BIG DATA ENGINEERING AND CLOUD COMPUTING ADOPTION MODEL IN UNITED ARAB EMIRATES LARGE BUSINESS ORGANIZATIONS

WALEED SAEED MAHMOUD MAHMOUD ALI

DOCTOR OF PHILOSOPHY

2023

Faculty of Information and Communication Technology

BIG DATA ENGINEERING AND CLOUD COMPUTING ADOPTION MODEL IN UNITED ARAB EMIRATES LARGE BUSINESS ORGANIZATIONS

Waleed Saeed Mahmoud Mahmoud Ali

Doctor of Philosophy

2023

BIG DATA ENGINEERING AND CLOUD COMPUTING ADOPTION MODEL IN UNITED ARAB EMIRATES LARGE BUSINESS ORGANIZATIONS

WALEED SAEED MAHMOUD MAHMOUD ALI

A thesis submitted
in fulfilment of the requirements for the degree of Doctor of Philosophy

Faculty of Information and Communication Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2023

DECLARATION

I declare that this thesis entitled “Big Data Engineering and Cloud Computing Adoption Model in United Arab Emirates Large Business Organizations” is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in the candidature of any other degree.

Signature : ……..

Name : Waleed Saeed Mahmoud Mahmoud Ali

Date : ……25/09/2023……

APPROVAL

I hereby declare that I have read this thesis and in my opinion, this thesis is sufficient in terms of scope and quality for the award of Doctor of Philosophy.

Signature :……

Supervisor Name : Professor Dr. Abd Samad bin Hasan Basari

Date :……25/09/2023……..……

DEDICATION

I am eternally grateful to all of my family members, especially my parents, for their unwavering love, support, and encouragement. There are no words to adequately express my gratitude to my parents for their sensitivity, guidance, and love. Not just while I was pursuing my Ph.D. But it's been that way for the rest of my life. My thanks also extend to my family for their unwavering support. I am grateful to all three of my children, Khaled, Zaidedine, and Omar, for motivating me to finish this Degree.

ABSTRACT

The adoption of cloud computing and Big Data Engineering (BDE) in large business organizations in the United Arab Emirates (UAE) can bring numerous benefits, including increased efficiency, flexibility, and scalability of computing resources. Many researchers have proposed that BDE can drive Cloud Computing Adoption (CCA). However, there is a need for more empirical research to determine the extent to which BDE drives CCA in practice. This research could help organizations better understand the relationship between these technologies and develop more effective strategies for adoption. Despite the potential benefits of BDE and CCA in large business entities in the UAE, there is a lack of research or studies on the topic. This has resulted in a limited representation of the adoption of these technologies in UAE. As a result, it is important for organizations to carefully evaluate their readiness for adoption, develop a clear strategy, and seek guidance from experienced service providers. Since developing technologies have strong ties to CCA and BDE, CCA has taken place in both hypothetical and business situations. Due to the technological advancements and changes in the current business landscape, it is important to examine how BDE can affect CCA and its implications. In order to stay relevant in BDE, business organizations must produce advanced results at every level of their organization. The CCA impact model was developed using BDE factors and two widely established models namely the Technology Acceptance Model (TAM) and Technology Organization-Environment (TOE). The CCA was extended by including variables related to BDE. Six independent variables were examined: usefulness, ease of use, security effectiveness, cost-effectiveness, intention to use and need for Big Data technology. A sample size of 250 was used to collect data from large business organizations in the UAE. After data cleaning and removing missing values, the sample size was 204. The data were analyzed using binary logistic regression. In the current business climate, it is important to examine BDE's impact on CCA, as well as the longer-term implications of BDE and CCA on organizations. BDE requires organizations to produce advanced results at every level of their organization. Based on the findings, CCA is predicted by perceived ease of use, perceived usefulness, security effectiveness, intention to use Big Data, and need to utilize Big Data technology and cannot be predicted by cost-effectiveness. The model correctly predicts whether the participant companies would use cloud computing 90.6% of the time and explained 9.4% of the variance in cloud computing adoption. The results of the study are believed can help managers decide whether to adopt cloud computing.

