Determinants of Big Data Adoption: An Empirical Study on Sri Lankan Firms

Sachini Nidusha Weerasinghe
Department of Management of Technology
Faculty of Business
University of Moratuwa, 10400
Sri Lanka
sachiniweerasinghe5@gmail.com

Nihal Jayamaha
Department of Operations & Engineering Innovations
School of Food and Advanced Technology
Massey University
Private Bag 11222, Palmerston North 4442
New Zealand
n.p.jayamaha@massey.ac.nz

Abstract

Big Data (BD) is a vital cog in Industry 4.0 technology that modern organizations adopt to gain competitive advantage. Although the factors that lead BD usage are documented in the literature, there is lack of understanding as to how these factors are related to one another in predicting and explaining BD usage. This study fills the gap by developing and empirically testing a theoretical model, using the Sri Lankan industry as the context. Technical, Organizational and Environmental (TOE) Framework is used as the theoretical lens for theory building. The seven constructs of the theoretical model were identified by sifting through a large volume of literature to identify the critical success factors (CSFs) of BD usage. The relationships between the constructs were also hypothesized from the extant literature. The model posited that Top Management Support drives the organization's Analytics Culture and Information Systems Competence to enhance its Analytics Capacity, to cause BD usage. In addition to Analytics Capability (the internal factor), the study posits that Competitive Pressure and Government Pressure also act as additional causal antecedents of BD usage. A questionnaire was developed to capture different manifestations of the seven constructs and partial least squares structural equation modelling (PLS-SEM) technique was used to analyze the data collected. In the main, the data supported the hypotheses underpinning the theory. Analytics Capability and Competitive Pressure found to have strong and medium effects on BD. The only exception was that the Government Pressure was found to have no effect on BD usage (p=0.332). Having demonstrated construct validity of all seven constructs, the study estimated the strengths of the theoretical relationships between the constructs; the latter also enabled the researcher to provide suggestions to top managers as to how BD usage can be increased in their firms. To the best of our knowledge, this study is the first to formulate a theory that explains how different organizational elements work together to improve the BD usage. The questionnaire and the practical implications of the empirical findings will be useful to BD organizations for continuous improvement.

Keywords
Big Data, Big Data Adoption, Industry 4.0, Big Data Organizations, TOE Framework

1. Introduction
The digital transformation is characterised by exchange of a wide variety of data in large amounts between machines and users at high speeds. Such data, commonly known as Big Data (BD), is the most pivotal cog in Industry 4.0 technologies (Lasi et al. 2014; Saturno et al. 2017). In addition to "data and knowledge management", there are other

constructs that are required to achieve businesses success: leadership, strategy, customer focus, human resource focus and process management (Flynn and Saladin 2001). However, unlike these other constructs, the meaning of "data and knowledge management" has changed profoundly with the advent of Industry 4.0 technologies (Curry et al. 2014; Nguyen 2018). The managerial problem of this study, and hence the overarching research question is: how can modern industrial organisations optimize the use of big data to achieve business success? The managerial problem/research question is addressed taking Sri Lanka as the context.

Heavy loads of data in an organization need to be treated with BD analytical tools to take data-based decisions in the modern competitive context. Many large and small to medium enterprises (SMEs) use BD tools in their organizations to make decisions. Yet, most of them are not fully utilizing the enablers of BD adoption to create value the BD is supposed to deliver. Organizations invest huge amount of capital to align with industry 4.0 concept via strategic plans, but it is doubtful whether the return on investment on BD adoption to take real time decisions can be justified in some of the organizations. Therefore, one needs to examine what factors enable BD usage in organisations that have made a strategic decision to use such data.

BD organisations (BDOs) need business domain knowledge to identify business cases that can (potentially) be solved using BD. On the other hand, BDOs need systems (e.g., hardware and protocols) in place to acquire and store data, as well as design data analysis processes to solve problems (sense making) (Curry 2016; Rim and Sonia 2019; Wirén et al. 2019). No matter how well-equipped an organization is to make sense out of BD, there are numerous factors that can either drive or restrain BD adoption and/or usage success (Adrian et al. 2017; Côrte-Real et al. 2019; Lunde et al. 2019; Maroufkhani et al. 2020). This is because organizations do not operate in vacuums, and there are both internarial environment factors (e.g., the top management support and technology competence) and external environmental factors (e.g., competitive pressure, and the regulatory environment) that affect BD adoption to achieve successful business outcomes (Sun et al. 2020; Surabhi & Sushil, 2019).

This research, which is an empirical study involving Sri Lankan BD user organizations, focusses on BD usage, and the enablers that lead to BD usage to create value.

1.1 Objectives

The aim of this research is to develop a causal predictive model that predicts and explains BD usage through the enablers of BD adoption. These objectives are derived from the gaps in the literature. More importantly, the literature is rather thin on attempts to explain how the enablers of BD adoption are related to one another in predicting an explaining the actual usage of BD in BDOs. Another notable gap in the literature is not knowing the interplay between the enablers of BD adoption and the actual BD usage in Sri Lankan BDOs. Thus, the three specific objectives of the study are to:

Objective 1: Determine the enablers of BD usage in BDOs

Objective 2: Develop and test a causal-predictive model that predicts and explains BD usage through a system of interconnected enablers of BD adoption.

