
# *LOCAL AND GLOBAL MEASURES FOR MEASURING PERFORMANCE OF BIG DATA ANALYTICS PROCESS*

**ISMAIL MOHAMED ALI**

**FSKTM 2020 10**

# LOCAL AND GLOBAL MEASURES FOR MEASURING PERFORMANCE OF BIG DATA ANALYTICS PROCESS

By

**ISMAIL MOHAMED ALI**

**Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy**

**October 2019**

**COPYRIGHT**

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

**LOCAL AND GLOBAL MEASURES FOR MEASURING PERFORMANCE OF BIG DATA ANALYTICS PROCESS**

By

**ISMAIL MOHAMED ALI**

**October2019**

**Chairman : Associate Professor Yusmadi Yah Jusoh, PhD**
**Faculty : Computer Science and Information Technology**

One pivotal aspect of big data is the process which handles it, mainly referred to as big data analytics (BDA) process. BDA process is an end-to-end process which consists of several stages including data acquisition, data preparation (integration and pre-processing), data analysis, visualization, and interpretation. More has been written about the quality of big data, its dimensions and algorithms applied on data to solve complex problems. However, fewer studies have focused on measuring the performance of BDA process. The success of big data analytics does not merely depend on the quality of data, but also on the performance of the process in which the data are collected, the way data are processed, and how it is presented to the users. Measuring the performance of this process could have enormous benefits in terms of better outcomes, satisfied customers, and evidence-based practices. Therefore, this study aims to identify the local measures that serve measuring the performance of the individual phases of the BDA process, and the global measures that holistically contribute to the performance of the BDA process, and to propose, accordingly, a performance measurement model.

A literature review was conducted, and a conceptual model was derived. Then, based on the conceptual model, a questionnaire was developed. Subsequently, a confirmation study that included an expert review, pilot study and survey was conducted. For the expert review, a questionnaire consisting of 49 items excluding demographic questions, and the conceptual model were sent to four subject-matter experts for verification. Based on the feedback of the experts, the questionnaire and the model were revised. The final survey which was distributed consisted of 48 questions. To ensure the reliability of the instruments, a pilot study was tested with 22 users in big data area. Afterwards, a survey was conducted with a larger population of big data analytics practitioners, and 100 responses were collected for analysis. Then,

a prototype was developed as a proof of concept. Two subject-matter experts viewed the prototype and confirmed that it was in alignment with the proposed model.

The results of confirmation study demonstrated the reliability and validity of the proposed model. The results also revealed the relationships among model constructs, namely: efficiency, effectiveness, technology, competency, and working conditions. In this regard, four out of seven hypotheses for this research were supported. Descriptive statistics was also used to provide a brief summary of the data in the study. Besides the confirmation study, the prototype was evaluated by experts. The results of the evaluation demonstrated the practicality of the proposed model in the real world and elucidated how it can assist organizations in measuring the performance of their big data systems.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

**PENGUKURAN TEMPATAN DAN GLOBAL UNTUK MENGUKUR PRESTASI PROSES ANALITIK DATA BESAR**

Oleh

**ISMAIL MOHAMED ALI**

**Oktober 2019**

**Pengerusi : Profesor Madya Yusmadi Yah Jusoh, PhD**
**Fakulti : Sains Komputer dan Teknologi Maklumat**

Salah satu aspek penting data besar ialah proses pengendalian, iaitu merujuk sebagai Proses Analitik Data Besar (BDA). Proses BDA adalah proses awal-hingga-akhir yang terdiri daripada beberapa tahap termasuk perolehan data, penyediaan data (integrasi dan pra-pemprosesan), analisis data, visualisasi, dan tafsiran. Terdapat banyak kajian mengenai kualiti data besar, dimensi dan algoritma yang digunakan pada data untuk menyelesaikan masalah yang kompleks. Walau bagaimanapun, kurang kajian terdahulu mengenai prestasi proses BDA. Kejayaan analisis data besar tidak hanya bergantung kepada kualiti data, tetapi juga prestasi proses di mana data dikumpulkan, cara memproses data, dan bagaimana ia disampaikan kepada pengguna. Pengukuran prestasi proses memberi manfaat yang besar dari segi hasil yang lebih baik, tahap kepuasan pelanggan, dan amalan berasaskan bukti. Oleh itu, kajian ini bertujuan untuk mengenal pasti ukuran prestasi bagi proses BDA dan mencadangkan model untuk pengukuran prestasi proses BDA.

Kajian literatur telah dijalankan, dan model konseptual dibangunkan. Seterusnya, berdasarkan model konseptual, soal selidik dibangunkan. Satu kajian pengesahan yang merangkumi kajian pakar, kajian rintis dan tinjauan telah dijalankan. Untuk ulasan pakar, soal selidik yang terdiri daripada 49 item yang tidak termasuk soalan demografi, dan model konseptual telah dihantar kepada empat pakar berkaitan untuk pengesahan. Berdasarkan maklum balas para pakar, soal selidik dan model telah disemak. Kaji selidik akhir terdiri kepada 48 soalan. Untuk memastikan kebolehpercayaan instrumen, satu kajian rintis diuji dijalankan kepada 22 orang pengguna data besar. Selepas itu, tinjauan dijalankan dalam populasi besar pengguna analitik data besar, dan 100 maklum balas telah dikumpulkan untuk dianalisa. Kemudian, prototaip telah dibangunkan sebagai bukti konsep. Dua pakar subjek mengesahkan prototaip adalah sejajar dengan model yang dicadangkan.

