Assessing Suitability of Applying Big Data Analytics within Small to Medium-Sized Businesses via an ROI-Based Graphic Model

Ryan Grizzle and Yanzhen Qu

***Abstract*— Big data has presented itself as a term, phenomena, and paradigm with the potential for many opportunities and challenges. The new potentials of big data seemingly continue to expand both in possibility and complexity. With efforts to exploit these new potentials requiring investments from businesses and organizations it becomes necessary to understand what value is gained from such efforts. Through the methodology of design science this study develops an artifact which incorporates the return on investment model which can be utilized by SMEs. The artifact provides an abstract process model for the use of value assessment of big data efforts by SMEs. This study finds through using test cases that the ROI model can be applied to a generalized artifact which guides the assessment of big data efforts. Further, it is found that through a graphical design the development of a simple and intuitive artifact can be accomplished.**

***Index Terms*—Big data, big data value, small and medium enterprises, SME, return on investment, ROI, design science.**

I. Introduction

The advent of big data has the potential to bring multiple changes to social, technical, and economic environments. Big data has been described as a term, phenomenon [1], [2] and paradigm [3]. One of the aspects to emerge from big data is the ability of businesses and organizations to leverage and capitalize on a large amount of data. However, big data efforts require resource investments by companies and organizations and can be costly based on several factors [4]. Given that big data can potentially yield benefits but has a cost it becomes necessary to understand both what big data is and the value creation potential of big data. Part of understanding how big data efforts can produce value is in understanding what processes in the design, development, and deployment of big data systems and solutions projects are gainful. Guidance in the process of developing big data solutions can inform business and organizational decision-makers in determining the potential value creation of projects and determine the cost, risk, and return of such projects. However, there seemingly remains complexities surrounding the understanding of how big data is defined and in the development of big data systems and solutions.

Manuscript received March 14, 2022; revised May 23, 2022.

Ryan Grizzle and Yanzhen Qu are an alumnus of Colorado Technical University, Colorado Springs CO 80907 USA (e-mail: Ryan.Grizzle@alumni.ctuonline.edu, yqu@coloradotech.edu).

II. Related Work

A. Big Data Understanding

While efforts have been made to clarify the definition of big data, discerning those definitions have arguably been met with several challenges and complexities. Ekbia, *et al*. [5] found that the purview of what constitutes big data lacks consensus for its definition, scope, and character. De Mauro, *et al*. [6], through a survey of existing definitions, derived the meaning as, "Big Data represents the Information assets characterized by such a High Volume, Velocity and Variety to require specific Technology and Analytical Methods for its transformation into Value"(p. 1). De Mauro, et al. [6] place big data conceptually at the intersections between information, technology, impact, and methods. Efforts have continued in the pursuit to develop an understanding of how to define big data, but a definitive and concrete definition seemingly remains aloof. It has been argued that big data has no single definition [7].

Throughout much of the research, the characteristics or properties of big data are often described as the V's of big data. However, the matter of what V's constitute the characteristics or properties of big data may arguably be indecisive and confusing. Some of the earlier accounts list 3 V's of big data as volume, variety, and velocity [8], [9], and these three V's remain primarily persistent throughout the research with regards to big data characteristics. However, as research has progressed the number of V's continues to grow from four [10] to five [11] and upwards of seven [12] and more. Matters in defining big data are made further complex as its definition is viewed to extend beyond just the characteristics or properties of the V's. Some researchers Demchenko, *et al*. [13] argue the properties known as the V's of big data by themselves are not sufficient in defining big data, incorporating them into a structural definition which includes new data models, analytics, infrastructure and tools, and sources and targets for data types and structures. Ylijoki and Porras [14] state that current definitions lack consideration for important aspects such as security, privacy, or its disruptive nature. In some cases, the argument has been made that the absence of a definitive definition of big data based on its characteristics, technology, or trends has led to present definitions of big data being constituted as a social phenomenon, analytical techniques, a process, or a data set

DOI: 10.7763/IJMO.2022.V12.812

[15]. Additional support for understanding big data generalization not only in its characteristics but through measuring its impact is given by Hajirahimova and Aliyeva [16], stating big data as having significance in assessing growth and performance in an information society but lacks generally accepted measurement indicators.

B. Big Data Systems

Given the size and scale that many systems can be nowadays, it is arguably in the best interests of companies and organizations to have some form of structure for big data systems and solutions. Big data system design can be costly due to infrastructure costs, steep learning curves for frameworks, and complexities in architectures [4]. While many common design elements exist such as the extract, transform, and load process [17] other factors such as component selection and interaction, environments, scenarios [17], [18] and the way big data is defined such as being related to the attributes of data, technical aspects, the means of overcoming data and technical challenges, or social impact [15] contribute to the complexities of system and solution designs. Big data system and solution development is also influenced by the emphasis on differing elements, even when the architectures are within the same domain. For example, some architectures in the domain of education emphasize more concrete aspects such as the infrastructure and software services Wu, *et al.* [19] while others address a more abstracted identification of the constituent systems Matsebula and Mnkandla [20]. At an even further level of abstraction others concentrate on conceptual definitions of the data and algorithms Michalik, *et al.* [7]. Design can also be influenced by the targeted data and the application of certain technologies. Yao, *et al*. [21] develop a Hadoop based system utilizing Mahout-based distributed recommendation engine while Tu, *et al.* [22] focus on a design based around the utilization of Spark to target wearable devices and mobile phones, home intelligent equipment, and vital signs monitoring instruments, both architectures can be considered within the medical domain. Big data systems and solutions are shown to be diverse in both the available technologies and methods which can be utilized in projects.

C. ROI as the Valuation Model

Determining the value of systems can prove to be a difficult task. The study conducted by Christina and Kelly [23] on knowledge management and big data within SME competitiveness of the agri-food sector expresses that while value can be created through data analytics it becomes difficult in quantifying value with firms' overall success. Seemingly some of the benefits provided by big data systems are hard to concretely measure and are relative to an individual organization's implementation of the system. Gupta and George [24] identified the three categories of tangible, human, and intangible elements of big data analytics as being significant to understanding and accomplishing success with projects. These elements of big data have similarities with those of information technology (IT) systems and information systems (IS) in contributing to the complexity of the related systems. One method which has presented itself as a means of measuring value within IS [25] and IT systems [26] is the use of return on investment (ROI). While arguably big data analytics may have some distinctions from being a direct parallel to IS and IT systems it also shares many similarities. Shim, *et al*. [27] outline in addressing issues and solutions related to big data systems that further understanding ROI of big data systems to be an important factor of consideration. The significance of ROI even at the more discrete level of big data elements is relatively important aspect. Abdullah, *et al.* [28] express that within big data projects managing data quality should incorporate the use of ROI estimations. While understanding the full and true value of big data can have degrees of complexities and nuance the ROI method of evaluation has shown to be one which is applicable. Gao, *et al.* [29] states that as part of critical success factors relating big data analytics projects should have some focus on ROI. The ROI methodology has shown to be applicable in IS, IT, and big data systems which have a degree of complexity. More specifically to big data the application of ROI as a methodology has presented itself to be of relevance within academic research in addressing the higher-level conceptual aspects, some of the discrete elements, and the success factors of big data analytics projects and initiatives.