Objective 3: Interpret the empirical test results on BD usage, from a practitioner's standpoint

2. Literature Review

Within the phenomenon industry 4.0, BD plays a major role in industrial revolution, and it is a necessary component in a business to achieve the competitive advantage (Perova-Antova and Ilieva 2018). The amount of data continuously produced in internet of things (IoT) needs to be properly filtered and translated to usable formats in the day-to-day business activities to take operational and strategic business decisions. BD is the key to improve the efficiency and value creation of businesses; BD opens opportunities and facilities new functionalities (Perova-Antova and Ilieva 2018).

Any concept or an object can be defined based on its characteristic or its uses and deliverables. The same applies to BD. Researchers have defined it according to its nature and uses. Most of BD definitions are based on the characteristic of BD: volume, velocity, and variety. BD are generally characterized as high volume, high velocity and/or high variety information assets (Faroukhi et al. 2020; Jcobs 2009; Laney 2001). However, BD are sometimes defined using its deliverables such as problem solving, value extraction, enabling analyzing, processing features, and enabling visualizing (Dubey et al. 2019; Perova-Antova and Ilieva 2018).

The concept BD emerged to distinguish data in everyday use to data that is high in volume and variety (Curry 2016; Hu et al. 2014). From 1970s to 2011, the volume of data holding capacities have changed from megabytes to gigabytes to terabytes to petabytes to exabytes (Hu et al. 2014). BD can exist in different formats: texts, videos, audios, signals; the data can be structured, semi-structured or unstructured (Curry et al. 2014; Gaurav et al. 2018). Due to large volume and diverse variety of data being involved, researchers have found that BD has other characteristics built into it. These are (apart from high volume, high speed and high variety) veracity and value (Curry 2016; Nguyen 2018).

The purpose of BD, based on the literature, is to create value, increase efficiency in business operations and facilitating new functionalities in BD related processes (Perova-Antova and Ilieva 2018). Even though BD has different characteristics or uses, the popularity of BD among the BD users comes from its applications in day-to-day activities. The process of making sense out of large volumes of scattered data to create value is known as BD value creation process. The traditional view of value, or more technically precisely customer value, is the relative worth of an offering based on customer perception (Vargo et al. 2008). Thus, from a buyer-firm perspective, the firm creates value-laden goods and services (at competitive prices) and customer consumes this value (Vargo et al. 2008). The BD value chain (BDVC) is a concept derived from Michael Porter's Value Chain concept (1985) very recently, through the collective wisdom of many authors; a BDVC encompasses the processes involved in BD value creation within a BD life cycle, for organizational decision making (Curry 2016; Faroukhi et al. 2019).

The amount of data to be handled by an organization becomes more advanced and complex with the 4th industrial revolution since organizations attempt to gain the maximum benefit from the available technology to achieve competitive advantage (Côrte-Real et al. 2019; Curry 2016). Hence, decisions are not easy to derive or obtained by using traditional regression and classification approaches at highest quality. Obtaining values and making important business decisions from BD will benefit an organization in many ways like cost and risk reduction, improving and sustaining quality, increasing efficiency and enhance customer satisfaction. Therefore, the adoption of BD in organizations is practicing in many industries depending on their strategies and goals with evolving huge amount of data.

Successful implementation of BD in an organization would lead to seamless BD value creation that in turn leads to accurate BD based decisions (Curry et al. 2014; Elia et al. 2020). The success factors or the enablers of BD implementation and their effect needs to be known to increase BD value creation. Therefore, identifying the determinants or enablers of BD adoption in an organization is beneficial for the management to align BD based decision making to gain competitive advantage.

The critical success factors (CSFs) method, which gained momentum in the 1980s was originally meant to aid managers in strategic planning to understand which critical factors (from trivial may) they should focus on to achieve corporate success (Bullen and Rockart 1981). The CSF method has since been applied to management subdomains such as quality management, new product development, and BD analytics.

Most common generalisable finding in the selected articles from the literature is the three contexts being used to classify the BD enablers: the organizational context, the technological context, and the environmental context. Each context accommodates multiple CSFs. A three-step iterative process was used to funnel-down the critical of the critical factors under each context. The funnel approach (Bryman 2012) is quite common in literature synthesis. Three-step funnelling process was used to isolate the most critical of the critical factors under organizational, technological, and environmental context. The resulted BD adoption determinants under each context is shown in Table 1 with abbreviations used.