Hasil kajian pengesahan menunjukkan kebolehpercayaan dan kesahihan model yang dicadangkan. Hasilnya juga menunjukkan hubungan antara model pembinaan, iaitu kecekapan, keberkesanan, teknologi, kompeten, dan keadaan kerja. Begitu juga, hasil penilaian prototaip menunjukkan model yang dicadangkan adalah praktikal di dunia nyata. Hal ini menjelaskan bagaimana ia boleh membantu organisasi dalam mengukur prestasi sistem data besar mereka.

## ACKNOWLEDGEMENTS

In the name of Allah, the Most Gracious and the Most Merciful. All praise to Allah for giving opportunity, strength and determination in completing this thesis.

First, I would like to extend my immense gratitude to my supervisor Assoc. Prof. Dr. Yusmadi Yah Jusoh for her guidance, supervision and support throughout my PhD research. Her invaluable ideas, advices and expertise helped me in the time of research and writing of this thesis.

Besides my supervisor, I wish to thank the rest of my thesis committee: Prof. Dr. Rusli Abdullah, Dr. Rozi Nor Haizan Nor and Assoc. Prof. Dr. Lilly Suriani Affendey, for their encouragement and insightful comments.

I would also like to express my sincere gratitude to my family, my mother: Hasna Mohamud Hasan, for her prayers, support and caring throughout my life, to all my brothers and sisters, specially my brother: Abukar Mohamed Ali for his generous contributions to my life.

My special thanks go to my wife Hamdi Mohamud Osoble and my children who were a source of happiness and motivation through all ups and downs.
I extend my thanks to the management of SIMAD University for offering me a scholarship for my PhD.

Lastly, I would like to thank all the people who helped me in one way or another to finish my PhD thesis.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

**Yusmadi Yah Jusoh, PhD**
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

**Rusli Abdullah, PhD**
Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

**Rozi Nor Haizan Nor, PhD**
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

**Lilly Suriani Affendey, PhD**
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

_____
**ZALILAH MOHD SHARIFF, PhD**
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

**Declaration by graduate student**

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software

Signature: _____ Date:_____

Name and Matric No: Ismail Mohamed Ali, GS47634

**Declaration by Members of Supervisory Committee**

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: _____
Name of Chairman of Supervisory Committee: Associate Professor Dr. Yusmadi Yah Jusoh

Signature: _____
Name of Member of Supervisory Committee: Professor Dr. Rusli Abdullah

Signature: _____
Name of Member of Supervisory Committee: Dr. Rozi Nor Haizan Nor

Signature: _____
Name of Member of Supervisory Committee: Associate Professor Dr. Lilly Suriani Affendey

**TABLE OF CONTENTS**

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF APPENDICES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| ASP | Active Server Pages |
| AVE | Average Variance Extracted |
| BD | Big Data |
| BDA | Big Data Analytics |
| BI | Business Intelligence |
| COMP | Competency |
| CR | Composite Reliability |
| CRISP-DM | Cross Industry Standard for Data Mining |
| EFFEC | Effectiveness |
| EFFIC | Efficiency |
| ETL | Extract, Transform, and Load |
| IS | Information System |
| KDD | Knowledge Discovery in Database |
| SPA | Store, Process, and Access |
| SPO | Structure-Process-Outcome |
| TECH | Technology |
| UI | User Interface |
| WC | Working Conditions |
| FS | Feedback Score |
| AVG | Average |
| BDAPMM | BDA Process Performance Measurement Model |

# CHAPTER 1

# INTRODUCTION

## 1.1 Background

Big data is an information asset with a high volume, velocity and variety as defining characteristics, as well as specific technology and analytical methods used for harnessing such information and transforming it into economic value that impacts companies and societies. As its recognition grows, more researchers are directing their attention to big data, not only from a technical perspective, but also, socio-technical aspects which bring people, process, and technology into play. In information system (IS), the broadened scope of big data research includes big data analytics (BDA), big data infrastructure, and transformational impact (Goes, 2014). BDA, as one of the IS research directions, involves the data, the tools and techniques for data processing and analytics, and most importantly, the process, termed as BDA process, which connects all things together.

BDA process produces the knowledge and the insights that businesses need. The efforts to improve and optimize this process have arguably a sound justification. Especially, the efforts that enlighten the specific skill sets to perform the BDA process related activities, technology to enhance the BDA process execution, supportive work environment, and performance measurement ways for spotting performance gains and gaps.

The BDA process is not free from challenges, including those related to how to capture, integrate, transform, and analyse data, and convey the results (Sivarajah et al., 2017). Also, heterogeneity, lack of structure, error-handling, privacy, timeliness, provenance, and visualization all exist through the BDA process from data acquisition to interpretation (Alguliyev et al., 2017), and hence affect its performance.

Performance evaluation is also a perspective from which to look at big data. According to Veiga et al. (2018), evaluating the performance of BDA systems is the usual way of getting information about the expected execution time of analytics applications. The challenges lie in how to evaluate the performance of these applications and determine the factors that affect their quality (Villalpando et al., 2014). The challenges of performance evaluation are echoed by the benefits of performance evaluation including understanding the sources of performance degradation and discovering improvement opportunities. Having mentioned performance concept, the other thing is to know how it can be conceptualized into BDA settings.