D. Small to Medium Enterprises and Big Data

The current body of research specifically regarding SMEs and big data seems to be relatively small to that of big data in general. Some research expresses a still current need regarding various aspects of SMEs and big data such as the impact of big data on performance [30], how SMEs extract and utilize data [23], the applicability of big data to SMEs [31], and a need for case studies [32]. The wide diversity of SMEs in operations, environments, and capabilities may make it difficult to definitively and distinctly develop a universal understanding of their abilities regarding big data capabilities. Soroka, *et al.* [33] found in their study of Welsh manufacturing SMEs that while there may be some demand for big data analysis current solutions may not be viable for the SMEs and the SMEs seem to be ill-prepared and ill-equipped for big data analytics. Pirola, *et al.* [34] state in their study of Italian SMEs that the SMEs have a lack of knowledge regarding big data technologies and a lack of digital strategy, vision, and action plan. Noonpakdee, et al. [35] find that 90% of the SMEs within the service sector of Thailand involved in their study had IT infrastructure that was not ready and that the SMEs also lacked skilled enough staff. Sune Dueholm and Jensen [36] express in their research on Danish SMEs that to the extent where big data is used in 75 percent of the cases value is created but respondents were unsure of how big data is used across their companies and how value is created in their differing business processes. Iqbal, *et al*. [32] in their research on Pakistani SMEs state that a very low level of big data understanding exists but that this is not unexpected as this is prevalent within a large fraction of SMEs in both developing and developed countries.

Small to medium enterprises (SME) may face several barriers in making use of big data analytics due to a multitude of differing factors [37]. Christina and Kelly [23] express that for SME’s it can be challenging to accumulate, analyze and accurately interpret data often due to a lack of skill, time, and resources, and from a resource-based perspective, this presents limitations for SMEs to utilize big data. The lack of skilled professionals as a barrier for SMEs to utilize big data is expressed by various other researchers [32], [38]. *Mangla, et al.* [39] state a lack of financial resources as a significant obstacle regarding SMEs' adoption of big data analytics. Some researchers have suggested the use of cloud-based technologies as a means to circumvent some of the obstacles faced by SMEs in big data adoption [40], [41] but even within using cloud-based technologies obstacles to SME adoption still exist [42].

E. Big Data Value

Big data has proved to have many challenges both technically and organizationally [43]-4[5]. Research into organizational challenges indicated that big data was an integrated part of an organization and its adoption and utilization were impacted by more than just technical aspects. LaValle, *et al.* [46] expressed that insight gained from big data stems from more than just technical analysis but integrates organizational understanding and that some of the largest barrier's organizations face are less related to data and technology and more related to managerial and cultural. One aspect identified as an organizational challenge needing further research was the need for dynamic system designs based on understanding the needs of both users and technologies [44]. Kaisler, *et al.* [44] express that additional design challenges are created because end-users will often not be the system designers.

A concept that was adapted and applied to big data analytics as a means of understanding value was the *value-chain*, a concept introduced by Michael E. Porter in 1985 [47]. The application of the *value-chain* served as a framework that could provide a structured understanding of big data analytics at differing levels. Miller and Mork [48] provided recommendations for applying the *value-chain* concept to big data analytics abstract and provided overarching guidance at an organizational level addressing the three main categories of data discovery, data integration, and data exploitation. The *value-chain* concept also proved to be able to apply to system design as well. Hu, et al. [49] demonstrated the use of the *value-chain* as a means of mapping system architecture, focusing on the four elements of generation, acquisition, storage, and processing. Chen, et al. [50] identified the technical challenges of the value-chain concept identifying that further theoretical research was needed, with one of the fundamental problems of big data being a need for a rigorous and holistic definition of big data, a structural model of big data, a formal description of big data, and a theoretical system of data science.

III. Problem Statement, Hypothesis Statement, and Research Question

A. Problem Statement

The problem is that there is still a lack of complete understanding regarding big data's effective use [2], [51], value creation [52], [53], and use in small and medium enterprises [31], [54], [55].

B. Hypothesis Statements

H_1: The assessment of artifact processes guided by big data success factors and assessed using the ROI model can help determine the benefits of big data analytics for SMEs.

H_0: The assessment of artifact processes guided by big data success factors and assessed using the ROI model cannot help determine the benefits of big data analytics for SMEs.

C. Research Question

Q1: What processes which are guided by big data analytics success factors and assess the value creation through the ROI model are needed by SMEs to determine if big data analytics is beneficial?

IV. Methodology

The purpose of the study is to conduct design science research, which results in the creation of a process artifact in the context of value that outlines the value assessment of big data analytics systems and solutions, intended for use by small and medium enterprises (SME). The current problem is that there is a lack of complete understanding regarding big data's effective use [2], [51], value creation [52], [53], and specifically to this research its use in small and medium enterprises [31], [54], [55]. Though potential benefits for big data within SMEs exist [56] the diverse nature of operational environments and complexity of big data systems drive a need for further knowledge of big data systems solutions for small and medium enterprises. The development, testing, and refinement of solutions are needed to allow for effective decision-making regarding big data systems and projects.

A. Research Method and Design

The research method selected for this study is design science. The complexities of big data systems and projects and their unique circumstances and environments of operations dictate the need for the creation and reevaluation of solutions to identify and refine applicable designs and processes to aid in the effectiveness and success of big data projects by small and medium enterprises. Design science research allows for and integrates the ongoing and continued reevaluation of artifacts. Since design science seeks to further the knowledge base and produce solutions in the form of an artifact it is applicable and well suited for this study as a methodology. More specifically with regards to the aforementioned needs of big data solutions within SME's design science methodologies integrate environmental factors and previously established knowledge basis into the development of solutions which undergo an iterative design cycle of the development and evaluation of an artifact [57]. Additionally, design science also allows for the creation of an artifact to take place within a grounded context of a literature review [58].