Table 1. Filtered BD Adoption Determinants in Different Contexts

Context	BD Adoption Enablers/Determinants	Abbreviation
Organizational	Top Management Support Analytics Culture	TMS ANC
Technological	IS Competence Analytics Capability	ISC AC

Environmental	Competitive Pressure Government Pressure	CPP GOP

The filtering of published studies has resulted in a conceptual framework that stands analogous to the Technology-Organization-Environment (TOE) framework advanced by Tornatzky and Fleischer (1990). The TOE framework is a general theory on technology adoption (technological innovation decision making) at firm level. In the broadest sense, the TOE framework argues that the organizational, technological, and environmental factors drive an organization to adopt (new) technology (Awa et al. 2017; Baker 2012). In our study, the technology concerted is the adoption of BD. Of the six constructs isolated as enablers of BD usage—Top Management Support, Analytics Culture, IS Competence, Analytics Capability, Competitive Pressure, and Government Pressure—the last two, namely Competitive Pressure and Government Pressure are external factors which do not have a direct effect on the remaining four factors, which are all internal factors. The four internal factors however are causally related to one another. For example, a Top Management Support change must exist to observe a change in the other three internal factors. Seven hypotheses were formulated using extant literature to build the theoretical model.

H1: Top Management Support has a positive effect on the Analytics Culture.

H2: Top Management Support has a positive effect on IS Competence.

H3: IS Competence has a positive effect on Analytics Capability.

H4: Analytics Culture has a positive effect on Analytics Capability.

H5: Analytics Capability has a positive effect on BD usage.

H6: Competitive Pressure has a positive effect on BD usage.

H7: Government Pressure has a positive effect on BD usage.

3. Methods

The minimum sample size guidelines of Cohen (1992), based on his power analysis, was used to determine the minimum sample size (n = 75) to test the regression paths. The data collected (data collection details in the next section) were screened for examining unusual observations, frequency distribution, response bias and common method bias (CMB). CMB is the bias (variance of the construct scores) occurring due to the way in which a survey has been administered, rather than due to the actual variation owning to the constructs; this is an issue that can sometimes be a problem in self-administered surveys (MacKenzie and Podsakoff 2012). The Harman's single factor test, which involves principal components analysis (PCA), was used to test for CMB (Chang et al. 2010; Podsakoff et al. 2003). The PCA was performed using Minitab 19 software.

The partial least squares structural equation modelling (PLS-SEM) technique was used to test the hypotheses using SmartPLS v3 software package (Ringle et al 2015). SmartPLS v3 has built in algorithms to test construct reliability (e.g., outputting the Cronbach's alpha coefficient of each construct) and validity ratios (e.g., the average variance extracted to test convergent validity and HTMT ratio of correlations to test discriminant validity) (Hair Jr et al. 2021; Ringle et al. 2015).

4. Data Collection

A questionnaire was designed to collect data on the seven constructs of the theoretical model. The questionnaire contained 22 Likert style, agreement-seeking statements derived from prior literature (a seven-point scale was used). Before distributing the survey questionnaire, it was pilot tested/screened for readability, consistency, and quality, using six respondents who were experts from different BD organizations and academia. Pilot testing did not flag any necessity to revise the questionnaire.

The data were collected online by inviting the respondents to participate in the study by responding to the questionnaire, which was administered using the Google Forms platform. Survey questionnaire was sent to 180 middle managers belonging to 70 BD organizations. It was ensured that a particular BD firm would not be over-represented. However, larger firms were found to have more BD projects in place, and therefore more BD users. Social media platforms were used to launch the survey, which resulted in 80 responses (this included the six responses received in the pilot study phase).

5. Results and Discussion

Demonstrating Absence of CMB, Scale Reliability and Validity

Four cases were found to contain unusual observations. These cases were deleted prior to theory testing, resulting in 76 usable cases. The PCA performed responses on the 22 survey items extracted five factors returning an Eigenvalue > 1.0 (Kaiser criterion). The five components retained extracted 75.8% of the variability of the 22 dimensions (indicators). More importantly, the first principal component extracted only 36.1% of the total variability, which less than the 40% upper bound value prescribed by Podsakoff et al. (2003) to demonstrate no apparent CMB, based on Harman's single factor test.

5.1 Numerical Results

In SmartPLS, construct reliability is measured using Cronbach's Alpha and Composite Reliability coefficients (Hair Jr et al., 2021). Based on guidelines issued by Nunnally (1978), for basic research, a reliability coefficient in excess of 0.7 is desired. The reliability coefficients reported in Table 2 clearly indicate that the scales easily pass the 0.70 lower-bound cut-off value for a reliable scale. In addition, Table 2, shows the average variance extracted (AVE) for each construct. The AVE of a construct indicates the variability a construct extracts from its indicators relative to measurement error (Fornell and Larcker 1981). An AVE > 0.50 indicates that there is a strong relation between the construct and its assigned indicators to demonstrate convergent validity (Fornell and Larcker 1981; Hair Jr et al. 2021). The AVE values shown in Table 2 clearly demonstrate that the construct operationalization passes the AVE > 0.50 criterion for convergent validity.