MODEL PENGGUNAAN KEJURUTERAAN DATA RAYA DAN PENGKOMPUTERAN AWAN DALAM ORGANISASI PERNIAGAAN BESAR DI EMIRIAH ARAB BERSATU

ABSTRAK

Penggunaan pengkomputeran awan dan Kejuruteraan Data Raya (BDE) dalam organisasi perniagaan besar di Emiriah Arab Bersatu (UAE) boleh membawa banyak faedah, termasuk peningkatan kecekapan, fleksibiliti dan kebolehskalaan sumber pengkomputeran. Ramai penyelidik telah mencadangkan bahawa BDE boleh memacu Penggunaan Pengkomputeran Awan (CCA). Walau bagaimanapun, terdapat keperluan untuk penyelidikan yang lebih empirikal untuk menentukan sejauh mana BDE memacu CCA dalam amalan. Penyelidikan ini boleh membantu organisasi memahami dengan lebih baik hubungan antara teknologi ini dan membangunkan strategi yang lebih berkesan untuk diterima pakai. Walaupun potensi manfaat BDE dan CCA dalam entiti perniagaan besar di UAE, terdapat kekurangan penyelidikan atau kajian mengenai topik tersebut. Ini telah menyebabkan representasi terhad penggunaan teknologi ini di UAE. Akibatnya, adalah penting bagi organisasi untuk menilai dengan teliti kesediaan mereka untuk diterima pakai, membangunkan strategi yang jelas dan mendapatkan bimbingan daripada penyedia perkhidmatan yang berpengalaman. Memandangkan teknologi membangun mempunyai hubungan yang kukuh dengan CCA dan BDE, CCA telah berlaku dalam kedua-dua situasi hipotetikal dan perniagaan. Disebabkan oleh kemajuan teknologi dan perubahan dalam landskap perniagaan semasa, adalah penting untuk mengkaji cara BDE boleh menjejaskan CCA dan implikasinya. Untuk kekal relevan dalam BDE, organisasi perniagaan mesti menghasilkan keputusan lanjutan di setiap peringkat organisasi mereka. Model impak CCA telah dibangunkan menggunakan faktor BDE dan dua model yang digunakan secara meluas iaitu Model Penerimaan Teknologi (TAM) dan Teknologi Organisasi-Persekitaran (TOE). CCA telah diperluaskan dengan memasukkan pembolehubah yang berkaitan dengan BDE. Enam pembolehubah tidak bersandar telah diperiksa: kebergunaan, kemudahan penggunaan, keberkesanan keselamatan, keberkesanan kos, niat untuk digunakan dan keperluan untuk teknologi Data Raya. Saiz sampel sebanyak 250 telah digunakan untuk mengumpul data daripada organisasi perniagaan besar di UAE. Selepas pembersihan data dan mengeluarkan nilai yang hilang, saiz sampel ialah 204. Data dianalisis menggunakan regresi logistik binari. Dalam iklim perniagaan semasa, adalah penting untuk mengkaji kesan BDE terhadap CCA, serta implikasi jangka panjang BDE dan CCA terhadap organisasi. BDE memerlukan organisasi untuk menghasilkan keputusan lanjutan pada setiap peringkat organisasi mereka. Berdasarkan penemuan, CCA diramalkan oleh persepsi kemudahan penggunaan, persepsi kegunaan, keberkesanan keselamatan, niat untuk menggunakan Data Raya, dan perlu menggunakan teknologi Data Raya dan tidak boleh diramalkan dengan keberkesanan kos. Model ini dengan betul meramalkan sama ada syarikat peserta akan menggunakan 90.6% masa pengkomputeran awan dan menjelaskan 9.4% varians dalam penggunaan pengkomputeran awan. Hasil kajian dipercayai boleh membantu pengurus membuat keputusan sama ada untuk mengguna pakai pengkomputeran awan.

ACKNOWLEDGEMENTS

I'd want to convey my deepest respect, gratitude, appreciation, and genuine gratitude to my main supervisor, Professor Dr. Abd Samad bin Hasan Basari, for his persistent direction, support, and encouragement, all of which have benefited me greatly. His guidance equipped me with the resources and abilities I needed to complete my research work. In addition, I'd like to express my gratitude to my co supervisor, Associate Professor Ts. Dr. Zeratul Izzah binti Mohd Yusoh, for her assistance. After that, I'd like to express my gratitude to the numerous people who assisted me with my research by providing aid, support, and encouragement. All companies and individuals in the UAE who contributed to the empirical research by sharing their experiences and corporations in supplying the necessary data for the completion of this thesis must be thanked.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