B. Population and Samples

This studies intendent is to produce an artifact derived from literature reviews and intended for the utilization of developing big data systems. The study does not utilize human subjects and does not take samplings from a population to collect or analyze any data or results for the study. The artifact is intended to be generalized and evaluated utilizing test cases built from academic literature reviews. This study has no targeted individuals or organizations but is intended to produce an artifact, which can be utilized generally by any SME business, organization, or entity in developing a big data system or solution.

C. Artifact Creation

The production and refinement of an artifact is the central focus of this study. The artifact focused on within this study is the developed process model which guides big data systems and solutions development through value assessment using ROI. The form of ROI evaluation that is utilized within this artifact is

$$ROI = \frac{(Gain\ from\ Investment - Cost\ of\ Investment)}{(Cost\ of\ Investment)}. \quad (1)$$

The artifact treats big data system development as an ongoing effort that incorporates feedback into processes for iterative system development cycles. The processes within the artifact incorporate business and organizational needs and are evaluated based upon return on investment potential for the individual processes and the project. The generalized form of the proposed artifact is given in figure 1.

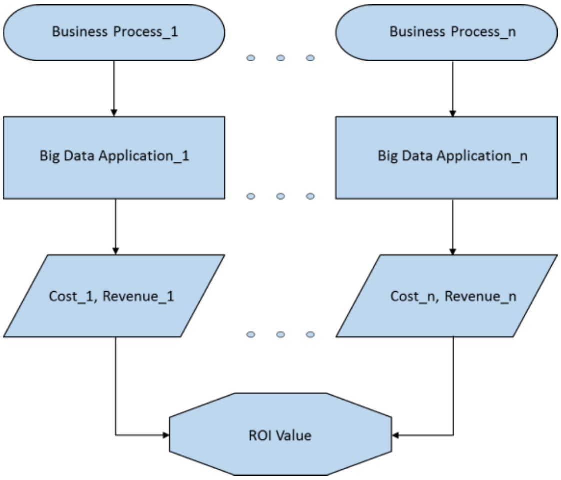

Fig. 1. Proposed Artifact

The artifact intends to provide multiple steps or processes within the development of a big data analytics system or solutions and evaluate the return on investments (ROI) within each of those steps/processes. The artifact is developed guided by the selected design science model provided by Hevner [57] which can be found in Fig. 2.

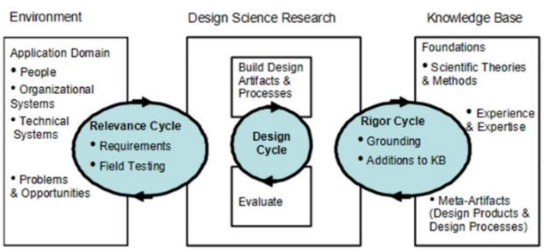

Fig. 2. Guiding design science model.

The reasoning for the selection of the design science model to guide the development of the artifact was due to its inclusion of the inputs of environment and knowledge base and their alignment with this study's intent.

Within design science, both the process for the development of the artifact and the development of the artifact itself is typically iterative [59]. The review of academic literature during the cycle of relevance within the selected design science model establishes drivers for the artifact. The extracted drivers inform what factors are relevant for evaluation within the artifacts process steps. The environment drivers and knowledge base provided by academic literature informed the structuring and development of the artifact's elements.

D. Data Collection and Artifact Evaluation Procedures

Design science seeks to advance the knowledge base through the production of solutions in the form of an artifact which typically undergoes an ongoing iterative life cycle in both the process for the artifact development and the artifact itself [59]. Iterations of the artifact's lifecycle involve building then evaluating the artifact while continuously integrating drivers from environments and rigor from the knowledge base. Hevner, *et al.* [59] posited that the iterative evaluation of an artifact within design science serves to identify weaknesses within an artifact and direction for refinement and reassessment processes being described for future research. Hevner, *et al.* [59] also argue that an artifact must be internally consistent. The focus of the evaluation in this study is to identify and test the artifacts' logical capabilities and consistencies and the application of the ROI formula within it.

The artifact is evaluated with the utilization of three simulated test cases. Simulated test case development is informed by academic literature. The processes of the artifact will be evaluated by applying those processes to the simulated test cases. The determination of what constitutes the value assignments within the artifact for a business process, big data technology/solution, and cost/revenues are subjective as the assignment of these elements would be determined by the user. For example, an SME may determine that a valid business process for utilizing big data analytics is marketing and advertising while another may not. Also, two SMEs may select the business process of marketing and advertising but assign different big data technologies/solutions to that business process such as social media or in-house algorithmic development. While the objective of the test cases is to test the logical capabilities of the artifact consideration in determining the assignment of the individual elements for the business processes and the big data technologies/solutions is given in this study. Because of the subjectivity of test case elements, this research reviewed academic literature to establish a better notion and basis for the assignment of these elements. Some authors have expressed a lack of case studies regarding big data and SMEs [32] and throughout the investigation of the literature, this study has come to the same conclusion. While an extensive body of knowledge regarding case studies relating to big data and SMEs is lacking some literature existed and was used for guidance in developing test cases. The academic material reviewed to establish a basis for the artifact's element assignments was largely limited to material that dealt with SMEs and big data as these are the primary focus of the artifact's application. While some explicit examples were given within the literature reviewed in most cases the literature reviewed did not make specific mention of big data technology/solutions and businesses processes pairings but typically referred to them categorically such as marketing being a business process and social media being a big data technology/solution. Some strong examples for element assignment could be derived from the literature such as the case of the assignment of social media and marketing. Some research finds that social media is an important technology impacting big data technology growth [38] and identified as being used within the marketing context of SMEs [30], [60], [61], as having positive marketing value for SMEs [62] and has seen increasingly greater use by SME's [63], [64]. While other examples found within the reviewed literature were more generalized such as the use of customer relationship management (CRM) and CRM systems [33], [36], [65]-[67] where CRM was capable of supporting multiple business functions and utilizing multiple technologies [68]. Similar to CRM big data technologies/solutions were not necessarily isolated to a single business use case within SMEs but instead capable of having a diversified range of applications such as Google Analytics [61], [62], [68] and Salesforce [66], [68], [69]. Hopkins and Hawking [70] provided insights into the use of specific big data technologies like SAP HANA and the businesses process within the logistics industry. Santoro, et al. [71] state that in their case study for retailer's survey participants expressed the benefits of data exploitation in business operations such as logistics and operations but did not specify the technologies used. The works of Hopkins and Hawking [70] and Santoro, *et al*. [71] demonstrate that two differing industries may have the same business processes. It is presumed for the purpose of testing that the big data tools/solutions utilized for businesses processes can be applied to those same business processes in other industries but utilized in different ways. For the test cases, some of the elements were considered common across domains for the purposes of testing while others were considered more likely to be specific in their implementation to a domain in order to demonstrate the extensibility of the artifact. The business processes considered common throughout the differing test cases are marketing and advertising, customer relations management (CRM), consumer analytics, and data driven decision-making.