Table 2. Scale Reliability and Convergent Validity Statistics

Construct	Reliability Evidence		Convergent Validity Evidence
	Cronbach's Alpha	Composite Reliability	AVE
Analytics Capability (AC)	0.943	0.963	0.898
Analytics Culture (ANC)	0.797	0.908	0.831
Big Data Usage (BDU)	0.875	0.882	0.602
Competitive Pressure (CPP)	0.795	0.874	0.698
Government Pressure (GOP)	0.906	0.941	0.842
Information Systems Competence (ISC)	0.801	0.882	0.714
Top Management Support (TMS)	0.878	0.925	0.804

The most robust and advanced approach of testing discriminant validity in PLS-SEM is the Heterotrait-Monotrait ratio of correlations (HTMT) approach (Henseler et al. 2015). Therefore, HTMT ratios of correlations were assessed to test the discriminant validity of the measurement system representing the constructs. Based on the HTMT criterion, discriminant validity is shown if the HTMT ratios of correlations happen to be less than 0.850 (Henseler et al. 2015). Table 3 shows the HTMT ratios of correlations of each construct. Figures shown in Table 3 demonstrate that the HTMT criterion for discriminant validity is met by the constructs, as operationalized through their indicators.

Table 3. The Heterotrait-Monotrait Ratio of Correlations of the Constructs

Construct	AC	ANC	BDU	CPP	GOP	ISC	TMS
AC							
ANC	0.713						
BDU	0.258	0.348					
CPP	0.393	0.564	0.338				
GOP	0.291	0.348	0.236	0.640			
ISC	0.814	0.754	0.352	0.301	0.193		
TMS	0.603	0.731	0.237	0.315	0.191	0.513	

5.2 Graphical Results

The parameters estimated by SmartPLS is shown in Figure 1. Figure 1 shows the estimated relationships between the constructs and their measures (i.e., factor loadings) as well as relationships between constructs in the form of standardized structural regression coefficients (the regression coefficients are reported in standardised form because the scores of the constructs are in standardised form). The important parameter estimates to look for are the size and sign of the structural regression coefficients (path coefficients), the estimated R^2 values of the endogenous constructs (shown inside the circles in Figure 1) and the *statistical significance* of structural regression coefficients (path coefficients) shown separately (Table 4).

