DOI:

BIG DATA ANALYTICS USAGE BY BANKING SECTOR IN EUROPEAN COUNTRIES

PHAN DAI THICH

DOCTORAL DISSERTATION

UNIVERSITY OF MISKOLC

2023

DOI:

BIG DATA ANALYTICS USAGE BY BANKING SECTOR IN EUROPEAN COUNTRIES

PHAN DAI THICH

DISSERTATION SUBMITTED TO UNIVERSITY OF MISKOLC IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Ph.D. IN ECONOMICS

HANTOS ELEMÉR DOCTORAL SCHOOL OF BUSINESS, MANAGEMENT AND REGIONAL SCIENCES

UNIVERSITY OF MISKOLC

2023

DECLARATION

I, Phan Dai Thich, hereby declare that this dissertation submitted for my Ph.D. in Economics represents my own work. All sources used in researching this dissertation are fully acknowledged, and all quotations are properly identified in the reference section.

Signed: Phan Dai Thich

Miskolc, 2023

ACKNOWLEDGEMENT

Many people have contributed to the successful completion of this thesis.

First, I would like to express my gratitude to my parents and my parents-in-law for their great support in helping me pursue my dreams. Despite the long distance, their encouragement has been a great source of motivation for me to work harder. I would like to express my appreciation to my beloved wife for her unconditional love and support throughout my 4 years of PhD study. I especially dedicate this success to my young daughters, Khanh Nhien and Khanh Dan.

Secondly, I am particularly thankful to Professor Kovács Levente for his invaluable guidance and advice throughout my program. Without his academic and personal support, I would not have been able to complete this thesis. I am also grateful to Professor Emeritus Miklós Herdon and Dr. Pál Zsolt for their enthusiastic support and insightful opinions.

I am grateful for the support of the Stipendium Hungaricum Program, the Hungarian Banking Association (HBA), and the Banking Association of Central and Eastern Europe (BACEE). Additionally, I would like to express my deep appreciation to the administrative staff and lecturers at the University of Miskolc.

Dissertation Title	Big data analytics usage by banking sector in European Countries
Author	Phan Dai Thich
Supervisor	Prof. Dr. Kovács Levente

EXECUTIVE SUMMARY

Digital transformation impacts every aspect of economies and businesses worldwide. Its influence can be seen in digital economies and e-Government, making it one of the essential pillars of a nation's development strategy. The effects of digital transformation are evident in the business world, and companies that do not embrace it could be putting their existence at risk. Digital technologies have revolutionized the business landscape by streamlining processes and creating new opportunities for communication and collaboration. With the help of digital technologies, enterprises can now access big volumes of information quickly and easily and innovate products and services faster than ever before. Among the emerging digital technologies, big data analytics is considered the game changer in business, is the fourth paradigm of science, and is the next frontier for innovation, competition, and productivity. As one of the new technologies shaping the banking sector's future, big data offers numerous benefits to banks since it enables banks to analyze customer behavior better and support better decision-making that can improve their operations.

The existing studies on big data analytics in the banking sector are very limited, but most published studies are limited to big data's benefits and prospects. Despite the importance of big data technology, there remains a paucity of empirical evidence on the understanding of the role of big data analytics (BDA) capabilities in the performance of banks. This thesis seeks to fill the gap in the literature by proposing a model that combines the resource-based view (RBV), the technology organization environment framework (TOE), and dynamic capability theory (DC). It empirically investigates the effects of BDA usage on market performance and risk management performance.

Based on the comprehensive literature review, the thesis proposed its developed model to investigate the factor influencing BDA usage and its impact on banks' performance in 27 European Union Member States (EU-27). The study utilized the most developed statistical method, partial least square–structural equation modeling (PLS-SEM), to examine the relationship between BDA infrastructure, BDA skills, big data quality, top management supports, BDA dynamic capabilities, and market performance, risk management performance. Data was collected through survey questionnaires. Finally, 114 bank-level

responses were used for the quantitative analysis. PLS-SEM was examined in both the measurement model and structural model. In assessing the measurement model, the indicator reliability, internal consistency reliability, composite reliability, construct reliability, convergent, and discriminant validity have been checked. In assessing the structural model, the results from the bootstrapped procedure released the important findings from this thesis.

The results obtained from PLS-SEM analysis show that BDA skills were the most influential factor contributing to BDA usage. On other hand, the impact of top management support, BDA infrastructure and big data quality on BDA usage was found to be insignificant. Another noteworthy finding from this study is that the relationship between BDA skills and banking performance was fully mediated by BDA usage. Furthermore, the statistical analyses revealed that BDA infrastructure had a significant direct impact on risk management performance in banks. The thesis also discovered that the utilization of BDA in banks could foster the development of innovative dynamic capabilities, subsequently, enhancing the banks' comparative advantages. It was found that BDA usage had a significant positive impact on both market performance and risk management performance.

The thesis provides the first empirical evidence of the impact of BDA usage on the banks' performance. This is the first published result demonstrating the benefits of BDA on risk management performance in the banking sector. The result of this thesis provides practical solutions to the banking sector and contributes solid evidence on the role of dynamic capabilities in the big data context.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS AND ABBREVIATIONS

AI	Artificial intelligence
API	Application programming interface
ATMs	Automated teller machines
BD	Big data
BI	Business intelligence
CB-SEM	Covariance-based Structural Equation Modeling
CR	Composite reliability
DC	Dynamic capabilities theory
DL	Deep learning
DOI	Diffusion of innovation
DT	Digital technologies
ERP	Enterprise resource planning
GoF	Goodness of Fit
HTMT	The heterotrait-monotrait ratio
ICT	Information and Communication Technologies
IoTs	Internet of Things
IT	Information technology
KBV	Knowledge-Based View
ML	Machine learning
PLS-SEM	Partial least square–structural equation modeling
RBV	Resource-based view
Scopus	Elsevier's Scopus
SMEs	Small and medium enterprises
SQL	Structured query language
TAM	Technology Acceptance Model
TOE	Technology organization environment
TTF	Task-Technology Fit Theory
UTAUT	The unified theory of acceptance and use of technology
VIF	Variance inflation factor
WoS	Clarivate Analytics' Web of Science

CHAPTER 1: INTRODUCTION

1.1 Background of the study

Digital transformation impacts every aspect of economies and businesses worldwide. Its influence can be seen in digital economies and e-Government, making it one of the essential pillars of a nation's development strategy. The future of the European economy is focused on the twin transition of green and digital transformation (European Commission, 2022). In order to achieve this objective, the European Union (EU) has set out a digital target for 2030 which emphasizes the importance of investing hundreds of Euros million into key digital technologies and enhancing residents' digital skills.

The effects of digital transformation are evident in the business world, and companies that do not embrace could be putting their existence at risk. Digital technologies (DT) have revolutionized the business landscape through streamlining processes and creating new opportunities for communication and collaboration. With the help of digital technologies, enterprises can now access big volumes of information quickly and easily, as well as innovate products and services faster than ever before. Enterprises across all industries are leveraging DT to improve productivity, reduce costs, and capitalize on the newest technology development. By using DT, enterprises are able to stay competitive and successful in business's ever-changing and fast-paced world. The impacts of digital technology can be seen across most industries, particularly manufacturing, healthcare services, transportation, and banking services.

In the banking sector, financial technology firms' growth and development are evident. With the power of new technologies such as artificial intelligence (AI), machine learning (ML), and big data (BD), these newcomers are bringing more efficient financial solutions, faster responses to customer need, and becoming a threat to traditional banks. Moreover, the financial industry faces even more competition than ever as big-tech giants like Amazon, Apple, Google, and Alibaba enter the field.

Among these emerging digital technologies, big data analytics is arising as the game changer in business (Fosso Wamba et al., 2017), as the fourth paradigm of science (Strawn, 2012), as a next frontier for innovation, competition, and productivity (Manyika et al., 2011). As one of the new technologies shaping the future of the banking sector, big data offers numerous benefits to banks since it enables banks to analyze customer behavior better and support better decision-making that can improve their operations. Big data technology can be incredibly advantageous for the banking sector since it enables banks to analyze customer behavior better and support better decision-making that can improve their operations. Banks can also utilize big data technology to detect fraud quickly and accurately and gain insights into customer preferences that can then be used to create more personalized products and services. Further, big data technology allows banks to build smarter and more efficient models to identify and target the right customers with relevant offers. Overall, big data technology is a powerful tool to help banks significantly improve their efficiency and profitability. Raguseo (2018) pointed out that the advantages of utilizing big data among studied enterprises are tremendous, and can be divided into four categories: strategic, transactional, informational, and transformational benefits. Despite the numerous advantages associated with big data analytics projects, many enterprises have found that managing them has become the most pressing problem (Raguseo, 2018), thus big data is not widely used in enterprises (Frank et al., 2019). Sun et al. (2018) identified 26 important factors that significantly impact the adoption of big data by enterprises. The capacity to analyze vast amounts of data, extract meaningful insights, and derive substantial value from big data poses a fundamental challenge for enterprises (Niebel et al., 2019). More effort and knowledge are needed from top managers to discover the benefits and risks related to big data usage at their enterprises.

The significance of big data technology for the banking sector is increasingly attracting the interest of different stakeholders, such as financial institutions and central banks (Irving Fisher Committee, 2021). In 2015, only a small proportion of central banks (about 10%) were placed on deepening their understanding of big data; however, by 2020, this had increased to around 65% (Irving Fisher Committee, 2021). In addition to the complexities of using big data, which requires coordination, administration, human resources, and infrastructure, the ability to take advantage of it differs between central banks from advanced economies and emerging markets. The former has more experience in this area, while the latter still faces many challenges (Irving Fisher Committee, 2021). Statista (2023) revealed that the market value of BDA is projected to rise significantly in the coming years, reaching above 650 billion dollars in 2029. However, the development of big data analytics in financial institutions in Europe is still at its early stage (EBA, 2020), which means more in-depth studies on various aspects of the effective use of big data are necessary.

The research on the use of big data analytics at the enterprise level is varied. However, in the banking and financial sector, it is scattered and limited to certain geographic markets such as the United States, China, the United Kingdom, and India (Nobanee et al., 2021). Research on big data in banking mainly focuses on customer analytics to increase customer retention and achieve competitiveness (Nobanee et al., 2021) and emphasizes the impact of big data analytics on banking operations. In a study on sustainability capabilities, Ali et al. (2020) combined the four factors, including green supply chain management, green human resource practices, commitment, and big data integration, to investigate the relationship between these factors and bank performance in terms of financial and environmental performance. The results from this study confirmed the role of green practices in supply chain management and human resources, combined with big data association, leading to an improved financial and environmental performance at commercial banks. Although not directly referring to big data analytics, Zelenka & Podaras (2021) used case studies from the banking sector to highlight the importance of data quality in business intelligence for creating tactical knowledge that effectively impacts decision-making. Risk management is an area of interest due to the rising cyber impacts on banking operations. Meanwhile, the banking sector is considered one of the primary data-driven industries in terms of applying new information technologies for management (Grover et al., 2018). A recent study, Milojević & Redzepagic (2021) briefly introduced how using AI, ML, deep learning (DL), and big data analytics in risk management practices at banks. Furthermore, this research proposed a framework for using these cutting-edge techniques in building and developing risk management methods. The use of ML, DL and AI are increasingly being promoted in risk management at banks. In the future big data analytics' contribution to the use of AI, ML, and especially DL is undisputed; it can create greater efficiency in risk management (Milojević & Redzepagic, 2021). Combining big data technology and Monte-Carlo method, Manuylenko et al. (2021) proposed a model for evaluating the potential strategic innovation risk in the banking sector. Previous researchers also highlighted the prospect of adopting big data technology in support of the era of bank of future. However, empirical studies on big data usage in the banking sector remain scarce.

Individuals may use new technologies such as mobile banking or fintech applications due to perceived usefulness, perceived ease of use (as suggested by the Technology Acceptance Model), and other factors, such as social influence or facilitating conditions (as indicated by the unified theory of acceptance and use of technology – UTAUT). However, at the firm level, the decision to use new technology like big data analytics is considered a strategic decision (Maroufkhani, Wan Ismail et al., 2020), which not only focuses on the factors that determine its adoption but also how to use it to enhance business innovation, create BDA capabilities and increase enterprise competitiveness (M. Gupta & George, 2016). The main goal is to identify how to leverage this new technology in the banking industry, particularly big data analytics, to benefit businesses. Therefore, this thesis focuses on the aspects of establishing a strong big data analytics capability in order to boost the banks' competitive advantages.

1.2 Research aims of the study

The existing studies on big data analytics in the banking sector are very limited, but most published studies are limited to its benefits and prospects. Despite the importance of big data technology, there remains a paucity of evidence on the understanding of the role of BDA capabilities in the performance of banks. To the best of my knowledge, no published studies in this sector show empirical evidence on how factors from technology organization environment framework (TOE), resource-based view (RBV), and dynamic capabilities theory (DC) could be combined to lead to banks' competitive advantage. This study would be the initial contribution to this aspect. This thesis seeks to fill the gap in the literature by proposing a model that combines RBV, TOE, and DC theory. It empirically investigates the effects of BDA usage on market performance and risk management performance. In this study, the use of big data analytics has a significant effect on banking performance in two ways. Firstly, it enables banks to understand their customers better, develop more competitive products quickly, and provide more personalized services. Secondly, it improves risk management, allows for better monitoring of transactions, more effective fraud prevention, and better-quality risk assessment and management in daily operations. Therefore, there are three primary aims of this study:

R1: to examine the relevant theories in context of big data analytics.

R2: to investigate empirically which factors influence the usage of big data analytics.

R3: to examine the relationship between big data analytics usage and banking performance.

1.3 Research questions

This study endeavors to assess the usage of big data analytics (BDA) within the context of enterprise level, encompassing various facets of the business. To deal with the overarching objectives, this study aims to address three significant research questions, to resolve the three primary aims articulated earlier.

The primary research question pertains to the examination of pertinent theories concerning the implementation of big data analytics at the enterprise level. Numerous theoretical frameworks exist that can be employed to assess the factors that influence big data analytics and evaluate its impact on banking performance. Hence, the first research question is formulated as follows:

RQ1. What are the relevant theories in context of big data analytics usage at enterprise level?

The second research question aims to investigate the factors that influence the successful utilization of big data analytics within the banking sector. In this thesis, the author selects four factors from the RBV and TOE framework as the determining factors of BDA usage. Among these variables, top management support has been frequently mentioned in previous studies, highlighting its significance as a managerial factor/ resource. Meanwhile BDA skills and BDA infrastructure are identified as fundamental factors, emphasizing the importance of human resources and tangible resources in the utilization of any digital technology. Lastly, big data quality is considered a contextual factor in the research of BDA topic. Consequently, the principal second research question and four sub-research questions are formulated as follows:

RQ2. What are the determining factors of big data analytics usage in the EU-27 member states' banking sector?

RQ2.1 What is the impact of top management support on big data analytics usage in the EU-27 member states' banking sector?

RQ2.2 What is the impact of BDA skills on big data analytics usage in the EU-27 member states' banking sector?

RQ2.3 What is the impact of BDA infrastructure on big data analytics usage in the EU-27 member states' banking sector?

RQ2.4 What is the impact of big data quality on big data analytics usage in the EU-27 member states' banking sector?

The third research question aims to evaluate the impact of big data analytics (BDA) usages on banking performance. The utilization of BDA for various purposes and at different levels yields distinct effects on banking performance. Notably, emerging technologies such as BDA engender intense competition within the banking sector, prompting banks to assert market dominance through novel and superior products and services. Simultaneously, the study will also examine risk management performance, which constitutes a unique characteristic within the banking industry. Consequently, the third primary research question, along with its two sub-research questions, is structured as follows:

RQ3. Does big data analytics usage influence banking performance in the EU-27 member states' banking sector?

RQ3.1 Does big data analytics usage influence market performance in the EU-27 member states' banking sector?

RQ3.2 Does big data analytics usage influence risk management performance in the EU-27 member states' banking sector?

1.4 Methodology of the study

The holistic methodological approach taken in this thesis is a mixed methodology based on the case study, qualitative, and quantitative approach:

The case-study method was adopted to provide an in-depth analysis of the digital transformation process at two banks in Hungary.

The systematic literature review was utilized to understand better what has been done in BD topic in enterprise level, especially BDA's importance and factors affecting the use of big data analytics in the banking system practice. Furthermore, analyzing data from many sources and interviewing experts from the banking industry also helps add accuracy and reliability to the questionnaire.

The quantitative method is utilized to test the research hypotheses. This is also the main research method used in this thesis. The partial least square–structural equation modeling (PLS-SEM) analysis was selected in this study to test the relationship between the variables in the research model.

The combination of the three research methodologies above was used to answer the research question in this thesis.

1.5 Structure of the thesis

The present thesis encompasses a comprehensive framework comprised of seven distinct chapters, as depicted in Figure 1 below.

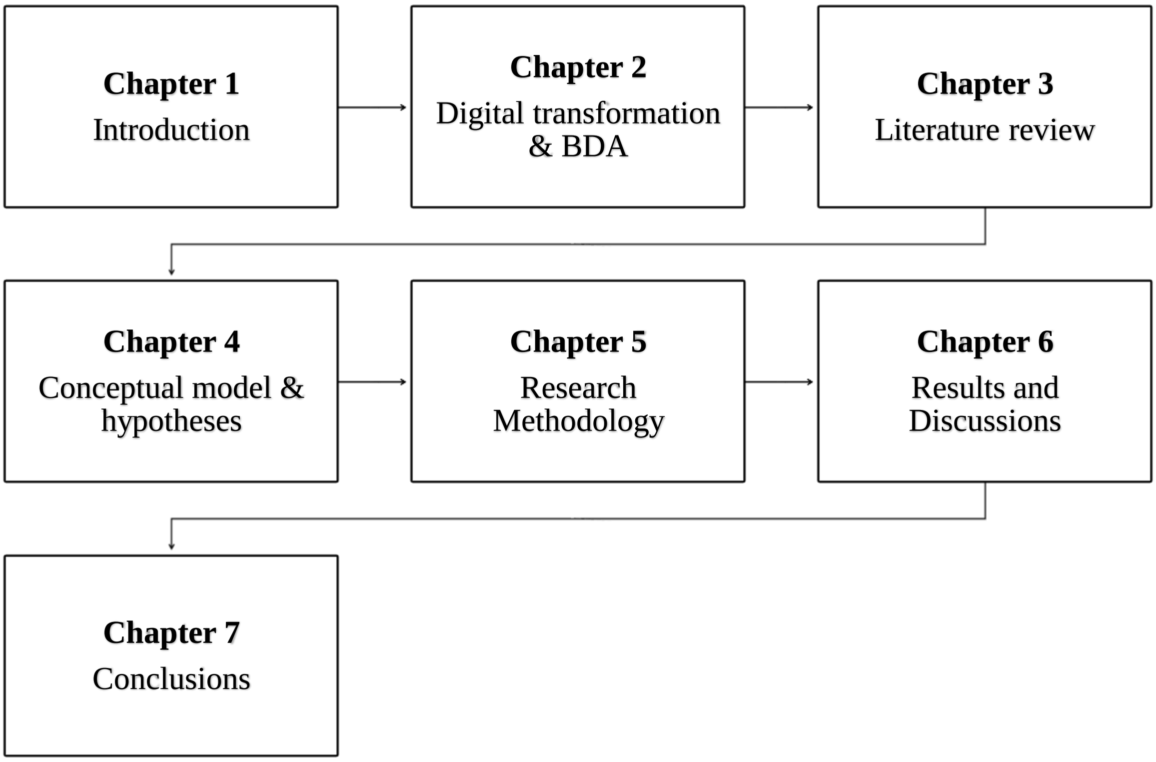


Figure 1: Research structure of the thesis
Source: Own edition

Chapter 1 is the introduction of the thesis. In this chapter, the author explains the digital transformation trend in economies and the significance of utilizing digital technologies to enrich banking performance. The research aims and hypothesis are also presented in this chapter, followed by a brief introduction of the scope and methodology.

Chapter 2 presents more in-depth analysis about the digital transformation progress in the banking sector. After that, this thesis highlights some dominant frameworks for digital transformation in the banking industry. Moreover, two case studies of digital transformation in two Hungarian commercial banks are examined. Finally, the last two sections of chapter 2 present the understanding of big data analytics and the current trend of using this technology in the European region.

Chapter 3 is concerned with the literature review part. A comprehensive literature review was conducted. By laying out the relevant theories in using digital technology research, this chapter highlights the advantages and drawbacks of each of these single theories.

Chapter 4 develops the proposed research model and presents research hypotheses. The research model proposes the relationship between 4 factors: top management support, BDA infrastructure, BDA skills, big data quality and BDA usage. This chapter also presents the relationship between BDA usage and bank performance, including risk management and market performance.

Chapter 5 is the methodology part. This chapter presents the data material used for the quantitative analysis of this thesis. Moreover, the reason for selecting the PLS-SEM methodology is also mentioned. The data collection method is presented in this chapter.

Chapter 6 presents the results and discussion. The section begins by examining the main results of the PLS-SEM method, including assessing the measurement and structural model. After a discussion of the results of the quantitative analysis, this chapter highlights the key findings and compares the results of this thesis with previous studies.

Finally, **Chapter 7** emphasizes this study's theoretical contributions to the research field's current knowledge and the practical implications for business. The limitations and suggestions for future research are also mentioned.

CHAPTER 2: INTRODUCTION TO DIGITAL TRANSFORMATION AND BIG DATA ANALYTICS IN BANKING SECTOR

2.1 Introduction

This chapter presents more in-depth information about the digital transformation progress in the banking sector. After analyzing the digital transformation development in the banking sector, this chapter highlights some dominant frameworks for digital transformation in the banking industry. Moreover, two digital transformation case studies in two Hungarian commercial banks are examined. Finally, the last two sections of chapter 2 present the understanding of big data analytics and the current trend of using this technology in the EU-27 Member States.

2.2 Development of digital transformation in banking sector

A digital transformation implementation is a comprehensive approach to revising strategy and business models through digital technology (Besson & Rowe, 2012). However, the terms digitization and digitalization often need clarification and have various interpretations among academics. It is, therefore, important to clarify the meaning of these words.

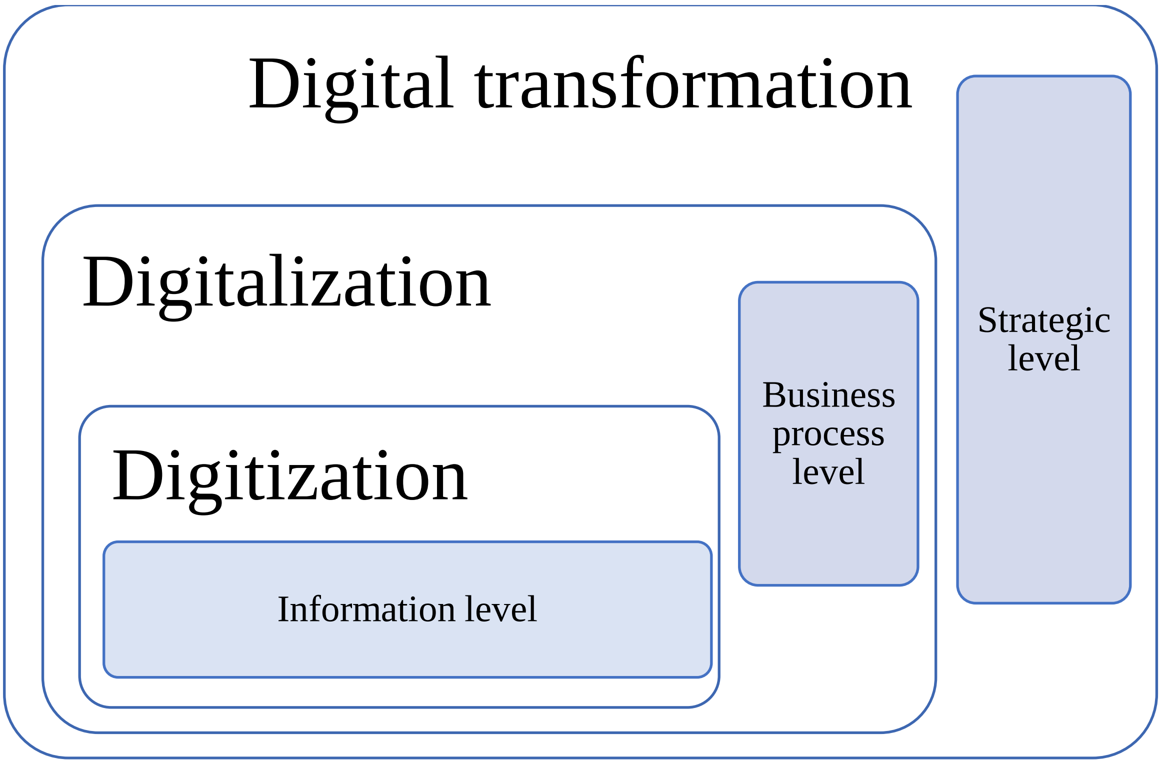


Figure 2: The scope of related terms
Source: Own edition

Digitization is the process of converting analog information into digital information, aiming to create value for stakeholders (Schallmo & Williams, 2018). On the other hand, digitalization is transforming business processes using digital technology (Schallmo & Williams, 2018). Digital transformation, meanwhile, is the alteration of the business model to better serve customers through the use of digital technology (Bloomberg, 2018; Schallmo & Williams, 2018). Despite the emergence of a range of definitions of digital transformation, Nwankpa and Roumani (2016), Libert et al. (2016), and Stief et al. (2016) have concluded that Digital transformation refers to the implementation and application of digital technology and Information and Communication Technologies (ICTs) to transform existing business operations into new processes, products/services, or business models, sometimes involving the creation of new digital products. Kane et al. (2015) emphasize that digital transformation does not lie in digital technologies itself but in how organizations organize and implement it. Janssens (2019) provides a broader definition of digital transformation as a shift to a new management model and philosophy that encourages innovation and new business models, as well as the utilization of digital technologies to improve the experience of both internal and external customers.

The banking sector is no exception to the trend of utilizing technology in operations. For many years, the use of new technologies has been commonplace in this sector; however, the last decade has seen a significant rise in adoption, largely due to the emergence of the internet and the penetration of smartphones, which have enabled a broader range of banking services. Given the potential opportunities and difficulties posed by digital technologies, there is a growing recognition of the need for a comprehensive strategy for digital transformation in banking. To provide an overview of digital transformation in the banking sector, researchers have proposed various digital maturity models and digital frameworks.

Cuesta et al. (2015) identify three stages of digital transformation in the banking sector: the introduction of new channels and digital products, technological adaptation, and organizational culture transformation. Changes in competitive circumstances drive these stages and involve large investments in digital technology. Matt et al. (2015) propose a digital transformation framework that combines four primary dimensions: the use of technologies, changes in value creation, structural changes, and financial aspects. They argued that the use of technologies is a strategic priority in digital transformation, as it helps banks add new products and services or introduce new business models and opportunities.

Furthermore, the emergence of new technologies and value chains necessitates changes in structural operations. Additionally, financial aspects are the foundation for implementing the remaining three dimensions. In recent years, mobile banking apps, online savings, and virtual cards have become increasingly popular, and Fintech and big tech firms have provided significant competition in the financial market. Thus, Backbase (2020) presents a digital-first framework to demonstrate how traditional banks can remain competitive in the face of digital newcomers. This framework consists of four components: omnichannel banking, smart banking, modular banking, and open banking. Meanwhile, Melnychenko et al. (2020) classified the revolution of digital banking into three stages. The first stage focused on using ATMs and call centers to expand distribution channels. The second stage saw the introduction of mobile, internet, and social networks, with banks launching mobile applications to diversify and personalize services. Finally, the third stage is characterized by the increasing use of technologies such as AI, chatbots, big data, and blockchain, which are applied to all aspects of banking operations.

The banking sector's digital transformation process is correlated with and mirrors the advancement of banking technology. Consequently, the transformation and the implementation of new technologies will result in the introduction of new services and products in the bank. However, success in digital transformation is only guaranteed for some. Numerous factors, such as those related to management, technology, and external stakeholders, can influence the success of digital transformation in banks. Building on the research conducted by Tran et al. (2022), the next section focuses on the part digital technologies play in the digital transformation of two commercial banks in Hungary.

2.3 Digital transformation in banking sector

2.3.1 Case Study 1

Bank A is one of Hungary's leading financial service providers, with over 100 branches and 5000 employees. It prioritizes the agricultural sector and small and medium enterprises (SMEs), offering a range of financial services (such as AgroDevelopment, Restart Investment, Crisis Loan, Agricultural Direct Payments Pre-Financing, Agricultural Investment Loan with interest rate subsidy, and Agro-Enterpreneur Overdraft Facility) and online financial tools (a Land subsidy calculator, Management calculator, Agricultural machinery financing calculator, and Farmland calculator) to customers. Digital transformation at this bank happened first in strategic planning. To support its digital transformation strategy, the bank has implemented 25 digital projects to improve customer experience and enhance banking operations. The utilization of digital technologies is not only to augment internal process efficacy but also to foster an innovative culture. This is evidenced by the bank's adoption of digital technology to train staff and managers and the implementation of internal standards for the communication process within the corporation. Furthermore, a change in administration program has been implemented to reduce the time to launch new products/services and increase customer satisfaction. Additionally, the bank has established an innovation center to stimulate an innovative culture and bolster the bank's competitiveness. The bank has adopted new technologies like big data analytics and artificial intelligence to develop novel business models and automate and robotize operational processes. Furthermore, it has supplied customers with easily comprehensible information, such as direct kiosks and online videos, and has collaborated with start-up companies to add more personalized features. This bank has made a contribution to society by decreasing paperwork and paper consumption, diminishing business travel and emissions, and enhancing financial literacy through the implementation of digital technologies (Tran et al., 2022).

