltz & Shamshurin 2016a)
5	**Governance**	Data Standardization, Legal, Data And Information Quality, System Quality, Data Availability And Quality	(Saltz & Shamshurin 2016a; Cato et al. 2016; Kim & Park 2017)

This section describesthe fivecategories 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. Thefive categories related to the research questionforthis 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 asorganizational 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 havepositive 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

Thiscategory 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 practicesundertaken 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 Dataoftenhas 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 aSystematic LiteratureReview of the existing literature related to theCritical 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.
2. AlexandrosLabrinidis& H. V. Jagadish, 2012. Challenges and Opportunities with Big Data. Proceedings of the VLDB Endowment, pp.1–15.
3. Caralli, R.A., 2004. The Critical Success Factor Method : Establishing a Foundation for Enterprise Security Management. , (July).
4. Cato, P., Golzer, P. &Demmelhuber, W., 2016. An investigation into the implementation factors affecting the success of big data systems. Proceedings - 2015 11th International Conference on Innovations in Information Technology, IIT 2015, pp.134–139.
5. Chen, H.M. et al., 2016. Amazon in the air: Innovating with big data at Lufthansa. Proceedings of the Annual Hawaii International Conference on System Sciences 2016, 2016–March, pp.5096–5105.
6. Evers, J.M., 2014. Critical Success Factors Of Business Intelligence And Big Data Analysis. Master Thesis.
7. Eybers, S. &Hattingh, M.J., 2017. Critical success factor categories for big data: A preliminary analysis of the current academic landscape. 2017 IST-Africa Week Conference (IST-Africa), pp.1–11. Available at: http://ieeexplore.ieee.org/document/8102327/.
8. Félix, B.M., Tavares, E. &Cavalcante, N.W.F., 2018. Critical success factors for Big Data adoption in the virtual retail: Magazine Luiza case study. RevistaBrasileira de Gestao de Negocios, 20(1), pp.112–126.

9. Galetto, M., 2018. What is Data Management? - NGDATA. Available at: https://www.ngdata.com/what-is-data-management/.
10. Gao, J., Koronios, A. &Selle, S., 2015a. Towards A Process View on Critical Success Factors in Big Data Analytics Projects. Twenty-first Americas Conference on Information Systems, pp.1–14.
11. Gómez, L.F. &Heeks, R., 2016. Measuring the Barriers to Big Data for Development: Design-Reality Gap Analysis in Colombia's Public Sector, Available at: http://www.digitale-chancen.de/transfer/downloads/MD280.pdf.
12. Halaweh, M. &Massry, A. El, 2015. Conceptual Model for Successful Implementation of Big Data in Organizations. Journal of International Technology and Information Management, 24(2), pp.21–29.
13. Hayen, R.L., D.Rutashobya, C. & Vetter, D.E., 2007. An Investigation Of The Factors Affecting Data Warehousing Success. International Association for Computer Inform ation Systems, VIII(2), pp.547–553.
14. Khan, S. et al., 2017. A survey on scholarly data: From big data perspective. Information Processing and Management, 53(4), pp.1339–1351.
15. Kim, M.K. & Park, J.H., 2017. Identifying and prioritizing critical factors for promoting the implementation and usage of big data in healthcare. Information Development, 33(3), pp.257–269.
16. Kitchenham, B., 2007. Guidelines for performing Systematic Literature Reviews in Software Engineering. Software Engineering Group School of Computer Science and Mathematics, p.65.
17. Kitchenham, B., 2004. Procedures for performing systematic reviews.
18. Koronios, A., Gao, J. &Selle, S., 2014. Big Data Project Success – a Meta Analysis. Pacis.
19. Lavalle, S. et al., 2011. Big Data, Analytics and the Path From Insights to Value. , (52205).
20. Nieder, N., 2016. Effective Big Data Management : A Development Of Critical Success Factors And An Analysis Of Firms' Capabilities In The Automotive Industry.
21. Saltz, J.S. &Shamshurin, I., 2016a. Big data team process methodologies: A literature review and the identification of key factors for a project's success. Proceedings - 2016 IEEE International Conference on Big Data, Big Data 2016, pp.2872–2879.
22. Schüll, A. &Maslan, N., 2018b. On the Adoption of Big Data Analytics: Interdependencies of Contextual Factors. In International Conference on Enterprise Information Systems (ICEIS 2018). pp. 425–431.
23. TechTerms, 2012. Data Management Definition. Available at: http://www.techterms.com/definition/data_management.
24. Veeneman, W. et al., 2018. PETRA : Governance as a key success factor for big data solutions in mobility. Research in Transportation Economics, 69(May), pp.420–429. Available at: https://doi.org/10.1016/j.retrec.2018.07.003.
25. Wang, Y., Kung, L.A. & Byrd, T.A., 2016. Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. Technological Forecasting and Social Change, 126, pp.3–13. Available at: http://dx.doi.org/10.1016/j.techfore.2015.12.019.
26. Yeoh, W. &Popovic, A., 2015. Extending The Understanding Of Critical Success Factors For Implementing Business Intelligence Systems. Journal of the association for information science and technology, 3(2), pp.80–90.