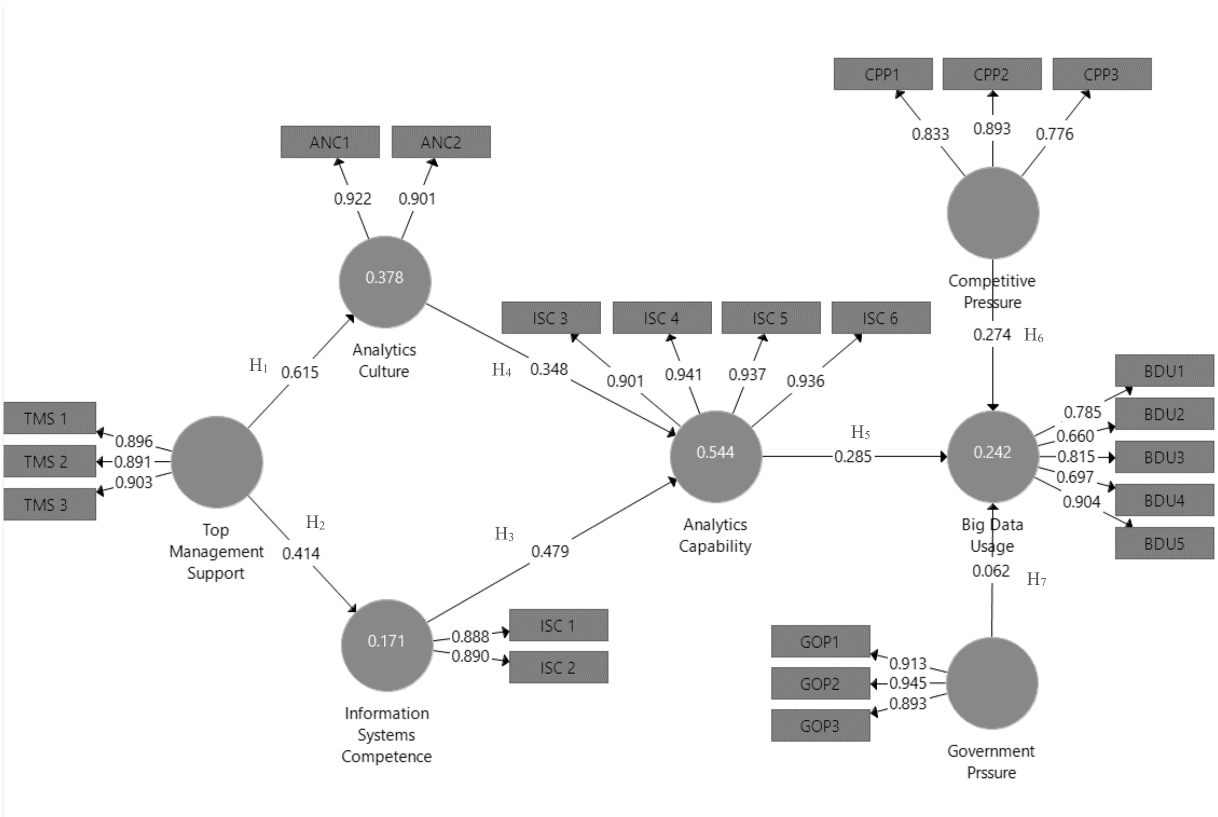

Figure 1. Parameter estimates of the measurement model and the structural model

5.3 Proposed Improvements

Based on the hypothesis test results (section 5.4), the following six sets of proposed improvement suggestions can be made to increase BD usage in problem solving: (1) Top Management Support is needed for the organisation to develop its capability through its Analytics Culture and IS Competence; (2) Both the Analytics Culture and IS Competence are equally important for increasing the Analytics Capability; (3) Top Management Support has a practically significant total effect on the Analytics Capacity and therefore it pays for the Top Management to focus their attention on improving the Analytics Culture and IS Competence; (4) Top Management must be aware that improving the Analytics Capability is one thing and expecting this Analytics Capability to result in more and more BD use is another thing. The results show that the latter does not quite happen as much as the Top Management might anticipate. This is because there is only a moderate positive relationship between the Analytics Capability and BD Usage; (5) While there is competition among firms in Sri Lanka, this competition does not necessarily manifest in the BD space. This finding may be useful to top managers in strategically selecting BD projects; (6) The Sri Lankan government does not seem to impose any pressure on firms, in the BD space. This is expected even before the fact because the Sri Lankan government has many other priorities to deal with, at the present point in time.

5.4 Validation

PLS-SEM adopts a nonparametric approach to establish the statistical significance of the estimated parameters for hypothesis testing and other purposes. Consequently, Smart-PLS was prompted to use the well-known nonparametric method "bootstrapping" to calculate the *T* statistics (and hence the *p* values) of the structural regression coefficients. In this regard, 5000 resamples were generated via bootstrapping to calculate the T values and thereby the p values using a one-tailed test to calculate the *p* values. Table 4 shows the hypothesis developed and obtained standardised regression coefficients, T values, p values and comments on the respective result. Results shown in Table 4 suggest that at 0.05 level of significance, only H_6 and H_7 are not supported by the data, although H6 is weekly supported by data ($p < 0.10$).

Table 4. The Statistical Significance of the Structural Relationships

Structural Relationship	Associated Hypothesis	Standardised Regression Coefficient	T Value	p Value	Comment
TMS → ANC	H_1	0.615	5.805	0.000	H_1 strongly supported
TMS → ISC	H_2	0.414	3.071	0.001	H_2 strongly supported
ISC → AC	H_3	0.479	3.727	0.000	H_3 strongly supported
ANC → AC	H_4	0.348	2.771	0.003	H_4 strongly supported
AC → BDU	H_5	0.285	1.667	0.048	H_5 supported
CPP → BDU	H_6	0.274	1.410	0.080	H_6 weakly supported
GOP → BDU	H_7	0.062	0.434	0.332	H_7 not supported
Note: This study takes the significance level (alpha risk) as 0.05 as it is customary to do so.					

Testing and Interpreting H1: Top Management Support has a positive effect on the Analytics Culture

The standardized regression coefficient of 0.615 ($p < 0.001$) in the Top Management Support → Analytics Culture relationship indicates a strong causal relationship. Top Management Support explains 37.8% of the variability of the Analytics Culture. This is deemed a large positive causal effect because an R2 of 37.8% is a large effect (any $R2 \geq 25.93\%$) in social and behavioral science research (Cohen 1992). In a business, the top management tend to support initiatives (in this study, increasing BD usage to create more value) that have the potential to improve bottom-line results. What top management often overlook is the much-need context (in this study the Analytics Culture and IS Competence) that is required to achieve bottom-line outcomes, or more generally, stakeholder results. This study showed that creating an Analytics Culture is very important in causing BD usage, which in turn arguably leads to stakeholder outcomes, including bottom-line results.

Testing and Interpreting H2: Top Management Support has a positive effect on IS Competence

The standardized regression coefficient of 0.414 ($p = 0.001$) in the "Top Management Support → IS Competence" relationship resulting in an R2 of 17.1% implies a medium effect size. This is based on the prescriptions provided by Cohen (1992) on the effect size ($R2 \geq 13.04\%$ being a medium effect). Comparing the strength of the said relationship (the R2 value) with the "Top Management Support → Analytics Culture" relationship, it becomes clear that although there is statistically significant "Top Management Support → IS Competence" relationship ($p = 0.001$), this relationship is weak relative to the "Top Management Support → Analytics Culture" relationship ($p < 0.001$ and R2 = 37.8%). This finding makes both theoretical sense and practical sense, as explained below.