Performance measurement is the process of quantifying the efficiency and effectiveness of action (Neely et al., 1995). In a BDA systems' perspective, the performance consists of the front-end and back-end performance properties (Liu, 2014). Back-end performance is related to the performance of the BDA system's functions, where front-end performance focuses on user experience and their satisfaction with the results of analytics. This suggests the need for considering both efficiency and effectiveness in measuring the performance of BDA systems.

Efficiency is related to the system's availability and performance over effort, whereby effectiveness is concerned with the impact of information on assisting users to perform their work (Heo & Haan, 2000). Therefore, resource utilization, time related metrics and capacity (Villalpando et al, 2014; Brunnert et al., 2014), as well as throughput, response time, latency (Onyeabor & Ta'a, 2018) can be regarded as performance measures in BDA systems, whereas measures such as satisfaction, timeliness, usefulness and result representation are used to scrutinize the system's success from a user's perspective. It means individual and organizational objectives are attained along with the system's objectives.

A BDA system, being viewed as a process, encompasses two perspectives: firstly, a system's perspective where data are being acquired, pre-processed, integrated, and analysed, and secondly, a user perspective, where results are presented and interpreted into the business context. These two ends are where the performance of BDA should be observed. In addition, performance is not always assumed as a standalone entity. It is determined by the capability of the system to which it belongs. In big data, there are a number of factors that enhance big data analytical capability, among them are human capital (both technical skills and managerial skills) and the technology that handles the volume and speed of big data (Mikalef et al., 2017). Lastly, this research strives putting the above concepts all together meaningfully, does it tackle issues in big data is an important question ahead.

## 1.2 Problem Statement

Big data is the input handled and processed through the BDA process and the output of this process will be the extracted knowledge which is transformed into a variety of business benefits. Throughout this process, a number of challenges exist, whether they are related to capturing, integrating, transforming, and analysing data, or whether they pertain to conveying the results to users (Sivarajah et al., 2017), and leveraging the generated insights and knowledge into real-life applications. Developers and data scientists deal with these challenges applying mathematical and statistical methods as well as analytical tools to BDA settings. Despite such a developer-centred approach, a major need arises for organizations to know the performance of their BDA systems and how these systems contribute to their business. Performance measures that cater for the performance of the individual BDA process phases are important considerations.

The existing literature highlights performance measurement in big data. For example, software product quality concepts such as performance efficiency and reliability are associated with performance measurement of big data applications (Villalpando et al., 2014). Performance efficiency concepts suit the efficiency of BDA process, where reliability can be roughly categorized under the effectiveness of BDA process.

Studies also examine the performance comparison among existing frameworks such as Hadoop, Spark, and Flink (Veiga et al., 2016; 2018). However, end-users rarely gain benefits from the research findings provided by these studies. Hence, the work remains to be done with respect to the development of tools that can provide more useful insights to users (Veiga et al., 2018). In this regard, performance insights can benefit most when users can act on them to make decisions. It means performance is measured and performance gaps and gains are revealed, then the results are exploited in performance improvement initiatives.

Quality assessment for both big data and process quality is another aspect featured in the current literature of big data (Onyeabor & Ta'a, 2018). According to Janssen et al. (2017), the quality of big data is not merely dependent on data, but also on the process in which the data are collected and the way they are processed. Assessing the quality of the process handling big data covers data processing and analysis phases (Serhani et al., 2016), but needs to consider the way big data is acquired and how it is interpreted and visualized to the users. This conforms to the existing literature which indicates that BDA performance should incorporate both back-end performance for BDA system's functions, and front-end performance for user experience and user satisfaction (Liu, 2014). Factors, like human skills needed for performing the work, and the technology required for handling the size, the variety, and the speed of data accumulated by organizations (Mikalef et al., 2017; Gupta & George, 2016), have paramount importance to the capability of BDA process. Therefore, this research addresses the following issues:

- The key issue lies in identifying performance measures that measure not only the functions of BDA systems, but also determine the user satisfaction with the BDA systems.
- For achieving durable performance, the need also arises to understand the factors that contribute to the performance of the BDA process. This is the vital and the largely neglected side of the BDA process because measuring performance can be significant when performance can be improved.

## 1.3 Research Questions

1. What are the measures for measuring the performance of BDA process?
2. What are the factors that contribute to the performance of BDA process?
3. How to measure the performance of BDA process?

## 1.4 Research Objectives

This research is aimed at achieving the following objectives:

1. To propose a performance measurement model for BDA process will be the main objective of this research. In order to achieve the main objective, the following subobjectives are derived:
2. To identify the local measures for measuring the performance of the BDA process in terms of BDA system's functions and user satisfaction.
3. To identify the global measures that contribute to the performance of the BDA process.

## 1.5 Research Contribution

The contribution of this research is considered from both theoretical and practical perspectives. The primary theoretical contribution lies in the BDA process performance measurement model being proposed. The model depicts inter-relationships among the participating constructs. Similarly, performance measures, metrics and performance-contributing factors being elaborated in this thesis will be invaluable contributions strengthened by findings of empirical findings and judgements of subject-matter experts. The theoretical contribution will hopefully be an invaluable contribution to the body of knowledge of big data and will represent a precious input to future big data research endeavours.