2.3.2 Case Study 2

Bank B is a member of a leading banking and insurance business institution in Europe, headquartered in a Western European Country. According to an independent rating agency, the level of digitalization at the parent bank is second only to a fully digital bank in the Belgian market, demonstrating the success of its digital transformation. Hungary is one of the six core markets for the bank. The parent bank has invested 1.5 billion EUR in a 3-year innovation and digital transformation strategy from 2017-2020, which uses an omnichannel approach to optimize customer experiences and create more integrated and seamless interactions between channels. The bank has also added a new dimension to its organizational culture to promote cooperation, support for innovative ideas, and solidarity within the group. A Chief Innovation Manager and an innovation board have been appointed to execute the digital transformation plan. Team-building groups have been created to promote learning and sharing new ideas and experiences. The bank has invested in AI and data analysis to simplify processes and make faster and more accurate decisions. It has also launched digital assistants, upgraded its banking platform, and cooperated with fintech companies and Vloggers to provide banking services and knowledge. The bank has complied with the Payment Service Directive 2 (PSD2) and built legal and infrastructure frameworks to allow employees to work from home safely during the Covid-19 time (Tran et al., 2022).

2.3.3 Some features of using digital technologies from two case studies

Through two case studies, it can be realized that the application of digital technologies in commercial banks is an integral part of their digital transformation strategy, thus demonstrating the importance and positive contribution of digital technologies to digital transformation. Furthermore, these two case studies demonstrate that digital technologies can be used to achieve identified objectives:

- Enhancing the quality of services/products provided.
- Researching and developing new products.
- Testing new processes to improve work efficiency and reach customers.

Nevertheless, the level of using digital technologies has only been at the intermediate level, indicating that numerous essential elements influence the efficacy of digital technologies, particularly the significant impact from human resources, ranging from upper management to personnel with superior expertise.

2.4 Understanding of using big data analytics

2.4.1 Definition and characteristic of big data analytics

Massive amounts of data are generated worldwide (Kambatla et al., 2014). Specifically, big tech corporations such as Meta or YouTube produce a huge quantity of data each minute, with an even larger and more diverse selection than ever before. As a result, structural data, like text or numerical datasets; semi-structural data, such as voice recordings, videos, and images; and non-structural data, such as posts and comments on social media platforms, are all seeing a surge in frequency. Data is continuously growing at a rapid pace, making it impractical to precisely determine what qualifies as "big data". Consequently, despite its frequent usage in literature, there are varying definitions of big data within the academic sphere (Zhou et al., 2014). However, most of the previous literature has sought to describe big data by examining its inherent characteristics. One of the first researchers to define big data was Laney (2001), who noted that "big data" term is characterized by three distinct characteristics, called three V: volume, velocity, and variety. Whereas volume refers to the huge amount of data and velocity refers to the speed at which these data are generated, variety refers to data characterized by the various types. Similarly, according to the European Commission's definition, big data encompasses substantial quantities of data that are rapidly generated from diverse and numerous sources (European Commission, 2023). In a study published on the Journal of Big Data, Batko & Ślęzak (2022) simply suggest "big data" as “the massive amount of data sets that cannot be stored, processed, or analyzed using traditional tools”. Supporting this view, in comparing big data to traditional data, IBM (2023) notes that data is considered big when it exceeds “the ability of traditional data processing to capture, manage, and process the data”.

Based on the aforementioned definitions, researchers emphasize the significance of massive data through various characteristics associated with big data. These characteristics make the existing technology impossible to manage big data (Dumbill, 2013). The widely accepted definition of big data, as posited by IBM researchers, entails the incorporation of three fundamental components referred to as the “three V’s”: volume, variety and velocity (Zikopoulos et al., 2012). These characteristics serve as fundamental standards for distinguishing between “big data” and “traditional data” (Miloslavskaya & Tolstoy, 2016). Over time, the original three V characteristic has been expanded by different authors based on the specific context (F. Y. Wang, 2012). For instance, Zhou et al. (2014) introduced additional characteristics such as the veracity of data, while L. Zhou et al. (2017) supplemented the value of data. According to Abdalla (2022), the 5V characteristic of big data has been identified by previous researchers. However, it is worth noting that in certain cases, the 5V characteristics may expand to include additional dimensions such as visualization and variability, as suggested by Sivarajah et al. (2017). Table 1 below describes the popular 7V characteristic of big data in the banking industry.

Table 1: The seven characteristics of big data

Big data characteristics	Definition and example	Supporting literature
Volume	Is defined as the huge amount of data generated by various sources. Size of information stored in terabytes and petabytes exceeds the storage capability limits of normal operations. For example, commercial banks may encounter challenges when dealing with transaction data for risk identification, as the volume of data can be exceedingly large prior to processing.	Laney (2001), IBM (2023), O’Leary (2013), Abdalla (2022), Batko & Ślęzak (2022), Zhou et al. (2014), L. Zhou et al. (2017), Wahyudi et al. (2018).
Variety	Refer to the various types of data which often mentioned about the new types of data created by portable devices, sensors, or social platforms. Data types include structured information such as the plain information stored at the spreadsheets in banks, and much larger data sets are unstructured information which collected from	Laney (2001), IBM (2023), O’Leary (2013), Abdalla, (2022), Batko & Ślęzak (2022), Zhou et al. (2014), L. Zhou et al. (2017).

	text, video, images, or voices. These new types of data require much more effort and technological advancement to extract valuable information for the banking sector.	
Velocity	Refer to the speed at which data is generated. For example, in the banking sector, the real-time data from financial/credit card transactions could be useful to identify the fraud.	Laney (2001), IBM (2023), O'Leary (2013), Abdalla, (2022), Batko & Ślęzak (2022), Zhou et al. (2014), L. Zhou et al. (2017), Wahyudi et al. (2018).
Veracity	Refers to the accuracy/quality (about the authenticity, reputation, availability, accountability) of data. This characteristic highlights the problems in using data from many sources which may be unclear about the accuracy/ ownership of data.	Zhou et al. (2014), Baig et al. (2019), Abdalla (2022), Batko & Ślęzak (2022), Wahyudi et al. (2018)
Visualization	Refer to ability to interpret data and resulting insights, increase the perceptions.	Abdalla (2022), Batko & Ślęzak (2022), Wahyudi et al. (2018)
Variability	Refer to the inconsistency (the change of meanings and interpretations based on its context) of data. This characteristic poses a significant challenge for banks when processing text data from social media. The meaning of a word used in a particular context can vary when used in a different context.	Batko & Ślęzak (2022), Wahyudi et al. (2018)
Validity	Refers to the adherence of data generation to procedures and regulations. The banking sector is required to adhere strictly to governmental regulations (such as privacy law) as well as internal regulations. Therefore, the assessment of data validity must be conducted with great care to avoid potential issues arising from invalid data in the future.	Wahyudi et al. (2018)

Source: Own edition

The present locus of debate regarding big data resides in the potential implications associated with the adoption of big data within firms and businesses (Niebel et al., 2019). For managers in large firms, one of the most notable aspects of big data lies in the potential opportunities and benefits it offers, as well as the associated infrastructure requirements (Schultz, 2013). To maximize the advantages derived from big data, the analytic process assumes a paramount role. There is consensus among scholars that BDA is the next level of traditional business intelligence (BI) tools (Gillon et al., 2014). Business intelligence has undergone a prolonged evolution and is easy to use in each functional department. In contrast, the development of big data analytics is still in its early stages; requiring a significant reliance on human skills (combining business skills, data knowledge, and computer skills) to build company-specific, custom-made solutions (Debortoli et al., 2014).

Recently, terms like AI and ML have gained significant attention beyond the realm of computer science terminology. This expanded usage is primarily due to the increasing application of AI and ML in various business driven by the widespread utilization of big data (Milojević & Redzepagic, 2021). The integration of BDA and AI has potential to enhance firms' computational capabilities and facilitate autonomous decision-making, thereby improving the efficiency of the supply chain. This synergy between BDA and AI offers numerous benefits in sustainable manufacturing and circular economy capabilities (Bag et al., 2021). In a broader context, according to Frank et al. (2019), the integration of AI and BDA is considered a pivotal component in smart manufacturing within Industry 4.0. The data collected from sensors can be utilized for predictive analysis, specifically in identifying potential machine failures and production overload. This application ultimately enhances the efficiency of quality control processes. In the banking and finance sectors, machine learning algorithms such as classification trees and artificial neutral networks are extensively utilized in risk management activities such as credit scoring and fraud detection (Milojević & Redzepagic, 2021). The substantial support from big data allows AI to become more intelligent and contribute significantly across various fields (Pejić Bach et al., 2020). Placed within the context of the digital transformation trend and Industry 4.0, the role of big data has become more important than ever before.

Batko & Ślęzak (2022) define BDA as "techniques and tools used to analyze and extract information from big data that unable to do by traditional tools". By utilizing big data platforms such as Apache Hadoop and Spark, BDA enables the processing of data that surpasses the limitations of conventional tools in terms of data storage, analysis and management (Zhu et al., 2019). From the perspective focused on the benefits that big data brings to businesses, Fosso Wamba et al. (2015) define BDA as a systematic process encompassing the collection, analysis, utilization and interpretation of data across various functional divisions. The primary objective of this process is to acquire actionable insights, generate business value, and attain competitive advantages (Fosso Wamba et al., 2015). Supporting this view, according to IBM (2023), BDA is defined as the utilization of advanced analytic techniques on extensive and heterogeneous data sets, encompassing structured and unstructured data, which can be processed through either streaming or batch methods. This process enables enterprises to acquire novel insights, leading to substantial improvements in the speed and quality of decision-making. This research utilized the definition of big data analytics proposed by Fosso Wamba et al. (2015) as it emphasizes the noteworthy impact of big data on enterprise performance.

2.4.2 Using big data analytics in EU-27 Member States

This section presents the practical usage of big data analytics in the EU-27 Member States. Data was collected from EUROSTAT offices (online data code: ISOC_EB_BD and ISOC_EB_BDN2). This dataset includes information on all 27 countries in the European region at the enterprise level. The number of employees indicates the enterprise's size, with large enterprises having 250 or more employees and small and medium-sized enterprises (SMEs) having 10-249 employees. The utilization of BDA was analyzed from two perspectives: by sources of BDA and by industrial sectors.

a) Using BDA in EU-27 by enterprises (Large enterprises and SMEs)

In general, the use of BDA in enterprises within the EU-27 region is still limited. For example, regarding the data sources used for BDA (at Figure 3), nearly 7% of whole enterprises (including SMEs and large enterprises) use BD sources from geolocation and social media, compared with 3.3% of enterprises using BD sources from smart devices.

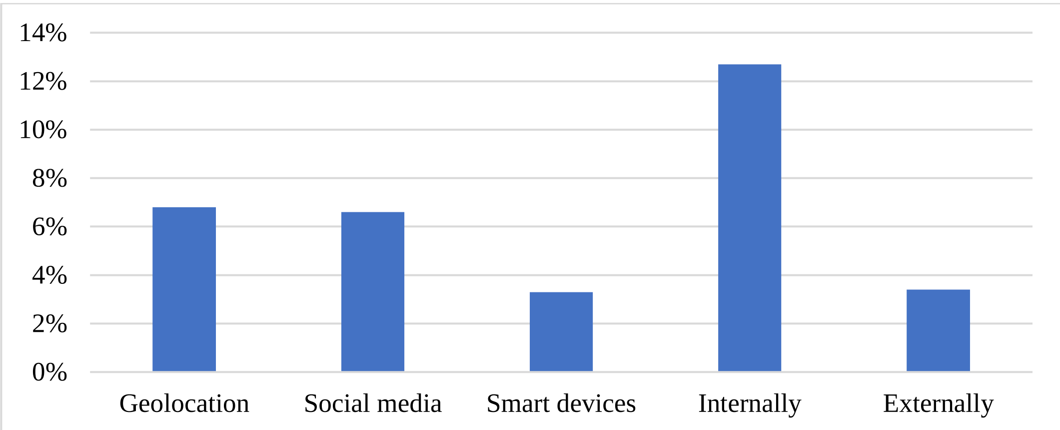


Figure 3: The usage of BDA in SMEs and large enterprises by data sources
Source: Own edition

Furthermore, there is a difference in using big data analytics between SMEs and large enterprises. BD from smart devices is used most by large enterprises (17%). At the same time, for SMEs, the data for BD analysis is mainly collected from geolocation and social media sources.

Primarily, enterprises use in-house experts to analyze big data (12.7%), while only 3.4% employ third parties to explore BD. Again, this is consistent with both large enterprises and SMEs (as seen in Figure 3).

Figure 4 shows that two distinct trends in the utilization of BDA among countries in the EU-27 are evident. Comparative data among the EU-27 countries suggests a likely positive correlation between BDA usage among large enterprises and SMEs (as seen in Figure 4). First, countries with high adoption rates of BDA by large enterprises are those with high adoption rates of BDA by SMEs. The countries with the highest rates of BDA adoption (group 1) are primarily those with the most advanced economies in Western and Northern Europe, such as Denmark, Netherlands, Malta, Belgium, Finland, Ireland, Sweden, France, Luxembourg, and Germany. Conversely, countries in Group 2 and Group 3, mostly from the East and South European regions with transitioning economies or lower developed economies, have lower average BDA adoption rates than the European average.

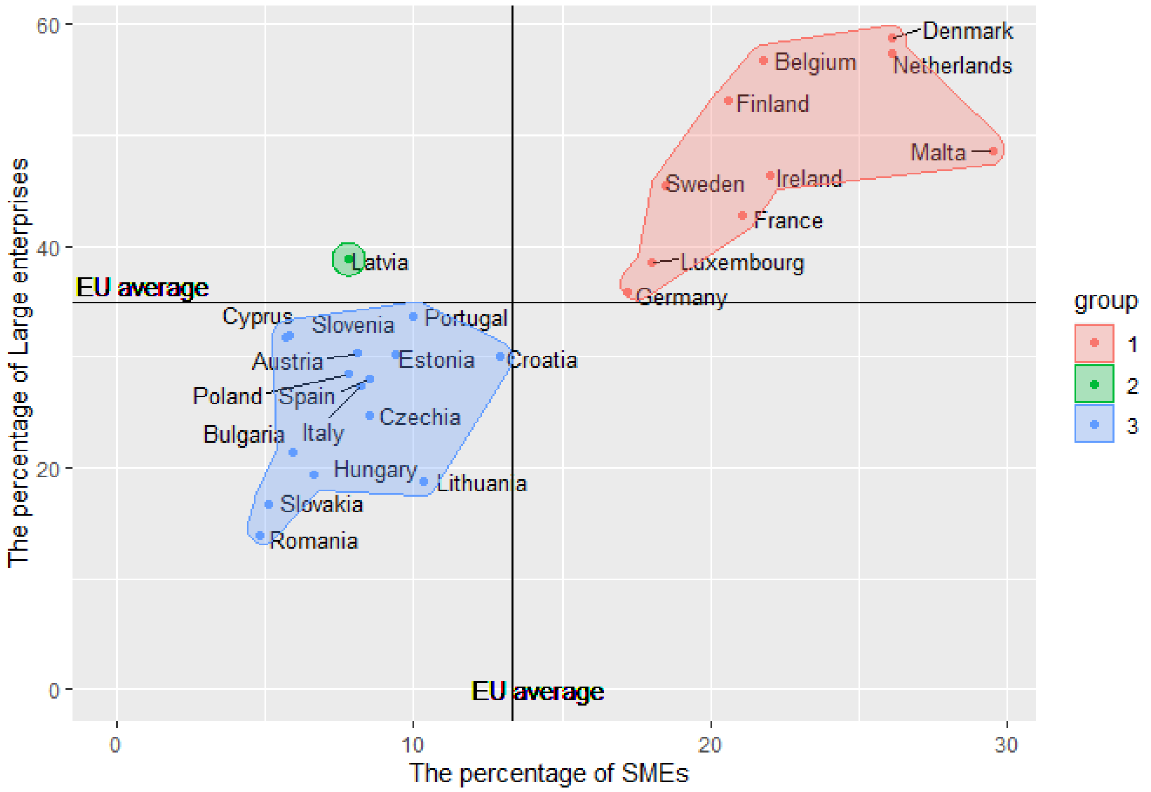


Figure 4: The percentage of enterprises using big data in EU-27 Member States

Source: Own edition

The differences in the use of BDA between enterprises' sizes are evident. The rate of BDA usage in large enterprises is much higher than that of SMEs (Figure 5). On average, the rate of large enterprises using BD is 35% in the EU-27 Member States, while this figure is around 13% for SMEs. Large enterprises, with their greater resources in terms of personnel, technology, and need for larger BD solutions, are most likely to adopt big data analytics (Pejić Bach et al., 2020).

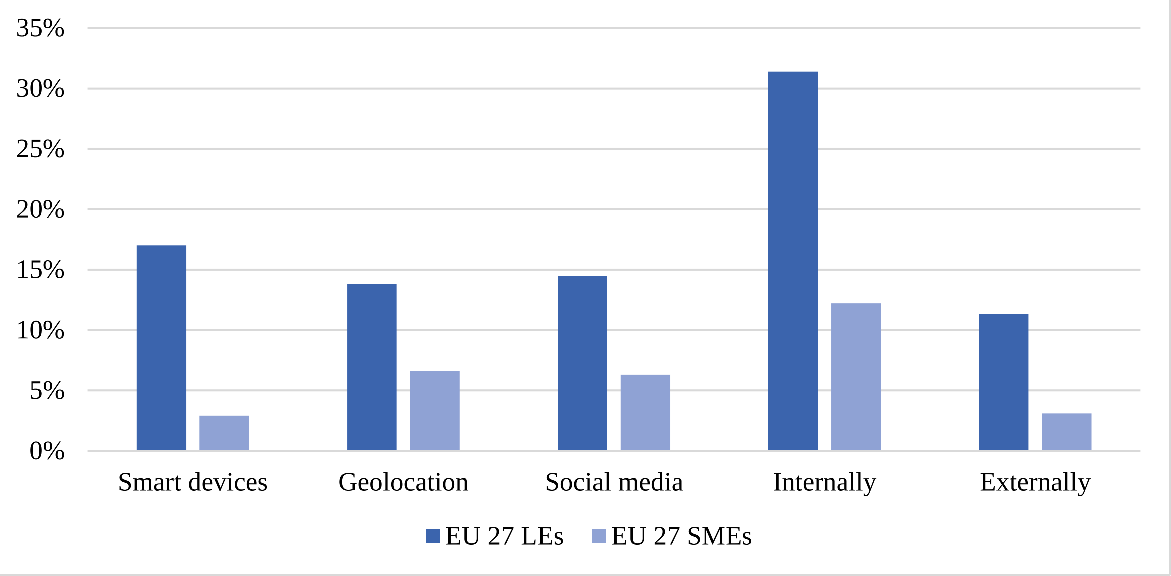


Figure 5: The usage of BDA by the types of enterprises
Source: Own edition

b) Using BDA in EU-27 Member States by sector

Figure 6 illustrates the rate of big data utilization across industrial sectors and data sources.

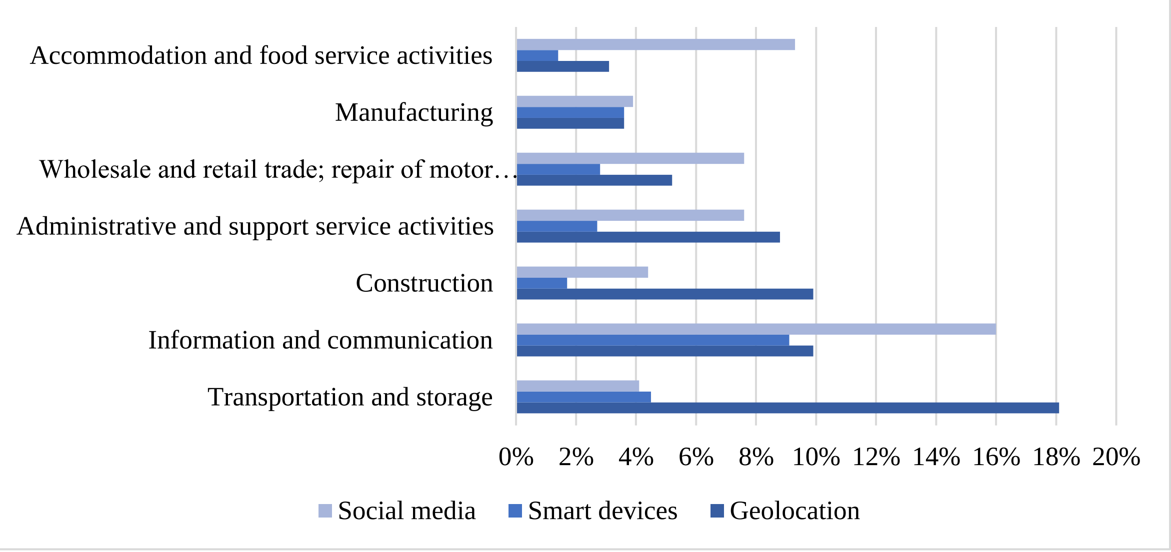


Figure 6: The usage of big data in EU-27 Member States by sectors
Source: Own edition

Geolocation and social media data are the most widely utilized big data sources, as evidenced by the sectors' frequent utilization. The highest utilization of geolocation data is seen in the transportation and storage sector (18.1%), while the highest utilization of social media data is seen in the information and communication sector (16%).

2.5 The use of big data analytics in banking sector

2.5.1 The BDA market in banking sector

In comparison with other industries, the banking sector has some advantages in using big data technology. This sector has a high demand for data utilization. Recent statistics from Capgemini (2020) indicates that the banking sector is at the forefront in using data for decision-making. After long-term development, banking sector has recorded and stored more customer data than any sector (Giebe et al., 2019).

Globally, international companies offer this sector a variety of big data solutions, including fraud detection, customer data management, social media analytics, etc. In 2019, banking emerged as the top industry globally in terms of utilizing big data analytics. It accounted for 13.9 percent of the total revenues generated by big data analytics worldwide. Following banking, manufacturing, services, and the public sectors were also significant contributors to the revenue generated from big data analytics (IDC, 2019).

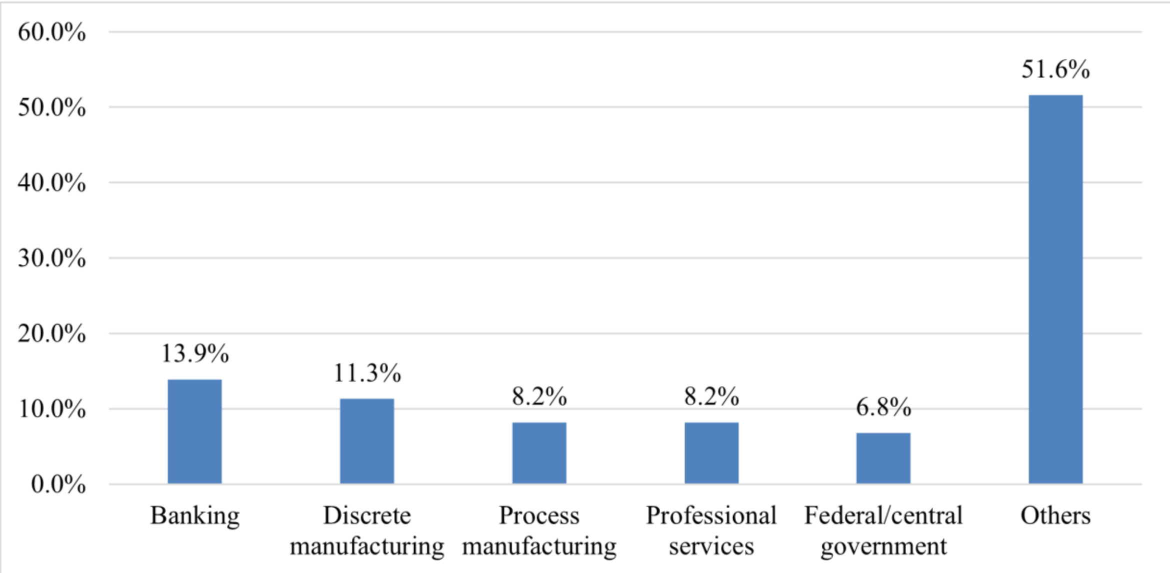


Figure 7: The big data analytics revenue by sectors in 2019
Source: Data from Statista

On global scale, a recent reveal from Statista has showed that Splunk, Oracle, IBM emerged as the leading providers of BDA software, capturing a substantial 11%, 9%, and 6% of the marketshare, respectively (Wikibon, 2018). Further details about the market dominance of various big data and analytics providers refer to Appendix 4. In the banking sector, the active participation of large companies makes this BDA market more competitive, with big tech who leading in providing BDA solutions such as Oracle Corporation, IBM, SAP SE, Aspire systems Inc, Alteryx Inc (Mordor Intelligence, 2023). Over the next five years, North America is forecasted to be the largest market in BDA solutions; meanwhile, the Asia Pacific region will witness the fastest growth in using BDA in the banking sector (Mordor Intelligence, 2023). Obviously, the banking sector plays the most active role in the market when it comes to using BDA solutions. For example, in India, banks spend the highest percentage, occupying 18% of the BDA market, surpassing all other sectors (IDC, 2021). Similarly, Italian banks account for 28% of the market share in terms of BDA spending (Osservatori Digital Innovation, 2021). These figures indicate that the banking sector has been and continues to invest heavily in BDA to leverage the advantages it brings.

2.5.2 The process of BDA usage in banking sector

Typical big data analytics architecture involves significant steps such as data collection, ingestion, preparation, analyses, and presentation (Abednego, 2021).

Data source: Data collected by banks can come from multiple sources. These include operational data (such as transactions, account activity, mails/ call centers, logs, trades), historical/live market data (such as instruments, quotes, or index values) and external data (Web API, web searches, social networks).

Data ingestion: Data ingestion is the process of obtaining raw data from different sources and ensuring that all necessary data is accounted for. Although it involves significant resources, external services like GCP, AWS, or Microsoft Azure are often used for their various data ingestion tools (Hadoop, Kalfa, Storm), which can be tailored to the specific needs of a project. All the data in this step is stored in a data repository known as a data lake. The term “data lake” can be defined as an extensively expandable storage repository that securely retains an immense volume of unprocessed data in its original format (Miloslavskaya & Tolstoy, 2016). As a results, data lakes typically store unstructured and easily accessible data, utilizing dynamic analytical applications and a flat architecture that distinguishes them from data warehouses (Miloslavskaya & Tolstoy, 2016).

Hadoop is a distributed framework for storing and processing data on clusters of commodity hardware (Abdalla, 2022).

Kafka is an event streaming platform for high-performance data pipelines, streaming analytics, data integration, and mission-critical applications that run on a distributed environment. It separates data streams from systems, allowing them to be stored and reused elsewhere.

Storm is a distributed real-time processing system designed for unlimited streaming data and is open source (Abdalla, 2022). It can be used for real-time analytics, online machine learning, continuous computation, and ETL activities.

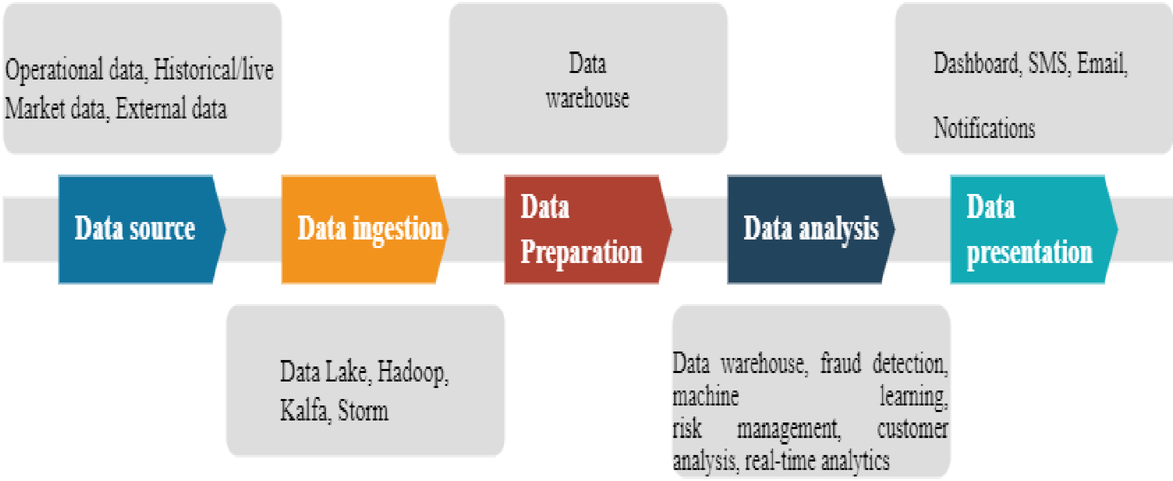


Figure 8: The typical big data analytics in banking sector.
Source: Own edition

Data preparation: Data stored in a data lake includes all kinds of data that are not yet classified or ready for analysis. Therefore, data preparation consists of the Extract, Transform, and Load (ETL) process. This process entails extracting data from a legal system, transforming it into a consistent format, enhancing data quality, and loading it into a target database (IMB, 2023). From here, the data can be ready for different analysis targets and stored in data warehouses. Data warehouse refers to a relational database system used for storing, analyzing, and reporting functions (Oracle, 2023). Data warehouse benefits enterprises in many aspects such as: generating integrated, stable, and ready data for analysis, reducing the volume of data allowing fast analysis (Oracle, 2023). Some commercial analytics tools such as AWS Kinesis Analytics, Azure Data Explorer, and GCP DataPrep can be used for data preparation. RedShift, Azure SQL, and BigQuery are also available for data warehouse services.