E. Data Analysis Procedures

During the evaluation of the artifact specified values for each test case are assigned to the generalized from of the artifact. The result of whether the artifacts steps logically function within the test case is evaluated as either being able to be applied to the test case or inapplicable to the test case. Determination of applicability is based upon the artifacts generalized form and its elements capabilities in accommodating specifically assigned values which are respective to the categories of business process, big data application, and cost/revenue and contextualized by the specific test case. Analysis of the applicability of the utilization of the ROI formula within the artifact is conducted by inputting the cost/revenue values assigned within each test case. Overall determination of the applicability of a processes model which incorporates the use of the ROI formula for the use of big data value assessment is determined by if the artifact is capable of yielding a calculable ROI result after value assignments are given to the artifacts elements in each test case.

V. Results of Artifact Evaluation

Data collection was accomplished by applying the artifact to three test cases representing potential small to medium enterprise business domains with varying numbers of business processes. The application of the process model to the test cases both evaluated the model as an artifact and provided testing for the hypothesis. The test hypothesis of H1 and H0 are as follows:

H1: The assessment of artifact processes guided by big data success factors and assessed using the ROI model can help determine the benefits of big data analytics for SMEs.

H0: The assessment of artifact processes guided by big data success factors and assessed using the ROI model cannot help determine the benefits of big data analytics for SMEs.

Presented here is the third test case which evaluates the artifact within the context of a banking company. The common business processes remain the same as the other two test cases, but the potentially different business processes are expanded to three additional business processes which are not included in the previous two test cases. The test case is presented in Fig. 3.

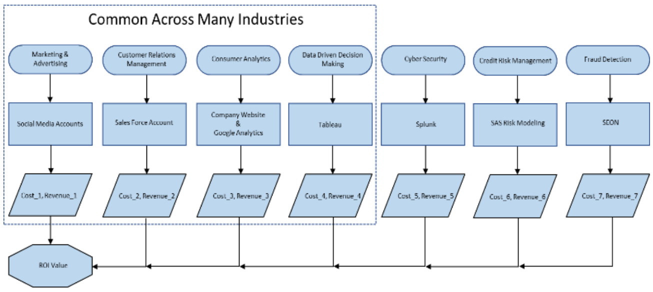

Fig. 3. Banking company.

Table I below provides a description and the specifications for the elements found in Fig. 5 which are grouped together within the context that they are being viewed as being common across many Industries.

TABLE I: Banking Company (Common Across Industries)

Name of Element	Specification
Marketing & Advertising	Used in establishing a presence within communities.
Social Media	Establishing community outreach and sentiment development in communities.
Cost_1, Revenue_1	Costs include the maintenance of social media accounts while revenue comes from the generation of new accounts and product sales. Cost_1 = $700, and Revenue_1 = $10000
Customer Relations Management	Aggregation of customers transactional data used in predictive analytics for targeted product offers.
Salesforce.com Account	Using commercial vendors could prevent large systems infrastructure costs.
Cost_2, Revenue_2	Costs come from the Salesforce.com account while the revenue come from optimizing product lines and offerings. Cost_2=$1200, Revenue_2=$15000
Consumer Analytics	Utilization of data analytics with an SME bank's website.
Company Web Site & Google Analytics	Analytics from website can be incorporated into analytics conducted on CRM systems data.

Cost_3, Revenue_3	Cost is maintenance of the company website while revenue represent the increase in product sales extended from the CRM systems. Cost_3=$3000, Revenue_3=$20000
Data-Driven Decision Making	Executive decision-making using data and data analytics.
Tableau	Tableau provides the means for SME banks to analyze internal data to visualize current costs and operations.
Cost_4, Revenue_4	Costs represent maintaining a Tableau account while revenue represents optimization in business operations. Cost_4=$2000, Revenue_4=$15000

Table II below provides a description and the specifications for the elements found in figure 5 which are grouped together within the context that they are being viewed as unique, separate from those common across many industries.

TABLE II: BANKING COMPANY (UNIQUE)

Name of Element	Specification
Cyber Security	Cyber security using big data analytics for real-time assessment and response to cyber threats.
Splunk	Splunk is a vendor which allows for the capture, storage, aggregation, and analysis of machine data.
Cost_5, Revenue_5	Costs represent the one-time purchase of a Splunk perpetual license and staff to conduct analysis. Revenues could be calculated approximating operational losses which may occur from cyber incidents. Cost_5=$150000, Revenue_5=$100000
Credit Risk Management	Use of big data in credit risk analysis beyond current capabilities allowing for better modeling of such risks.
SAS Risk Modeling	SAS is a software service provider that allows for modeling used in risk management.
Cost_6, Revenue_6	Costs represent the SAS risk modeling product annual fee plus personnel for analytics while the revenue generated may be viewed as the savings generated from resulting and in a reduction in lending that results in defaulted loans annually. Cost_6=$58700, Revenue_6=$90000
Fraud Detection	Optimization of fraud detection using big data analytics platforms and service providers.
SEON	SEON offers cloud services for the automation of fraud detection.
Cost_7, Revenue_7	Costs represent the annual fees for maintaining SEON account services while revenue can include both the reduction in costs of fraud detection operations and a reduction in fraud events. Costs_7=$14400, Revenue_7=$75000

The time frame in cost/revenue values are obtained can be assigned as needed by the SME, for example those tabulated over or a year or a month are equal valid for assignment within the model. The ROI calculations for the banking company based on the values from Table 3 and 4 are

$$\text{ROI Value } = \frac{(\$325000-\$230000)}{(\$230000)} = 41.30\%. \quad (2)$$

A. Findings

When evaluated by the varying test cases the model's abstract and generalized nature allows it to be applied to each test case with the capability to encapsulate multiple business processes. The model remained agnostic of any business domain, business process, or big data application. The inclusion of more specific use case businesses cases in addition to the common business processes across domains also demonstrated that the model was capable of being horizontally expanded as needed to encompass further branches without those incorporations acting as an impediment to the functionality of the previously established instantiations of the elements.