IS Competence captures technology readiness, technology competence, and technology knowledge. While Top Management Support is required for the organization to gain IS Competence, because allocating resources is a top management activity, Top Management Support is also needed in BD project management activity to act as a catalyst for skills development and organizational learning. Despite all this, the results imply that Top Management Support ($\beta = 0.414$) is not so critical to IS Competence as to Analytics Culture ($\beta = 0.615$). In any organization, the top management is instrumental in establishing and reinforcing both the organizational culture and the occupational cultures (in this study, more specifically, the Analytics Culture) (Al-Sai et al. 2020; Cameron and Quinn 2011). Thus, a stronger causal relationship Top Management Support → Analytics Culture is theoretically justifiable, although a detailed study may shed light on this differential.

Testing and Interpreting H3: IS Competence has a positive effect on Analytics Capability
The theorical basis of H3 is that more IS competent an organization is, the more it enhances its Analytics Capability. The Analytics Capability is needed by an organization to gain access to speedy information, traceability, and to possess the required technology readiness for BD adoption (Mikalef et al. 2019). The standardized regression coefficient of 0.479 ($p < 0.001$) associated with H3 suggests that the causal relationship IS Competence → Analytics Capability is statistically and practically significant (an extremely low p value and a practically very high regression coefficient respectively).

Testing and Interpreting H4: Analytics Culture has a positive effect on Analytics Capability
The shared values, norms, beliefs, and behavior patterns (the Analytics Culture) can achieve constancy of purpose (thinking an organization as an interconnected system) towards improvement. The standardized regression coefficient of 0.348 ($p = 0.003$) associated with H4 suggests that the causal relationship Analytics Culture → Analytics Capability is statistically and practically significant (low p value and a practically high regression coefficient respectively).

The Analytics Capability is being explained by two constructs: the Analytics Culture and IS Competence. The R2 associated with the Analytics Capacity (54.4%) shown in Figure 1 is the joint squared correlation between the Analytics Capacity and its two causal constructs. An R2 of 54.4% indeed signifies a very large causal effect, based on Cohen's guidelines (Cohen 1992) mentioned earlier.

The results imply that the IS Competence → Analytics Capability causal relationship ($\beta = 0.479$) is stronger than the Analytics Culture → Analytics Capability causal relationship ($\beta = 0.348$). This seems to suggest that technical skills (IS Competence) might be bit more important than social skills (Analytics Culture) in a BD analytics context, which sounds somewhat surprising, because normally soft skills are more strategic to an organization than hard skills. Further research is needed to make more reliable interpretations.

Testing and Interpreting H5: Analytics Capability has a positive effect on BD Usage
Data supporting H5 implies that the Analytics Capacity as a cause of BD usage is active (meaning this cause has an effect). This is encouraging from an organizational perspective. The theory and empirical support of the hypotheses implies that in order to improve BD usage, a BD organization must improve its Analytics Capability via its improved Analytics Culture and IS Competence, both being driven by the Top Management Support. However, because the standardized regression coefficient associated with H5 is modest (= 0.285), the total effect of "Top Management Support" on BD usage resulted in a small value (= 0. 117).

Testing and Interpreting H6: Competitive Pressure has a positive effect on BD Usage
In this study, H6 returns a weakly significant relationship ($\beta = 0.274$ and $p = 0.080$). It is argued that this weak relationship is returned not because there is less competition among commercially active organizations as such private sector organizations in Sri Lanka (indeed this is not the case), but because there is currently limited scope to use BD for competitive advantage.
H7: Government Pressure has a positive effect on BD Usage
Governments impose policies, rules and regulations upon firms on matters such as data security, data storage, and data processing (Dal-woo et al. 2015; Lunde et al. 2021; Maroufkhani 2020; Sun et al. 2020; Verma and Mumbai 2019). In addition, governments impose regulations on firms on such matters as fair trade and green economics, which usually warrant firms to use BD to find new ways of conducting business. The nonsignificant result (H7 was not supported

by the data, β = 0.062 retuning a p value of 0.332) suggests that none of the above-mentioned forms of government pressure has any effect on the Sri Lankan industry.

6. Conclusion

The three objectives were achieved, and it is argued that the study attempts to make an academic contribution by way of formulating and testing a theory to explain the BD usage phenomenon. Previous studies posited several enables/determinants of BD usage. They were presented merely as predictors (independent variables) that predict BD usage. This study identified the critical predictors (through a sifting process) and organised them in a causal predictive manner. This is to create a path model that not only to predicts BD usage, but also explains how BD usage is caused. The path model/theory is useful for both the academia and practitioners alike to understand or distinguish successful BD implementation from unsuccessful BD implementation.

The validated survey instrument (the questionnaire) means that managers can use this instrument for continuous improvement and benchmarking purposes. Managers can conduct self-assessments to take baseline measurements (scores) for each CSF (determinant or enabler) of BD usage and keep improving their scores by taking appropriate action. This is the first practical contribution of the research. It is acknowledged that improving the scores of some CSFs will be more difficult than the others. For example, improving the score of the Analytics Culture could be more challenging than improving the score of IS Competence because a technical resource can be acquired more easily (e.g., via recruitment) than a soft, intricately bundled, resource (specifically, the culture) which is deep rooted. It is suggested that the study be replicated in larger economies to improve the external validity of the findings.