Data analysis: This step is the fundamental core in the architecture of big data analytics (Zhu et al., 2019). Data can be analyzed for specific projects such as fraud detection, machine learning, risk management, customer analysis, and real-time analytics. Depending on the objectives of each project, banks will construct different analysis models through using machine learning theory and the power of AI. The projects are then built, tested, and saved in the data warehouse.

Data presentation: Finally, the data analyzed and stored at the data warehouse can be presented for various purposes, including sending SMS alerts to users about unusual credit card usage or presenting data on real-time dashboards.

CHAPTER 3: THE LITERATURE REVIEW

3.1 Introduction

This chapter presents the literature review part. A holistic approach to literature review was applied to follow the stream of the research topic. A systematic literature review conducted aims to recognize what prior research has been done on big data analytics usage at the enterprise level and identify areas where further research is required. The literature review process also assists this study in filling existing research gaps and enriching the academic sphere.

3.2 Research on big data analytics in enterprise level

3.2.1 Classification of research on big data analytics in enterprise level

This study conducted a systematic literature review by utilizing scientific literature as data for its analyses (Kumpulainen & Seppänen, 2022). Identifying and selecting prior studies is essential to guarantee the literature review's quality. Concurrently, this method should ensure that noteworthy prior studies are not disregarded in the comprehensive review. The procedure of examining studies on the usage of big data analytics at the firm level is depicted in the figure below. The systematic literature review is based on combining two large databases, Elsevier's Scopus (Scopus) and Clarivate Analytics' Web of Science (WoS), which are currently among the most prestigious scientific databases. The coverage of natural sciences and engineering by WoS is extensive, whereas the coverage of social sciences by Scopus is comparatively higher (Mongeon & Paul-Hus, 2016). In a literature review, the literature data serves as the input for the review process. To ensure the review is conducted in a systematic manner, the literature data collection must also be conducted systematically (Kumpulainen & Seppänen, 2022). This study conducted data collection through the following steps: data search, data selection, and citation data acquisition (Kumpulainen & Seppänen, 2022).

During the data search steps (Step 1), the author found relevant studies based on keyword searches on Scopus and WoS, respectively. The search strategy encompasses three sets of keywords. The initial set comprises the term "big data" while the second set comprises terms pertaining to entities such as "firm", "organization" and "company". The third set of keywords encompasses terms such as "use", "usage" and "adoption". The inclusion of the second and third sets aims to signify the level of the research objective. These terms effectively facilitate the identification of studies focusing on the application of big data analytics in business contexts, thus bypassing research only centered on information technology (IT). A total of 465 studies were identified, including 355 studies from Scopus and 110 studies from the Web of Science databases.

Next, during the data selection stage, three sub-steps were implemented to include or exclude relevant studies. Firstly, specific selection criteria were established, which exclusively considered studies published in English, of specific article types, and before December 2022. Following this step, a total of 133 studies were included, with 79 from Scopus and 54 from WoS. Subsequently, five duplicated studies were identified from both Scopus and WoS. Consequently, 128 studies were remaining for the next step, involving the screening of abstracts. Abstract screening further resulted in the exclusion of studies that focused on public health and information technology. Duplicate studies present in the two databases were also removed.

Ultimately, 85 articles specifically addressing the big data adoption at firm level were utilized for the purpose of conducting a systematic literature review.

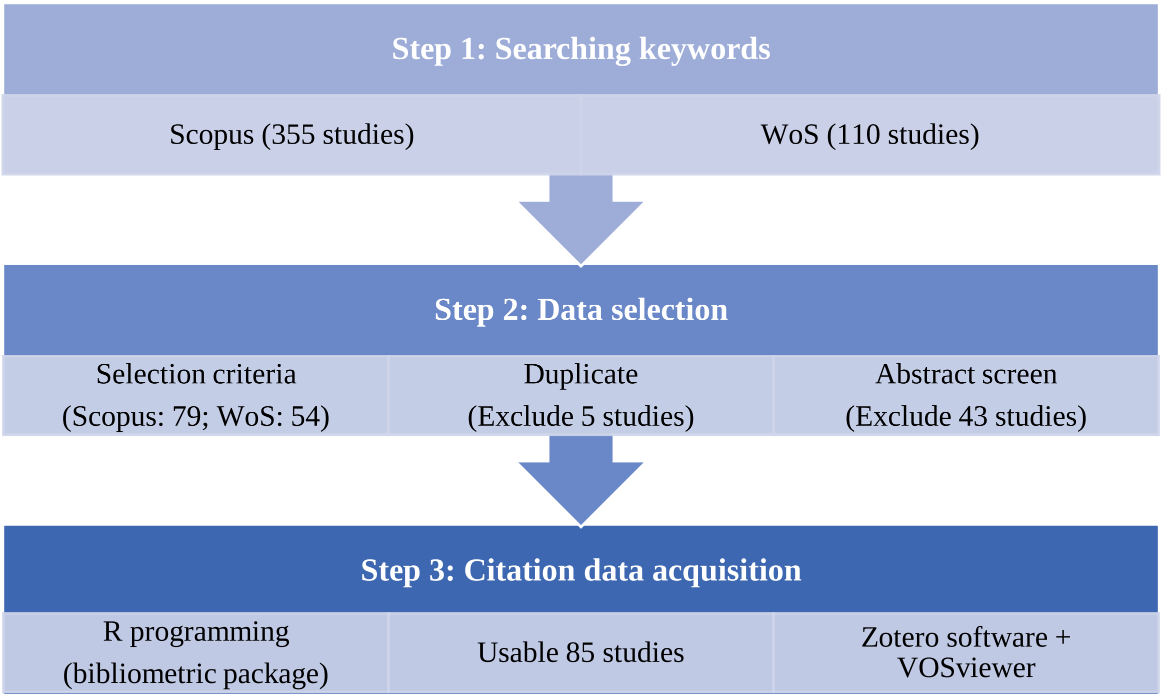


Figure 9: The step of collecting literature data

Source: Own edition

Note: In step 1, search query on Scopus (KEY("big data") AND KEY("firm" OR "organization" OR "enterprise" OR "company" OR "bank" OR "financial institution") AND KEY("adoption" OR "use" OR "usage" OR "implement") AND (LIMIT-TO (DOCTYPE,"ar")) AND (EXCLUDE (PUBYEAR,2023)) AND (LIMIT-TO (LANGUAGE,"English")), and on WOS (AK=("big data" AND ("firm" OR "organization" OR "company" OR "bank" OR" financial institution" OR "enterprise")) AND AB=("use" OR "usage" OR "adopt" OR "implement").

The literature review in this study is presented in two parts. The first part shows the overview of bibliometric information from previous studies. The main aims of the first part highlight the distribution of previous studies by produced countries, investigated publication years, high impacted authors, and journals. The first part is conducted by utilizing R programming with a bibliometric package (biblioshinny). The second part of the comprehensive literature review is based mainly on content analysis, and the graphic demonstration is presented using Zotero and VOSviewer software.

3.2.2 Distribution of previous studies by year of publication and source

Figure 10 indicates that research on big data analytics has been gaining attention among researchers in recent years.

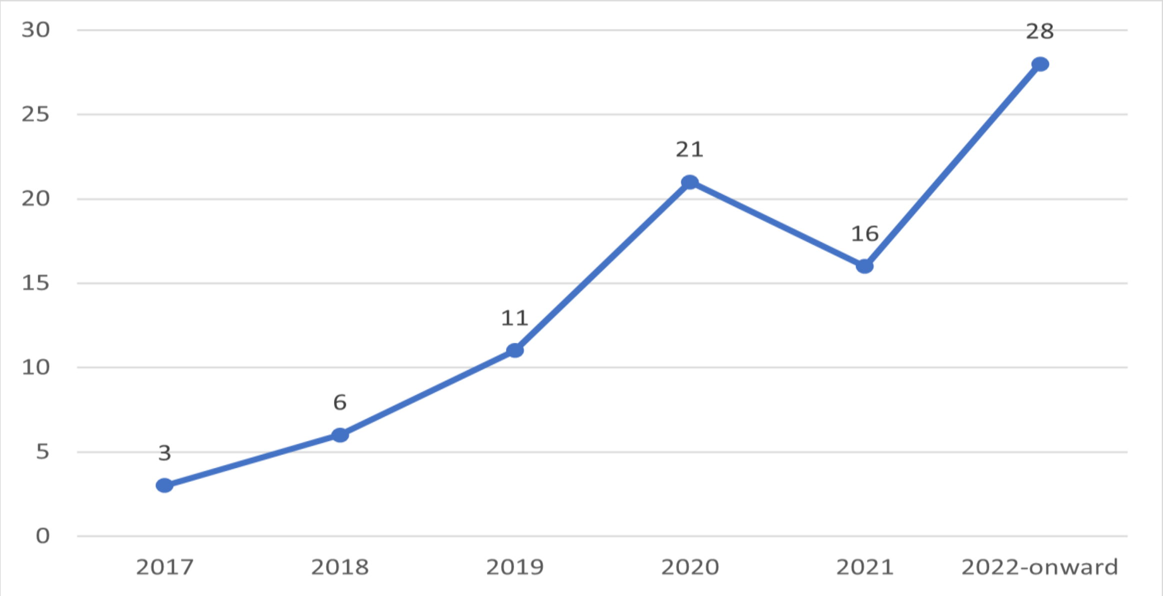


Figure 10: The annual publication on big data analytics
Source: Own edition

Between 2017 and 2022, 85 studies were conducted on this topic. The amount of research on big data at the organizational level has dramatically increased. From 2017 to 2019, there were only approximately seven studies per year. However, the last three years have witnessed a strong increase in the number of studies, with an average of about 23 published research annually. Despite the application of this new technology, research on BD in enterprises remains relatively modest.

Studies have predominantly been published in business, information management, and information systems journals (Figure 11). Of these, six journals have had the most published articles, accounting for approximately 30% of the total number of studies. These journals are the Journal of Business Research (6 studies), Management Decision (4 studies), Technological Forecasting and Social Change (6 studies), and International Journal of Information Management (3 studies), International Journal of Production Economics (3 studies), Journal of Computer Information Systems (3 studies).

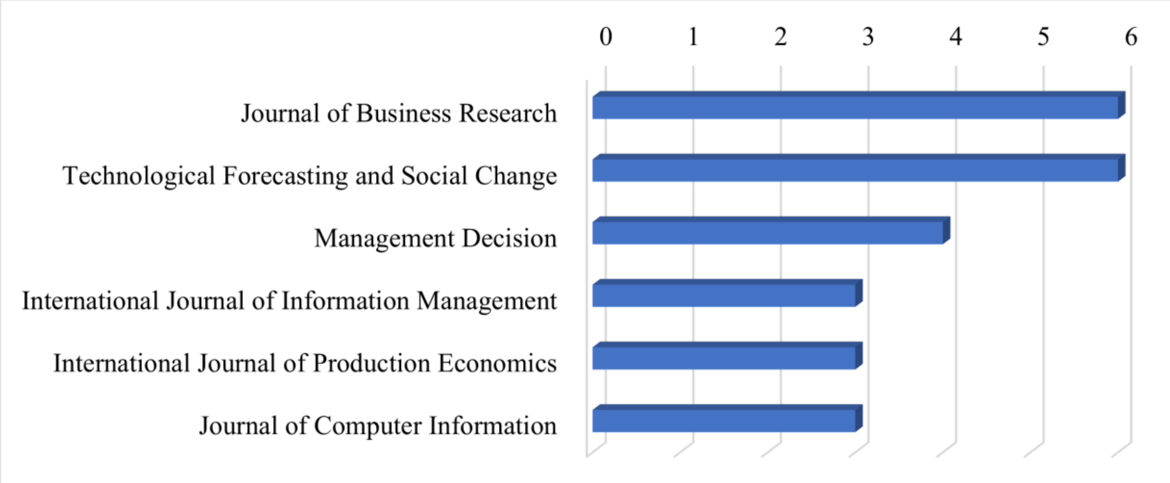


Figure 11: The most popular journal (only presenting journals with more than 2 articles)
Source: Own edition

3.2.3 Distribution of the relevant authors and authors' impact

Among the authors researching this topic, there are six authors who have published more than two articles.

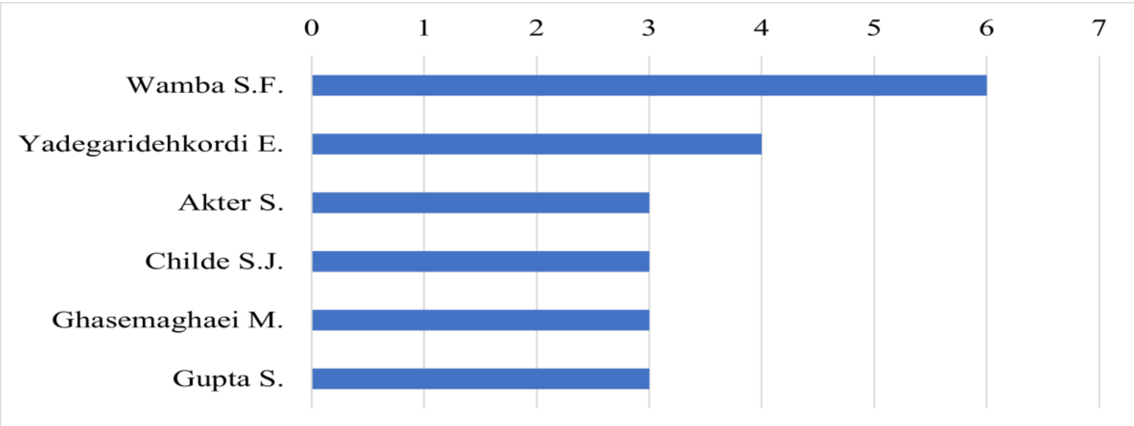


Figure 12: Top 6 productive authors overtime (2017-2022)
Source: Data collected from Scopus and WOS, presented by author

Wamba S.F. has become the most popular researcher as he has the most research articles on big data at the firm level, with six articles, followed by Yadegaridehkordi E. with four papers. Akter S., Childe S.J., Ghasemaghaei M., and Gupta S. each have three papers. Meanwhile, the highest impacted article based on the number of citations belongs to Frank et al. (2019) which presented the structure of Industry 4.0 technologies in manufacturing companies.

Table 2: The most cited articles

Titles	Year	Authors	Total Citations
Industry 4.0 technologies: Implementation patterns in manufacturing companies	2019	Frank et al.	1000
Big data analytics and firm performance: Effects of dynamic capabilities	2017	Wamba et al.	727
Big data and predictive analytics for supply chain and organizational performance	2017	Gunasekaran et al.	521
Role of institutional pressures and resources in the adoption of big data analytics powered artificial intelligence, sustainable manufacturing practices and circular economy capabilities	2021	Bag et al.	166
Modelling quality dynamics, business value and firm performance in a big data analytics environment	2017	Ji-fan Ren et al.	162
The Effect of Big Data and Analytics on Firm Performance: An Econometric Analysis Considering Industry Characteristics	2018	Müller et al.	148
Digital academic entrepreneurship: The potential of digital technologies on academic entrepreneurship	2019	Rippa & Secundo	134

Note: only articles with more than 100 citations are shown

Source: Own edition

3.2.4 Distribution of previous studies by countries' scientific production

The result from bibliometrix shows that the leading position in paper production (based on authors' affiliations) is mainly the story of Europe countries, the US, and the Asia continent.

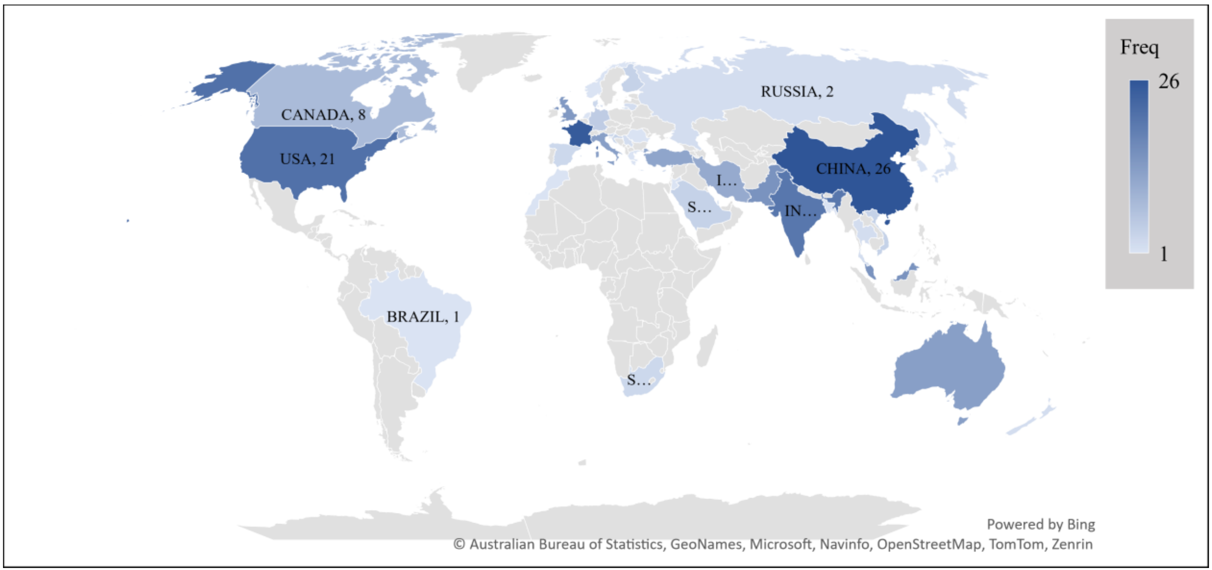


Figure 13: The distribution of Country Scientific Production
Source: Own edition

Note: The figure created based on the acquisition citation data obtained from two prominent research databases, namely Scopus and Web of Science, the author subsequently employed a Bibliometric package to gather information about the affiliations of the previous authors involved in the discussed subject matter.

Among the top 10 productive countries, half is concentrated in Anglo-Saxon countries (such as France, the US, the UK, Italia, and Australia), and another half is concentrated in developing countries in Asia (such as China, India, Malaysia, Pakistan, and Iran). The participation of researchers from Africa and South America is particularly rare.

3.2.5 Distribution of previous studies by investigated countries

Of the 85 studies conducted, about 50% studies were investigated in Asian countries (41 studies). In Asia, four countries, such as China (12 studies), Pakistan (8 studies), India (8 studies), and Iran (6 studies), were dominant in researching big data at the organizational level. There were 21 studies in Europe (not including studies on a global scale), with half conducted in France (5 studies), the UK (3 studies), and Spain (2 studies). Studies conducted in North America, primarily in the United States (5 studies), are the primary focus of American research. No research has been conducted in South America and Africa, except in Brazil, South Africa, and Egypt. The remaining studies were investigated in a dispersed manner across Southeast Asia, Europe, and Arabic countries.

3.2.6 The evolution of topic trend

The figure below displays the central topics of previous studies from 2017 to 2022, as revealed by overlay visualization analysis (from VOSviewer). The results indicate that research topics concerning the utilization of big data have evolved recently (Figure 14). Around 2019, researchers focused on applying the resource-based view theory to evaluate the role of big data in creating business value and increasing customer satisfaction for firms. Subsequently, researchers devoted considerable attention to topics related to data analytics management. Many studies employed the dynamic capabilities theory to assess the impact of big data on a firm's competitive advantage and innovation performance. Recently, there has been much attention from researchers on the research direction focusing on evaluating the BDA capabilities, the role of human resources, and the combination of big data analytics with AI, cloud computing, Internet of Things (IoTs).

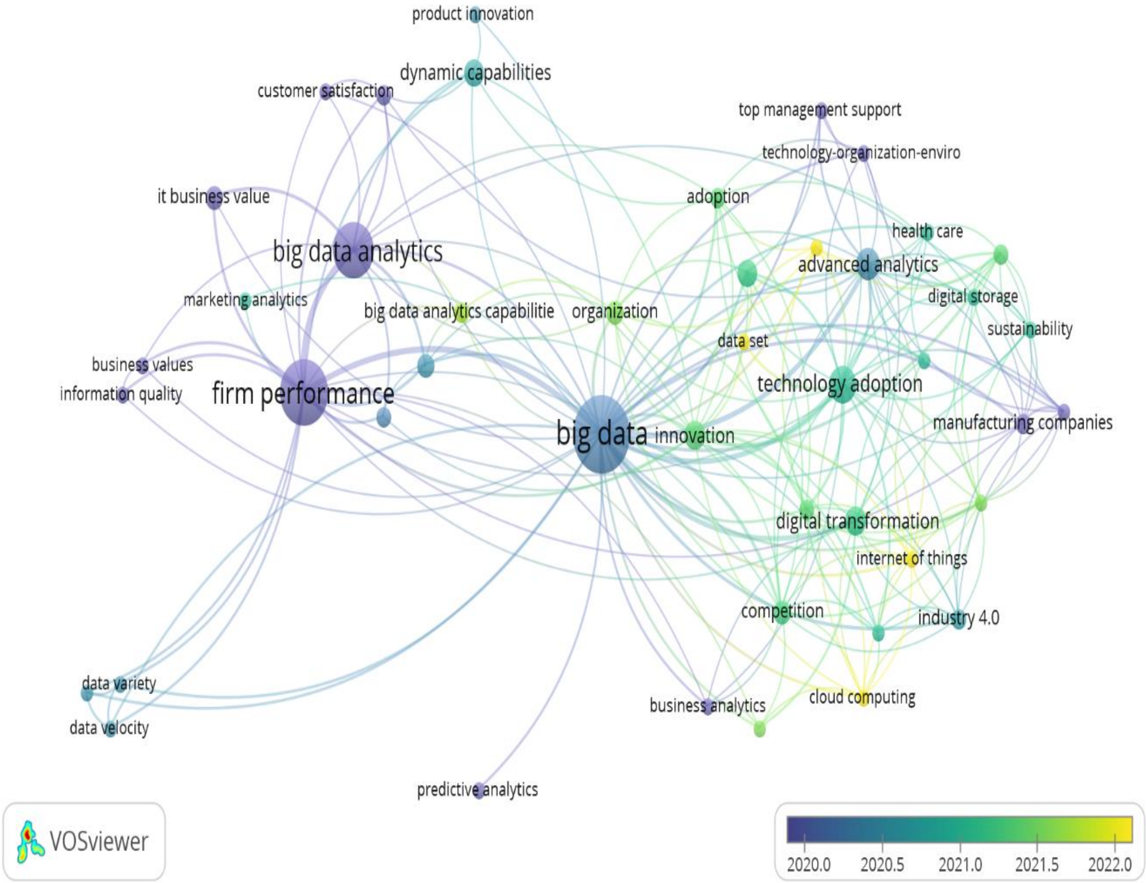


Figure 14: The evolution of topic trend
Source: Own edition

Note: The data utilized for the generation of this visualization was extracted from the keywords present in previously published articles. The author employed VOSviewer software to perform the analysis, specifically selecting a minimum threshold of two keywords while excluding structural terms such as "literature review," "conceptual framework," "survey," and "future research directions."

3.2.7 Distribution of previous studies by sectors

Out of eighty-five studies, thirty-two have focused on big data for specific industries, with the most attention being given to manufacturing (twelve studies), information technology and telecommunications (five studies), hospitality-tourism-transportation (three studies), healthcare (three studies), and service & retail (three studies). The remaining studies are dispersed across various sectors, including construction, education, and mining. Meanwhile, the majority of studies are based on cross-sector data.

Significantly, 14 studies have been conducted on particular firm sizes, with most researchers focusing on the application of BDA in SMEs. Of these, 11 studies have been dedicated to micro, small, and medium enterprises (Yu et al., 2022; Gurău & Ranchhod, 2020; Maroufkhani P., Tseng et al., 2020; Maroufkhani P., Wan Ismail et al., 2020; Saleem et al., 2020; Yadegaridehkordi et al., 2020; Başak et al., 2022; Bhatti et al., 2022); while only one study has been conducted on start-up companies (Behl, 2022). In contrast, only two studies have been conducted on large enterprises (Mahmoud Ali et al., 2022; Begenau et al., 2018) and a study on medium and large enterprises (Raguseo & Vitari, 2018).

3.2.8 Distribution of studies by research method

Figure 15 illustrates the research methods utilized in previous studies concerning this topic. The empirical method was the dominant research method for examining the usage/adoption of big data at the enterprise level. In addition, other methods were used occasionally in a few studies, including case studies (6 studies), conceptual research (7 studies), and review (4 studies).

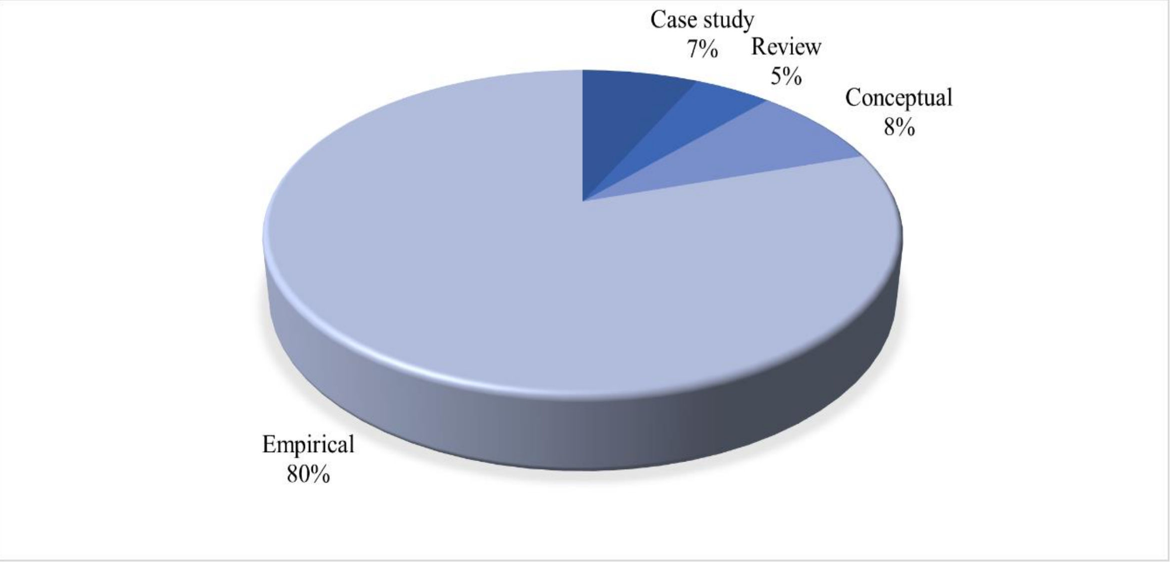


Figure 15: The distribution by research method
Source: Own edition

3.3 Distribution of studies by research area

Cluster analysis conducted using VOSviewer network visualization revealed two main themes for studies on the utilization of big data in enterprises: (1) technology adoption and (2) the impact of big data on firm performance.

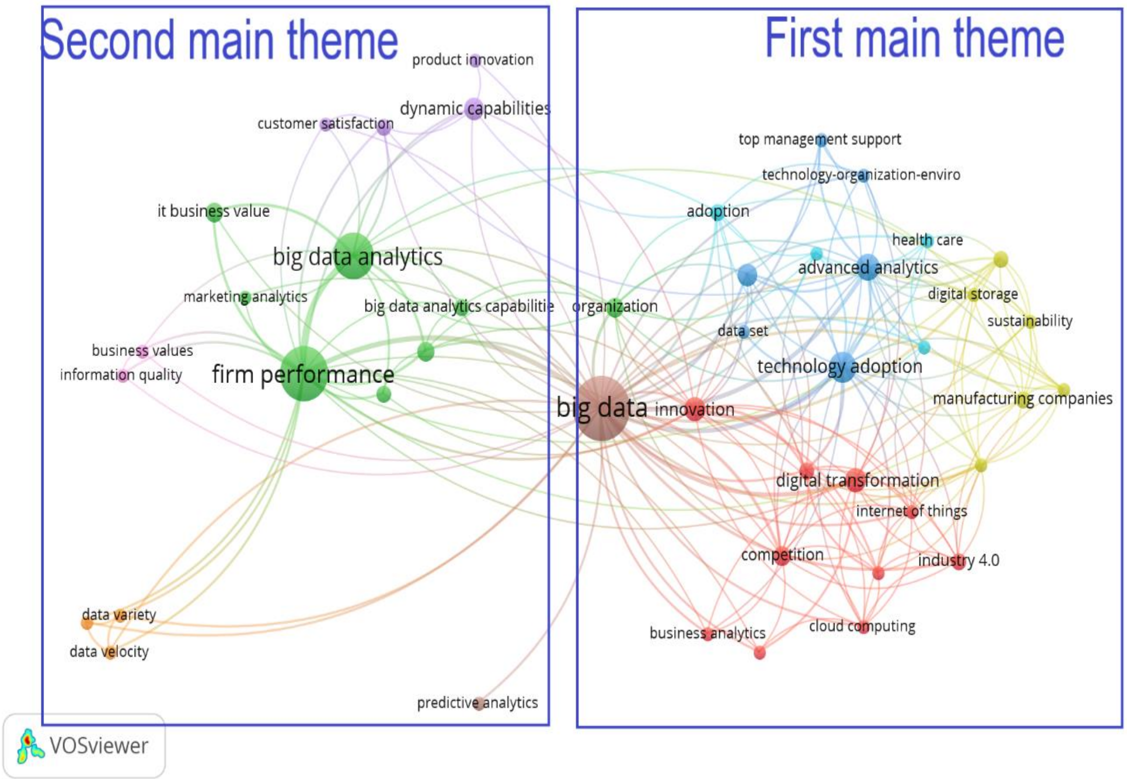


Figure 16: The classification of research area based on cluster analysis.
Source: Own edition

3.3.1 First main theme: technology adoption

Previous studies on this cluster have emphasized the significance of digital technologies, with big data technology emerging as a key technology. An empirical study from Vărzaru (2022) noted the beneficial effect of digital technologies on the accounting performance of healthcare organizations. Additionally, the adoption of digital technologies, such as big data, is a strategic decision of great importance for manufacturing (Frank et al., 2019) and telecommunication firms (Moumtzidis et al., 2022).