Regarding the testing of the two hypotheses. It is found that when evaluated by the test cases the model can be applied and can yield an ROI value calculation. The ROI model was able to be logically integrated into the artifact. The artifact with the ROI model integration could receive inputs to its elements and yield a ROI value. The findings of the study where that H_1 was found to be true and that the ROI model could be incorporated into a process guide and that the evaluation of the process guide and the artifact remained logically consistent and functional when differing and varying inputs were applied. H_0 was rejected as it is predicated on the premise that the ROI model could either not be integrated into the artifact or if integrated could not provide a means of assessment. As previously stated, the ROI model was capably integrated into the artifact and that evaluation of the artifact yielded calculable ROI results leading to the rejection of H_0.

The test cases were developed using specified business processes and big data applications for concrete evaluation of the model, but these were not derived from case studies or real-life application of the model. Arguably a test case encapsulating any potential combination of business process and big data application within any business domain could be synthesized and applied in the evaluation of the model and the model still be applicable and yield a return on investment (ROI) assessment. In test cases 1 and 3, values were assigned that resulted in a positive ROI, however, this does not have to be the case. The artifact can just as easily be used to evaluate big data solutions which are not yielding a positive ROI such as in test case 2 and still provide the benefit to SMEs in developing a holistic understanding of big data efforts. The artifact by design is intended to be capable of being applied and evaluated in a generalized fashion with some form of ROI being yielded. The artifact is also relatively simplistic by design so that it may be used within SMEs by non-technical individuals. The generalized and simplistic nature of the artifact allows for SME's who may lack resources, such as technical personnel, the ability to easily take specific parameters of their business processes and chosen technologies and input them into the generic structure of the artifact. In this study, the model evaluation and data generation take place under hypothetical parameters. However, whether the parameters set forth by a test case are derived from actual case studies or are hypothetical is arbitrary for the scope of this study as the relevant factor is not the values and parameters applied to the model but that values and parameters can be applied to the model, and the model yield a result.

In addition to allowing for the calculation of return on investment the model also provides a framework by which companies can identify redundancies in business processes and big data tools/solutions. Identifying redundancies in both processes and tools may provide SMEs the opportunity to aggregate and consolidate resources needed for big data projects reducing the costs of those projects. For example, within the Transport Company example both the big data tools/solutions used for the business processes of data-driven decision making which are Tableau and logistics management which is SaaS Business Intelligence Tools are business intelligence tools. It may be feasible in such a scenario that only one of the software is needed to support both functions. The artifact may provide the means to not only identify such reductions in overhead post big data projects but may also provide a means of allowing for pre-big data project planning and projecting reducing the initial costs of those projects.

The artifact also exposed that many of the operations that may be conducted by SMEs for varying businesses processes may have interrelated elements and possibly even dependencies with one another. Particularly related to the more common functionalities found across many of the SME industries may be the relationships between CRM systems and enterprise management. In the three test cases, analyses from multiple business processes were presented as drivers to a better executive decision. The drivers for data-driven decision-making stemmed from incites gained through the various analytics conducted from other businesses processes allowing for optimization of business operations. Through recording and evaluating the inputs given to the artifact SMEs can be provided a means of developing a better understanding of how big data efforts can be incorporated for current and future enterprise efforts. The ability to better understand big data potentials within the SMEs may provide a means of extending competitiveness by more effectively leveraging the ability of SMEs to be more flexible and dynamic within their markets.

Results of the evaluation of the artifact show that it has the capability of allowing SMEs to better understand what elements contribute as success factors to big data efforts and what efforts detract from them within the context of calculating ROI. In a second test case for a retail company value resulting in a negative ROI where assigned. The second test case demonstrates that the investments made do not make business sense according to the ROI evaluation. The retail company may elect to remove or consolidate costly big data elements to achieve a positive ROI. For example, the retail company Cost_2, Cost_3, and Cost_4 include data analytics personnel wages. The cost/revenue pairs of Cost_2 = $42000, Revenue_2 = $15000 and Cost_4 = $42000, Revenue_4 = $25000 negatively impact the ROI while the pair Cost_3 = $55000, Revenue_3 = $90000 positively impact the ROI. The retail store may wish to remove big data efforts for the cost/revenue pairs of Cost_2/Revenue_2 and Cost_4/Revenue_4 while retaining the pair Cost_3/Revenue_3. This demonstrates the artifacts ability to intuitively and simply expose the value of big data efforts and allow for decision making based upon quantitative evaluations.

VI. Conclusion

This study was conducted to develop and evaluate a process artifact used for the analysis of value creation of big data analytics and to test the validity in the use of the ROI model within the artifact. The artifact was evaluated using hypothetical test cases. The findings of the study showed that the ROI models use within the artifact were valid and that the artifact's internal logic remained consistent when varying inputs given using differing test case scenarios were applied. It was found that the abstract nature of the artifact allowed for formulaic evaluation of scenarios and not be inhibited by specific parameters. The applicability of the study provided benefits both in the form of academic and practical applications. Academically the study furthered the body of knowledge as it relates to the value creation of big data, more specifically the study furthered investigations into understanding big data elements and relationships within the context of applying the ROI model to a process artifact. The practical contribution of the study is established through the development of an artifact that can be freely utilized by small and medium enterprises to develop a further understanding of their specific systems and per their specific use cases due to the artifact's abstract nature. The artifact produced within this study exists within the early development portion of the design life cycle. Future research efforts through further design science methodologies and case studies could extend the artifact. Further extension of the artifact through the methodologies of design science and case studies can expand the body of knowledge relating to big data value creation through mapping relationships of big data elements both internally to one another and those relationships existing between the elements and external factors.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

R. Grizzle identified the need of developing a graphic model to simplify the analysis of value creation of the big data analytics tools used in various business functions within small to medium-size businesses. Both authors worked together to develop the graphic model. R. Grizzle created the test cases to validate the usability of the model, and wrote the paper, while Y. Qu supervised and suggested improvements to it. Both authors had approved the submitted version.

References

[1] B. Chroneos-Krasavac, J. Soldic-Aleksic, and G. Petkovic, "The big data phenomenon: The business and public impact," *Industrija,* vol. 44, no. 2, pp. 117-144, 2016.

[2] F. P. S. Surbakti, W. Wang, M. Indulska, and S. Sadiq, "Factors influencing effective use of big data: A research framework," *Information & Management,* vol. 57, no. 1, p. 103146, 2020/01/01/ 2020, doi: https://doi.org/10.1016/j.im.2019.02.001.

[3] R. Patgiri and A. Ahmed, "Big Data: The V's of the Game Changer Paradigm," in *Proc. 2016 IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, 12-14 Dec. 2016 2016, pp. 17-24.

[4] M. M. Bersani, F. Marconi, D. A. Tamburri, A. Nodari, and P. Jamshidi, "Verifying big data topologies by-design : a semi-automated approach," *Journal of big data,* vol. 6, no. 1, pp. 1-23, May 2019 2019-05-18 2019, doi: http://dx.doi.org/10.1186/s40537-019-0199-y.