AI	-	Artificial Intelligence
API	-	Application Programming Interface
ATU	-	Attitude Towards Using
AVE	-	Average Variance Extracted
BDE	-	Big Data Engineering
BI	-	Business Intelligence
BIM	-	Building Information Modelling
BIU	-	Behavioural Intention to Use
CC	-	Cloud Computing
CCA	-	Cloud Computing Adoption
CCM	-	Cloud Computing Model
CIOs	-	Chief Information Officers
CMR	-	Customer-Managed Relationship
CPR	-	Computer-based Patient Record
CPU	-	Central Processing Unit
CSP	-	Cloud Service Provider
DOI	-	Diffusion of Innovation
DSV	-	Discriminant Validity
ECR	-	Earnings Credit Rate

EE	-	Effort Expectancy
EHRs	-	Electronic Health Record
EMR	-	Electronic Medical Record
EMRS	-	Emergency Medical Rescue Services
ETL	-	Extract, Transform, and Load
GCC	-	Gulf Cooperation Council
GPU	-	Graphical Processing Unit
HDFS	-	Hadoop Distributed File System
IaaS	-	Infrastructure-as-a-Service
IBA	-	Information Base Applications
IBT	-	Institution-Based Trust
ICT	-	Information and Communication Technology
IDC	-	International Data Corporation
IoT	-	Internet of Things
IMD	-	In-Memory Databases
JVM	-	Java Virtual Machines
MM	-	Motivational Model
NBD	-	National Bank of Dubai
NIST	-	Institute of Standards and Technology
PaaS	-	Platform-as-a-Service
PE	-	Performance Expectancy
PEU	-	Perceived Ease of Use
PHR	-	Personal Health Record
PMRI	-	Patient Medical Record Information

PU	-	Perceived Usefulness
RDBMS	-	Relational Database Management System
SaaS	-	Software-as-a-Service
SD	-	Standard Deviation
TAM	-	Technology Acceptance Model
TOE	-	Technology Organization Environment
TPA	-	Third-Party Evaluators
UAE	-	United Arab Emirates
UTAUT	-	Unified Theory of Acceptance and Use of Technology
VPN	-	Virtual Private Network
YARN	-	Yet Other Resource Arbitrator

LIST OF PUBLICATIONS

1. Ali, W. S. M. M., Basari, A. S. H. and Yusoh, Z. I. M., 2022. Logistic Regression Model to Examine the Impact of Big Data Engineering for Cloud Computing adoption in UAE. *International Journal of Engineering Trends and Technology*, 70 (12), pp. 415-420.

2. Ali, W. S. M. M., Basari, A. S. H., Yusoh, Z. I. M. and Doheir, M., 2023. Pilot Study on the Adoption of Cloud Computing in UAE Large Business Organizations, *International Journal of Engineering Trends and Technology*, 71(3), pp. 17-22.

CHAPTER 1

INTRODUCTION

1.1 Introduction

In today's world, Informational and Communication Technology (ICT), architecture, infrastructure, analysis method and data management are not improving with the growth of data. The cloud computing, big data, social media and mobile broadband all are examples of ICT transformation (Khan and Al-Yasiri, 2018). The past decade has proven that consolidating IT resources within a large organization can lead to better performance and better management of ICT and services. It all started with Service Oriented Architecture (SOA), which provides a conceptual framework for developing web services. Then came the rise of online computing, which moved to cloud platforms and clouds (Chen et al., 2020). Big Data Engineering (BDE) have become the mainstay of scientific computing. This has recently begun to change as organizations that provide online services such as search, online advertising and social media generate and analyse large volumes of data. In addition, organizations are getting more data from free and registered public sources, not only from public bodies but also from non-governmental organizations, all of which differ in structure, quality and quantity (Rajabion et al., 2019).

The convergence of big data engineering and cloud computing (CC) is mainly because the respective technology paradigms emphasize important information, flexibility and accessibility. CC is the result of the development of distributed computing technology, made possible by fast and inexpensive networking, fast-selling hardware, high-performance

optics, and interactive networking technologies (Wang et al., 2018). CC has become a powerful framework for solving large and complex IT problems and transforming the way IT infrastructure is set up and deployed. Another important goal is to use computers as a solution for processing large amounts of data. Although cloud computing is widely adopted by many organizations, the exploration of big data Engineering in the cloud is still in its infancy (Sun et al., 2018). The Big data engineering is becoming popular due to the rise of cloud computing technologies that support data-intensive computing.