The implementation of big data technology as part of digital technologies in an enterprise necessitates a comprehensive framework and strategic planning. Frank et al. (2019) and Kayabay et al. (2022) have both addressed the strategic use of digital technologies in enterprises. Frank et al. (2019) asserted that the use of digital technologies at the enterprise level is a combination of front-end and based technologies. Front-end technologies, such as smart manufacturing, smart products, smart supply chain, and smart working, are complemented by base technologies, including IoTs, cloud services, BD, and analytics (Frank et al., 2019). Frank et al. (2019) found that smart manufacturing is the most adopted front-end technology. Meanwhile, companies find it difficult to implement base technologies, as big data and analytics are not widely used in the studied sample. The findings of this study raise a series of questions regarding the factors that influence the utilization of big data analytics in enterprises at a strategic level. Kayabay et al. (2022) explained that, in the context of data accumulation over time, traditional enterprises must strategically transform into data-driven enterprises in order to capitalize on the benefits of data. However, in order for organizations to successfully transition into data-driven enterprises, they must acknowledge and overcome various challenges related to data, organization, technology, and strategy, as emphasized by Kayabay et al. (2022). In recognition of these challenges inherent in the transformation process, Kayabay et al. (2022) put forth a Data Science Road mapping framework that encompasses strategic, data, technological, and organizational aspects. This framework serves as a guide to support established companies in their journey towards becoming data-driven entities.

In addition, digital technologies not only present challenges and advantages to established industries but also bring these same opportunities to those new to the field, such as digital entrepreneurs. In accordance with the findings of Frank et al. (2019), Rippa & Secundo (2019) affirmed that digital technologies, such as big data analytics, internet of things (IoT), 3-D printing, and cloud services, have a pervasive influence on all facets of organizations. Recognizing the shortage of research on the role of digital technologies in academic entrepreneurship, Rippa & Secundo (2019) proposed a novel concept of digital academic entrepreneurship. This concept is akin to a roadmap, aiding entrepreneurs in understanding the role of digital technologies, the various forms of emergent digital academic entrepreneurship, the stakeholders engaged through digital technologies, and the processes of academic entrepreneurship.

Nevertheless, each digital technology has distinct features that require individual assessment. For instance, data quality is essential when utilizing BDA; however, service quality has a considerable impact on the effectiveness of IoTs technology, which in turn significantly influences the intention to use this technology (Moumtzidis et al., 2022). Bag et al. (2021) suggested that the implementation of BDA in conjunction with AI in manufacturing companies could improve supply chain management and circular economy capabilities. To realize this positive impact, however, firms should concentrate on upgrading intangible resources and workforce skills (Bag et al., 2021).

The implementation of big data in enterprises has attracted considerable scholarly attention due to its significant impact on enterprises, requiring careful consideration of its benefits and difficulties. Fanelli et al. (2023) conducted a review of 20 relevant articles to assess the contribution of BDA to the healthcare sector. This review identified four distinct categories of BDA contributions: quality of care, quality of services, crisis management, and data management. Notably, Saeed et al. (2022) found that a one-point increase in big data utilization is associated with a 27.4% increase in firm legitimacy, with the effect being more pronounced for firms in highly competitive industries.

Previous studies have confirmed that the usage of BD contributes to various aspects of enterprises. However, enterprises' success rate using BDA is lower than expected. Big Data is not a one-size-fits-all model; success in its application depends on numerous factors. This explains why most of the prior research in this area has been dedicated to exploring the factors that affect the adoption of big data.

This research direction is largely based on the TOE framework. Verma & Chaurasia (2019) conducted a study to examine the various factors influencing adopters and non-adopters of BDA. The results indicated that relative advantage, complexity, and competitive pressure had a significant effect on both groups. However, for BDA adopters, other factors such as compatibility, top management support, technology readiness, and organizational data environment significantly influenced big data analytics adoption. Meanwhile, Park & Kim (2019) highlighted that the perceived benefits of big data are the most influential factors in accepting BDA. Additionally, experts and firms have acknowledged technological capabilities, financial investment capabilities, and data quality and integration as important contributors to BDA adoption. Chaurasia & Verma (2020) found that perceived strategic values of BDA have a significantly positive impact on BDA in the construction sector.

Maroufkhani, Wan Ismail, et al. (2020) investigated the adoption of BDA in 112 Iran SMEs and found that although three dimensions in the TOE framework contribute significantly to creating business value from BDA, technology and organizational dimension have more significant impacts. This research agreed that SMEs should consider the effective utilization of BDA as a capacity to have a beneficial effect on company performance.

Combining TOE and diffusion of innovation theory (DOI), Baig et al. (2021) found that all studied factors, including relative advantage (RA), complexity, compatibility, top management support (TOP), financial resources, human expertise and skills, competitive pressure, security and privacy, and government policies, had a significant influence on the BDA adoption in the education sector in Pakistan. Nevertheless, it was remarkable that IT infrastructure did not influence the adoption of BDA. Bakici et al. (2022) conducted a qualitative study in which they interviewed experts in the field of project management to propose a framework of factors that influence the adoption of big data. TOE was also used as the foundational theory for the other El-Haddadeh et al. (2021) and Youssef et al. (2022) studies.

A limited amount of research has employed TAM and UTAUT. For instance, Gurău & Ranchhod (2020) combined TAM and the practice-lens perspective in their study. Through a case study analysis of a small firm in the transportation sector, Gurău & Ranchhod (2020) discovered that the implementation of BDA applications in small firms is based on the company's and its employees' specific needs, and the technology only becomes functional through the direct involvement of human agents. This research supports the importance of entrepreneurial managers who take proactive actions while attempting to reduce market risk and increase competitive advantage. Cabrera-Sánchez & Villarejo-Ramos (2020) used the UTAUT model to explore the determinant of big data technologies in the retail sector. The study found five factors influencing the intention to use the BD technique: performance expectancy, social influence, resistance use, facilitating conditions, and opportunity cost.

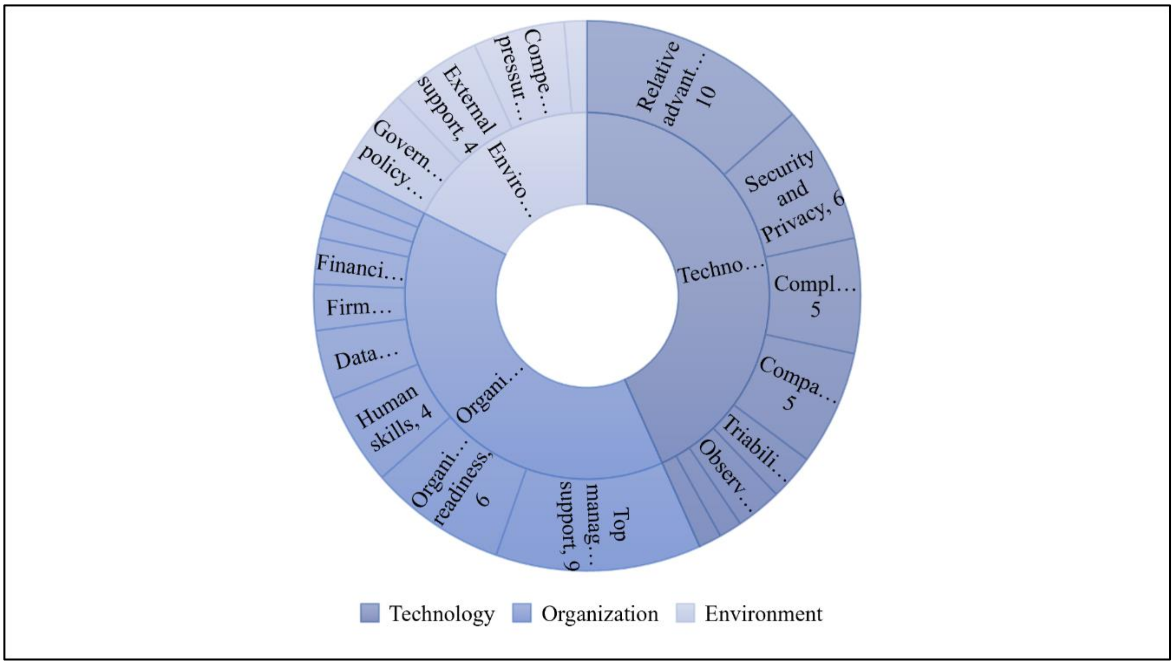


Figure 17: The significant occurrence of factors from TOE's dimension
Source: Own edition

Generally, the significant impact of various TOE factors has been found in previous studies. However, the use of BD depends on many factors, making it more complex than simply implementing a software program or specific technologies, such as mobile banking services and digital financial services. Baig et al. (2019) conducted a review of 20 studies on the adoption of big data and identified a total of 42 significant factors in four dimensions: technology, organization, environment, and innovation, which have an impact on the adoption of big data. Figure 17 presents the most significant mentioned factors among previous studies.

Among these studied factors, top management support, relative advantage, security and privacy, compatibility, complexity were the significant dominant factors in previous studies. The emergence of numerous factors associated with the advantages of BDA, such as relative advantage, perceived benefits, perceived usefulness, and perceived strategic value, indicates that the benefits of BDA are the most critical considerations for businesses when implementing BDA. Many studies highlighted the importance of BDA's relative advantage, benefits, strategic value, usefulness, and perceived performance (El-Haddadeh et al., 2021; Lutfi et al., 2022; Shahbaz et al., 2020). In a research study aimed at achieving sustainable development goals through the adoption of BDA, El-Haddadeh et al. (2021) discovered that the perceived benefits, the complexity of BDA, and top management support had a significant impact on the decisions to adopt BDA in enterprises within the United Kingdom. Likewise, Lutfi et al. (2022) discovered that within hotel industry, the recognition of the advantages associated with big data could serve as the sole motivation for adopting this technology, even if other technical factors do not meet the necessary criteria. Logically, firms would only utilize big data if they derived more advantages. For instant, in the education sector, big data help organizations have greater control over work to accomplish tasks more quickly and increase data storage (Baig et al., 2021). This finding was also reported by Park & Kim (2019) and Maroufkhani, Wan Ismail, et al. (2020). In the healthcare sector, BDA helps employees improve their job performance (Shahbaz et al., 2020). Additionally, security and privacy are the primary determinants of the implementation of BDA. These concerns are seen as the primary impediment to the retail industry's adoption of BDA (Youssef et al., 2022). Even in the hotel sector, security and privacy are the most significant inhibitors of using BD (Yadegaridehkordi et al., 2020). Prior studies have found that other factors from DOI theory, such as compatibility, complexity, observability, and trialability, often affect BDA adoption.

Previous studies have identified several factors from the organizational dimension that influence the use of BDA, with top management support being particularly important. Other factors include organizational readiness, human skills, data quality, financial resources, and firm size (Maroufkhani, Wan Ismail, et al., 2020; Verma & Chaurasia, 2019; Youssef et al., 2022). Previous studies have not often discussed the environmental dimension. Yet, some factors were identified as having significant impacts on the utilization of BDA, such as competitive pressure, external support, government policy, and support.

3.3.2 Second main theme: the impact of big data on firm performance

This research theme investigates big data's effects on enterprises in terms of business value, supply chain, innovation performance, and BDA capabilities. Research on the business value aspect, Ji-fan Ren et al. (2017) posited that big data can act as a catalyst to facilitate more efficient business operations. This paper investigated the quality dynamics in a big data environment associated with increased business value and firm performance. It was found that the business value of big data mediates the positive correlation between system and information quality and firm performance. Furthermore, Raguseo & Vitari (2018) found that investments in big data analytics generate business value, which may lead to financial benefits for a firm. This business value can be divided into four categories: transactional value, strategic value, transformational value, and informational value. Nevertheless, Ghasemaghaei (2019) contended that a willingness to invest in BDA does not necessarily ensure success in deriving value from it. Rather, Ghasemaghaei (2019) proposed that firms must be structurally and psychologically prepared. Structural readiness comprises IT infrastructure, tool functionality, BDA skills, and data size. Psychological readiness is determined by the extent of the mindset and beliefs to proactively seek business solutions based on existing IT resources.

Supply chain management is a field that derives advantages from the utilization of big data analytics. Research on the supply chain performance aspect, Gunasekaran et al. (2017) highlighted the role of top managers in constructing BDA assimilation capability. Therefore, top managers' commitment to obtaining resources (Connectivity and Information Sharing) and constructing BDA assimilation capability would improve supply chain and organizational performance. In addition, Gu et al. (2021) discovered that supplier development did not initially impact firm performance, but the relationship became significant as BDA was adopted.

In research on the innovation performance aspect, Yu et al. (2022) found that the impact of green dynamic capability on green innovation adoption was more robust when moderated by BDA capabilities. Niebel et al.'s (2019) analysis of German firm-level data within a knowledge production function framework suggested that big data analytics are important factors in determining the likelihood of a firm becoming a product innovator and the market success of product innovations in both the manufacturing and service sectors. Furthermore, investment in IT-specific skills is necessary to realize these results.

By collecting survey questionnaires, many previous researchers have also found evidence of the impact of big data analytics on firm performance. Chatterjee et al. (2022) discovered that BDA has an indirect influence on financial and operational performance, which is mediated by business processes and dynamic capabilities. Rakic et al. (2022) found that digital services based on BDA had the most significant effect on firms' revenue in manufacturing firms. A survey from 159 Indian firms by Mishra et al. (2019) also found that BDA adoption directly impacts firm performance, while BDA adoption is influenced by IT deployment and human resources capabilities.

Thirathon et al. (2022) found that utilizing BDA could be beneficial for managers in the decision-making process. The authors note that managers in small firms tend to possess more quantitative skills than those in larger firms, thus leading to decisions being based more on analysis than those of managers in larger firms.

A study by Fosso Wamba et al. (2019) discovered that the excellence of quality dimensions propels superior firm performance. Therefore, IT and talent quality are the most important for BDA quality. Furthermore, Ghasemaghaei (2021) supported that evaluating BDA quality should not only evaluate as a single variable; it is necessary to evaluate the characteristics of big data separately for better understanding. Ghasemaghaei (2021) found that the influence of big data on firm performance is contingent upon the characteristics of big data. This study revealed that four characteristics of big data (i.e., volume, velocity, variety, and veracity) had no significant impact on firm performance. However, these characteristics of big data influence the data value, which has a significant impact on firm performance.

The contributions of big data and big data analytics to enterprises are numerous; however, they vary depending on the characteristics of each industrial sector, firm size, and the adoption time frame. According to Begenau et al. (2018), large firms have an advantage in the modern economy due to their increased economic activity and longer firm history. This allows them to produce more data, which can be used to improve investors' forecasts and reduce equity uncertainty. As processor speed rises, large firms benefit more from data analysis, reducing their cost of capital and enabling them to grow larger. (Müller et al., 2018) found that the returns from BDA varied depending on the firm's industry, with IT-intensive and highly competitive industries being the most successful in extracting value from BDA assets. Interestingly, Huang et al. (2020) suggested that the impact that the implementation of big data has on firm performance varies depending on the firm level or sector. A global study by Huang et al. (2020) which analyzed data from 174 enterprises from 2010-2014, found that big data implementation can positively affect financial performance and market value but not productivity. Furthermore, the effect is not stronger for first movers. Using panel data from 176 firms, a study by Raguseo et al. (2020) found that firm size is a significant factor in determining the profitability of firms when BDA is utilized. The authors suggest that the success of BDA solutions in terms of firm profitability may be contingent upon the availability of resources. Consequently, the authors posit that dynamic capabilities for deploying resources can complement the firm's ability to select resources in the value-creation process through BDA. BDA's value lies in its ability to accumulate or grow the available resources.

Despite the significant BDA contributions, recently, the question of how to enhance its contributions in order to yield greater efficiency for businesses has become increasingly important yet needs to be better studied Sabharwal & Miah (2021). In order to answer this question, recent studies have expanded on the topic of BDA capability. In accordance with the previous literature on IT and BDA, many researchers investigated the classification of BDA capabilities. Anwar et al. (2018) argued that BDA capabilities consist of BD technological and personal capabilities. Combining these two capabilities allows enterprises to use information concerning markets and customers, which enables them to recognize external trends and variations and consequently create more suitable products and services that grant them a sustainable competitive advantage. Akter et al. (2019) further added information capabilities to BDA capabilities. They suggested that enterprises should attempt to acquire the resource base to construct dynamic capability (including dynamic technology, talent, and information capability). These three dynamic capabilities would lead to sustained firm performance. Here, Akter et al. (2019) emphasized the role of human resources, stating that hiring personnel with strong analytics abilities and preparing the current workforce through official training programs would ensure the success of developing BDA capability. Similarly, S. Gupta et al. (2019) suggested that combining key resources such as data, managerial abilities, and technical abilities will enable enterprises to construct BD predictive capability, which positively affects firm performance and operational performance. According to Razaghi & Shokouhyar (2021), firms could only convert BDA into improved business performance by focusing on BDA management capability. Furthermore, BDA management capability consists of four components: planning, decision-making, coordination, and control, which enable firms to respond quickly and effectively to changes in the industry and their environment (Razaghi & Shokouhyar, 2021).

Building on RBV, DC and contingency theory, Gu et al. (2021) conducted a study to explore the relationship between BDA capability and firm performance. The authors argued that the usage of BDA in supply chain could be considered as a dynamic capability of firm. By analyzing 108 firms, the study confirmed that BDA capabilities have a direct positive impact on both firm performance and supplier development. Additionally, the study uncovered evidence that the dynamic capability generated through the use of BDA mediated the relationship between supplier development and firm performance.

3.4 Review of relevant theories

This part initially reviewed the most employed theories for explaining the factors influencing BDA usages, such as the technology acceptance model, diffusion of innovation, and technology-organization-environment framework. Subsequently, theories on the impact of BDA usage on banking performance, including the resource-based view, knowledge-based view, and dynamic capabilities theory, are discussed. By laying out these relevant theories in using digital technology research, this section highlighted the advantages and disadvantages of each.

3.4.1 Technology Acceptance Model

The technology acceptance model (TAM) theory was introduced firstly by Davis (1989), which concluded two important factors influencing users' intention behavior: perceived usefulness (PU) and perceived ease-of-use (PEOU). Davis et al. (1989) defined perceived usefulness as the degree to which an individual believes using a particular technology would enhance their job performance. Perceived ease-of-use refers to the extent to which an individual believes using technology is effortless. According to Davis (1989), perceived usefulness and perceived ease of use of new technology positively impact the intention to adopt it. From an individual perspective, the TAM theory has become widely used to explain why banks' customers adopt mobile banking technology (Shaikh & Karjaluoto, 2014).

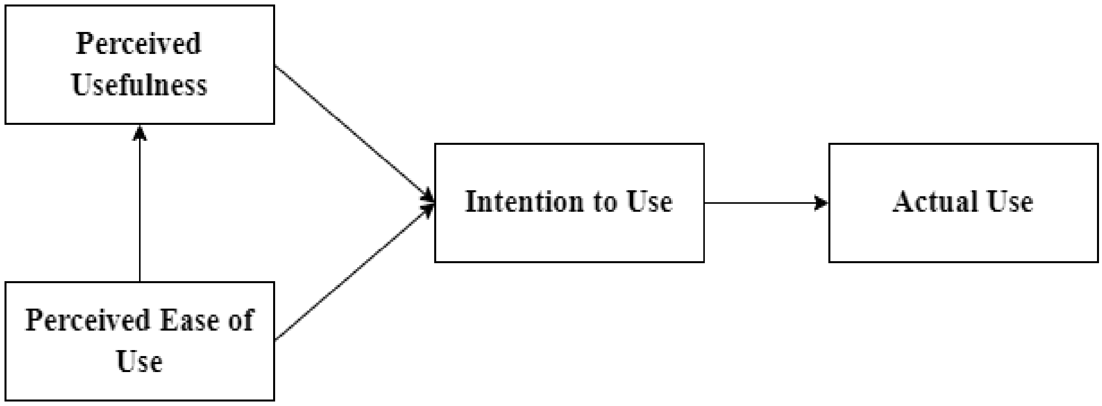


Figure 18: Technology Acceptance Model
Source: David (1989)

Much research on innovative technologies topics has been based on the TAM theory. In addition, numerous authors have extended TAM by incorporating various external factors, resulting in TAM2 (Extended technology acceptance model) (Venkatesh & Davis, 2000). The extended technology acceptance model enhances the comprehensibility of TAM theory by introducing new variables such as social influence and cognitive instrumental processes (Venkatesh & Davis, 2000). Many studies at the firm level utilize TAM theory as their foundation for examining the adoption behavior of new technologies, even though initially, this theory was used at the individual level. Gangwar et al. (2015) discovered that the technology and organization elements from the TOE framework have a significant effect on the TAM factors, which in turn positively affect cloud computing adoption from a firm perspective when studying the factors that influence its adoption using a combination of TAM and TOE.

3.4.2 Diffusion of Innovation

To explain an innovation's conceptual content, Rogers (1995) pointed out that an idea, practice, or object that is new and patent to the adoption unit can be clarified as an innovation. The diffusion of innovation theory (DOI) from Rogers (1995) provides a valuable account of how innovation diffuses. As noted by Rogers (1995), the diffusing process of a technological innovation often takes a long time is influenced by many factors from the moment it was created to the time of acceptance by individuals and society. According to Rogers (1995), adopters of innovation could be classified into 5 types, including: innovators, early adopters, early majority, late majority, and laggards.

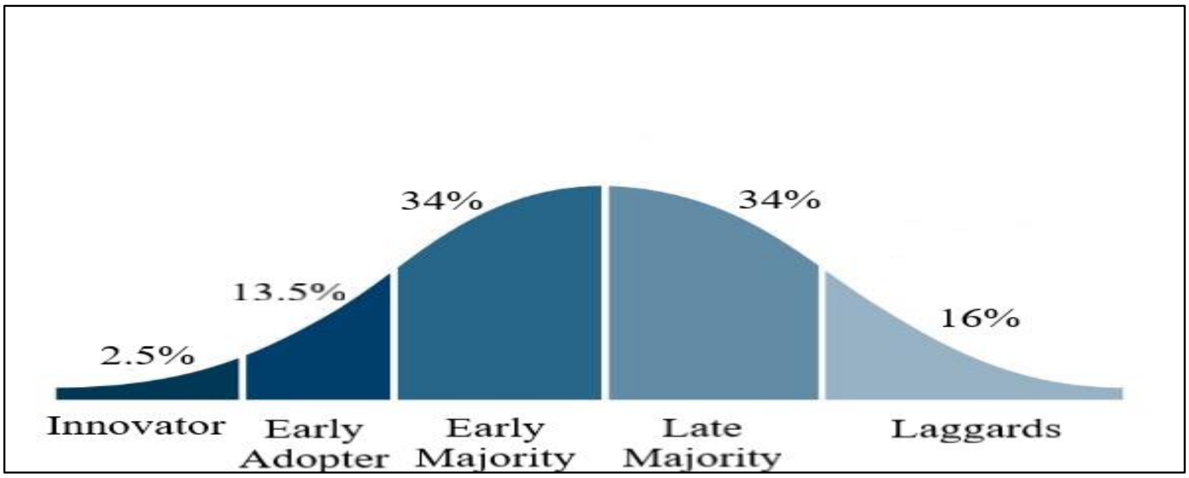


Figure 19: Diffusion of Innovation Curve
Source: Rogers (1995)

For innovation to be widely applied among individuals or organizations, it is necessary to increase the rate of diffusion of an innovation (Rogers, 1995). According to the DOI theory, Rogers (1995) highlighted five critical factors related to innovation's characteristics: relative advantage, compatibility, complexity, trialability, and observability. Among these characteristics, relative advantage, compatibility, trialability, and observability are the drivers to speed up innovation, whereas complexity is considered the barrier that should be increased to increase innovation diffusion in organizations. Relative advantage implies the extent to which an innovative application is identified as more beneficial than the other solution (Rogers, 2003). Much empirical research has found the critical role of relative advantage in the field of innovation adoption. There are many criteria to estimate the relative advantage in the domain of big data applications. For example, organizations adopting big data solutions can enhance their service quality, achieve more business opportunities, raise their competitiveness (Sun et al., 2018). Compatibility implies the extent to which an innovation is perceived as consistent with the present values, past experiences, and needs of likely adopters (Rogers, 2003). Organizations would likely adopt a specific big data application compatible with their technological infrastructure (Sun et al., 2018). Complexity refers to the extent to which an innovation is perceived as hard to understand and use (Rogers, 2003). Big data applications' complexity was a barrier affecting the organization's adoption decision (Sun et al., 2018). Rogers (2003) suggested that relative advantage, compatibility, and complexity were the dominant factors when explaining diffusion rate. Existing research recognizes DOI's critical role in explaining the factors influencing the diffusion rate of technological innovations. Al-Jabri & Sohail (2012) analyzed the data from mobile banking customers and concluded that factors from DOI theory could explain 42.8% of mobile banking adoption. The study also found that relative advantage, compatibility, and observability have a significantly positive influence on the customers' decision to adopt MB, whereas trialability and complexity have no impact on mobile banking adoption. Recently, Verma & Chaurasia (2019) found that based on the combination of TOE and DOI, two factors from DOI (relative advantage and complexity) significantly influenced the adoption of big data analytics, while compatibility had an insignificant impact.

3.4.3 Technology-Organization-Environment Framework

The technology-organization-environment framework (TOE) was introduced in 1990 by Tornatzky & Fleischer (1990). After that, the TOE framework has become a widely-used theory in adopting information technology to explain IT adoption at the firm level (Lai et al., 2018; Verma & Chaurasia, 2019). According to Tornatsky & Fleischer (1990), many factors that impact the adoption of innovation in firms can be grouped into three main contexts: technological, organizational, and environmental. *The technological context* indicates all the applicable technologies in usage at firms or in the marketplace. Each type of new technology will have a different impact on organizations. Baker (2012) described two types of innovations, namely: competence-enhancing technologies and competence-destroying technologies. The first one will gradually change the firm's resources and improve the existing technologies in the enterprise.

Meanwhile, the latter will quickly replace, eliminate the existing resources, and transform to new technologies leading to a dramatic transformation of the firm's resources. Therefore, managers need to consider technological characteristics when adopting an innovation. *The organizational context* indicates the typical features or inner resources of organizations. This context includes communicating style in firms, reporting relationships, working structures, and top management support toward innovative changes. Meanwhile, *the environmental context* indicates factors from the external environment such as competition in the industry, the level of use of new technology from customers, suppliers, third parties, and the relevant regulatory aspect.

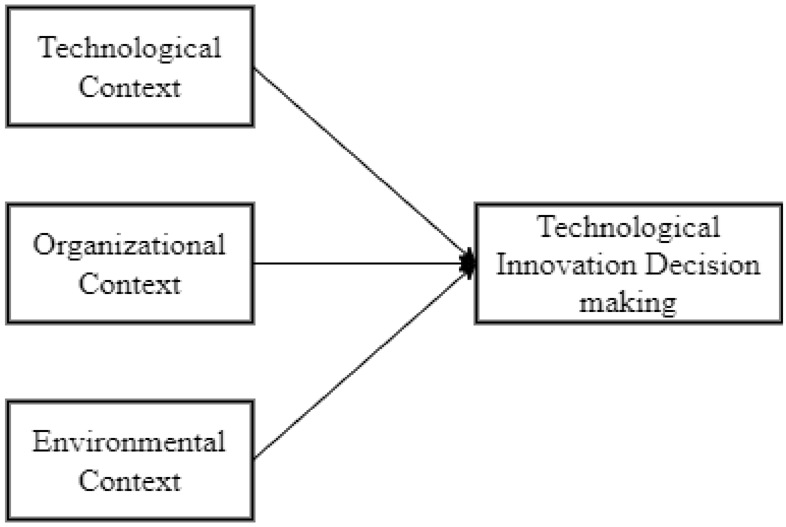


Figure 20: Technology-organization-environment framework
Source: Tornatzky & Fleischer (1990)

Many studies based on TOE assess the factors influential in the adoption of information technology and information systems in firms (Hsu et al., 2014), such as cloud serve adoption (Hsu & Lin, 2016); big data adoption (Maroufkhani, Wan Ismail, et al., 2020; Park & Kim, 2019). Recently, a study that conducted by Kulkarni & Patil (2020) investigated blockchain technology in banking services by adopting the TOE framework. This study has shown that all studied factors (relative advantage, compatibility, perceived cost, security, firm scope, learning culture, top management support, competitive pressure, government policy, and consumer readiness) have a significant influence on adopting blockchain technology in Indian banks. These factors can explain 67.4% of the variance of the dependent variable. Similarly, a study from Gangwar (2018) extended the TOE framework by adding a context construct - security- when examining big data adoption drivers in Indian firms. Some others attempted to combine TOE with other Information systems (IS) theories to add a deeper understanding and higher explanation to the model. A cloud enterprise resource planning (ERP) adoption study from AlBar & Hoque (2019) combined the TOE framework with DOI theory. This study found that TOE factors such as competitive environment, ICT infrastructure, regulation environment, ICT skills, and top management supports significantly impact the cloud ERP in Saudi Arabia companies. Meanwhile, some factors from DOI, such as observability and relative advantage, were significantly important in cloud ERP. Similarly, Oliveira & Martins (2010) also combined TOE with Iacovou et al. (1995) model to investigate the adoption of e-business. In a study of 112 SMEs in Iran's manufacturing sector, Maroufkhani, Wan Ismail, et al. (2020) identified that the adoption of BDA could enhance the financial and market performance of enterprises; however, only technological and organizational dimensions have a significantly indirect impact on market performance through mediating of BDA adoption.