[5] H. Ekbia *et al.*, "Big data, bigger dilemmas: A critical review," *Journal of the Association* for *Information Science and Technology,* vol. 66, no. 8, pp. 1523-1545, 2015.

[6] A. De Mauro, M. Greco, and M. Grimaldi, "A formal definition of Big Data based on its essential features," *Library Review*, vol. 65, no. 3, pp. 122-135, 2016 2019-12-20 2016, doi: http://dx.doi.org/10.1108/LR-06-2015-0061.

[7] P. Michalik, J. Štofa, and I. Zolotová, "Concept definition for big data architecture in the education system," in *Proc. 2014 IEEE 12th International Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 23-25 Jan. 2014 2014, pp. 331-334, doi: 10.1109/SAMI.2014.6822433.

[8] P. Russom, "Big data analytics," *TDWI best practices report, fourth quarter*, vol. 19, no. 4, pp. 1-34, 2011.

[9] D. Laney, "3D data management: Controlling data volume, velocity and variety," *META group research note*, vol. 6, no. 70, p. 1, 2001.

[10] A. Abbasi, S. Sarker, and R. H. Chiang, "Big data research in information systems: Toward an inclusive research agenda," *Journal of the Association for Information Systems*, vol. 17, no. 2, p. 3, 2016.

[11] Y. Demchenko, P. Grosso, C. D. Laat, and P. Membrey, "Addressing big data issues in scientific data infrastructure," in *Proc. 2013 International Conference on Collaboration Technologies and Systems (CTS)*, 20-24 May 2013 2013, pp. 48-55, doi: 10.1109/CTS.2013.6567203.

[12] M. A. U. D. Khan, M. F. Uddin, and N. Gupta, "Seven v's of big data understanding big data to extract value," in *Proc. 2014 Zone 1 Conference of the American Society for Engineering Education*, 3-5 April 2014 2014, pp. 1-5, doi: 10.1109/ASEEZone1.2014.6820689.

[13] Y. Demchenko, E. Gruengard, and S. Klous, "Instructional Model for Building Effective Big Data Curricula for Online and Campus Education," in *Proc. 2014 IEEE 6th International Conference on Cloud Computing Technology and Science*, 15-18 Dec. 2014 2014, pp. 935-941, doi: 10.1109/CloudCom.2014.162.

[14] O. Ylijoki and J. Porras, "Perspectives to definition of big data: a mapping study and discussion," *Journal of Innovation Management*, vol. 4, no. 1, pp. 69-91, 2016.

[15] H. K. Yaseen and A. M. Obaid, "Big Data: Definition, Architecture & Applications," *JOIV: International Journal on Informatics Visualization*, vol. 4, no. 1, 2020.

[16] M. S. Hajirahimova and A. S. Aliyeva, "About big data measurement methodologies and indicators," *International Journal of Modern Education and Computer Science*, vol. 9, no. 10, p. 1, 2017.

[17] A. Mohammad, H. Mcheick, and E. Grant, "Big data architecture evolution: 2014 and beyond," presented at the Proceedings of the fourth ACM international symposium on Development and analysis of intelligent vehicular networks and applications, Montreal, QC, Canada, 2014.

[18] J.-H. Yu and Z.-M. Zhou, "Components and Development in Big Data System: A Survey," *Journal of Electronic Science and* Technology, vol. 17, no. 1, pp. 51-72, 2019/03/01/ 2019, doi: https://doi.org/10.11989/JEST.1674-862X.80926105.

[19] Z. Wu *et al.*, "Towards building a scholarly big data platform: Challenges, lessons and opportunities," in *IEEE/ACM Joint Conference on Digital Libraries*, 8-12 Sept. 2014 2014, pp. 117-126, doi: 10.1109/JCDL.2014.6970157.

[20] F. Matsebula and E. Mnkandla, "A big data architecture for learning analytics in higher education," in *Proc. 2017 IEEE AFRICON*, 18-20 Sept. 2017 2017, pp. 951-956, doi: 10.1109/AFRCON.2017.8095610.

[21] Q. Yao, Y. Tian, P.-F. Li, L.-L. Tian, Y.-M. Qian, and J.-S. Li, "Design and development of a medical big data processing system based on Hadoop," *Journal of Medical Systems*, vol. 39, no. 3, pp. 23-23, 2015, doi: 10.1007/s10916-015-0220-8.

[22] Y. Tu, Y. Lu, G. Chen, J. Zhao, and F. Yi, "Architecture Design of Distributed Medical Big Data Platform Based on Spark," in *Proc. 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, 24-26 May 2019 2019, pp. 682-685, doi: 10.1109/ITAIC.2019.8785620.

[23] O. C. Christina and S. Kelly, "Facilitating knowledge management through filtered big data: SME competitiveness in an agri-food sector," (in English), *Journal of Knowledge Management*, vol. 21, no. 1, pp. 156-179, 2017 2020-11-17 2017, doi: http://dx.doi.org/10.1108/JKM-08-2016-0357.

[24] M. Gupta and J. F. George, "Toward the development of a big data analytics capability," *Information & Management*, vol. 53, no. 8, pp. 1049-1064, 2016/12/01/ 2016, doi: https://doi.org/10.1016/j.im.2016.07.004.

[25] A. Botchkarev, P. Andru, and R. Chiong, "A Return on Investment as a Metric for Evaluating Information Systems: Taxonomy and Application," *Interdisciplinary Journal of Information, Knowledge & Management*, vol. 6, 2011.

[26] M. Mogollon and M. Raisinghani, "Measuring roi in e-business: A practical approach," *Information Systems Management*, Article vol. 20, no. 2, p. 63, Spring2003 2003, doi: 10.1201/1078/43204.20.2.20030301/41472.10.

[27] J. P. Shim, A. M. French, C. Guo, and J. Jablonski, "Big data and analytics: Issues, solutions, and ROI," *Communications of the Association for Information Systems*, vol. 37, no. 1, p. 39, 2015.

[28] N. Abdullah, S. A. Ismail, S. Sophiayati, and S. M. Sam, "Data quality in big data: a review," *International Journal of Advances in Soft Computing and its Applications*, vol. 7, no. 3, pp. 17-27, 2015.

[29] J. Gao, A. Koronios, and S. Selle, "Towards a process view on critical success factors in big data analytics projects," in *Proc. Twenty-First Americas Conference on Information Systems*, Puerto Rico, 2015, pp. 13-15.