References

Adrian, C., Abdullah, R., Atan, R. and Jusoh, Y. Y., Factors influencing to the implementation success of big data analytics: A systematic literature review. Paper presented at the 2017 *International Conference on Research and Innovation in Information Systems (ICRIIS),* 2017.

Al-Sai, Z. A., Abdullah, R. and Husin, M. H., Critical Success Factors for Big Data: A Systematic Literature Review. *IEEE Access*, 8, 118940-118956. 10.1109/ACCESS.2020.3005461, 2020.

Awa, H. O., Ojiabo, O. U. and Orokor, L. E., Integrated technology-organization-environment (T-O-E) taxonomies for technology adoption. *Journal of Enterprise Information Management*, 30(6), 893-921. 10.1108/JEIM-03-2016-0079, 2017.

Baker, J., The Technology–Organization–Environment Framework. In Y. K. Dwivedi, M. R. Wade and S. L. Schneberger (Eds.), *Information Systems Theory: Explaining and Predicting Our Digital Society*, Vol. 1 (pp. 231-245). New York, NY: Springer New York. 10.1007/978-1-4419-6108-2_12, 2012.

Bryman, A., *Social research methods* (4th ed.). New York, NY: Oxford University Press Inc., 2012.

Bullen, C. V. and Rockart, J. F., *A primer on critical success factors*. Cambridge, MA: Center for Information Systems Research, Sloan School of Management, Massachusetts Institute of Technology, 1981.

Cameron, K. S. and Quinn, R. E., *Diagnosing and changing organizational culture: Based on the competing values framework*. New York, NY: John Wiley & Sons, 2011.

Chang, S.-J., van Witteloostuijn, A. and Eden, L., From the Editors: Common method variance in international business research. *Journal of International Business Studies*, 41(2), 178-184. 10.1057/jibs.2009.88, 2010.

Cohen, J., A power primer. *Psychological Bulletin*, 112(1), 155-159, 2010.

Côrte-Real, N., Ruivo, P., Oliveira, T. and Popovič, A., Unlocking the drivers of big data analytics value in firms. *Journal of Business Research,* 97, 160-173, 2019.

Curry, E., The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches. Paper presented at the *New Horizons for a Data-Driven Economy*, 2016.

Curry, E., Kikiras, P., Freitas, A., Domingue, J., Thalhammer, A., Lasierra, N., . . . Ngonga, A., *Big Data Technical Working Groups White Paper*, 2014.

Dubey, R., Gunasekaran, A., Childe, S. J., Blome, C. and Papadopoulos, T., *Big Data and Predictive Analytics and Manufacturing Performance: Integrating Institutional Theory, Resource-Based View and Big Data Culture*. Oxford: John Wiley & Sons Ltd, 2019.

Elia, G., Polimeno, G., Solazzo, G. and Passiante, G., A multi-dimension framework for value creation through Big Data. *Industrial Marketing Management*, 90, 617-632, 2020.

Faroukhi, A., El Alaoui, I., Gahi, Y. and Amine, A., A Multi-Layer Big Data Value Chain Approach for Security Issues. *Procedia Computer Science*, 737–744, 2020.

Flynn, B. B. and Saladin, B., Further evidence on the validity of the theoretical models underlying the Baldrige criteria. *Journal of Operations Management*, 19(6), 617-652, 2001.

Fornell, C. and Larcker, D. F., Structural Equation Models with Unobservable Variables and Measurement Error: Algebra and Statistics. *Journal of Marketing Research*, 18(3), 382-388. 10.1177/002224378101800313, 1981.

Gaurav, D., Yadav, J. K. P. S., Kaliyar, R. K. and Goyal, A., An Outline on Big Data and Big Data Analytics. Paper presented at the 2018 *International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, 2018.

Hair Jr, J. F., Hult, G. T. M., Ringle, C. M. and Sarstedt, M., *A primer on partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, CA: Sage publications, 2021.

Henseler, J., Ringle, C. M. and Sarstedt, M., A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135, 2015.

Hu, H., , Y. W., Chua, T.-S. and L, X., Toward Scalable Systems for Big Data Analytics: A Technology Tutorial. *IEEE Acesss*, 2, 652-687, 2014.

Jcobs, A., *The Pathologies of Big Data*. Volume 52(Issue 8), pp 36–44, 2009.

Laney, D., *3D Data Management: Controlling Data Volume, Velocity and Variety*. META Group Research Note, 2001.

Lasi, H., Fettke, P., Kemper, H.-G., Feld, T. and Hoffmann, M., Industry 4.0. *Business & Information Systems Engineering*, 6(4), 239-242. 10.1007/s12599-014-0334-4, 2014.