The advantage of TOE: The TOE approach facilitates the identification and evaluation of the factors influencing the decision to adopt innovative technologies in a comprehensive manner. Furthermore, the TOE's flexibility enables the researcher to select factors according to each technology's specific characteristics and context. The TOE framework is in line with Rogers' (2003) theory of innovation diffusion, which acknowledges five technological attributes as determinants of any adoption decision: relative advantage (RA), complexity (CX), compatibility (CM), observability, and trialability. Consequently, the TOE framework elucidates the adoption of innovation and a substantial number of empirical studies have been conducted in various IS domains.

The disadvantages of TOE: The TOE is rarely used to explain the impact of innovation on firm performance. Rather, elements of the TOE are investigated as requirements for the successful implementation of innovation, thus constituting firm capability/resources, which consequently influence firm performance.

3.4.4 Resource-based View

The resource-based view (RBV) is the most popular and effective theory in investigating the impact of information technology on firm performance. The RBV was originally first addressed by Penrose (1959) who viewed firms as repositories of physical and human resources. Based on the RBV, Barney (2014) argued that the firm's performance depends on the attributes/ quality of its resources and capabilities. The RBV theory was based on two main assumptions about resources (Barney, 1991). Firstly, the strategic resources of firms in an industry are heterogeneous. Secondly, these strategic resources cannot be easy to mobilize among firms. The RBV posits that firms are composed of a distinct set of idiosyncratic resources and capabilities (Grant, 1996). Resources are central to the resource-based view (Curado & Bontis, 2006); however, there are different types of resources varying across industries and evolving over time, thus rendering the concept of resources dynamic. Researchers proposed different classifications of resources. Barney (1991) has provided a classification of firms' resources, including physical, human, and organizational resources. Grant (2010) suggested that the firm's resources include tangible resources (such as financial, physical resources), intangible resources (such as organizational culture), and human resources (such as employee skills). Barney (1991) noted that not all these resources are strategic resources. Theoretically, firms use strategic resources that are valuable, rare, inimitable, and nonsubstitutable to create their sustainable competitive advantage (Barney, 1991; Eisenhardt & Martin, 2000). According to the RBV, the difference in performance among firms within an industry could be explained using their unique resources and capabilities.

Studies that apply RBV often refer to a firm's resources and capabilities from various perspectives. For example Ji-fan Ren et al. (2017) evaluate the quality of resources that BDA brings to the firm. Drawing on RBV and the information systems success theory, Ji-fan Ren et al. (2017) investigated the relationship between system quality, information quality, business value, and financial performance in a big data environment. The study concluded that when firms apply big data analytics, the system quality and information quality would positively impact the business value and then would have a positive impact on firm performance. A study of 219 automotive and allied manufacturers in South Africa examines the reasons for which firms engaging in manufacturing sector adopt big data analytics-powered artificial intelligence (Bag et al., 2021). Drawing on institutional theory and the resource-based view, the results of this study emphasized the influence of institutional pressures on resources and their implications for the implementation of big data analytics-powered artificial intelligence (Bag et al., 2021).

Meanwhile, other researchers have also employed RBV in studies such as Ghasemaghaei (2018); Gunasekaran et al. (2017); M. Gupta & George (2016). RBV is among the most-adopted theory in investigating big data analytics' influence on firm performance (Maroufkhani et al., 2019).

However, some of RBV's shortcomings should be noticed. Some assumptions about heterogenous resource in resource base theory has been criticized for being static, and to be unlikely in dynamic markets and the RBV fails to address the influence of market dynamism and firm evolution over time (C. L. Wang & Ahmed, 2007). For example, one of the main resource in big data technology is data which would not be a unique resource for a firm because it can be achieved by others (Braganza et al., (2017).

3.4.5 Knowledge-based view

The resource-based view serves as the foundation for the Knowledge-Based View (KBV) theory, which emphasizes the strategic importance of knowledge to a firm (Grant, 1996). KBV is concerned with "the nature of coordination within the firm, organizational structure, the role of management and the allocation of decision-making right, determinants of firm boundaries, and the theory of innovation" (Grant, 1996, pp 110). Similar to the resource-based view, the knowledge-based view does not explicitly define knowledge but rather outlines characteristics of knowledge in management, such as transferability, aggregation potential, and appropriability (Grant, 1996). In KBV, knowledge-related assets are intangible resources and long-lasting assets, which could ensure a sustainable competitive advantage for firms. Firms can gain more benefits from knowledge through organizational learning and sharing process.

3.4.6 Dynamic Capabilities

The dynamic capabilities expanded from a resource-based view of the firm and attempted to explain how firms achieve competitive advantages in dynamic markets (Eisenhardt & Martin, 2000). According to Teece et al. (1997), dynamic capabilities emphasize two important aspects of competitive advantage that have been neglected in previous strategy studies. Term “dynamic” means changing the business environment; dynamic also refers to the capacity to renew competencies to fit with the volatile business environment (Teece et al., 1997). “Capabilities” emphasizes the key role of strategic management, including three important stages: adapting, integrating, and reconfiguring all resources, skills, and functional competencies in responding to changing environments. Dynamic capabilities are considered an emerging and potentially integrative approach to understanding the newer competitive advantage sources (Teece et al., 1997). RBV pays attention to resources, while dynamic capabilities emphasize organizational and strategic routines (Eisenhardt & Martin, 2000). The dynamic capabilities term is defined as the firm’s capacity to integrate, build, and reconfigure internal and external competencies to handle fast-changing circumstances (Teece et al., 1997). After decades, many researchers also proposed their definition of dynamic capabilities. To clarify what dynamic capabilities is, Ambrosini & Bowman (2009) argued four specific points: (1) dynamic capabilities are not an ad hoc problem-solving event or a spontaneous reaction. (2) the use of dynamic capabilities is intentional, deliberate. (3) while dynamic capabilities are concerned with strategic change, they are not a synonym for it. (4) dynamic capabilities describe intentional efforts to change the firm’s resource base. To be more specific, the term “capability” in “dynamic capability” should not be separated from the adjective “dynamic.”. In the resource-based view, capabilities are either processed by which the resources are utilized or resources in the general sense. A dynamic capability is not a capability in the resource-based view sense; a dynamic capability is not a resource. In the RBV, resources are the focal point, and firms strive to acquire and select resources that will give them a competitive advantage. On the other hand, dynamic capabilities emphasize the development of capabilities to maximize the potential of the resources (Curado & Bontis, 2006). A *dynamic capability* is a process that impacts resources. Dynamic capabilities are about developing an adequate resource base. Dynamic can refer to change in the resource base, to the renewal of resources. Dynamic capabilities have a different approach in compared to the competitive force's framework (which concerned with market entry, entry deterrence and positioning), the game-theoretic models (which concerned with the interaction among rivals' move), the resource-based perspectives (which concerned with firm-specific assets). Whereas dynamic capabilities are building a framework which is based on the distinctive and difficult-to-replicate advantages. It means that to become a strategic capability, a capability must meet 3 characteristics: customer orientated, unique and difficult to replicate. To determine a firm's dynamic capabilities, Teece et al. (1997) divided it into 3 categories: process, positions, and path.

Organizational and managerial processes refer to the way things are done in the firm. *Integration:* it is particularly important for a firm to integrate external activities, internal activities, and technologies through some activities such as: strategic alliances, buyer-supplier relations, technology collaboration. The way firm manages this kind of production, collaboration would create a difference in firms' competence in comparison with other firms. It can be explained that the changing environment requires firms to change technology, productive system, therefore, it requires to integrate and coordinate many units inside firms because of high interdependency. In practical terms, there are a lot of big companies that failed because of the mismatch in organizational processes. *Learning:* as Teece et al. (1997) referred to the process by which repetition and experimentation enable tasks to be performed better and more quickly and new production opportunities to be identified. There are some important characteristics of learning. Firstly, learning is related to organizational and individual skills. Secondly, the organizational knowledge created by coordinated/communicated search procedures resides in new patterns of activity. *Reconfiguration and Transformation:* this is the ability of a firm to sense the need to reconfigure the firm's asset structure, and to accomplish the necessary internal and external transformation. As Teece et al. (1997) mentioned that this ability can be improved by frequent practice. In context of reconfiguration and transformation, if a firm has ability to scan the environment, to evaluate markets and competitors it can accomplish the change.

Position. One thing that is important to the strategic management of a firm is business assets. In context of dynamic capability, Teece et al. (1997) referred business assets to the assets related to difficult-to-trade knowledge assets, reputational and relational assets. *These assets include* technical assets, complementary assets, financial assets, and locational assets.

Path: Refers to the strategic alternatives available to the firm and the attractiveness of the opportunities. *Path dependencies:* it is said that "where a firm can go is a function of its current position and the paths ahead". *Technological opportunities:* this term refers to how far and how fast an area of industrial activity can proceed is in part due to the technological opportunities that lie before it. If firms engage more in R&D or primary research, the recognition of such opportunities can be closer.

In big data research, many authors applied dynamic capabilities theory as a foundation theory. S. Gupta et al. (2019) proposed a model based on dynamic capabilities theory to investigate the relationship between big data analytics capability, cloud enterprise resource planning services and firm performance in India. X. Chen & Zhang (2016) studied the effects of environmental uncertainty on the relationship between relative advantage and perceived credibility to mobile health services adoption in China. Meanwhile, other researchers attempted to combine dynamic capabilities theory with other such as: task-technology fit theory (TTF) (Ghasemaghaei et al., 2017). By employing the DC theory and TTF theory, Ghasemaghaei et al. (2017) concluded that the usage of data analytics work as a type of dynamic capability, then this capability would influence firms' agility.

3.4.7 Comparing relevant theories

The TAM model is one of the most popular models in studying user behavior intentions toward new technologies. However, the TAM model is selected for studies on the individual level. For firm-level research, the TAM model is rarely employed for the following reason. Firstly, the TAM model, originally composed of perceived usefulness and perceived ease of use, seems to underestimate the complexity of applying technology in enterprises. Secondly, two factors from the TAM model, perceived usefulness, and perceived ease of use, have been mentioned in theories such as TOE and DOI. According to the TOE framework, for businesses, many factors influence the decision to use new technology, including technologies, organization, and environmental dimensions. As explained by DOI, businesses need to consider five factors: relative advantage, compatibility, complexity, trialability, and observability. Accordingly, relative advantage refers to the benefits of the new technology to adopted firms. This factor has the same meaning as perceived usefulness in the TAM model. Complexity refers to how hard it is to adopt innovative technology in firms. This factor has the opposite meaning of perceived ease of use in TAM theory.

Theoretically, the simultaneous use of TOE and DOI theories often occurs in previous studies. The TOE provides a broader view; meanwhile, the elements in the DOI are frequently described as representing the technology and organizational dimension in the TOE framework (Lai et al., 2018). Practically, the TOE framework offers flexibility in research. Depending on the specific type of technology or the study's subject, researchers can remove or add the related variables to the studied model. However, these theories are often applied to explain factors influencing firms' decisions to use new technology.

Therefore, to assess the impact of technology on the firm's outcome, researchers often use resource theories such as resource-based, knowledge-based, and dynamic capabilities.

3.5 Summary of literature review and research gap

3.5.1 Summary of literature review

The topic of big data is receiving significant attention from researchers worldwide, evidenced by the increasing number of studies and high citation numbers. Research on the usage of big data analytics at the enterprise level is primarily published in specialized journals focusing on information management, business, and information systems. Researchers actively engaged in this topic primarily concentrate their efforts in Europe, the United States, and Asian countries, which are prominent regions with extensive studies conducted in this field. In recent times, there has been an increasing emphasis on research endeavors concerning the BDA capabilities, human resources implications, and the integration of big data with various other digital technologies.

With respect to the sectors under investigation, an estimated one-third of previous studies were devoted to specific industries including manufacturing, IT and telecommunications, healthcare, hospitality, tourism, transportation, and services. Nevertheless, a substantial portion of these studies tend to assess cross-sector applications. Survey methodology continues to dominate as the preferred research method.

Regarding research themes, this study identifies two major areas that have been previously investigated. The first research area is technology adoption, which explored the role of big data technology in strategic planning, digital transformation implementation, supply chain management, and firm legitimacy. Many studies focused on evaluating the factors influencing the adoption of big data analytics using theories such as TOE, DOI, and TAM. The second research area is the impact of big data on firm performance. Studies in this area assessed the effects of big data analytics on competitive advantage across various dimensions such as business value, supply chain, decision-making, innovation performance, and capabilities.

3.5.2 Identifying the current research gaps

Although the number of studies on big data analytics at the enterprise level is increasing significantly and attracting growing attention from researchers globally, the literature review above identified some current research gaps that future studies need to address.

Firstly, despite the big data would be a foundational technology contributing significantly to the effective application of machine learning and artificial intelligence in enterprises, there is a lack of research on the integration of these technologies with big data. A prominent example of such integration is the use of AI to enhance data quality, detect and eliminate dirty data prior to making data-based decisions (O'Leary, 2013). Nevertheless, previous studies examining the combined effects of AI and big data on enterprise performance remain limited.

Secondly, research on big data analytics at the enterprise level primarily focused on the manufacturing sector, with a lack of investigation into the use of BDA in banks. Within the context of digital transformation in the banking sector, this study would address this research gap by providing empirical evidence concerning the influence of using big data analytics (BDA) on banking performance.

Thirdly, the identification of appropriate research theories that align with the characteristics of BDA is an ongoing debate. In particular, understanding the use of BDA in the context of digital transformation in the banking sector, which is strongly associated with the goal of generating new business model, innovative services/products, and offerings, raises the need to assess whether the use of BDA leads to the creation of innovative capabilities that positively impact firm performance. This research also aims to provide additional evidence for future studies utilizing the resource-based view and dynamic capabilities theory in the context of digital technologies.

CHAPTER 4: CONCEPTUAL MODEL AND RESEARCH HYPOTHESIS

4.1 Introduction

This chapter develops the proposed research model and presents research hypotheses. The research model proposes the relationship between four factors: top management support, BDA infrastructure, BDA skills, big data quality, and BDA usage. Additionally, this chapter presents the relationship between BDA usage and bank performance, including risk management and market performance.

4.2 Conceptual Model

According to the literature review, many studies have been conducted to analyze the factors that influence big data usage in enterprises. These factors are varied and complex. In particular, there is evidence from prior studies indicating that big data analytics capabilities are shaped by organizational leaders' abilities to acquire key resources, such as infrastructure and human skills. Big data quality (a specific context-based factor from big data analytics) was suggested to be included in BDA research (S. Gupta et al., 2019). BDA capabilities created from the usage of BD is considered as a organizational dynamic capability (innovative capabilities) through the combination of key resource factors to produce new services/ products or new way of working within enterprises (S. Gupta et al., 2019).

The banking industry is widely regarded as a leading field in the implementation of information technology. However, the establishment of the necessary infrastructure for new technologies, such as big data, is essential. The adoption of new technology like big data analytics necessitates the acquisition of new skills; as Phan & Tran (2022) have noted, BDA skills are a combination of quantitative, computer, and business skills in the banking industry. BDA technology differs from other application software in its ability to be highly personalized, making it suitable for the conditions of each bank. The success of this technology is largely dependent on the expertise of BDA experts and staff. Furthermore, in the context of big data technology, the quality of the data is an essential factor; the more prepared businesses are for the characteristics of big data, such as high volume, high velocity, high veracity, and real data, the more opportunities they will gain from utilizing these data. Furthermore, larger data sets create possibilities for developing new services, products, and innovative applications. Accordingly, the combined factors between TOE and RBV affect the use of BDA in banks. In which BDA infrastructure is the element taken from the TOE framework. In RBV, BDA infrastructure is referred to as technical or intangible resources. This resource would have an impact on the BDA capability which was referred to as the usage of BDA. Meanwhile BDA skills are in the organizational dimension from TOE. According to RBV, human skills are considered as a type of human resource. Top management support is in the organization dimension in TOE. In RBV, the support from top management is another type of interrelation, managerial resource, which contributes to creating firm capabilities. The role of top managers in identifying and acquiring the winning resources becomes very important for businesses, even creating the economic performance of firms (Curado & Bontis, 2006). Throughout this study, the term big data analytics usage refers to innovative capabilities based on deploying big data analytics, such as creating new methods, products, and products in the bank. In digital transformation, enterprises attempt to achieve innovative things by combining and exploring BDA human resources, data, technology resources, and management expertise. As a result, these enterprises could achieve a sustainable competitive advantage. Thus, this process of using BDA should become a dynamic capability within enterprises (Braganza et al., 2017).

This thesis proposes a theoretical model based on the integration of the TOE framework, the resource-based view, and the dynamic capability theory, consisting of two main components.

Four factors (independent variables) include top management support, BDA infrastructure, BDA skills, and big data quality. These four exogenous variables are hypothesized to impact BDA usage significantly.

Three dependent variables are BDA usage, risk management, and market performance. BDA usage is hypothesized to impact risk management performance and market performance significantly.

4.3 Hypotheses on the factors influencing BDA usage

4.3.1 Top management support

Liang et al. (2007) highlight that the commitment of senior-level management encompasses both the beliefs and active involvement of top-level managers. The beliefs of top-level management indicate the managers' perspectives on the advantageous aspects of IT innovation, whereas the participation of top-level management showcases their support through the development of visions, strategies, goals, and standards for IT innovation. The beliefs and active involvement of senior-level management have a substantial influence on how organizations adopt and incorporate IT transformations. According to Mikalef and Krogstie (2020), without management support, it is difficult for firms to achieve higher levels of innovation capability. Resources are not limited to tangible assets but also encompass the interrelationships within firms (Curado & Bontis, 2006). Thus, in line with the RBV, this study considers top management support as an intangible resource. With its expertise and foresight, top management support for BDA provides a foundation for organizations to develop BDA capabilities. It is suggested that organizations need top management support to be more efficient, flexible and able to respond to the dynamic environment beyond relying solely on their IS infrastructure (Ragu-Nathan et al., 2004). When top-level managers possess a comprehension of the benefits associated with BDA, they are inclined to endorse its utilization through various means, such as establishing the necessary infrastructure, enhancing BDA competencies, and providing financial support (Lai et al., 2018). The utilization of BDA necessitates the collection, analysis, and comprehension of data from diverse enterprise functions, thereby making top-level management support instrumental in addressing and resolving communication and coordination challenges (D. Q. Chen et al., 2015; Verma & Chaurasia, 2019), while concurrently mitigating conflicts and resistance (Gangwar, 2018). Another study reaffirms the strategic significance of top-level management support in facilitating the adoption and implementation of BDA (D. Q. Chen et al., 2015). Previous studies suggested that with higher support from the top managers, there is a higher likelihood that enterprises will use BDA. Therefore, in this thesis, the hypothesis is proposed that:

H1: Top management support will have a positive impact on the usage of BDA.

4.3.2 Big data infrastructure

Previous studies highlighted the importance of technological readiness, technological competence, or organizational readiness, which refers to the preparedness of firms in terms of both resources and skills, to effectively adopt and utilize new technologies (Maroufkhani et al, 2020; Wang et al., 2010). In the context of BDA, the IT infrastructure play a vital role in providing the necessary technical foundation for the seamless implementation of BDA initiatives (Lai et al., 2018; Park & Kim, 2019). Small and medium-sized enterprises (SMEs) require sufficient technical resources, as the absence of such resources impedes their ability to implement BDA effectively (Maroufkhani, Wan Ismail, et al., 2020). Furthermore, Shirazi et al. (2022) discovered that using big data analytics tools during the aggregation and analysis stages ensures the success of new products. Therefore, BDA infrastructure should be considered the foundation for using BDA; they include hardware, software, cloud computing, high-performance computers, and large-capacity storage (Wang et al., 2021). For example, Deutsche Bank has also invested heavily in infrastructure platforms for BDA, such as data labs or Hadoop platforms, which allow this bank to digest and provide advanced analytics to get business insight from big data (Grover et al., 2018). Therefore, this thesis supports the idea that:

H2: BDA infrastructure will have a positive impact on the usage of BDA.

4.3.3 Big data analytics skills

The successful implementation and sustained management of complicated BDA projects necessitate the possession of proficient knowledge and skills by the staff involved (Ali et al., 2021; Gangwar, 2018). Grossman and Siegel (2014) noted that BDA techniques encompass the fusion of data analysis, business knowledge, and IT proficiencies. Personnel engaged in BDA projects should demonstrate the competencies to handle emerging technologies, including natural language processing, text mining, video/voice/image analytics, and visual analytics (Schultz, 2013). Investment in skilled big data scientists and engineers has become the most essential premise in using BDA (Müller et al., 2018). Park & Kim (2019) proposed that acquiring the necessary big data management and analytic competency can be accomplished through training and by leveraging the expertise of external experts. Verma & Chaurasia (2019) concurred that employees or data scientists should employ advanced data science practices to gain a comprehensive understanding of the business domain, thereby ensuring compliance with BDA requirements and delivering actionable business outcomes. According to Debortoli et al. (2014), there is a significant different between traditional business intelligent skill and big data analytics skills. Although both of them focus on supporting decision making, BDA jobs require more quantitative and machine learning skills (Debortoli et al., 2014; Phan & Tran, 2022). Similarly, the possession of BDA skills emerges as a key factor motivating enterprises to adopt and implement BDA solutions (Maroufkhani et al, 2020; Verma & Chaurasia, 2019; Wamba et al., 2017). Therefore, it is highly likely that BDA skills have a positive impact on the usage of BDA in the banking sector.

H3: BDA skills will have a positive impact on the usage of BDA.

4.3.4 Big data quality

When banks encounter the features of big data, they typically employ BDA to optimize the benefits brought by the data (Ghasemaghaei, 2018). In today's society, the internet and mobile phones enable banks to engage with customers more frequently and gather a greater amount of data. This data amplifies in terms of volume, velocity, and variety, presenting itself in structured, semi-structured, or unstructured formats. Furthermore, advancements in technology and partnerships with third-party entities will empower banks to gather a wider range of data from various sources, encompassing different types of data. Research conducted by Lai et al. (2018) indicated that as data grows in volume, is generated at a faster pace, and is stored in a more diverse manner, an increasing number of enterprises tend to adopt big data analytics. Through the utilization of the 3V's characteristics of BD (volume, velocity, and variety) as a capability within a company, Ghasemaghaei (2018) discovered a noteworthy positive influence of big data utilization on firm performance. However, enterprises are primarily concerned with quality of data and the integration of BD into the usage of BDA, as noted by Park & Kim (2019). Data quality, as defined by Kwon et al. (2014), pertains to the coherence and soundness of the collected data. This implies that banks that possess higher data quality will augment their utilization of BDA. In the context of this study, BD quality refers to characteristics of big data, specifically denoting the size of data in terms of volume, velocity, and variety. Therefore, this research hypothesizes that:

H4: Big data quality will have a positive impact on the usage of BDA.

4.4 Hypotheses on the impact of big data analytics usage on banking performance

Big data analytics benefits banks beyond the sole function. It covers many aspects of banking operations, from analyzing customers, bringing products and services to them, improving existing operations, and creating new risk management methods (Nobanee et al., 2021).

4.4.1 BDA usage and market performance

Study from Y. Wang et al. (2021) confirmed that using financial technologies could help commercial banks improve their business model, service efficiency, attract more customers, enhance risk management. Ghasemaghaei's (2018) research revealed the advantages of BDA in enhancing the quality of products and services and enhancing the overall customer experience. Many prominent banks are utilizing BD to gain insights into customers' awareness, perceptions, and satisfaction (Schultz, 2013). For instance, financial institutions such as banks possess the capability to analyze unstructured or semi-structured data, enabling them to discern customer needs or concerns through various means such as website clicks or voice recordings obtained from call centers (Schultz, 2013). The Bank of America, in particular, utilizes BD with the primary objective of enhancing the quality of customer information (Schultz, 2013). By gaining a deeper understanding of their customers, banks can expedite the exploration of new markets, accelerate the introduction of novel products or services, achieve a higher success rate in launching these offerings, and attain a larger market share relative to their competitors. Leveraging substantial volumes of real-time data, along with diverse data types, empowers organizations to deliver superior products/services while augmenting operational efficiency beyond that of their rivals (Ali et al., 2020; Ghasemaghaei, 2018). Confirming the positive impact of BDA on organizational market performance a cross-sectional study conducted in India by S. Gupta et al. (2019) supports this assertion. Therefore, it was hypothesized that:

H5: BDA usage will have a positive impact on market performance.

4.4.2 BDA usage and risk management performance

Risk management constitutes a pivotal functional distinction between the banking sector and non-financial enterprises (Aebi et al., 2012). Consequently, advancements in technology are poised to exert a significant influence on the risk management domain within banking operations. Using new financial technologies with advanced algorithms and big data availability allows traditional banks improve for screening borrowers, predicting customer default (Y. Wang et al., 2021). Specifically, the proliferation of applications using BDA technologies is anticipated to confer advantages upon risk managers in the banking sector, enabling them to make more informed decisions while reducing associated costs (Härle et al., 2016). For example, the utilization of BDA facilitates the analysis of customers' information, thereby aiding banks in making precise determinations concerning service provisions such as retail lending and the detection of financial crimes. Melnychenko et al. (2020) and Y. Wang et al. (2021) identified that big data technologies allow banks to reduce operational risks, improve fraud management, and risk assessment.

H6: BDA usage will have a positive impact on risk management performance.

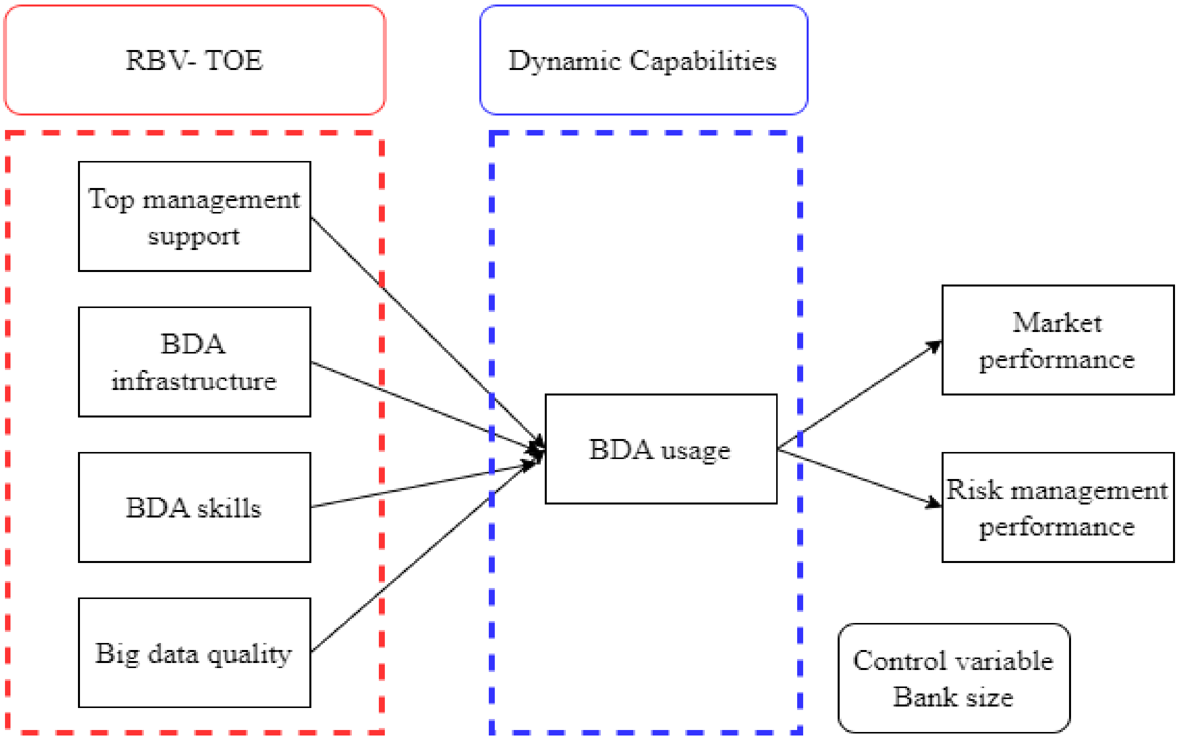


Figure 21: The conceptual model of this thesis
Source: Own edition

The proposed model, presented in Figure 21, illustrates the direct and indirect relationships among the aforementioned variables. The model displays the direct relationship between BDA infrastructure, BDA skills, big data quality, top management support and BDA usage, as well as the direct relationship between BDA usage and market performance and risk management performance. Furthermore, this model also establishes the mediating role of BDA usage.