The practical contribution of this research lies in the implementation of the proposed model on measuring the performance of the BDA process. This will help organizations understand the performance of their big data analytics systems.

Holistically, the research is expected to contribute to the success of big data analytics projects, thereby realizing the promise of big data in improving decision making, optimizing business processes, and creating new business models. Achieving this will undoubtedly increase business success and contribute to the digital economy at large.

## 1.6 Research Scope

This study examines the performance of the BDA process. The performance measures will be investigated. Similarly, the study will investigate performance- contributing factors for the BDA process. Based on the findings of the literature to be reviewed, a conceptual model will be developed. The model will be validated through a confirmation study. The study is only committed to measuring performance of the BDA process. In the data collection process, the BDA practitioners will be our target respondents.

## 1.7 Organization of the Thesis

This thesis is organized as follows:

Chapter 1 provides the introduction and the background of the study. It describes the research problem, and the objectives of the research. The research scope and contribution are also explained in this chapter.

Chapter 2 is the literature review that provides a review and discussions of existing works relevant to this research. The information in this chapter is sourced from journals, conference proceedings, and books, as the main references.

Chapter 3 discusses the research methodology followed in this study. The research methodology consists of five phases, namely reviewing the existing literature, conducting confirmation study, proposing a model, developing and testing a prototype, and verifying the model with experts.

Chapter 4 describes the confirmation study which presents descriptive statistics of the survey results, the measurement model and the structural model of the study, and discussion on research findings.

Chapter 5 elaborates the proposed model, describes its constructs and indicators, and illustrates their relationships. It also maps the model's constructs and their indicators to BDA process, and explains the rationale of how they work together. The chapter also elaborates formulas, rating scales, and the implementation framework for the proposed model.

Chapter 6 presents results and discussions based on the prototype evaluation. The prototype was evaluated using the feedback from subject-matter experts.

Chapter 7 concludes the research, highlights the research contributions and the limitations, and describes directions for future work.

## 1.8 Summary

This chapter presented the introductory information about this research, including problem statement, objectives and scope. It also covered the potential contributions that the research makes theoretically and practically. Finally, the organization of this thesis was explained.

## REFERENCES


Adrian, C., Abdullah, R., Atan, R., & Jusoh, Y. Y. (2017, July). Factors influencing to the implementation success of big data analytics: A systematic literature review. In *Research and Innovation in Information Systems (ICRIIS), 2017 International Conference on* (pp. 1-6). IEEE.

Agrawal D, Bernstein P, Bertino E et al (2011). Challenges and opportunities with big data 2011-1. Cyber Center Technical Reports, (Paper 1). Retrieved from https://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1000&context=ccte ch

Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment?. *International Journal of Production Economics*, *182*, 113-131.

Alguliyev, R. M., Gasimova, R. T., & Abbasli, R. N. (2017). The obstacles in big data process. *International Journal of Modern Education and Computer Science*, *9*(3), 28-35.

Alter, S. (2013). Work system theory: overview of core concepts, extensions, and challenges for the future. *Journal of the Association for Information Systems*, *14*(2), 72-121.

Adrian C, Sidi F, Abdullah R, Ishak I, Affendey LS, Jabar M. A. (2016). Big data analytics implementation for value discovery: A systematic literature review. *Journal of Theoretical & Applied Iinformation Technology*, *93*(2), 385-393.

Amaravadi, C. S., & Lee, I. (2005). The dimensions of process knowledge. *Knowledge and Process Management*, *12*(1), 65-76.

Assunção, M. D., Calheiros, R. N., Bianchi, S., Netto, M. A., & Buyya, R. (2015). Big Data computing and clouds: Trends and future directions. *Journal of Parallel and Distributed Computing*, *79*, 3-15.

Bally, L., Brittan, J., & Wagner, K. H. (1977). A prototype approach to information system design and development. *Information & Management*, *1*(1), 21-26.

Bayrak, E. A., & Kirci, P. (2019). Intelligent big data analytics in health. In *Early Detection of Neurological Disorders Using Machine Learning Systems* (pp. 252-291). IGI Global.

Benitez-Amado, J., Henseler, J., & Castillo, A. (2017). Development and update of guidelines to perform and report partial least squares path modeling in Information Systems research. *Proceedings of the 21st Pacific Asia Conference on Information Systems, Langkawi, Malaysia*, pp. 1-15.

Beynon-Davies, P., Tudhope, D., & Mackay, H. (1999). Information systems prototyping in practice. Journal of Information Technology, 14(1), 107-120.

Biuk-Aghai, R. P., Kou, W. T., & Fong, S. (2016, May). Big data analytics for transportation: Problems and prospects for its application in China. In *2016 IEEE Region 10 Symposium (TENSYMP)* (pp. 173-178). IEEE.

Blasini, J., & Leist, S. (2013). Success factors in process performance management. *Business Process Management Journal*, *19*(3), 477-495.

Blomkvist, J., & Holmlid, S. (2011). Existing prototyping perspectives: considerations for service design. In *4th Nordic Design Research Conference (NorDes 2011), 29-31 May, Helsinki, Finland* (pp. 31-40). School of Art & Design, Aalto University.

Boncea, R., Petre, I., Smada, D. M., & Zamfiroiu Anagrama, A. (2017). A maturity analysis of big data technologies. *Informatica Economica*, *21*(1), 60-71.