Lunde, T. Å., Sjusdal, A. P. and Pappas, I. O., Organizational Culture Challenges of Adopting Big Data: A Systematic Literature Review. In I. O. Pappas, P. Mikalef, Y. K. Dwivedi, L. Jaccheri, J. Krogstie, and M. Mäntymäki (Eds.), *Digital Transformation for a Sustainable Society in the 21st Century* (pp. 164-176). Cham: Springer International Publishing, 2019.

MacKenzie, S. B. and Podsakoff, P. M., Common Method Bias in Marketing: Causes, Mechanisms, and Procedural Remedies. *Journal of Retailing*, 88(4), 542-555, 2012.

Maroufkhani, P., Wan Ismail, W. K. and Ghobakhloo, M., Big data analytics adoption model for small and medium enterprises. *Journal of Science and Technology Policy Management*, 11(4), 483-513. 10.1108/JSTPM-02-2020-0018, 2020.

Mikalef, P., Boura, M., Lekakos, G. and Krogstie, J., Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, 98, 261-276, 2019.

Nguyen, T. L., A Framework for Five Big V s of Big Data and Organizational Culture in Firms*. 2018 IEEE International Conference on Big Data (Big Data)*, 5411-5413, 2018.

Nunnally, J. C., *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill, 1978.

Perova-Antova, D. and Ilieva, S., Using Big Data Value Chain to Create Government Education Policies. Paper presented at the *ACM Computer Systems and Technologies conference*, Bulgaria, 2018.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y. and Podsakoff, N. P., Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879-903. 10.1037/0021-9010.88.5.879, 2003.

Porter, M. E., *Competitive Advantage- creative and Sustaining*: The Free Press, 1985.

Rim, L. and Sonia, M., Toward a Conceptualization of Big Data Value Chain: From Business Problems to Value Creation. In M. Karim & A. Wassim (Eds.), *Business Transformations in the Era of Digitalization* (pp. 76-92). Hershey, PA, USA: IGI Global. 10.4018/978-1-5225-7262-6.ch005, 2019.

Ringle, C. M., Wende, S. and Becker, J.-M., SmartPLS 3 (Version 3). Bönningstedt: SmartPLS. Retrieved from https://www.smartpls.com, 2015.

Saturno, M., Pertel, V. M., Deschamps, F. and Loures, E., Proposal of an automation solutions architecture for Industry 4.0. Paper presented at the *24th International Conference on Production Research*, Poznan, Poland, 2017.

Schein, E. H., *Organizational culture and leadership*. San Francisco, CA: Jossey-Bass, 1992.

Sun, S., Hall, D. J. and Cegielski, C. G., Organizational intention to adopt big data in the B2B context: An integrated view. *Industrial Marketing Management*, 86, 109-121, 2020.

Surabhi, V. and Sushil, C., Understanding the Determinants of Big Data Analytics Adoption. *Information Resources Management Journal (IRMJ)*, 32(3), 1-26. 10.4018/IRMJ.2019070101, 2019.

Tornatzky, L. G. and Fleischer, M., *The Processes of Technological Innovation*. Lexington, MA: Lexington Books, 1990.

Vargo, S. L., Maglio, P. P. and Akaka, M. A., On value and value co-creation: A service systems and service logic perspective. *European Management Journal*, 26(3), 145-152, 2008.

Wirén, M., Mäntymäki, M. and Najmul Islam, A., Big Data Value Chain: Making Sense of the Challenges. Paper presented at the *18th Conference on e-Business, e-Services and e-Society (I3E)*, Trondheim, Norway, 2019, 2019-09-18.

Biographies

Sachini Weerasinghe is a senior autonomation engineer working in the apparel industry in Sri Lanka. She has hands on experience in industry 4.0, industrial engineering and lean manufacturing over five years. Sachini holds a BSc Engineering, first class honours degree from University of Moratuwa, Sri Lanka. She will be graduating with her Master of Business Administration in Management of Technology in December 2022 from the same university. Sachini will be leaving to New Zealand to read her PhD on green buildings—a research that has the potential to use IoT and big data to create new solutions. She is interested in industry 4.0, IoT and sustainable engineering research areas.

Nihal Jayamaha is a senior lecturer attached to Massey University, New Zealand. Nihal received his PhD in 2008 from the same university, with an endorsement in Technology Management. Nihal's teaching and research interests cover quality engineering, quality improvement, sustainability, quality management models and systems, and latent variable modelling techniques (e.g., structural equation modelling). Prior to becoming an academic, Nihal worked in the utility industry in the UAE, and in Sri Lanka for around 20 years in various capacities, including as an electrical engineer, senior technical auditor, and a project manager. Nihal has published in prestigious academic journals such as the *International Journal of Production Research*, *Journal of Industrial Ecology*, *Environmental Science & Technology*, *Total Quality Management & Business Excellence, International Journal of Lean Six Sigma, International Journal of Quality and Reliability Management*, *Benchmarking: An International Journal*, *The TQM Journal*, and *Measuring Business Excellence*.