4.5 The mediating role of big data analytics usage

Establishing hypotheses is crucial when examining mediating effects. In this stage, the study adheres to the transmittal approach recommended by Rungtusanatham et al. (2014) in developing 8 mediating hypotheses that link 4 characteristics of BDA to two bank performance variables through the mediation of BDA usage. Meanwhile to assess the relationship between these characteristics of BDA and banking performance, 8 direct hypotheses were also developed. As suggested by Hair et al. (2019), the outcomes derived from both the direct and mediating effects can aid in discerning the various types of mediating relationships existing between variables.

4.5.1 On the relationship between top management support and bank performance

Understanding the complexity of using BDA, top manager would have more support on preparing the compatible resources for banks (Verma & Chaurasia, 2019). This capacity to discover and acquire resources is regarded as a type of managerial resources (Curado & Bontis, 2006). According to the resource-based view, previous study supported the direct impact of top management support on firm performance (Curado & Bontis, 2006; Sheikh et al., 2018). This relationship emphasizes the responsibility of top managers in managing both external and internal factors that contribute to the effective operations of an enterprise. In the banking sector, the endorsement of utilizing BDA by top-level managers can be regarded as a demonstration of the managerial prowess of banks. This, in turn, leads to the creation of competitive advantages for banks in the market and enhances the efficiency of their risk management practices. Therefore, this thesis hypothesized that:

H1a: Top management support will have positive impact on the market performance.

H1b: Top management support will have positive impact on the risk management performance.

Meanwhile, managerial skills play a significant role in creating BDA capability at banks (Ali et al., 2021). Top management supports is also recommended as a factor to be included in studies on BDA at the firm level (Fosso Wamba et al., 2017). Sheikh et al. (2018) found that in case of the textile industry, top management support has a stronger impact on the firm performance through using electronic marketing. In the context of big data technology, managers are responsible for finding and developing BDA capabilities, stimulating investment in BDA infrastructure, and persisting in constructing data resources for banks. This factor supports the development of innovative dynamic capabilities within enterprises when using BDA, then influencing the competitive advantages of firms in terms of improving market performance and risk management performance. Therefore, this thesis hypothesized that:

H7: Top management support will have indirect impact on the market performance through mediating by the usage of BDA.

H8: Top management support will have indirect impact on the risk management performance through mediating by the usage of BDA.

4.5.2 On the relationship between BDA infrastructure and bank performance

Based on the resource-based view theory, various previous study confirmed that IT infrastructure have direct impact on firm performance (Byrd et al., 2008). Higher level of BDA infrastructure allows banks more capacity to analysis, understand the complex patterns hindering inside of big data. Therefore, BDA infrastructure would have a direct impact on bank performance. Therefore, this thesis hypothesized that:

H2a: BDA infrastructure will have positive impact on the market performance.

H2b: BDA infrastructure will have positive impact on the risk management performance.

In addition, there are many cases when enterprises move to build complex BDA infrastructure, the situation creates an island of information to managers. There is too much data needed to interpret to change into business outcomes. Powell & Dent-Micallef (1997) found that solely relying on information technology does not guarantee an increase in the competitive advantage of firms. Nevertheless, evidence from high-performing firms confirmed that information technology can influence firm performance by mediating through aspects such as business culture or supplier relationships (Powell & Dent-Micallef, 1997). Zhan & Tan (2020) suggested that BDA infrastructure serves better through fostering the development of new services/product/process, then it would be more practical to gain benefits from big data. Therefore, this thesis hypothesized that:

H9: BDA infrastructure will have indirect impact on the market performance through mediating by the usage of BDA.

H10: BDA infrastructure will have indirect impact on the risk management performance through mediating by the usage of BDA.

4.5.3 On the relationship between BDA skills and bank performance

From the aspect of resource based view, BDA skills is considered as a human capital resource which play as the strategic resource to improve firm performance (R.M. Grant, 2010). Particularly, within the context of digitalization, human capital exhibits a noteworthy positive correlation with both economic and ICT development in the regions of the EU-27 (Tran et al., 2023). Mikalef et al. (2019) discovered that in banking and financial sector, data and technical skills in BDA have become a crucial factor in distinguishing firm performance. Ghasemaghaei (2018) found the direct impact of analytical skills of employees on firm performance. It is likely the fact that in context of big data analytics, knowledge and experiences of BDA employees allow enterprises to transform data into business insight and improve the decision making and competitive advantages. Moreover, the growing utilization of big data analytics in risk management within banking necessitates proficiency in programming, statistical techniques, mathematics, and technological advancements (Dicuonzo et al., 2019). This combination of expertise is essential for effectively converting vast and diverse datasets into valuable decision-making insights (Dicuonzo et al., 2019). Therefore, this thesis hypothesized that:

H3a: BDA skills will have positive impact on the market performance.

H3b: BDA skills will have positive impact on the risk management performance.

Furthermore, this study investigates further the impact of BDA skills to firm performance including market performance and risk management performance at banks. The digital skills within a firm prompts the development of new dynamic capabilities, instigating consequential adjustments in the business model, thereby fostering enhancements in overall firm performance (Vial, 2019). The result of acquiring BDA skills should target innovative services and products, then the usage of BDA would help banks much faster/better/more efficient in delivering their services than competitors.

H11: BDA skills will have indirect impact on the market performance through mediating by the usage of BDA.

H12: BDA skills will have indirect impact on the risk management performance through mediating by the usage of BDA.

4.5.4 On the relationship between big data quality and bank performance

In the banking industry, data is collected from various sources such as ATM, mobile banking applications, or internet banking transactions. This information gives a comprehensive view of customer behavior and helps the bank identify suspicious transactions, predict operational losses and make better relevant risk measurements/ decisions (Härle et al., 2016). By utilizing big data analytics, banks gain an advantage by being able to create real-time visualizations, enhancing the diversity and functionality of online and mobile service products (Davenport et al., 2012). Therefore, this thesis hypothesized that:

H4a: Big data quality will have positive impact on the market performance.

H4b: Big data quality will have positive impact on the risk management performance.

Data quality will improve the process of launching new banking services as well as play as a fundamental role in calculating and measuring risks (Härle et al., 2016). Then these innovative services/products, running in real-time and offering better visualization with more information, will help banks create distinct and superior products compared to their competitive counterparts. The findings from Ghasemaghaei & Calic (2020) supported that firms using big data characteristics to generate new ideas can improve firm performance, therefore, this study highlighted the mediating role of innovative BDA capabilities on the relationship of big data quality and firm performance. Therefore, this thesis hypothesized that:

H13: Big data quality will have indirect impact on the market performance through mediating by the usage of BDA.

H14: Big data quality will have indirect impact on the risk management performance through mediating by the usage of BDA.

CHAPTER 5: RESEARCH METHODOLOGY

5.1 Introduction

This chapter presents the data material used for the quantitative analysis of this thesis, as well as the rationale for selecting the Partial Least Squares - Structural Equation Modeling (PLS-SEM) methodology. Following the data collection part, the main part of the PLS-SEM method will be discussed, including the assessment of the measurement and structural models.

5.2 Research philosophy

In research methodology, the two most mentioned schools are the philosophies of positivism and post-positivism. The methodological distinctions between quantitative and qualitative research are based on the philosophical traditions of positivism and post-positivism, respectively (Polit & Beck, 2006). Some differences between these research philosophies should be mentioned. Positivism believes that there is an objective reality that is not affected by human behavior. The collected data that is observable, distinguishable, and quantifiable should be used to gain true knowledge (Crossan, 2003). Therefore, positivism supports the idea that mathematics and formal logic can be used to make analytical statements about the world we observe, and that induction can be used to create generalizations and laws (Crossan, 2003). Meanwhile, the post-positivist approach attempts to better understand phenomena through qualitative analysis (Crossan, 2003). In this thesis, the author follows positivism, which is the most popular philosophy in Information systems (Alavi & Carlson, 1992). Moreover, using quantitative method allows researchers to understand the phenomenon without researcher bias, and this method improves the generalizability and reproducibility of conclusions.

5.3 Survey research methodology

The survey research approach is the dominant research method in the study using the enterprise-level BDA. This approach refers to the whole process from collecting data from many enterprises to analyzing data using statistical tools (Gable, 1994). In addition, the survey research method allows researchers to test the research hypothesis and produce generalizable results. This study utilized the survey research method for several reasons. Firstly, in order to answer the research question, this study based upon the literature review and big data context to build and propose research model. The research model presents the relationship among constructs (dependent and independent variables). Then, this study used the PLS-SEM technique to test research hypotheses; therefore, using the survey research method is appropriate (Creswell, 2009). Secondly, collecting data from target population is impossible for most researchers due to the time-consuming, expensive, and unrealistic nature of the facts. Therefore, using survey methods is the most appropriate research method.

5.4 Survey design

According to suggestion from Creswell (2009), in survey research method, there are three stages of this method that should be followed carefully, including: survey designing, collecting data, and analyzing data. This research utilized primary data from a survey questionnaire to evaluate the relationship between 4 independent variables and 3 dependent variables. Following the literature review, a survey questionnaire was formulated and distributed to firm-level respondents.

The survey consists of four sections. The first section is the research introduction, which includes the purpose and author information of the research. The second section is the demographic section, which includes four questions about respondents' age, types of banks, functional department, and total assets of their working banks. The main sections of the survey questionnaire (sections 3 and 4) used a 5-Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree) (see at Appendix A). The third section is the practice of using big data analytics section, which asks respondents to assess the preparation and level of using big data analytics at their banks, with 22 questions.

Table 3: Constructs, items and supporting theories in the research model

Constructs	Items	Supporting theories	Supporting references
Top management support	TOP1→4	RBV-TOE	D.Q. Chen et al (2015), Lai et al. (2018)
BDA infrastructure	INF1→5	RBV-TOE	M. Gupta & George (2016)
BDA skills	SKILL1→4	RBV-TOE	M. Gupta & George (2016)
Big data quality	DQ1→4	RBV-TOE	M. Gupta & George (2016)
BDA usages	USAGE1→5	DC	D.Q. Chen et al (2015), S. Gupta et al. (2019)
Market performance	MP1→4	DC	Ji-fan Ren et al. (2017); Raguseo & Vitari (2018)
Risk management performance	RP1→4	DC	Dicuonzo et al. (2019)

Source: Own edition

Five questions (5 items) adapted from D. Q. Chen et al. (2015) and own development (the 5th item) to assess the BDA usage. Five questions (5 items), adapted from M. Gupta & George (2016), to assess BDA infrastructure. Four question (4 items), adapted from (M. Gupta & George, 2016), to assess BDA skills. Four question (4 items), adapted from D. Q. Chen et al. (2015); Lai et al. (2018), to measure the top management support construct. Four questions (4 items), adapted from M. Gupta & George (2016) and own development, to assess big data quality construct. The final section requires respondents to assess their bank performance in terms of risk management and market performance with eight questions. In which, four questions (4 items), adapted from (Ji-fan Ren et al., 2017; Raguseo & Vitari, 2018), used to assess market performance, and four questions (4 items), adapted from Dicuonzo et al. (2019) to assess risk management performance. The usage of BDA was evaluated using a five-Likert scale, with a score of 1 indicating that the organization never uses BDA and a score of 5 indicating frequent use. Other questions from section 3 and 4 used the 5-Likert scale ranged from 1 (strongly disagree) to 5 (strongly agree). This scale is commonly employed to explore topics that have not been extensively studied. Finally, this study considers bank size as an important control variable. The research incorporates a control for bank size, as larger banks are more likely to exhibit a positive correlation with superior performance in utilizing BDA compared to smaller banks. Bank size is measured by the total assets of banks, with a score of 1, 2, or 3 indicating a small, medium, or large bank, respectively.

5.5 Selecting sample size

This study aims to investigate the factors influencing BDA usage and its impact on banking performance in Europe 27 countries. Therefore, in this study, population is the number of credit institutions in EU-27. According to the latest report from European Banking Federation (2022), there was 5263 credit institutions operating in EU-27.

In order to confirm the number of samples for this study, Sekaran & Bougie (2016) recommend a sample size of greater than 30 and less than 500 for PLS method. Additionally, G* Power software (Faul et al., 2007) was utilized to calculate the minimum sample size. With a level of confidence at 95%, a power estimate of 0.80 (Gefen et al., 2011), an effect size of f2= 0.15 (medium), and six predictors, the minimum sample size to test this model was determined to be 95. In addition, Reinartz et al. (2009) suggested that Partial Least

Squares (PLS) is the more suitable approach when there are 100 responses to attain adequate statistical power, given the quality of the measurement model.

5.6 Data collection

After building a survey for the research, the researcher intended to send it to banks by mail. However, printing the survey questionnaires and sending them by mail directly to banks has many limitations, such as being time-consuming and not environmentally friendly. Therefore, the second option - online distribution - was selected. By creating on the google form platform, surveys could be made conveniently and delivered quickly to target respondents.

A pilot survey was sent to four experts, including two professors and two bank experts, to receive their feedback on the structural appropriateness and the clarity of the language in the survey. To assess the reliability and validity of the survey questionnaire, a pilot test was conducted with 15 experts who had experience in digital transformation and big data projects in the banking sector. This process happened from 1st September 2022 to 20 September 2020. After receiving feedback from these experts, the language of the survey was revised, and the validity and reliability of the survey were verified.

Subsequently, a final survey questionnaire was generated using Google form and send to the qualified respondents via a professional network and emails between 1 October 2022 and 31 December 2022. The target respondents of this study were individuals with experience in digital transformation and big data projects at their respective banks, most of whom held mid-level managerial positions, such as department heads or team leaders. A total of 1055 potential respondents working at 538 credit institutions in EU-27 member states were contacted, out of which 128 responses were received, yielding a response rate of 12.1%. Due to missing values in performance assessment, 114 bank-level responses from 105 banks in the European Union were eventually used for analysis. Table 4 presents the data of the respondents, where two-thirds of them belong to the age range of 30-49 years. The majority (91.23%) of the respondents are employed in commercial banks, and 63% work in data management and information technology departments. In addition, 32.46%, 25.44%, and 42.10% of the respondents work in banks with total assets of over 100 billion EUR, between 50-100 billion EUR, and less than 50 billion EUR, respectively.

Table 4: Respondent's demographic information

Characteristic	Criteria	Frequency (*n* = 114)	Percentage (%)
Age	20-29 years old	25	21.9%
	30-49 years old	78	68.42%
	50 years old or older	11	9.68%
Bank types	Commercial bank	104	91.23%
	Cooperative bank	5	4.39%
	Others (Saving banks, …)	5	4.39%
Department	General management	7	6.14%
	Information Technology	32	28%
	Data management, Business Intelligent	40	35%
	Risk management/Controlling/ AML/ Compliance/	16	14.04%
	Marketing/ Customer services	10	8.77%
	Others:	9	8.05%
Total Assets	More than EUR 100 billion	37	32.46%
	Between EUR 50 - 100 billion	29	25.44%
	Less than EUR 50 billion	48	42.10%

Source: Own edition

5.7 Data analysis

Table 5 shows the descriptive statistic result of data used in the data analysis.

Table 5: The description of statistical data

	No.	Missing	Mean	Median	Min	Max	Std.Dev.
INF1	1	0	3.210526	3	1	5	1.251362
INF2	2	0	3.763158	4	1	5	1.214242
INF3	3	0	3.482456	4	1	5	1.199059
INF4	4	0	3.210526	3	1	5	1.208185
INF5	5	0	3.298246	3	1	5	1.218804
TOP1	6	0	3.464912	3.5	1	5	1.249282
TOP2	7	0	3.368421	3.5	1	5	1.184436
TOP3	8	0	3.245614	3	1	5	1.244893
TOP4	9	0	3.131579	3	1	5	1.272675
SKILL1	10	0	3.04386	3	1	5	1.31963
SKILL2	11	0	3.342105	3	1	5	1.246794
SKILL3	12	0	3.54386	4	1	5	1.145521
SKILL4	13	0	3.421053	4	1	5	1.247261
DQ1	14	0	3.210526	3	1	5	1.178522
DQ2	15	0	3.789474	4	1	5	1.092806
DQ3	16	0	3.192982	3	1	5	1.302667
DQ4	17	0	3.008772	3	1	5	1.392112
USAGE1	18	0	3.22807	3	1	5	1.303739

USAGE2	19	0	3.377193	4	1	5	1.27876
USAGE3	20	0	3.350877	3.5	1	5	1.247883
USAGE4	21	0	3.210526	4	1	5	1.346734
USAGE5	22	0	3.72807	4	1	5	1.221381
MP1	23	0	2.921053	3	1	5	1.220364
MP2	24	0	3.149123	3	1	5	1.277667
MP3	25	0	3.219298	3	1	5	1.202937
MP4	26	0	3.385965	3	1	5	1.193706
Size	27	0	1.903509	2	1	3	0.86187
RP1	28	0	3.570175	4	1	5	0.872077
RP2	29	0	3.280702	3	1	5	1.026281
RP3	30	0	3.578947	4	1	5	0.93973
RP4	31	0	3.482456	3	1	5	1.015216

Source: Own edition

There was no missing value in data. Constructs namely BDA infrastructure (INF 1:5), top management support (TOP 1:4), BDA skills (SKILL 1:4), big data quality (DQ 1:4), BDA usage (USAGE 1:5), Market performance (MP 1:4), risk management performance (RP 1:4) have maximum and minimum value at 5 and 1 respectively. Meanwhile, Size variable was measured by only one item, maximum value at 3 and minimum value at 1.

CHAPTER 6: RESULT AND DISCUSSION

6.1 Assessing measurement model

In the most recent studies, Structural Equation Modeling (SEM) is the main method used to determine the relationship between variables in the research model of BDA usage topic. SEM method has a number of attractive features: such as allowing assess measurement model and the structural model simultaneously; enabling to assess the direct and indirect effects among variables (Hair et al., 2017).

Studies utilizing SEM have identified two methods as being the most advantageous: Covariance-based Structural Equation Modeling (CB-SEM) and Partial Least Square-Structural Equation Modeling (PLS-SEM) (Hair et al., 2019). Each of the two have own advantages and disadvantages, however, PLS-SEM seems to be outstanding in many cases (Hair et al., 2019).

Table 6: The comparison between CB-SEM and PLS-SEM

Criteria	CB-SEM	PLS-SEM
Purpose	For confirmation of establish theory	For both exploratory and confirmation research.
The statistical objectives	Estimate model parameters that minimize the differences between the observed sample covariance matrix and the estimated covariance matrix.	To maximize the variance explained in the dependent variables.
Measurement philosophy	Based on common factor model	Based on the composite model
Statistical method	A parametric statistical method	Non-parametric method (relay on bootstrapping to derive standard error estimates
Sample size	Only for larger sample size (N>= 100)	For both larger and smaller sample size (N<= 100)
Data	Normally distributed data	Normally distributed data & non-normally distributed data
Structural relationship	Structural relationships are examined by the size and significance of the beta coefficients.	
	Goodness of fit (GoF) is the appropriate measure to evaluate the measurement and structural model. (Based on Chi-square statistic, CFI, GFI and RMSEA)	PLS-SEM has predictive capabilities (based on R^2 value). However, there is no established GoF measure for PLS-SEM

Source: Own edition

This research employed partial least square–structural equation modeling (PLS-SEM) to examine the relationship between four factors (namely BDA infrastructure, BDA skills, top management support, and big data quality) and the dynamic capabilities of using BDA. PLS-SEM is based on maximum likelihood, which has recently been deemed a favored research method in behavior social research (Fan et al., 2016). It is important to note the advantages of PLS-SEM (see more at Table 6). Firstly, it has been demonstrated to be an effective tool for exploratory research on novel relationships that have not been previously studied (Fan et al., 2016). Secondly, PLS-SEM is recommended for studies with limited sample sizes (Hair et al., 2013; Wetzels et al., 2009). This study utilizes the SEMinR package on the R environment to implement PLS-SEM (Ray et al., 2021).

In order to evaluate PLS-SEM results, the measurement model was evaluated by assessing the relevant different criteria (Hair et al., 2019). These criteria suggested are examining the indicator reliability, internal consistency reliability, composite reliability and construct reliability (Hair et al., 2019). Indicator reliability refer to the communality of an indicator (Hair et al., 2019). A benchmark that could be used to check indicator reliability are indicator loadings and indicator loading square. While indicator loadings are suggested to be higher than recommended value of 0.708, indicator loading square are recommended to be higher than 0.5 (Hair et al., 2019). The statistical result showed that all indicator loadings (ranging between 0.790 to 0.927) were higher than 0.708, except DQ2 with loading factor of 0.625. Hair et al. (2019) and Hair et al. (2022) suggested that indicators with loadings lower than 0.708 should be removed before further steps. Table 7 indicated that all indicator loadings ranged from 0.758 to 0.927, well above the recommended value of 0.708, and the minimum value of indicator loading square was at 0.574. It means that all indicators were reliable (Hair et al., 2019).

Next step of assessing the measurement model is evaluating the internal consistency reliability by checking the composite reliability (CR). This criterion refers to the degree that indicators within the same construct are associated with the others (Hair et al., 2022). From table 7, all composite reliability values were higher than 0.856 which means all constructs are satisfied with a high level of reliability. More conservatively, the Cronbach alpha and rhoA were also checked to establish the internal consistency reliability (Hair et al., 2022). All Cronbach's alpha values were higher than the recommended value of 0.7 and rhoA values ranged between Cronbach's alpha and CR values. In order to evaluate the convergent validity of each construct, the average variance extracted (AVE) was used to examine.

According to (Hair et al., 2019), AVE refers to the grand mean value of each indicator' squared loadings within the construct. It is suggested that the AVE of each construct should be higher than the minimum acceptable of 0.5 (Hair et al., 2022). From Table 7, all AVE values were satisfied with ranging from 0.663 to 0.804.

Table 7: The measurement model result

Variables	Indicator	Loading factor	Factor Loading Square	Cronbach's alpha	Composite reliability (CR)	Average variance extracted (AVE)	rhoA
BDA infrastructure	INF1	0.861	0.742	0.92	0.94	0.759	0.931
	INF2	0.811	0.658				
	INF3	0.873	0.762				
	INF4	0.911	0.83				
	INF5	0.896	0.803				
BDA skills	SKILL1	0.837	0.701	0.909	0.937	0.787	0.914
	SKILL2	0.889	0.79				
	SKILL3	0.927	0.859				
	SKILL4	0.894	0.8				
Big data quality	DQ1	0.813	0.662	0.749	0.856	0.663	0.768
	DQ3	0.81	0.656				
	DQ4	0.822	0.676				
Top management support	TOP1	0.902	0.814	0.919	0.942	0.803	0.929
	TOP2	0.882	0.779				
	TOP3	0.912	0.832				
	TOP4	0.889	0.79				
USAGE	USAGE1	0.868	0.753	0.904	0.929	0.723	0.906
	USAGE2	0.851	0.725				
	USAGE3	0.84	0.706				
	USAGE4	0.854	0.729				
	USAGE5	0.837	0.701				
Risk management	RP1	0.895	0.801	0.901	0.931	0.771	0.915
	RP2	0.872	0.761				
	RP3	0.84	0.706				
	RP4	0.903	0.815				
Market performance	MP1	0.758	0.574	0.894	0.927	0.763	0.898
	MP2	0.91	0.828				
	MP3	0.915	0.836				
	MP4	0.901	0.812				
Size	Size	1	1	1	1	1	1

Source: Own edition

The last step to assess the measurement model is examining the discriminant validity. The research used both two criteria: The Fornell and Larcker criterion (Fornell & Larcker,

1981) and the heterotrait-monotrait ratio (HTMT) (Henseler et al., 2015) to assess discriminant validity. According to Fornell & Larcker (1981), AVE of each construct should be higher than the squared correlation between constructs. Table 8 demonstrates that all square root of AVE values on the diagonal (*Italic*) were higher than construct inter-correlation.

Table 8: The Fornell and Larcker criterion result

	INF	SKILL	TOP	DQ	USAGE	Size	MP	RP
BDA infrastructure (INF)	*0.871*							
BDA skills (SKILL)	0.797	*0.887*						
Top management support (TOP)	0.686	0.732	*0.896*					
Big data quality (DQ)	0.587	0.554	0.537	*0.814*				
BDA usage (USAGE)	0.612	0.632	0.503	0.466	*0.850*			
Bank size (Size)	0.087	0.172	0.039	0.050	0.220	*1.000*		
Market performance (MP)	0.486	0.528	0.460	0.358	0.543	0.189	*0.873*	
Risk management performance (RP)	0.496	0.427	0.349	0.384	0.471	0.150	0.631	*0.878*

Note: The square root of AVE on the diagonal (*Italic*) and construct correlations on the lower triangle.

Source: Own edition

To be more conservative, Hair et al. (2022) suggested that the HTMT values should be used to check the discriminant validity. Henseler et al. (2015) recommended that the HTMT value should be lower than the threshold value of 0.9 which confirms discriminant validity is not present. From Table 9, the HTMT values were lower than the threshold of 0.85. Only one HTMT value of the pair of two constructs: BDA infrastructure and BDA skills was 0.868, but this value was close to 0.85 and far below 0.9, therefore, this study accepted this value.

Table 9: The heterotrait-monotrait ratio result

	INF	SKILL	TOP	DQ	USAGE	Size	MP	RP
INF								
SKILL	0.868							
TOP	0.736	0.798						
DQ	0.697	0.664	0.639					
USAGE	0.663	0.694	0.543	0.552				
Size	0.086	0.180	0.124	0.091	0.232			
MP	0.532	0.591	0.500	0.426	0.599	0.200		
RP	0.545	0.470	0.379	0.428	0.512	0.155	0.694	

Source: Own edition

Conservatively, this study tested the significance of all HTMT values to confirm whether these values are significantly different 1 or lower than the cut-off value 0.9 (Sarstedt et al., 2022). The results of this booth strapped procedure confirmed the discriminant validity of all studied constructs.

6.2 Assessing Structural model

6.2.1 Evaluating the collinearity issues

The first step in assessing the structural model was evaluating if collinearity issues exist. Hair et al. (2022) recommended that the variance inflation factor (VIF) should be calculated and then compared this value with the suggested threshold of five. Collinearity could not occur if the VIF values are lower than 3 (Hair et al, 2022). Checking results from Table 10 suggested that all VIF values were acceptable and lower than the recommended value of 3.

Table 10: The VIF values of structural model

	VIF
USAGE	
BDA infrastructure	3.110
BDA Skills	3.361
Top management support	2.369
Data quality	1.622
Market performance	
USAGE	1.051
Size	1.051
Risk management performance	
UAGE	1.051
Size	1.051

Source: Own edition

Only BDA infrastructure's VIF value (3.110) and BDA skills' VIF value (3.361) were suggested to occur the collinearity issues, however, these numbers were close to three. Then, it was confirmed that the collinearity problem was not likely a critical issue in the structural model.

6.2.2 Evaluating the path relationships

The next step to assess the structural model is evaluating the significance and relevance of structural model. Results from statistical analysis (Table 11) after the bootstrapping at

10000 samples (Hair et al., 2022; Streukens & Leroi-Werelds, 2016) provided the path coefficient estimates, t-value and level of confidence intervals (CI).

Table 11: Results of structural model

Hypotheses	**Main Model**	**Path coefficients**	**t-value**	**2.5% CI**	**97.5% CI**	**Result**
H1	TOP → USAGE	-0.001	-0.009	-0.218	0.220	Rejected
H2	INF → USAGE	0.251	1.738	-0.048	0.518	Rejected
H3	**SKILL → USAGE**	**0.369****	**2.928**	**0.123**	**0.620**	**Accepted**
H4	DQ → USAGE	0.116	1.282	-0.056	0.302	Rejected
H5	**USAGE → RP**	**0.460*****	**5.800**	**0.297**	**0.606**	**Accepted**
H6	**USAGE → MP**	**0.527*****	**7.265**	**0.379**	**0.665**	**Accepted**
	Control variable	Path coefficients	t-value	2.5% CI	97.5% CI	
	Bank size → MP	0.073	0.965	-0.075	0.218	
	Bank size → RP	0.049	0.541	-0.124	0.226	

Note: *** represents significant at 0.001 level, ** represents significant at 0.01, * represents significant at 0.05 level.

Source: Own edition

The significance of each path could be assessed by employing either the t-values of path coefficients or the confidence intervals as suggested by Sarstedt et al. (2022). If a value of zero falls within the range of 2.5% to 97.5% in the confidence interval, it demonstrates the presence of a significant path coefficient (Sarstedt et al., 2022). The initial analysis examined the effect of top management support, BDA infrastructure, BDA skills, and big data quality on the BDA usage. The results, as shown in Table 11, demonstrated the results of bootstrapping process with path coefficient, t-value and confidence intervals. The findings showed that the impact of top management support was insignificant with a path coefficient of -0.001. The impact of BDA infrastructure on the BDA usage was insignificant with a path coefficient of 0.251. The result from table 11 did support the relationship between BDA skills and the BDA with path coefficient of 0.369. Meanwhile, the impact of data quality on the BDA usage was insignificant with path coefficient of 0.116. Therefore, the study confirmed hypothesis H3, while H1, H2, H4 were rejected. This implies that, among the four factors under investigation, only BDA skills exert a statistically significant influence in the BDA usage.

The second set of analyses investigated the impact of BDA usage on banking performance. As shown in table 11, BDA usage had a significant positive impact on both market performance and risk performance, thus H5 & H6 were accepted.