Boudreau, M. C., Gefen, D., & Straub, D. W. (2001). Validation in information systems research: a state-of-the-art assessment. *MIS quarterly*, 1-16.

Bravo, E. R., Santana, M., & Rodon, J. (2015). Information systems and performance: the role of technology, the task and the individual. *Behaviour & Information Technology*, *34*(3), 247-260.

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4-7.

Browne, J., Devlin, J., Rolstadas, A., & Andersen, B. (1997). Performance measurement: the ENAPS approach. *International Journal of Business Transformation*, *1*, 73-84.

Brunnert, A., Vögele, C., Danciu, A., Pfaff, M., Mayer, M., & Krcmar, H. (2014). Performance management work. *Wirtschaftsinformatik*, *56*(3), 197-199.

Chandio, A. A., Tziritas, N., & Xu, C. Z. (2015). Big data processing techniques and their challenges in transport domain. *ZTE Communications*, (1), 10

Chen, H. M., Kazman, R., & Haziyev, S. (2016, January). Agile big data analytics development: An architecture-centric approach. In *2016 49th Hawaii International Conference on System Sciences (HICSS)* (pp. 5378-5387). IEEE.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile networks and applications*, *19*(2), 171-209.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. *Modern methods for business research*, *295*(2), 295-336.

Cohen, J. 1988. Statistical power analysis for behavioral sciences (2nd Edition), Erlbaum, Hillsdale, USA.

Cornford, T., Doukidis, G. I., & Forster, D. (1994). Experience with a structure, process and outcome framework for evaluating an information system. *Omega*, *22*(5), 491-504.

Chua, Y. P. (2012). *Mastering research methods*. Mcgraw-Hill Education.

Daniel, B. (2015). B ig D ata and analytics in higher education: Opportunities and challenges. *British journal of educational technology*, *46*(5), 904-920.

Davamanirajan, P., Kauffman, R. J., Kriebel, C. H., & Mukhopadhyay, T. (2006, January). Systems design, process performance and economic outcomes. In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)* (Vol. 8, pp. 168c-168c). IEEE.

Davenport, T. H., Barth, P., & Bean, R. (2012). How Big Data Is Different. *MIT Sloan Management Review*, *54*(1), 43.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.

De Mauro, A., Greco, M., & Grimaldi, M. (2015, February). What is big data? A consensual definition and a review of key research topics. In *AIP conference proceedings* (Vol. 1644, No. 1, pp. 97-104). AIP.

Debortoli, S., Müller, O., & vom Brocke, J. (2014). Comparing business intelligence and big data skills. *Business & Information Systems Engineering*, *6*(5), 289-300.

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, *3*(1), 60–95.

Delone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: a ten-year update. *Journal of Management Information Systems*, *19*(4), 9–30.

DeLone, W. H., & McLean, E. R. (2016). Information systems success measurement. *Foundations and Trends® in Information Systems*, *2*(1), 1-116.

Demchenko, Y., Gruengard, E., & Klous, S. (2014, December). Instructional model for building effective big Data curricula for online and campus education. In *2014 IEEE 6th International Conference on Cloud Computing Technology and Science*(pp. 935-941). IEEE.

Diouf, P. S., Boly, A., & Ndiaye, S. (2017, November). Performance of the ETL processes in terms of volume and velocity in the cloud: State of the art. In *2017 4th IEEE International Conference on Engineering Technologies and Applied Sciences (ICETAS)* (pp. 1-5). IEEE.

Dumas, M., van der Aalst, W., & Ter Hofstede, A. (Eds.). (2005). *Process aware information systems* (Vol. 1). Chichester: Wiley.

Elgendy, N., & Elragal, A. (2014, July). Big data analytics: a literature review paper. In *Industrial Conference on Data Mining* (pp. 214-227). Springer, Cham.

Elragal, A., & Klischewski, R. (2017). Theory-driven or process-driven prediction? Epistemological challenges of big data analytics. *Journal of Big Data*, *4*(1), 19.

Fitzgerald, L., Johnston, R., Brignall, T. J., Silvestro, R., & Voss, C. (1991). *Performance measurement in service businesses* (Vol. 69). London: Chartered Institute of Management Accountants, 6-31.

Furht, B., & Villanustre, F. (2016). Introduction to big data. In *Big data technologies and applications* (pp. 3-11). Springer, Cham.

Gachet, A., & Haettenschwiler, P. (2003, September). Developing intelligent decision support systems: A bipartite approach. In *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems* (pp. 87-93). Springer, Berlin, Heidelberg.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137-144.

Gartner IT Glossary - Big Data. Retrieved Jun 17, 2019, from http://www.gartner.com/it-glossary/big-data

Gheni, A. Y., Jusoh, Y. Y., Jabar, M. A., & Ali, N. M. (2017). Measuring the Global Virtual Teams (GVTs) Performance: Confirmation Study. *Indian Journal of Science and Technology*, *10* (8), 1-16.

Glavic, B. (2012). Big data provenance: Challenges and implications for benchmarking. In *Specifying big data benchmarks* (pp. 72-80). Springer, Berlin, Heidelberg.

Goes, P. B. (2014). Editor's comments: big data and IS research. *Mis Quarterly*, *38*(3), iii–viii.

Górny, A. (2017). The use of working environment factors as criteria in assessing the capacity to carry out processes. In *MATEC Web of Conferences* (Vol. 94, p. 06011). EDP Sciences.