[30] M. Nasrollahi, J. Ramezani, and M. Sadraei, "The Impact of Big Data Adoption on SMEs' Performance," (in English), *Big Data and Cognitive Computing*, vol. 5, no. 4, p. 68, 2021 2021-12-23 2021, doi: http://dx.doi.org/10.3390/bdcc5040068.

[31] S. Shah, C. B. Soriano, and A. Coutroubis, "Is big data for everyone? the challenges of big data adoption in SMEs," in *Proc. 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2017: IEEE, pp. 803-807.

[32] M. Iqbal, S. H. A. Kazmi, A. Manzoor, A. R. Soomrani, S. H. Butt, and K. A. Shaikh, "A study of big data for business growth in SMEs: Opportunities & challenges," in *Proc. 2018 International* Conference *on Computing, Mathematics and Engineering Technologies (iCoMET)*, 3-4 March 2018 2018, pp. 1-7, doi: 10.1109/ICOMET.2018.8346368.

[33] A. Soroka, Y. Liu, L. Han, and M. S. Haleem, "Big Data Driven Customer Insights for SMEs in Redistributed Manufacturing," *Procedia CIRP*, vol. 63, pp. 692-697, 2017/01/01/ 2017, doi: https://doi.org/10.1016/j.procir.2017.03.319.

[34] F. Pirola, C. Cimini, and R. Pinto, "Digital readiness assessment of Italian SMEs: a case-study research: IMS," *Journal* of *Manufacturing Technology Management*, vol. 31, no. 5, pp. 1045-1083, 2020 2021-06-03 2020, doi: http://dx.doi.org/10.1108/JMTM-09-2018-0305.

[35] W. Noonpakdee, A. Phothichai, and T. Khunkornsiri, "Big data implementation for small and medium enterprises," in *Proc. 2018 27th Wireless and Optical Communication Conference (WOCC)*, 30 April-1 May 2018 2018, pp. 1-5, doi: 10.1109/WOCC.2018.8372725.

[36] M. Sune Dueholm and P. Jensen, "Big data in the Danish industry: application and value creation," (in English), *Business Process Management Journal*, vol. 23, no. 3, pp. 645-670, 2017 2021-03-17 2017, doi: http://dx.doi.org/10.1108/BPMJ-01-2016-0017.

[37] M. Willetts, A. S. Atkins, and C. Stanier, "Barriers to SMEs Adoption of Big Data Analytics for Competitive Advantage," in *Proc. 2020* Fourth *International Conference On Intelligent Computing in Data Sciences (ICDS)*, 21-23 Oct. 2020 2020, pp. 1-8, doi: 10.1109/ICDS50568.2020.9268687.

[38] L. Rajabion, "Application and Adoption of Big Data Technologies in SMEs," in *Proc. 2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, 12-14 Dec. 2018 2018, pp. 1133-1135, doi: 10.1109/CSCI46756.2018.00219.

[39] S. K. Mangla, R. Raut, V. S. Narwane, Z. Zuopeng, and P. priyadarshinee, "Mediating effect of big data analytics on project performance of small and medium enterprises," (in English), *Journal of Enterprise Information Management*, vol. 34, no. 1, pp. 168-198, 2021 2021-12-31 2021, doi: http://dx.doi.org/10.1108/JEIM-12-2019-0394.

[40] N. R. Vajjhala and E. Ramollari, "Big data using cloud computing-opportunities for small and medium-sized enterprises," *European Journal of Economics and Business Studies*, vol. 2, no. 1, pp. 129-137, 2016.

[41] D. Kumar, S. Harsh Vardhan, and P. Verma, "Exploring suitability of cloud computing for small and medium-sized enterprises in India," (in English), *Journal of Small Business and Enterprise Development*, vol. 24, no. 4, pp. 814-832, 2017 2021-09-09 2017, doi: http://dx.doi.org/10.1108/JSBED-01-2017-0002.

[42] C. A. Ardagna, P. Ceravolo, and E. Damiani, "Big data analytics as-a-service: Issues and challenges," in *Proc. 2016 IEEE International Conference on Big Data (Big Data)*, 5-8 Dec. 2016 2016, pp. 3638-3644, doi: 10.1109/BigData.2016.7841029.

[43] V. Marx, "The big challenges of big data," *Nature*, vol. 498, p. 255, 06/12/online 2013, doi: 10.1038/498255a.

[44] S. Kaisler, F. Armour, J. A. Espinosa, and W. Money, "Big Data: Issues and Challenges Moving Forward," in *Proc. 2013 46th Hawaii International Conference on System Sciences*, 7-10 Jan. 2013 2013, pp. 995-1004, doi: 10.1109/HICSS.2013.645.

[45] T. Rabl, S. Gómez-Villamor, M. Sadoghi, V. Muntés-Mulero, H.-A. Jacobsen, and S. Mankovskii, "Solving big data challenges for enterprise application performance management," in *Proc. VLDB Endowment*, vol. 5, no. 12, pp. 1724–1735, 2012, doi: 10.14778/2367502.2367512.