Furthermore, evidence from the statistical analysis revealed the impact of the control variable and showed that bank size had insignificant impacts on market performance and risk management performance.

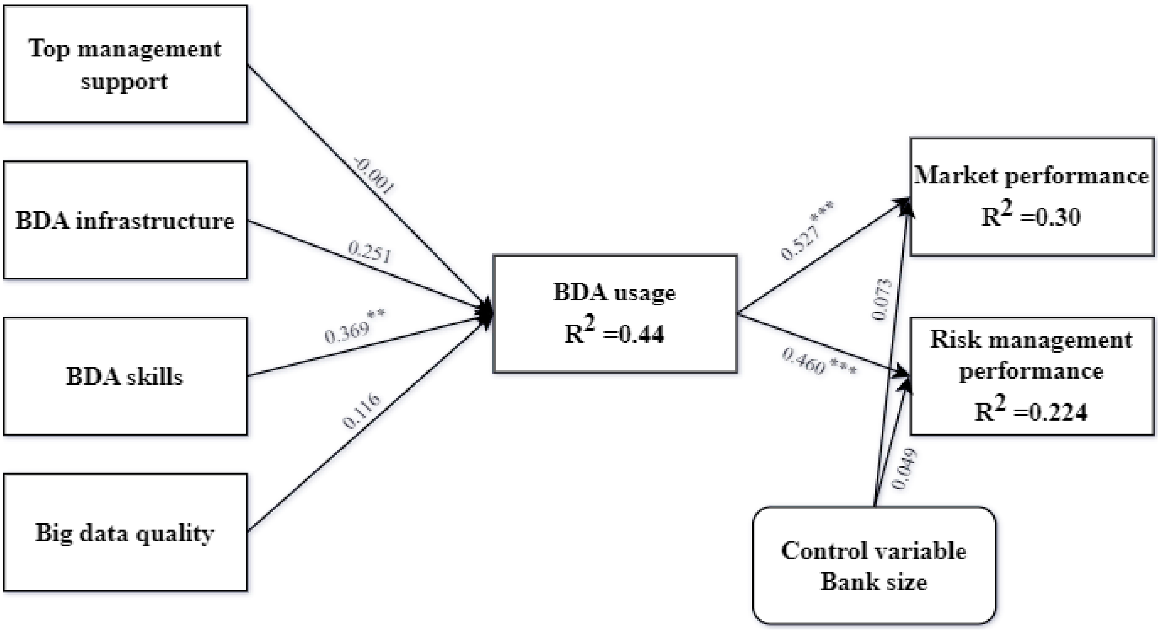


Figure 22: Result of path analysis
Source: Own edition

6.2.3 Assess the model's explanatory power and the goodness of fit

Although some researchers used some model fit measures in order to assess the model's quality. Hair et al. (2021) and Henseler & Sarstedt (2013) recommend that there have no established fit measures suggested for PLS-SEM. In order to assess the model' quality, researchers should rely on alternative measures such as assessing R^2 values / the model's predictive capabilities (Chin et al., 2012; Hair et al, 2021).

In this study, the R^2 values were used to examine the explanatory power of the studied model. According to Hair et al. (2019), the research model in social sciences possesses substantial, moderate, and weak explanatory power when R^2 values are at 75%, 50%, and 25%, respectively. From the statistical result, BDA infrastructure, Top management support, BDA skills, and big data quality explain 44% of the variance in BDA usage. These R^2 values can be considered as moderate. Furthermore, BDA usage and bank size explain 30% of the variance in market performance and 22.4 % of the variance in risk management performance. Therefore, these R^2 values showed that the research model had a moderate explanatory power.

As SEMinR package does not provide the GoF measures, this study followed the suggestion from Tenenhaus et al. (2005) to use the global criterion GoF which was calculated by the geometric mean of average commonality and the average R^2 (as demonstrated by the formula below). The Goodness of Fit (GoF) of 0.501 for the Partial Least Squares model was satisfactory (higher than the recommended value of 0.36 from Tenenhaus et al. (2005)), and this PLS model was able to explain 50.1% of the achievable fit.

GoF= $\sqrt{(average\ AVE * average\ \text{R2})} = \sqrt{(0.784 * 0.32)} = 0.501$

6.2.4 Evaluating the mediating effects

In order to investigate whether the relationship between top management support, BDA skills, BDA infrastructure, and big data quality on banking performance is mediated by BDA usage, this study examined mediation analysis, followed by suggestions from Hair et al. (2022) and Zhao et al. (2010). Table 12 provided summary statistics for the mediating relationship.

Table 12: Results of the mediating effect

Hypotheses	**Mediating effect (Indirect and direct relationship)**	**Path coefficients**	**t-value**	**2.5% CI**	**97.5% CI**	**Result**
H7	TOP → USAGE → MP	-0.000	-0.009	-0.115	0.122	No effect
H1a	TOP → MP	0.127	0.999	-0.126	0.374	
H8	TOP → USAGE → RP	-0.000	-0.009	-0.104	0.104	No effect
H1b	TOP →RP	-0.042	-0.343	-0.288	0.193	
H9	INF → USAGE → MP	0.132	1.738	-0.027	0.276	No effect
H2a	INF → MP	0.047	0.330	-0.227	0.337	
H10	INF → USAGE → RP	0.115	1.666	-0.022	0.168	Only direct
H2b	INF → RP	0.330	2.490	0.065	0.591	
H11	SKILL → USAGE → MP	0.195	2.531	0.060	0.360	Full mediating
H3a	SKILL → MP	0.186	1.155	-0.114	0.514	
H12	SKILL → USAGE → RP	0.169	2.568	0.053	0.313	Full mediating
H3b	SKILL → RP	-0.030	-0.201	-0.308	0.279	
H13	DQ → USAGE → MP	0.061	1.235	-0.029	0.168	No effect
H4a	DQ → MP	0.035	0.306	-0.193	0.248	
H14	DQ → USAGE → RP	0.053	1.170	-0.024	0.157	No effect
H4b	DQ → RP	0.125	1.137	-0.103	0.327	

Source: Own edition

In a mediating analysis conducted in this study, the first step involved checking the presence of a significant indirect relationship between independent variables and dependent variable (Zhao et al., 2010). Specifically, this study investigated the relationship between BDA infrastructures, BDA skills, top management support, big data quality and market performance & risk management performance, mediated by BDA usage, using the bootstrapping process. In the second step, the direct effects were assessed to determined the specific types of mediating effects, given the existence of indirect effects (Zhao et al., 2010).

Table 12 revealed that there was no significant indirect effect between BDA infrastructure and market performance & risk management performance. However, interestingly, this study discovered a significant direct relationship between BDA infrastructure and risk management performance, with path coefficient of 0.330.

Of particular importance, the study identified a significant mediating effect of BDA usage on the relationship between BDA skills and market performance and on the relationship between BDA skills and risk management performance. Meanwhile, the direct relationship between BDA skills and martket performance and risk management performance variables were found to be insignificant with coefficients of 0.186 and -0.03 respectively. According to the suggestion from Zhao et al. (2010), this situation was classified as full mediation. This means that the relationship between BDA skills and both market performance and risk management performance were completely mediated by BDA usage.

However, regarding the last variables of top management support and data quality, the study found two other important findings. First, both the indirect and direct relationship between top management support and banking performance were found to be insignificant. Second, both the indirect and direct relationship between data quality and banking performance were also determined to be insignificant.

6.3 Discussion

This section discussed the results of the quantitative analysis, emphasizing the significant findings and comparing and contrasting them with those of previous studies.

6.3.1 The relationship between top management support and BDA usage

The striking result of this study highlighted that top management support had no significant impact on BDA usage. This result contrasts with previous studies which have found that top management support had a significantly positive effect on the adoption (Lai et al., 2018; Verma & Chaurasia, 2019) and usage (D. Q. Chen et al., 2015) of BDA. Meanwhile, this finding supports evidence from previous observation by Braganza et al. (2017, p. 334), who found that “big data initiatives begin with issues business leaders consider important strategically”. This finding also is supported by previous study of Gurău & Ranchhod (2020). In project management, the role of managers (middle-level manager) play as the most influential factor in the big data adoption, because this position directly involve with the daily changes in using BDA (Bakici et al., 2022). Thus, although top management support found to be insignificant on the impact to BDA usage, this finding did not mean to underestimate the role of top management support, however, it could be explained by the fact that in banking sector, the top management supports often play an important role in kick-starting big data initiatives. They can initiate the benefits of big data by discussing with colleagues or big data experts within the organization. Furthermore, as decisions to adopt BDA in the early stages often involve third parties and require adherence to establishing standard security policies (Verma & Chaurasia, 2019), and adoption of BDA is a big decision so that the support of top management is of great importance at this stage (Baig et al., 2021). After this stage, leveraging BDA as a dynamic capability to develop new services/products, this stage will depend on human skills and capabilities. In addition, managers’ experiences and vision can influence big data usage by improving the BDA infrastructure, recruiting more BDA employees, or providing training to their current staff, all of which can lead to strategic outcomes. It extends the TOE framework in the banking sector, emphasizing the significance of top management support in the IS field as a catalyst and supporter of big data initiatives by providing the necessary BDA resources.

6.3.2 The relationship between BDA infrastructure and BDA usage

The results of this study indicate that BDA infrastructure has an insignificantly positive impact on BDA usage. This result is supported by Lai et al. (2018)’s finding but it is inconsistent with findings from Verma & Chaurasia (2019) which emphasizes the essential preparation of infrastructure for BDA adoption. This result can be attributed to the fact that banking is the most IT-driven sector, and banks in EU-27 regions have achieved a certain degree of infrastructure in using BDA. Moreover, many banks already have used 4V characteristic of big data. For instance, in order to exploit insight from real time data, banks invested in establishing their own big data centers to foster the availability of adopting BDA solutions. Therefore, the findings from this study suggested that BDA infrastructure is no longer a barrier to banks using big data technologies. This results may be different with findings from other sectors (Fosso Wamba et al., 2019).

6.3.3 The relationship between BDA skills and BDA usage

The result from table 11 suggested that BDA skills were the most important factor influencing BDA usage. This finding is consistent with the results of Akter et al. (2019) and Fosso Wamba et al. (2017), who highlighted that personnel expertise was the most significant component of big data analytics capability. Additionally, this finding supports the assertions of Debortoli et al. (2014) that big data projects are highly dependent on human capital. Similarity, Gurău & Ranchhod (2020) highlighted the importance of humans in the implementation of new technologies, which are seen as a means for organizations to achieve their business objectives. The results from this study highlighted the role of human skills in the BDA capabilities of banks. Data sources could be exploited efficiently and effectively in order to create innovative capabilities through highly qualified employees and well-educated staff.

6.3.4 The relationship between big data quality and BDA usage

The statistical data showed that BDA characteristic quality did not have significant influence on the BDA usage. It is interesting to note that solely using big data with 4V's characteristics did not create BDA dynamic capabilities for banks, then had no impact on banking performance. This finding is contrary to previous study which have suggested that big data is firm capability (Ghasemaghaei, 2018). It has been suggested that data asset and big data experts in IT-intensive sectors would allow enterprises achieve better performance (Müller et al., 2018). This interesting result seem to be consistent with previous study (Lai et al., 2018). The observed result may be due to the fact that banks' big data initiatives are still in their infancy, some small banks may not need all the features offered by complicated big data analytics applications and should only select the ones beneficial for their banks and market situation. This finding is also in accord with recent study indicating that there was no significant association between the data characteristics (volume, variety and velocity of data) and firm agility (Ghasemaghaei et al., 2017). Moreover, during the last three years, with the impact of the Covid-19 pandemic, banks may prioritize adopting big data analytics solutions to support smooth operations and home offices and provide service solutions to customers to minimize direct contact.

6.3.5 The relationship between BDA usage and market performance

The result from the statistical analysis showed that using BDA has a significantly positive influence on market performance in the banking sector. This result is consistent with findings from previous studies (Maroufkhani, Wan Ismail, et al., 2020; Raguseo, 2018; Yadegaridehkordi et al., 2020). Raguseo (2018) concluded that one of the main benefits from using big data is allowing companies response and improve their products and service faster than competitors. Moreover, using big data technologies helps companies react faster in the dynamic changing environment (Raguseo, 2018). This study's findings corroborate the conclusions of Manuylenko et al. (2021). This study found that the advantages of BDA are no longer merely potential; the quantitative results demonstrate that the utilization of BDA directly contributes to the bank's operations by creating dynamic and innovative capabilities, allowing the bank to deliver products and services faster, better, and with a higher success rate, thus enabling them to capture a larger market share. This result also accords with the earlier observation, which showed that data analytics usage only have significant positive impact on the firm agility in situation which firms have enough preparation in other resources such as human, technology and task requirements (Ghasemaghaei et al., 2017). When utilizing new technologies, the technology infrastructure and human capital were the most influential factors in the value-creation performance of the banking sector.

6.3.6 The relationship between BDA usage and risk management performance

The quantitative result highlighted that using BDA in creating innovative products/ services has significantly positive influence on risk management performance in the banking sector. According to Milojević & Redzepagic (2021), AI and ML would be likely to contribute more effective on the performance of risk management, meanwhile, big data analytics provide as the foundation which moderate this relationship stronger. This study provided empirical evidence on the direct impact of using big data analytics on the practices of risk management at banks in Europe. Through helping banks identify, measure, and controlling the related risk in daily operation, banks would likely to reduce the possibility of risk exposures. The result from this study is consistent with findings from previous study by Dicuonzo et al. (2019). Similarly, Y. Wang et al. (2021) who concluded that using new financial technologies such as big data analytics, AI, could improve commercial banks performance and enhance risk management performance.

6.3.7 The mediating effects of BDA usage

The result of the statistical analysis showed that the usage of BDA mediated the relationship between BDA skills and market performance and risk management performance. This finding provides evidence that supports the efficacy of dynamic capabilities in the context of big data. In this study, strategic resources such as human skills do not have a direct impact on market performance and risk management performance. The findings from this study is consistent with previous study (Powell & Dent-Micallef, 1997). There were no direct effects from BDA characteristics and banking performance, except the positive relationship between BDA infrastructure and risk management performance. In context of big data technology, this study discovered that having high-quality BDA staff could contribute to improving bank performance through creating innovative services/products or developing new methods for identifying fraud in banking operations. This finding is consistent with a previous study on dynamic capabilities by Gu et al. (2021).

The result from the mediating analysis highlighted that the relationship between BDA infrastructure, top management support and big data quality and banking performance was not significantly mediated through the usage of BDA. This finding suggests that the innovative services and products in the banking sector, which depend on the use of Big Data Analytics (BDA), are limited by regulatory factors. As a result, banks require a significant amount of time to achieve improved outcomes, even when they invest in BDA infrastructure. These results are in agreement with the ideas presented by Härle et al. (2016), which emphasized the possibility of complying with governmental regulations. Therefore, the indirect impact of enhancing infrastructure on market performance and risk performance through innovative services/products was not found significant in this research. This conclusion contradicts the findings of Ghasemaghaei (2018), who suggested that the use of complex BDA infrastructure does not directly affect organizational performance but rather moderates the relationship between BDA usage and organizational performance. Furthermore, the result of the direct relationship between BDA infrastructure and market performance is consistent with the findings from previous study (Ghasemaghaei, 2018; Powell & Dent-Micallef, 1997) found that solely IT resource has not produced outstanding enterprise performance. This study discovered that BDA infrastructure directly enhances banks' ability to improve risk management practices. This conclusion suggests that banks should invest more in BDA infrastructure in order to gain better results, such as identifying fraud quickly with real-time data and calculating risks faster.

The conclusions drawn from the mediating analysis further provide the evaluation of the resource-based view (RBV). It indicates that BDA infrastructure as tangible resource does not directly impact market performance but is beneficial for risk management performance. This reinforces the notion that, in the context of big data, the assumptions of RBV are no longer applicable. As companies provide numerous big data solutions to banks, infrastructure can be achieved through third parties, making the search for and utilization of big data no longer a hindrance for banks. Nevertheless, the dynamic capability theory suggests that the outstanding performance of banks can be better explained by the quality of their human resources and their innovative dynamic capability, facilitated by the usage of BDA. These factors are expected to have a positive impact on both aspects of bank performance.

CHAPTER 7: CONCLUSION

Chapter 7 presents the conclusion, limitations, and suggestions for future research, emphasizing this study's theoretical contributions to the research field's current knowledge and practical implications for business.

7.1 Theoretical contributions

This study combines and selects factors from the TOE and the RBV to identify the factors that influence the dynamic capabilities of big data analytics. The findings from this study make several significant contributions to current literature. The research results demonstrate that the combination of the TOE, the RBV, and dynamic capabilities provides a meaningful explanation for the competitive advantage of firms in terms of market performance and risk management performance. It can be concluded that the significant factors in the use of BDA are in line with the TOE and the RBV theory. The successful implementation of BDA requires a combination of strategic resources, such as BDA infrastructure, BDA skills, top management support, and big data quality. This study empirically demonstrates the importance of human resources (BDA skills) and innovative dynamic capabilities for the successful usage of BDA technology. Particularly, BDA skills are highlighted as the sole significant factor for effectively implementing digital technologies such as BDA in the banking sector.

Overall, this study investigates the relationship between four factors selected from the RBV and TOE. These factors include top management support, BDA infrastructure, BDA skills and big data quality, all of which are strongly associated within the context of big data. Additionally, the study examines BDA usage, which is defined as the innovative dynamic capabilities of enterprises across five aspects: (1) developing new products/ services, (2) implementing new business processes, (3) creating new customer relationships, (4) changing the way of doing business, (5) creating/building a new method for risk management / to detect fraud or financial crime prevention. Moreover, this study examines the impact of BDA usage on two aspects of banking performance: market performance and risk management performance. Specifically, market performance is evaluated based on the following aspects: (1) exploring new markets, (2) introducing new services/ products, (3) success rate of new products, and (4) market share. Risk management is assessed in terms of: (1) identifying risks, (2) calculating risk, (3) control and manage risk, and (4) reporting risk. Furthermore, this study also investigates the role of BDA usage as a mediator in the relationship between the four independent variables and banking performance.

The results obtained through the PLS-SEM method suggested that BDA skills are the most important and significant factor influencing BDA usage. Furthermore, the explanation of 0.44 indicates that BDA skills have a strong impact on BDA usage and can account for 44% of the variance in BDA usage. Additionally, it was found that BDA usage, as an innovative dynamic capability, positively affects both market performance and risk management performance, with the explanation powers of 0.3 and 0.224, respectively. These findings highlight the crucial role of utilizing BDA as a part of innovative dynamic capabilities in the banking sector.

Furthermore, the study reveals a mediating effect, indicating that through BDA usage, BDA skills positively impact banking performance. Interestingly, it was also found that BDA infrastructure has a significant direct impact on risk management performance in banks. This finding emphasizes the importance of infrastructure in effectively leveraging the benefits of big data technology.

This study makes some significant contribution:

The striking results from this study show that the top management support, BDA infrastructure and big data quality has no important impact on dynamic capabilities in using BDA. As a leading industry in the application of new technologies such as big data, the most challenging problem for banks is not upgrading their current infrastructure as well as their data quality but rather figuring out how and who can use big data analytics to create/build their innovative capabilities for the banking sector. These questions were answered by the results of this study.

It is noteworthy that senior management has a significant role in initiating big data analytics initiatives within banking institutions. At the first phrase of adopting big data analytics, the top management support plays an especially significant role. The support from the top management involves the leader's vision and efforts in creating a conducive atmosphere for the development of big data analytics within the bank. Therefore, big data analytics should be regarded as one of the bank's strategic objectives. Furthermore, due to the complexity of BDA, which involves multiple components and data from varieties of sources, there is a need for a powerful software and hardware infrastructure. Thus, the bank must entrust the responsibility of formulating BDA initiatives to a senior role, as this role would possess the necessary capabilities to facilitate better coordination of operations and the establishment of suitable governance policies to develop new ideas based on BDA results. Although during the phrase of using BDA to create innovative services/products, this study did not find the significant impact of the top management support, it does not mean we should underestimate this strategic role of top management.

The influence of the top manager is the initial impetus for implementing big data analytics (BDA) in the banking sector; however, the most significant factor in achieving a competitive advantage is the possession of BDA skills. This factor has the greatest impact on the bank's BDA capabilities, as the quality of training and working experience of BDA experts and staff is essential in building new products and services. This study evaluates BDA skills under the following aspects: (1) whether banks provide BDA training to our employees; (2) whether banks hire new employees that already have the BDA skills; (3) Our BDA staff has the right skills to accomplish their jobs; (4) whether bank staff has the proper education to fulfill their jobs. Furthermore, in context of big data technology, BDA skills should be considered as the key resource that monetizes the huge volume, variety, veracity, and velocity of collected data into innovative services and products for banks. This enables banks to enhance their market performance and risk management performance.

The findings of this research demonstrate that only BDA skills from resource-based views can serve as a basis for developing dynamic capabilities, thus enabling banks to gain a competitive advantage. In addition, this study provides a comprehensive evaluation of the application of BDA in banking, encompassing a wide range of areas. It is in line with current practice of not restrict big data analytics to a single functional department. BDA usage provides the bank with dynamic capabilities to support innovative production. Furthermore, the evaluation of BDA creating an innovative and dynamic capability is also appropriate for the context of digital transformation in banking operations.

The results from this empirical study provided solid evidence that using BDA should be considered a dynamic organizational capability at the enterprise level. Furthermore, in the banking context, this study found that using BDA allows banks to generate innovative capabilities by supporting new services/ products. These innovative capabilities, then, help banks enhance banks competitive advantages. In this study, the competitive advantage of banks was investigated with respect to risk management performance and market performance.

7.2 Practical implications

The insights gained from this study contribute to our understanding of using big data analytics in the banking sector in one of the most developed regions in the world.

The digital transformation process is happening at an increasingly faster pace as the speed of development and emergence of new technologies is unprecedented across various sectors in the national and international economies. New products, from an individual perspective, are continuously being renovated, such as smartphones, smartwatches, and smart glasses, with an increasing penetration rate in the markets. This undoubtedly creates a trend of changing consumer behavior in the future, including banks' customers.

From a business perspective, the explosion of new technologies and digital entrepreneurship, such as smart manufacturing, telecommunicating technologies, big data, and AI, cloud computing brings many opportunities for enterprises and banks. Therefore, both the supply and demand sides are facing rapid changes, demanding banks to create new products, services, and working methods by applying these digital technologies. One of the technologies that significantly impact banking operations is big data analytics. Research on big data analytics in the banking sector has provided some valuable insights.

First and foremost, ensuring the quality of human resources to meet the requirements of the BDA is crucial for effectively utilizing the advantages it offers. The capabilities of the BDA are evaluated in terms of training and recruitment. Using BDA in the banking industry necessitates technical skills and a basic background in banking and finance. Consequently, banks should consider providing training for existing employees in order to enhance their knowledge and improve the quality of their work. The combination of BDA skills is related to three fields of skills: programming, data, and business problem-solving skills. More importantly, banks can acquire appropriate BDA skills by recruiting experienced and fresh university graduates. Thus, collaborating and cooperating with innovation centers and universities becomes essential. These partnerships can contribute practical knowledge and create internship opportunities within banks, thereby facilitating the identification and development of talented employees.

Secondly, this study has identified a positive direct impact of BDA infrastructure on risk management performance. It is recommended that banks improve their infrastructure for using BDA, as this is likely to contribute to the enhancement of fraud detection capabilities through the aggregation of data from various sources with real-time data.

Thirdly, careful preparation of BDA platforms enables banks to maximize BDA capabilities. This is a crucial prerequisite, a strategic resource that banks need to take time to prepare. Although risk factors are not included in this study, this does not mean that managers are disregarding them. Rather, assessing risks commonly found in big data technology, such as privacy and security issues (Raguseo, 2018), should be carefully considered and evaluated during the preparation and testing stages.

Finally, although not included in this model, government support plays an important role in supporting the adoption of digital technologies at enterprises. The results from the study suggest many policy implications for governments and administrative managers. For policymakers, more focus should be placed on programs to enhance digital skills for the public and big data skills for younger generations who will work in related jobs in the future. BDA. More specifically, for universities, which provide the main workforce for the banking sectors, it is imperative to impart quantitative and programming skills, along with a comprehensive understanding of business knowledge, to students. Because BDA skills will require a good business background to generate questions, and quantitative and programming skills will help answer business requirements through better use of BDA software tools.

7.3 Limitations and directions for further research

Although the research has provided practical conclusions that complement the technology adoption theories to improve operational efficiency for enterprises, some study limitations must be acknowledged to guide future research.

Firstly, this study only evaluates the effectiveness of digital transformation activities in the banking sector. The factors affecting the development of dynamic capabilities from using big data analytics are pointed out in it. In the future, new technologies are often be interlinked, and big data analytics will go along with applying AI, ML, cloud computing, and chatbots in operations. Thus, evaluating big data analytics without considering the impact and usability of other technologies reduces the explanatory power of the research model. The author proposes future studies to expand the research model further to assess the overall impact of the application of technologies in the future of banking to different aspects of comparative advantages.

Second, stemming from the resource-based view and dynamic capabilities, this study evaluates internal factors such as human resource (BDA skills), infrastructure and management resources (top management support), and data resources that contribute to the effective building of BDA dynamic capabilities. One of the study's limitations is ignoring some influencing factors, such as data cultures, fintech cooperations, privacy and security issues, and governmental regulations. Research from Zelenka & Podaras (2021) suggests that creating a data culture can promote the exchange of knowledge within enterprises, thereby increasing a personal advantage in the enterprise. Grover et al (2018) suggested that the effect of BDA capability on business value creation could be moderated by data culture in the enterprise, which could delay or accelerate the impact of BDA capabilities depending on the development of data culture in the enterprise. Therefore, the author suggests expanding the assessment of new factors that can affect the practices of managing BDA to increase the model's explanatory power and strengthen the theories. Furthermore, while investigating the application of big data analytics in the banking sector through the survey methodology poses significant challenges, the author acknowledges that the sample size employed in this study is a limitation and suggests that further research should be undertaken using a larger sample.

Thirdly, further study should be carried out using alternative research methodologies, such as case studies or qualitative studies based on interviews, in addition to utilizing the survey method. While collecting surveys can yield more generalized findings, it is important to acknowledge that the challenges and problems associated with utilizing BDA to create banks' competitive advantage may vary across different banks. Therefore, employing qualitative methods can potentially uncover hidden results that may not be revealed through surveys alone.

Finally, the banking industry has maintained a high level of operational framework internationally. Therefore, to the greatest extent possible, this study only evaluated the use of big data analytics within banks in the EU-27 member states. Consequently, future research could expand the application of this research model to other areas. Furthermore, the present study regarded the EU-27 member states as an integrated banking market. Nonetheless, it is imperative to acknowledge that variations in economic conditions, educational attainment, and technological progress among these states may introduce deviations in the research outcomes. Consequently, it is recommended that forthcoming investigations incorporate these factors crucial considerations in their analyses.

REFERENCES

Abdalla, H. B. (2022). A brief survey on big data: technologies, terminologies and data-intensive applications. *Journal of Big Data, 9*(1), 107. https://doi.org/10.1186/s40537-022-00659-3

Abednego, K. (2021). Big Data Architecture for Fraud Detection and Prevention in Banking Industry. Retrieved January 15, 2023, from https://abedkristanto.medium.com/big-data-architecture-for-fraud-detection-in-banking-industry-afc22b89aeca

Aebi, V., Sabato, G., & Schmid, M. (2012). Risk management, corporate governance, and bank performance in the financial crisis. *Journal of Banking and Finance*, *36*(12), 3213–3226. https://doi.org/10.1016/j.jbankfin.2011.10.020

Akter, S., Fosso Wamba, S., Barrett, M., & Biswas, K. (2019). How talent capability can shape service analytics capability in the big data environment? *Journal of Strategic Marketing*, *27*(6), 521–539. https://doi.org/10.1080/0965254X.2018.1442364

Al-Jabri, I. M., & Sohail, M. S. (2012). Mobile banking adoption: Application of diffusion of innovation theory. *Journal of Electronic Commerce Research*, *13*(4), 379–391.

Alavi, M., & Carlson, P. (1992). A Review of MIS Research and Disciplinary Development. *Journal of Management Information Systems*, *8*(4), 45–62. https://doi.org/10.1080/07421222.1992.11517938

AlBar, A. M., & Hoque, M. R. (2019). Factors affecting cloud ERP adoption in Saudi Arabia: An empirical study. *Information Development*, *35*(1), 150–164. https://doi.org/10.1177/0266666917735677

Ali, Q., Salman, A., Yaacob, H., Zaini, Z., & Abdullah, R. (2020). Does Big Data Analytics Enhance Sustainability and Financial Performance? The Case of ASEAN Banks. *The Journal of Asian Finance, Economics and Business*, *7*(7), 1–13. https://doi.org/10.13106/jafeb.2020.vol7.no7.001

Ali, Q., Yaacob, H., Parveen, S., & Zaini, Z. (2021). Big data and predictive analytics to optimise social and environmental performance of Islamic banks. *Environment Systems and Decisions*, *41*(4), 616–632. https://doi.org/10.1007/s10669-021-09823-1

Ambrosini, V., & Bowman, C. (2009). What are dynamic capabilities and are they a useful construct in strategic management? *International Journal of Management Reviews*, *11*(1), 29–49. https://doi.org/10.1111/j.1468-2370.2008.00251.x

Anwar, M., Khan, S. Z., & Shah, S. Z. A. (2018). Big Data Capabilities and Firm's

Performance: A Mediating Role of Competitive Advantage. *Journal of Information & Knowledge Management*, *17*(04), 1850045. https://doi.org/10.1142/S0219649218500454

Backbase. (2020). *Banking 2025 Four pillars of the digital-first bank*. Retrieved from https://www.backbase.com/resources/banking-2025-whitepaper/

Bag, S., Pretorius, J. H. C., Gupta, S., & Dwivedi, Y. K. (2021). Role of institutional pressures and resources in the adoption of big data analytics powered artificial intelligence, sustainable manufacturing practices and circular economy capabilities. *Technological Forecasting and Social Change*, *163*(October 2020), 120420. https://doi.org/10.1016/j.techfore.2020.120420

Baig, M. I., Shuib, L., & Yadegaridehkordi, E. (2019). Big data adoption: State of the art and research challenges. *Information Processing & Management*, *56*(6), 102095. https://doi.org/10.1016/j.ipm.2019.102095

Baig, M. I., Shuib, L., & Yadegaridehkordi, E. (2021). A Model for Decision-Makers' Adoption of Big Data in the Education Sector. *Sustainability*, *13*(24), 13995. https://doi.org/10.3390/su132413995

Baker, J. (2012). The Technology–Organization–Environment Framework. In *Information Systems Theory*. Springer New York.