Graziano, A. M., & Raulin, M. L. (2006). Research Methods: A Process of Inquiry. 6. Mass.

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, *53*(8), 1049-1064.

Hair, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2014). A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM).

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing theory and Practice*, *19*(2), 139-152.

Hamilton, S., & Chervany, N. L. (1981). Evaluating information system effectiveness-Part I: Comparing evaluation approaches. *MIS quarterly*, 55-69.

Harrati, N., Bouchrika, I., Tari, A., & Ladjailia, A. (2016). Exploring user satisfaction for e-learning systems via usage-based metrics and system usability scale analysis. *Computers in Human Behavior*, *61*, 463-471

Heckl, D., & Moormann, J. (2010). Process performance management. In *Handbook on business process management 2* (pp. 115-135). Springer, Berlin, Heidelberg.

Heine, M. L., Grover, V., & Malhotra, M. K. (2003). The relationship between technology and performance: a meta-analysis of technology models. *Omega*, *31*(3), 189-204.

Heo, Jaeho and Haan, Ingoo, "Measuring performance of information systems in evolving computing environments: an empirical investigation" (2000). PACIS 2000 Proceedings. Paper 25. http://aisel.aisnet.org/pacis2000/25, 316-332.

Hoffmann, T. (1999). The meanings of competency. *Journal of European Industrial Training*, *23*(6), 275-286.

Hu, H., Wen, Y., Chua, T. S., & Li, X. (2014). Toward scalable systems for big data analytics: A technology tutorial. *IEEE Access*, *2*, 652-687.

Hussain, M., Al-Mourad, M. B., & Mathew, S. S. (2016, March). Collect, Scope, and Verify Big Data--A Framework for Institution Accreditation. In *2016 30th International Conference on Advanced Information Networking and Applications Workshops (WAINA)* (pp. 187-192). IEEE.

Iannacci, F., & Cornford, T. (2018). Unravelling causal and temporal influences underpinning monitoring systems success: A typological approach. *Information Systems Journal*, *28*(2), 384-407.

Ives, B., Olson, M., & Baroudi, J. J. (1983). The measurement of user information satisfaction. *Information Systems Working Papers Series, Vol*. 26, 785-793

Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., & Shahabi, C. (2014). Big data and its technical challenges. *Communications of the ACM*, *57*(7), 86-94.

Jaklic, J., Coelho, P. S., & Popovic, A. (2009). Information quality improvement as a measure of business intelligence system benefits. *WSEAS Transactions on Business and Economics*, *9*, 56-62.

Jackson, S. L. (2015). *Research methods and statistics: A critical thinking approach*. Cengage Learning.

Jansen-Vullers, M. H., Loosschilder, M. W. N. C., Kleingeld, P. A. M., & Reijers, H. A. (2007). Performance measures to evaluate the impact of best practices. In *Proceedings of Workshops and Doctoral Consortium of the 19th International Conference on Advanced Information Systems Engineering (BPMDS workshop)* (Vol. 1, pp. 359-368). Trondheim: Tapir Academic Press.

Janssen, M., van der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. *Journal of Business Research*, *70*, 338-345.

Jusoh, Y. Y., Hamdan, A. R., & Deraman, A. (2008). A case study on evaluating project achievement of ITSP implementation. *IJCSNS*, *8*(7), 54.

Kahrović, E. (2013). Design and implementation of business process performance measurement system. *Facta Universitatis–Series: Economics and Organization*, *10*(3), 283-299.

Kaur, A., Grover, P. S., & Dixit, A. (2019). Performance efficiency assessment for software systems. In *Software Engineering* (pp. 83-92). Springer, Singapore.

Kock, N., & Lynn, G. S. (2012). Lateral collinearity and misleading results in variance-based sem: an illustration and recommendations. *Journal of the Association for Information Systems*, *13*(7), 546-580.

Koronios, A., Gao, J., & Selle, S. (2014). Big data project success-A meta analysis. In *PACIS* (p. 376).

Kune, R., Konugurthi, P. K., Agarwal, A., Chillarige, R. R., & Buyya, R. (2016). The anatomy of big data computing. *Software: Practice and Experience*, *46*(1), 79-105.

Labrinidis, A., & Jagadish, H. V. (2012). Challenges and opportunities with big data. *Proceedings of the VLDB Endowment*, *5*(12), 2032-2033.

Lai, S. T., & Leu, F. Y. (2017, November). A critical quality measurement model for managing and controlling big data project risks. In *International Conference on Broadband and Wireless Computing, Communication and Applications* (pp. 777-787). Springer, Cham.

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META group research note*, *6*(70), 1.

Larson, D., & Chang, V. (2016). A review and future direction of agile, business intelligence, analytics and data science. *International Journal of Information Management*, *36*(5), 700-710.

Latham, G. P., & Pinder, C. C. (2005). Work motivation theory and research at the dawn of the twenty-first century. *Annu. Rev. Psychol.*, *56*, 485-516.

Lazaroiu, G. (2015). Employee motivation and job performance. *Linguistic and Philosophical Investigations, 14*, 97.

Lăzăroiu, G. (2015). Work motivation and organizational behavior. *Contemporary Readings in Law and Social Justice*, *7*(2), 66-75.