[46] S. LaValle, E. Lesser, R. Shockley, M. S. Hopkins, and N. Kruschwitz, "Big Data, Analytics and the Path From Insights to Value," (in English), *MIT Sloan Management Review,* vol. 52, no. 2, pp. 21-32, Winter Winter 2011 2017-11-17 2011.
[47] M. E. Porter and C. Advantage, "Creating and sustaining superior performance," *Competitive Advantage,* vol. 167, pp. 167-206, 1985.
[48] H. G. Miller and P. Mork, "From data to decisions: A value chain for big data," *It Professional,* vol. 15, no. 1, pp. 57-59, 2013.
[49] H. Hu, Y. Wen, T.-S. Chua, and X. Li, "Toward scalable systems for big data analytics: A technology tutorial," *IEEE access,* vol. 2, pp. 652-687, 2014.
[50] M. Chen, S. Mao, and Y. Liu, "Big data: A survey," *Mobile networks and applications,* vol. 19, no. 2, pp. 171-209, 2014.
[51] N. Rahman and F. Aldhaban, "Assessing the effectiveness of big data initiatives," in *Proc. 2015 Portland International Conference on Management of Engineering and Technology (PICMET)*, 2015: IEEE, pp. 478-484.
[52] P. Mikalef, I. O. Pappas, J. Krogstie, and P. A. Pavlou, "Big data and business analytics: A research agenda for realizing business value," *Information & Management,* vol. 57, no. 1, p. 103237, 2020/01/01/ 2020, doi: https://doi.org/10.1016/j.im.2019.103237.
[53] P. Mikalef, M. Boura, G. Lekakos, and J. Krogstie, "Big data analytics and firm performance: Findings from a mixed-method approach," *Journal of Business Research,* vol. 98, pp. 261-276, 2019/05/01/ 2019, doi: https://doi.org/10.1016/j.jbusres.2019.01.044.
[54] D. Sen, M. Ozturk, and O. Vayvay, "An overview of big data for growth in SMEs," *Procedia-Social and Behavioral Sciences,* vol. 235, pp. 159-167, 2016.
[55] Q. Shabbir Muhammad and W. Gardezi Syed Babar, "Application of big data analytics and organizational performance: the mediating role of knowledge management practices," *Journal of big data,* vol. 7, no. 1, Dec 2020 2020-08-03 2020, doi: http://dx.doi.org/10.1186/s40537-020-00317-6.
[56] P. Maroufkhani, I. Wan Khairuzzaman Wan, and M. Ghobakhloo, "Big data analytics adoption model for small and medium enterprises," (in English), *Journal of Science and Technology Policy Management,* vol. 11, no. 4, pp. 483-513, 2020 2021-05-28 2020, doi: http://dx.doi.org/10.1108/JSTPM-02-2020-0018.
[57] A. R. Hevner, "A three cycle view of design science research," Scandinavian *Journal of Information Systems,* vol. 19, no. 2, p. 4, 2007.
[58] B. Jan vom, A. Simons, K. Riemer, B. Niehaves, R. Plattfaut, and A. Cleven, "Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research," *Communications of the Association for Information Systems,* vol. 37, p. 9, 2015 2021-03-12 2015, doi: http://dx.doi.org/10.17705/1CAIS.03709.
[59] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," *MIS Quarterly,* vol. 28, no. 1, pp. 75-105, Mar 2004 2015-11-30 2004.
[60] K. Dam Nguyen Anh, T. Le Dinh, and W. Menvielle, "A systematic literature review of big data adoption in internationalization," (in English), *Journal of Marketing Analytics,* vol. 7, no. 3, pp. 182-195, Sep 2019 2020-11-17 2019, doi: http://dx.doi.org/10.1057/s41270-019-00054-7.
[61] M. Mattera, "SMEs transformation through usage and understanding of big data case study: Spanish restaurant industry," in *Proc. 2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA)*, 9-12 March 2018 2018, pp. 186-189, doi: 10.1109/ICBDA.2018.8367674.
[62] P. W. Handayani and W. Lisdianingrum, "Impact analysis on free online marketing using social network Facebook: Case study SMEs in Indonesia," in *Proc. 2011 International Conference on Advanced Computer Science and Information Systems*, 17-18 Dec. 2011 2011, pp. 171-176.
[63] J. W. Peltier and G. M. Naidu, "Social networks across the SME organizational lifecycle," *Journal of Small Business and Enterprise Development,* vol. 19, no. 1, pp. 56-73, 2012 2021-09-09 2012, doi: http://dx.doi.org/10.1108/14626001211196406.
[64] S. Robinson and H. A. Stubberud, "Social Networks And Entrepreneurial Growth," *International Journal of Management and Information Systems,* vol. 15, no. 4, pp. 65-70, Fourth Quarter Fourth Quarter 2011 2021-09-10 2011.
[65] J. Heilala, H. Helaakoski, R. Kuivanen, J. Kääriäinen, and L. Saari, "A review of digitalisation in the Finnish manufacturing SME companies," 2020.
[66] D. M. Schaeffer and P. C. Olson, "Big data options for small and medium enterprises," *The Review of Business* Information *Systems (Online),* vol. 18, no. 1, p. 41, 2014 2020-11-18 2014.
[67] A. Topalović and A. Azzini, "Data Mining Applications in SMEs: An Italian Perspective," *Business Systems Research,* vol. 11, no. 3, pp. 127-146, 2020 2021-09-10 2020, doi: http://dx.doi.org/10.2478/bsrj-2020-0031.
[68] B. Ogbuokiri, C. Udanor, and M. Agu, "Implementing bigdata analytics for small and medium enterprise (SME) regional growth," *IOSR Journal of Computer Engineering,* vol. 17, no. 6, pp. 35-43, 2015.
[69] R. S. Kalan and M. O. Ünalir, "Leveraging big data technology for small and medium-sized enterprises (SMEs)," in *Proc. 2016 6th International Conference on Computer and Knowledge Engineering (ICCKE)*, 20-20 Oct. 2016 2016, pp. 1-6, doi: 10.1109/ICCKE.2016.7802106.
[70] J. Hopkins and P. Hawking, "Big Data Analytics and IoT in logistics: a case study," (in English), *International Journal of Logistics Management,* vol. 29, no. 2, pp. 575-591, 2018 2021-09-10 2018, doi: http://dx.doi.org/10.1108/IJLM-05-2017-0109.
[71] G. Santoro, F. Fiano, B. Bertoldi, and F. Ciampi, "Big data for business management in the retail industry," *Management Decision,* vol. 57, no. 8, pp. 1980-1992, 2019 2021-09-09 2019, doi: http://dx.doi.org/10.1108/MD-07-2018-0829.

Ryan Grizzle is currently a system development specialits at Pantex Plant in Panhandle, Texas, USA. He has over seven years of systems planning, integration, and management. He received his Bachelor of Science in Electronics Engineering Technology from DeVry University, USA; Masters of Science in Software Engineering from Regis University, USA, and Doctorate of Computer Science from Colorado Technical University, USA.

Dr. Grizzle’s primary research interests include modeling and optimization of systems and system processes, big data analytics, and software engineering.

Yanzhen Qu received his B.Eng. degree in Electronic Engineering from Anhui University, China; M.Eng. degree in Electrical Engineering from the Chinese Science Academy, China; and Ph.D. degree in Computer Science from Concordia University, Canada.

Over his industrial professional career, he has served at the various executive level management positions responsible for Product R&D and the IT Operation at several multinational corporations. He has led his multinational engineering teams successfully developed several the world first very large real-time commercial systems and technologies. He is currently the Dean and a Professor of Computer Science, Engineering and Technology with Colorado Technical University, Colorado Springs, CO, USA. He is also the dissertation supervisor of many Computer Science doctoral students. He and his doctoral students have published several dozen scholarly articles, some of them received the best paper award at several IEEE international conferences.

Dr. Qu’s current research interests include data modeling, data science, cybersecurity and privacy, machine learning, e-learning technologies, software engineering, cloud computing, and affective computing. He has served as either the general chair, or the program chair, or a keynote speaker at many IEEE, ACM, IFIP international conferences. He is also an editorial board member of several peer-reviewed professional journals.