According to Hosseinian-Far et al. (2018), CC and BDE work best together and have a dialectical relationship. Open big data engineering is a necessary component of the accelerating growth of the information society we live in today, and cloud computing is a trend in technological advancement. The philosophy has changed from a traditional mainframe or service model to a data and operations model. These data types and analysis methods differ from traditional information types (Vajjhala and Ramollari, 2021). This study was based on survey data gathered from the experts or IT managers of large organisations in the United Arab Emirates. The study included variables like BDE, Technology-Organizational Environment (TOE) and Technology Acceptance Model (TAM).

1.2 Background

Big data engineering is a type of data analysis tools made possible by cutting-edge systems and architectures that enable quick data gathering, storage, and analysis. Data sources involve email, output from mobile phones, sensor-developed information, and output from social media, as well as typical corporate databases (Balachandran and Prasad, 2017). Data now comprise unstructured data, or data without a defined format,as well as organised database files. Big data necessitates a large amount of storage. The resources needed to work with big data engineering may still be expensive for large enterprises, even

as storage costs continue to decline. Network Attached Storage (NAS) is the foundation of a typical big data engineering and storage architecture (Balachandran and Prasad, 2017). The design of numerous NAS "pods," each of which has a number of storage devices linked to a NAS device, is necessary for a clustered NAS architecture. Subsequently, linking up to many NAS units is deemed essential for improved data sharing and browsing. The highly effective service-based computing system known as cloud computing has altered how IT infrastructure is installed and used. “Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service are three paramount related models (SaaS). However, this idea may also be useful to regard storing as a service or a database as a service (Ali et al., 2020).

Flexibility, affordability, minimum initial investment, quickness to market, and risk transfer are a few of the key elements that make cloud computing a popular paradigm for providing unique types of applications that are not commercially viable in major organisational infrastructure setups. Decision support systems and quick database management systems for workloads with lots of changes are essential elements of cloud architecture (Ming et al., 2018). In the database research community, scalable distributed data management has been a mainstay for more than three decades. Designing scalable systems that can manage heavy workloads and selective analysis workloads has received a lot of academic attention. Early concepts included parallel database systems and distributed databases for processing heavy workloads (Raut et al., 2019). While distributed database systems were not extremely successful or commercialized where different specialised metrics were used, the parallel databases went from prototype systems to big commercial systems. Ways to increase the availability and efficiency of Hadoop systems are also included.

Large organisations, on the other hand, employ data warehouses to manage vast volumes of data. Growing data volumes cannot be handled by storage infrastructures and solutions designed around them in a timely manner (El-Haddadeh, 2020). Calculations involving vast amounts of data can take many days, whereas calculations involving modest amounts of data only take a few seconds. It is required to provide a real-time or almost real-time reaction to massive amounts of data due to the increased requirements. Traditional data warehouses struggle with data management and analysis due to the Big Data volume. Data that fail to be processed by conventional database systems are referred to as big data, indicating that either too much data is present, the values are changing too fast, or the data do not adhere to the standards of a typical database management system (such as compliance) (Singh and Mansotra, 2019). We require extra CPU and memory resources due to the size and/or speed of the data, which are given by distributed processors and cloud-based storage. Large enterprises contemplating big data analytics technology have a good alternative for data storage: cloud computing. Cloud computing is the on-demand internet access to computer services, often offered with little effort by an outside organisation (Ali, 2019).

Three components make up a cloud computing service model. These are SaaS which enable organizations to access applications via internet-based hosts without managing supporting infrastructure. Businesses can use the platform as a Service (PaaS), utilizing internet-based hosts and their own customized programs and applications, control and programming languages (Srivastava and Nanath, 2017). In conclusion, businesses can purchase Internet hosts to provide infrastructure as a service (IaaS) by utilizing cloud computing, storage, networks, and other technology resources (Alismaili et al., 2020). Distributed computing is considered to be an important area for venture capital in the field of IT. Distributed computing is viewed as another peculiarity that gives freedoms to associations by offering huge assortments of effectively open virtual PC assets, and it has