Leyer, M., Heckl, D., & Moormann, J. (2015). Process performance measurement. In *Handbook on Business Process Management 2* (pp. 227-241). Springer, Berlin, Heidelberg.

Liebowitz, J. (2017). Thoughts on recent trends and future research perspectives in big data and analytics in higher education. In *Big data and learning analytics in higher education* (pp. 7-17). Springer, Cham.

Liu, Z. (2014, July). Research of performance test technology for big data applications. In *2014 IEEE International Conference on Information and Automation (ICIA)* (pp. 53-58). IEEE.

Lněnička, M. (2015). AHP model for the big data analytics platform selection. *Acta Informatica Pragensia*, *4*(2), 108-121.

Longo, A., & Motta, G. (2005, September). Design processes for sustainable performances: a model and a method. In *International Conference on Business Process Management* (pp. 399-407). Springer, Berlin, Heidelberg.

M., S. H. K. (2011). Standard deviation BT - International encyclopedia of statistical science. In M. Lovric (Ed.) (pp. 1378–1379). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-04898-2_535.

Merino, J., Caballero, I., Rivas, B., Serrano, M., & Piattini, M. (2016). A data quality in use model for big data. *Future Generation Computer Systems*, *63*, 123-130.

Mikalef, P., & Krogstie, J. (2019, April). Investigating the data science skill gap: an empirical analysis. In *2019 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1275-1284). IEEE.

Mikalef, P., Framnes, V. A., Danielsen, F., Krogstie, J., & Olsen, D. H. (2017). Big data analytics capability: antecedents and business value. In *Proceedings of the 21st Pacific Asia conference on information systems (PACIS)*.

Mikalef, P., Giannakos, M. N., Pappas, I. O., & Krogstie, J. (2018b, April). The human side of big data: Understanding the skills of the data scientist in education and industry. In *Global Engineering Education Conference (EDUCON), 2018 IEEE*(pp. 503-512). IEEE.

Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018a). Big data analytics capabilities: a systematic literature review and research agenda. *Information Systems and e-Business Management*, *16*(3), 547-578.

Miller, H. G., & Mork, P. (2013). From data to decisions: a value chain for big data. *It Professional*, (1), 57-59.

Mneney, J., & Van Belle, J. P. (2016, January). Big data capabilities and readiness of South African retail organisations. In *2016 6th International Conference-Cloud System and Big Data Engineering (Confluence)* (pp. 279-286). IEEE.

Monteiro, L. F. S., & de Oliveira, K. M. (2011). Defining a catalog of indicators to support process performance analysis. *Journal of Software Maintenance and Evolution: Research and Practice*, *23*(6), 395-422.

Morabito, V. (2015). Big data and analytics. *Strategic and organisational impacts*.

Neely, A., Gregory, M., & Platts, K. (1995). Performance measurement system design: a literature review and research agenda. *International journal of operations & production management*, *15*(4), 80-116.

Nunnally, J. C. (1978). Psychometric theory (2nd edit.) mcgraw-hill. *Hillsdale, NJ*, *416*.

Oliver, P. (2006). Purposive Sampling. *SAGE Publications.* Retrieved from http://methods.sagepub.com/reference/the-sage-dictionary-of-social-research-methods/n162.xml

Onyeabor, G. A., & Ta'a, A. (2018, June). A model for addressing quality issues in big data. In *International Conference of Reliable Information and Communication Technology* (pp. 65-73). Springer, Cham.

Oussous, A., Benjelloun, F. Z., Lahcen, A. A., & Belfkih, S. (2018). Big Data technologies: A survey. *Journal of King Saud University-Computer and Information Sciences*, *30*(4), 431-448.

Pak, H. D., Park, S. M., Jang, I. H., & Choe, Y. C. (2010). A measurement of non-linear relationship between is qualities and user satisfaction. In *AMCIS* (p. 504).

Parry, S. B. (1996). Just What Is a Competency?(And Why Should You Care?). *Training*, *35*(6), 58.

Pattanshetti, T., & Attar, V. (2017, January). Survey of performance modeling of big data applications. In *Cloud Computing, Data Science & Engineering-Confluence, 2017 7th International Conference on* (pp. 177-181). IEEE.

"Performance efficiency". Retrieved May 15, 2019 from: https://ISO25000.com/index.php/en/iso-25000-standards/iso-25010/59-performance-efficiency.

Petter, S., DeLone, W., & McLean, E. (2008). Measuring information systems success: models, dimensions, measures, and interrelationships. *European journal of information systems*, *17*(3), 236-263.

Pinsonneault, A., & Kraemer, K. (1993). Survey research methodology in management information systems: an assessment. *Journal of management information systems*, *10*(2), 75-105.

Pitt, L. F., Watson, R. T., & Kavan, C. B. (1995). Service quality: a measure of information systems effectiveness. *MIS quarterly*, 173-187.

Ponsard, C., Touzani, M., & Majchrowski, A. (2017). Combining process guidance and industrial feedback for successfully deploying big data projects. *Open Journal of Big Data (OJBD)*, *3*(1), 26-41.

"Process effectiveness and efficiency". Retrieved May 10, 2019 from: http://isoconsultantpune.com/iatf-169492016-process-effectiveness-and-efficiency

Ravichandran, T., & Lertwongsatien, C. (2005). Effect of information systems resources and capabilities on firm performance: a resource-based perspective. *Journal of Management Information Systems*, *21*(4), 237-276.

Rehman, M. H., Chang, V., Batool, A., & Wah, T. Y. (2016). Big data reduction framework for value creation in sustainable enterprises. *International Journal of Information Management*, *36*(6), 917-928.

"Resource utilization". Retrieved Jun 13, 2019 from: https://iso25000.com/index.php/en/iso-25000-standards/iso-25010

Rouse, W. B., Edwards, S. L., & Hammer, J. M. (1993). Modeling the dynamics of mental workload and human performance in complex systems. *IEEE transactions on systems, man, and cybernetics*, *23*(6), 1662-1671.

Rowley, J. (2014). Designing and using research questionnaires. *Management Research Review*, *37*(3), 308-330.

Russom, P. (2011). Big data analytics. *TDWI best practices report, fourth quarter*, *19*(4), 1-34.

Saltz, J. S. (2015, October). The need for new processes, methodologies and tools to support big data teams and improve big data project effectiveness. In *2015 IEEE International Conference on Big Data (Big Data)* (pp. 2066-2071). IEEE.

Serhani, M. A., El Kassabi, H. T., Taleb, I., & Nujum, A. (2016, June). An hybrid approach to quality evaluation across big data value chain. In *Big Data (BigData Congress), 2016 IEEE International Congress on* (pp. 418-425). IEEE.

Sheng, J., Amankwah-Amoah, J., & Wang, X. (2018). Technology in the 21st century: New challenges and opportunities. *Technological Forecasting and Social Change*.

Shull, F., Singer, J., & Sjøberg, D. I. (Eds.). (2007). *Guide to advanced empirical software engineering*. Springer Science & Business Media.

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of big data challenges and analytical methods. *Journal of Business Research*, *70*, 263-286.

Sommerville, I. (2011). Software engineering 9th Edition. *ISBN-10, 137035152*.

Spitzer, M., Nanic, I., & Ebner, M. (2018). Distance learning and assistance using smart glasses. *Education Sciences*, *8*(1), 21.

Statz, J. (2005). Measurement for process improvement. *Practical Software and Systems Measurement*, 1-17.

Straub, D., Boudreau, M. C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *Communications of the Association for Information systems*, *13*(1), 24.

Syamil, A., Doll, W. J., & Apigian, C. H. (2004). Process performance in product development: measures and impacts. *European Journal of Innovation Management*, *7*(3), 205-217.

Taleb, I., Serhani, M. A., & Dssouli, R. (2018, July). Big data quality: A survey. In *2018 IEEE International Congress on Big Data (BigData Congress)* (pp. 166-173). IEEE.

Tan, W., Shen, W., Xu, L., Zhou, B., & Li, L. (2008). A business process intelligence system for enterprise process performance management. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 38(6), 745-756.

Tangen, S. (2004). Performance measurement: from philosophy to practice. *International journal of productivity and performance management*, *53*(8), 726-737.

Thong, J. Y., & Yap, C. S. (1996). Information systems effectiveness: A user satisfaction approach. *Information Processing & Management*, *32*(5), 601-610.

"Throughput". Retrieved Jun 15, 2019 from: http://www.linfo.org/throughput.html

Tulasi, B. (2013). Significance of big data and analytics in higher education. *International Journal of Computer Applications*, *68*(14), 21-23.

"Understandability". Retrieved April 15, 2019 from: https://iso25000.com/index.php/en/iso-25000-standards/iso-25012/122-understandability.

Urbach, N., & Müller, B. (2012). The updated DeLone and McLean model of information systems success. In *Information systems theory* (pp. 1-18). Springer, New York, NY.

Urbach, N., & Ahlemann, F. (2010). Structural equation modeling in information systems research using partial least squares. *Journal of Information technology theory and application*, *11*(2), 5-40.

Veiga, J., Expósito, R. R., & Touriño, J. (2018). Performance evaluation of big data analysis. *Encyclopedia of Big Data Technologies*, 1-6.

Veiga, J., Expósito, R. R., Pardo, X. C., Taboada, G. L., & Tourifio, J. (2016, December). Performance evaluation of big data frameworks for large-scale data analytics. In *2016 IEEE International Conference on Big Data (Big Data)* (pp. 424-431). IEEE.

Villalpando, L. E. B., April, A., & Abran, A. (2014). Performance analysis model for big data applications in cloud computing. *Journal of Cloud Computing*, *3*(1), 19.

Vossen, G. (2014). Big data as the new enabler in business and other intelligence. *Vietnam Journal of Computer Science*, *1*(1), 3-14.

Wade, M., & Hulland, J. (2004). The resource-based view and information systems research: Review, extension, and suggestions for future research. *MIS quarterly*, *28*(1), 107-142.

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356-365.

Wang, J., Crawl, D., Purawat, S., Nguyen, M., & Altintas, I. (2015, October). Big data provenance: Challenges, state of the art and opportunities. In *2015 IEEE International Conference on Big Data (Big Data)* (pp. 2509-2516). IEEE.

Watson, H. J. (2014). Tutorial: Big data analytics: Concepts, technologies, and applications. *CAIS*, *34*, 65.

Wirth, R., & Hipp, J. (2000, April). CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining* (pp. 29-39). Citeseer.

Wong, K. K. K. (2013). Partial least squares structural equation modeling (PLS-SEM) techniques using SmartPLS. *Marketing Bulletin*, *24*(1), 1-32.