Bakici, T., Nemeh, A., & Hazir, O. (2022). Big Data Adoption in Project Management: Insights From French Organizations. *IEEE Transactions on Engineering Management*, 1–15. https://doi.org/10.1109/TEM.2021.3091661

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, *17*(1), 99–120. https://doi.org/10.1177/014920639101700108

Barney, J. B. (2014). How marketing scholars might help address issues in resource-based theory. *Journal of the Academy of Marketing Science*, *42*(1), 24–26. https://doi.org/10.1007/s11747-013-0351-8

Başak, S., Kılınç, İ., & Ünal, A. (2022). The effect of big data in transforming to learning organization a single-case study in IT sector. *VINE Journal of Information and Knowledge Management Systems*. https://doi.org/10.1108/VJIKMS-07-2021-0134

Batko, K., & Ślęzak, A. (2022). The use of Big Data Analytics in healthcare. *Journal of Big Data*, *9*(1), 3. https://doi.org/10.1186/s40537-021-00553-4

Begenau, J., Farboodi, M., & Veldkamp, L. (2018). Big data in finance and the growth of large firms. *Journal of Monetary Economics*, *97*, 71–87.

https://doi.org/10.1016/j.jmoneco.2018.05.013

Behl, A. (2022). Antecedents to firm performance and competitiveness using the lens of big data analytics: a cross-cultural study. *Management Decision*, *60*(2), 368–398. https://doi.org/10.1108/MD-01-2020-0121

Besson, P., & Rowe, F. (2012). Strategizing information systems-enabled organizational transformation: A transdisciplinary review and new directions. *The Journal of Strategic Information Systems*, *21*(2), 103–124. https://doi.org/10.1016/j.jsis.2012.05.001

Bhatti, S. H., Ahmed, A., Ferraris, A., Hirwani Wan Hussain, W. M., & Wamba, S. F. (2022). Big data analytics capabilities and MSME innovation and performance: A double mediation model of digital platform and network capabilities. *Annals of Operations Research*. https://doi.org/10.1007/s10479-022-05002-w

Bloomberg, J. (2018). Digitization, Digitalization, And Digital Transformation: Confuse Them At Your Peril. Retrieved from https://www.forbes.com/sites/jasonbloomberg/2018/04/29/digitization-digitalization-and-digital-transformation-confuse-them-at-your-peril/

Braganza, A., Brooks, L., Nepelski, D., Ali, M., & Moro, R. (2017). Resource management in big data initiatives: Processes and dynamic capabilities. *Journal of Business Research*, *70*, 328–337. https://doi.org/10.1016/j.jbusres.2016.08.006

Byrd, T. A., Pitts, J. P., Adrian, A. M., & Davidson, N. W. (2008). Examination of a path model relating information technology infrastructure with firm performance. *Journal of Business Logistics*, *29*(2), 161–187. https://doi.org/10.1002/j.2158-1592.2008.tb00091.x

Cabrera-Sánchez, J.-P., & Villarejo-Ramos, Á. F. (2020). Acceptance and use of big data techniques in services companies. *Journal of Retailing and Consumer Services*, *52*(July 2019), 101888. https://doi.org/10.1016/j.jretconser.2019.101888

Capgemini. (2020). Data-driven decision-making in organizations worldwide as of 2020, by sector. Retrieved June 7, 2023, from https://www.statista.com/statistics/1235436/worldwide-data-driven-decision-making-organizations-by-sector/

Chatterjee, S., Chaudhuri, R., Vrontis, D., & Thrassou, A. (2022). Impacts of big data analytics adoption on firm sustainability performance. *Qualitative Research in Financial Markets*. https://doi.org/10.1108/QRFM-01-2022-0005

Chaurasia, S. S., & Verma, S. (2020). Strategic determinants of big data analytics in the AEC sector: a multi-perspective framework. *Construction Economics and Building*, *20*(4), 63–81. https://doi.org/10.5130/AJCEB.v20i4.6649

Chen, D. Q., Preston, D. S., & Swink, M. (2015). How the Use of Big Data Analytics Affects Value Creation in Supply Chain Management. *Journal of Management Information Systems*, *32*(4), 4–39. https://doi.org/10.1080/07421222.2015.1138364

Chen, X., & Zhang, X. (2016). How environmental uncertainty moderates the effect of relative advantage and perceived credibility on the adoption of mobile health services by Chinese organizations in the big data era. *International Journal of Telemedicine and Applications*, *2016*. https://doi.org/10.1155/2016/3618402

Creswell, J. W. (2009). Mapping the field of mixed methods research. *Journal of Mixed Methods Research*, *3*(2), 95–108.

Crossan, F. (2003). Research philosophy: towards an understanding. *Nurse Researcher*, *11*(1), 46–55. https://doi.org/10.7748/nr2003.10.11.1.46.c5914

Cuesta, C., Ruesta, M., Tuesta, D., & Urbiola, P. (2015). The digital transformation of the banking industry. *BBVA Research*, 1–10. Retrieved from www.bbvaresearch.com

Curado, C., & Bontis, N. (2006). The knowledge-based view of the firm and its theoretical precursor. *International Journal of Learning and Intellectual Capital*, *3*(4), 367. https://doi.org/10.1504/IJLIC.2006.011747

Davenport, T. H., Barth, P., & Bean, R. (2012). How 'Big Data'is Different. *MIT Sloan Management Review*, *54*, 43–46.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, *35*(8), 982–1003. https://doi.org/10.1287/mnsc.35.8.982

Debortoli, S., Müller, O., & vom Brocke, J. (2014). Comparing Business Intelligence and Big Data Skills. *Business & Information Systems Engineering*, *6*(5), 289–300. https://doi.org/10.1007/s12599-014-0344-2

Dicuonzo, G., Galeone, G., Zappimbulso, E., & Dell'Atti, V. (2019). Risk management 4.0: The role of big data analytics in the bank secctor. *International Journal of Economics and Financial Issues*, *9*(6), 40–47. https://doi.org/10.32479/ijefi.8556

Dumbill, E. (2013). Making sense of big data. *Big Data*, *1*(1), 1–2.

EBA. (2020). *Report on Big Data and Advanced Analytics*.

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they? *Strategic Management Journal*, *21*(10–11), 1105–1121. https://doi.org/10.1002/1097-0266(200010/11)21:10/11<1105::AID-SMJ133>3.0.CO;2-E

El-Haddadeh, R., Osmani, M., Hindi, N., & Fadlalla, A. (2021). Value creation for realising the sustainable development goals: Fostering organisational adoption of big data analytics. *Journal of Business Research*, *131*(October 2020), 402–410. https://doi.org/10.1016/j.jbusres.2020.10.066

European Banking Federation. (2022). *Banking in Europe: EBF Facts & Figures 2022*.

European Commission. (2022). Towards a green, digital and resilient economy. Retrieved from https://ec.europa.eu/commission/presscorner/detail/en/IP_22_1467

European Commission. (2023). Shaping Europe's digital future.

Fan, Y., Chen, J., Shirkey, G., John, R., Wu, S. R., Park, H., & Shao, C. (2016). Applications of structural equation modeling (SEM) in ecological studies: an updated review. *Ecological Processes*, *5*(1), 19. https://doi.org/10.1186/s13717-016-0063-3

Fanelli, S., Pratici, L., Salvatore, F. P., Donelli, C. C., & Zangrandi, A. (2023). Big data analysis for decision-making processes: challenges and opportunities for the management of health-care organizations. *Management Research Review*, *46*(3), 369–389. https://doi.org/10.1108/MRR-09-2021-0648

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Fornell, C., & Larcker, D. F. (1981). Structural Equation Models With Unobservable Variables and Measurement Error : Algebra and Statistics. *Journal of Marketing Research*, *XVIII*(August), 382–388.

Fosso Wamba, S., Akter, S., & de Bourmont, M. (2019). Quality dominant logic in big data analytics and firm performance. *Business Process Management Journal*, *25*(3), 512–532. https://doi.org/10.1108/BPMJ-08-2017-0218

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, *165*, 234–246. https://doi.org/10.1016/j.ijpe.2014.12.031

Fosso Wamba, S., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356–365. https://doi.org/10.1016/j.jbusres.2016.08.009

Frank, A. G., Dalenogare, L. S., & Ayala, N. F. (2019). Industry 4.0 technologies: Implementation patterns in manufacturing companies. *International Journal of Production Economics*, *210*(September 2018), 15–26. https://doi.org/10.1016/j.ijpe.2019.01.004

Gable, G. G. (1994). Integrating case study and survey research methods: an example in information systems. *European Journal of Information Systems*, *3*(2), 112–126. https://doi.org/10.1057/ejis.1994.12

Gangwar, H. (2018). Understanding the Determinants of Big Data Adoption in India. *Information Resources Management Journal*, *31*(4), 1–22. https://doi.org/10.4018/IRMJ.2018100101

Gangwar, H., Date, H., & Ramaswamy, R. (2015). Understanding determinants of cloud computing adoption using an integrated TAM-TOE model. *Journal of Enterprise Information Management*, *28*(1), 107–130.

Gefen, Rigdon, & Straub. (2011). Editor's Comments: An Update and Extension to SEM Guidelines for Administrative and Social Science Research. *MIS Quarterly*, *35*(2), iii. https://doi.org/10.2307/23044042

Ghasemaghaei, M. (2018). Improving Organizational Performance Through the Use of Big Data Improving Organizational Performance Through the Use of Big Data. *Journal of Computer Information Systems*, *60*(5), 395–408. https://doi.org/10.1080/08874417.2018.1496805

Ghasemaghaei, M. (2019). Are firms ready to use big data analytics to create value? The role of structural and psychological readiness. *Enterprise Information Systems*, *13*(5), 650–674. https://doi.org/10.1080/17517575.2019.1576228

Ghasemaghaei, M. (2021). Understanding the impact of big data on firm performance: The necessity of conceptually differentiating among big data characteristics. *International Journal of Information Management*, *57*(October 2019), 102055. https://doi.org/10.1016/j.ijinfomgt.2019.102055

Ghasemaghaei, M., & Calic, G. (2020). Assessing the impact of big data on firm innovation performance: Big data is not always better data. *Journal of Business*

Research, *108*(April 2019), 147–162. https://doi.org/10.1016/j.jbusres.2019.09.062

Ghasemaghaei, M., Hassanein, K., & Turel, O. (2017). Increasing firm agility through the use of data analytics: The role of fit. *Decision Support Systems*, *101*, 95–105. https://doi.org/10.1016/j.dss.2017.06.004

Giebe, C., Hammerström, L., & Zwerenz, D. (2019). Big Data & Analytics as a sustainable Customer Loyalty Instrument in Banking and Finance. *Financial Markets, Institutions and Risks*, *3*(4), 74–88. https://doi.org/10.21272/fmir.3(4).74-88.2019

Gillon, K., Aral, S., Ching-Yung, L., Mithas, S., & Zozulia, M. (2014). Business analytics: Radical shift or incremental change? *Communications of the Association for Information Systems*, *34*, 287–296.

Grant, R.M. (2010). *Contemporary strategy analysis: Text and cases edition*. Wiley.

Grant, Robert M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, *17*, 109–122.

Grossman, R. L., & Siegel, K. P. (2014). Organizational Models for Big Data and Analytics. *Journal of Organization Design*, *3*(1), 20. https://doi.org/10.7146/jod.9799

Grover, V., Chiang, R. H. L., Liang, T. P., & Zhang, D. (2018). Creating Strategic Business Value from Big Data Analytics: A Research Framework. *Journal of Management Information Systems*, *35*(2), 388–423. https://doi.org/10.1080/07421222.2018.1451951

Gu, V. C., Zhou, B., Cao, Q., & Adams, J. (2021). Exploring the relationship between supplier development, big data analytics capability, and firm performance. *Annals of Operations Research*, *302*(1), 151–172. https://doi.org/10.1007/s10479-021-03976-7

Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S. F., Childe, S. J., Hazen, B., & Akter, S. (2017). Big data and predictive analytics for supply chain and organizational performance. *Journal of Business Research*, *70*, 308–317. https://doi.org/10.1016/j.jbusres.2016.08.004

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, *53*(8), 1049–1064. https://doi.org/10.1016/j.im.2016.07.004

Gupta, S., Qian, X., Bhushan, B., & Luo, Z. (2019). Role of cloud ERP and big data on firm performance: a dynamic capability view theory perspective. *Management Decision*, *57*(8), 1857–1882. https://doi.org/10.1108/MD-06-2018-0633

Gurău, C., & Ranchhod, A. (2020). Implementing Big Data Analytics in Small Firms: A

Situated Human Practice Approach. *Canadian Journal of Administrative Sciences / Revue Canadienne Des Sciences de l'Administration*, *37*(2), 111–121. https://doi.org/10.1002/cjas.1529

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Hair Jr., J. F., Matthews, L. M., Matthews, R. L., & Sarstedt, M. (2017). PLS-SEM or CB-SEM: updated guidelines on which method to use. *International Journal of Multivariate Data Analysis*, *1*(2), 107. https://doi.org/10.1504/IJMDA.2017.10008574

Härle, P., Havas, A., & Samandari, H. (2016). Risk - The future of bank risk management. *McKinsey&Co Corp.*, 1–7.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Henseler, J., & Sarstedt, M. (2013). Goodness-of-fit indices for partial least squares path modeling. *Computational Statistics*, *28*(2), 565–580. https://doi.org/10.1007/s00180-012-0317-1

Hsu, C. L., & Lin, J. C. C. (2016). Factors affecting the adoption of cloud services in enterprises. *Information Systems and E-Business Management*, *14*(4), 791–822. https://doi.org/10.1007/s10257-015-0300-9

Hsu, P.-F., Ray, S., & Li-Hsieh, Y.-Y. (2014). Examining cloud computing adoption intention, pricing mechanism, and deployment model. *International Journal of Information Management*, *34*(4), 474–488. https://doi.org/10.1016/j.ijinfomgt.2014.04.006

Huang, C.-K., Wang, T., & Huang, T.-Y. (2020). Initial Evidence on the Impact of Big Data Implementation on Firm Performance. *Information Systems Frontiers*, *22*(2), 475–487. https://doi.org/10.1007/s10796-018-9872-5

Iacovou, C. L., Benbasat, I., & Dexter, A. S. (1995). Electronic Data Interchange and Small Organizations: Adoption and Impact of Technology. *MIS Quarterly*, *19*(4), 465. https://doi.org/10.2307/249629

IBM. (2023). Big data and analytics support. Retrieved May 20, 2023, from https://www.ibm.com/docs/en/storage-scale-bda?topic=big-data-analytics-support

IDC. (2019). Share of big data and business analytics revenues worldwide in 2019, by industry. Retrieved June 7, 2023, from https://www.statista.com/statistics/616225/worldwide-big-data-business-analytics-revenue/

IDC. (2021). Market share of big data and analytics spending in India in 2021, by industry. Retrieved June 7, 2023, from https://www.statista.com/statistics/1264351/india-market-share-of-big-data-and-analytics-spending-by-industry/

Irving Fisher Committee. (2021). *Use of big data sources and applications at central banks*. *IFC Report no. 13. Basel: Irving Fisher Committee on Central Bank Statistics.* Retrieved from https://www.bis.org/ifc/publ/ifc-report-bigdata.pdf

Janssens, J. (2019). Digital Transformation journeys in a Digitized Reality. In *Advanced Methodologies and Technologies in Business Operations and Management* (pp. 282–294). IGI Global.

Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., & Childe, S. J. (2017). Modelling quality dynamics, business value and firm performance in a big data analytics environment. *International Journal of Production Research*, *55*(17), 5011–5026. https://doi.org/10.1080/00207543.2016.1154209

Kambatla, K., Kollias, G., Kumar, V., & Grama, A. (2014). Trends in big data analytics. *Journal of Parallel and Distributed Computing*, *74*(7), 2561–2573. https://doi.org/10.1016/j.jpdc.2014.01.003

Kane, G. C., Palmer, D., Philips Nguyen, A., Kiron, D., & Buckley, N. (2015). Strategy, Not Technology, Drives Digital Transformation. *MIT Sloan Management Review & Deloitte*, (57181), 27. Retrieved from http://sloanreview.mit.edu/projects/strategy-drives-digital-transformation/

Kayabay, K., Gökalp, M. O., Gökalp, E., Erhan Eren, P., & Koçyiğit, A. (2022). Data science roadmapping: An architectural framework for facilitating transformation towards a data-driven organization. *Technological Forecasting and Social Change*, *174*(February 2021), 121264. https://doi.org/10.1016/j.techfore.2021.121264

Kulkarni, M., & Patil, K. (2020). Block Chain Technology Adoption for Banking Services-Model based on Technology-Organization-Environment theory. *SSRN Electronic Journal*, 1–11. https://doi.org/10.2139/ssrn.3563101

Kumpulainen, M., & Seppänen, M. (2022). Combining Web of Science and Scopus datasets in citation-based literature study. *Scientometrics*, *127*(10), 5613–5631.

https://doi.org/10.1007/s11192-022-04475-7

Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, *34*(3), 387–394. https://doi.org/10.1016/j.ijinfomgt.2014.02.002

Lai, Y., Sun, H., & Ren, J. (2018). Understanding the determinants of big data analytics (BDA) adoption in logistics and supply chain management: An empirical investigation. *International Journal of Logistics Management*, *29*(2), 676–703.

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META Group Research Note*, *6*(70), 1.

Liang, H., Saraf, N., Hu, Q., & Xue, Y. (2007). Assimilation of Enterprise Systems: The Effect of Institutional Pressures and the Mediating Role of Top Management. *MIS Quarterly*, *31*(1), 59–87.

Libert, B., Beck, M., & Wind, Y. (2016). Questions to Ask Before Your Next Digital Transformation. *Harvard Business Review*, *60*(12), 11–13.

Low, C., Chen, Y., & Wu, M. (2011). Understanding the determinants of cloud computing adoption. *Industrial Management and Data Systems*, *111*(7), 1006–1023. https://doi.org/10.1108/02635571111161262

Lutfi, A., Al-Khasawneh, A. L., Almaiah, M. A., Alshira'h, A. F., Alshirah, M. H., Alsyouf, A., … Ali, R. Al. (2022). Antecedents of Big Data Analytic Adoption and Impacts on Performance: Contingent Effect. *Sustainability*, *14*(23), 15516. https://doi.org/10.3390/su142315516

Mahmoud Ali, W. S. M., Hasan Basari, A. S., & Mohd Yusoh, Z. I. (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), 415–420. https://doi.org/10.14445/22315381/IJETT-V70I12P239

Manuylenko, V. V., Borlakova, A. I., Milenkov, A. V., Bigday, O. B., Drannikova, E. A., & Lisitskaya, T. S. (2021). Development and Validation of a Model for Assessing Potential Strategic Innovation Risk in Banks Based on Data Mining-Monte-Carlo in the "Open Innovation" System. *Risks*, *9*(6), 118. https://doi.org/10.3390/risks9060118

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). *Big data: The next frontier for innovation, competition, and productivity*. *McKinsey Global Institute*.

Maroufkhani, P., Tseng, M. L., Iranmanesh, M., Ismail, W. K. W., & Khalid, H. (2020).

Big data analytics adoption: Determinants and performances among small to medium-sized enterprises. *International Journal of Information Management*, *54*(July). https://doi.org/10.1016/j.ijinfomgt.2020.102190

Maroufkhani, P., Wagner, R., Wan Ismail, W. K., Baroto, M. B., & Nourani, M. (2019). Big data analytics and firm performance: A systematic review. *Information (Switzerland)*, *10*(7), 1–21. https://doi.org/10.3390/INFO10070226

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

Matt, C., Hess, T., & Benlian, A. (2015). Digital Transformation Strategies. *Business & Information Systems Engineering*, *57*(5), 339–343. https://doi.org/10.1007/s12599-015-0401-5

Melnychenko, S., Volosovych, S., & Baraniuk, Y. (2020). Dominant ideas of financial technologies in digital banking. *Baltic Journal of Economic Studies*, *6*(1), 92. https://doi.org/10.30525/2256-0742/2020-6-1-92-99

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, *98*(July 2018), 261–276. https://doi.org/10.1016/j.jbusres.2019.01.044

Mikalef, P., & Krogstie, J. (2020). Examining the interplay between big data analytics and contextual factors in driving process innovation capabilities. *European Journal of Information Systems*, *29*(3), 260–287. https://doi.org/10.1080/0960085X.2020.1740618

Milojević, N., & Redzepagic, S. (2021). Prospects of Artificial Intelligence and Machine Learning Application in Banking Risk Management. *Journal of Central Banking Theory and Practice*, *10*(3), 41–57. https://doi.org/10.2478/jcbtp-2021-0023

Miloslavskaya, N., & Tolstoy, A. (2016). Big Data, Fast Data and Data Lake Concepts. *Procedia Computer Science*, *88*, 300–305. https://doi.org/10.1016/j.procs.2016.07.439

Mishra, D., Luo, Z., Hazen, B., Hassini, E., & Foropon, C. (2019). Organizational capabilities that enable big data and predictive analytics diffusion and organizational performance. *Management Decision*, *57*(8), 1734–1755. https://doi.org/10.1108/MD-03-2018-0324

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics*, *106*(1), 213–228.

https://doi.org/10.1007/s11192-015-1765-5

Mordor Intelligence. (2023). Big Data Analytics in Banking Market. Retrieved January 29, 2023, from https://www.mordorintelligence.com/industry-reports/big-data-in-banking-industry

Moumtzidis, I., Kamariotou, M., & Kitsios, F. (2022). Digital Transformation Strategies Enabled by Internet of Things and Big Data Analytics: The Use-Case of Telecommunication Companies in Greece. *Information*, *13*(4), 196. https://doi.org/10.3390/info13040196

Müller, O., Fay, M., & vom Brocke, J. (2018). The Effect of Big Data and Analytics on Firm Performance: An Econometric Analysis Considering Industry Characteristics. *Journal of Management Information Systems*, *35*(2), 488–509. https://doi.org/10.1080/07421222.2018.1451955

Niebel, T., Rasel, F., & Viete, S. (2019). BIG data – BIG gains? Understanding the link between big data analytics and innovation. *Economics of Innovation and New Technology*, *28*(3), 296–316. https://doi.org/10.1080/10438599.2018.1493075

Nobanee, H., Dilshad, M. N., Al Dhanhani, M., Al Neyadi, M., Al Qubaisi, S., & Al Shamsi, S. (2021). Big Data Applications the Banking Sector: A Bibliometric Analysis Approach. *SAGE Open*, *11*(4), 215824402110672. https://doi.org/10.1177/21582440211067234

Nwankpa, J. K., & Roumani, Y. (2016). IT Capability and Digital Transformation: A Firm Performance Perspective. In *International Conference of Information Systems,Dublin, Ireland*.

O'Leary, D. E. (2013). Artificial intelligence and big data. *IEEE Intelligent Systems*, *28*(2), 96–99. https://doi.org/10.1109/MIS.2013.39

Oliveira, T., & Martins, M. F. (2010). Understanding e-business adoption across industries in European countries. *Industrial Management and Data Systems*, *110*(9), 1337–1354. https://doi.org/10.1108/02635571011087428

Oracle. (2023). What is a data warehouse? Retrieved May 2, 2023, from https://www.oracle.com/database/what-is-a-data-warehouse/

Osservatori Digital Innovation. (2021). Distribution of the expenditure on big data analytics (BDA) in Italy in 2020, by industry. Retrieved June 7, 2023, from https://www.statista.com/statistics/1035547/big-data-analytics-expenditure-by-industry-italy/

Park, J. H., & Kim, Y. B. (2019). Factors Activating Big Data Adoption by Korean Firms. *Journal of Computer Information Systems*, *0*(0), 1–9. https://doi.org/10.1080/08874417.2019.1631133

Pejić Bach, M., Bertoncel, T., Meško, M., Suša Vugec, D., & Ivančić, L. (2020). Big Data Usage in European Countries: Cluster Analysis Approach. *Data*, *5*(1), 25. https://doi.org/10.3390/data5010025

Penrose, E. T. (1959). *The Theory of the Growth of the Firm*. Wiley: New York.

Phan, D. T., & Tran, L. Q. T. (2022). Building a Conceptual Framework for Using Big Data Analytics in the Banking Sector. *Intellectual Economics*, *16*(1), 5–23. https://doi.org/10.13165/IE-22-16-1-01

Polit, D. F., & Beck, C. T. (2006). Using research in evidence-based nursing practice. In *Essentials of nursing research : methods, appraisal, and utilization* (12th ed., pp. 457–494). Lippincott Williams & Wilkins.

Powell, T. C., & Dent-Micallef, A. (1997). Information technology as competitive advantage: the role of human, business, and technology resources. *Strategic Management Journal*, *18*(5), 375–405. https://doi.org/10.1002/(SICI)1097-0266(199705)18:5<375::AID-SMJ876>3.0.CO;2-7

Ragu-Nathan, B. S., Apigian, C. H., Ragu-Nathan, T. S., & Tu, Q. (2004). A path analytic study of the effect of top management support for information systems performance. *Omega*, *32*(6), 459–471. https://doi.org/10.1016/j.omega.2004.03.001

Raguseo, E. (2018). Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. *International Journal of Information Management*, *38*(1), 187–195. https://doi.org/10.1016/j.ijinfomgt.2017.07.008

Raguseo, E., & Vitari, C. (2018). Investments in big data analytics and firm performance: an empirical investigation of direct and mediating effects. *International Journal of Production Research*, *56*(15), 5206–5221. https://doi.org/10.1080/00207543.2018.1427900

Raguseo, E., Vitari, C., & Pigni, F. (2020). Profiting from big data analytics: The moderating roles of industry concentration and firm size. *International Journal of Production Economics*, *229*(January), 107758. https://doi.org/10.1016/j.ijpe.2020.107758

Rakic, S., Pero, M., Sianesi, A., & Marjanovic, U. (2022). Digital Servitization and Firm Performance: Technology Intensity Approach. *Engineering Economics*, *33*(4), 398–

413. https://doi.org/10.5755/j01.ee.33.4.29649

Ray, S., Danks, N. P., & Calero Valdez, A. (2021). R package seminr: Domain-speciic language for building and estimating structural equation models version 2.1.0.

Razaghi, S., & Shokouhyar, S. (2021). Impacts of big data analytics management capabilities and supply chain integration on global sourcing: a survey on firm performance. *The Bottom Line*, *34*(2), 198–223. https://doi.org/10.1108/BL-11-2020-0071

Reinartz, W., Haenlein, M., & Henseler, J. (2009). An empirical comparison of the efficacy of covariance-based and variance-based SEM. *International Journal of Research in Marketing*, *26*(4), 332–344. https://doi.org/10.1016/j.ijresmar.2009.08.001

Rippa, P., & Secundo, G. (2019). Digital academic entrepreneurship: The potential of digital technologies on academic entrepreneurship. *Technological Forecasting and Social Change*, *146*(November 2017), 900–911. https://doi.org/10.1016/j.techfore.2018.07.013

Rogers, E. (2003). *Diffusion of Innovation*. New York, NY: Free Press.

Rogers, E. M. (1995). Diffusion of Innovations: Modifications of a Model for Telecommunications. In *Die Diffusion von Innovationen in der Telekommunikation*. https://doi.org/10.1007/978-3-642-79868-9_2

Rungtusanatham, M., Miller, J. W., & Boyer, K. K. (2014). Theorizing, testing, and concluding for mediation in SCM research: Tutorial and procedural recommendations. *Journal of Operations Management*, *32*(3), 99–113. https://doi.org/10.1016/j.jom.2014.01.002

Sabharwal, R., & Miah, S. J. (2021). A new theoretical understanding of big data analytics capabilities in organizations: a thematic analysis. *Journal of Big Data*, *8*(1), 159. https://doi.org/10.1186/s40537-021-00543-6

Saeed, A., Riaz, H., & Baloch, M. S. (2022). Does big data utilization improve firm legitimacy? *Technological Forecasting and Social Change*, *182*(July), 121847. https://doi.org/10.1016/j.techfore.2022.121847

Saleem, H., Li, Y., Ali, Z., Mehreen, A., & Mansoor, M. S. (2020). An empirical investigation on how big data analytics influence China SMEs performance: do product and process innovation matter? *Asia Pacific Business Review*, *26*(5), 537–562. https://doi.org/10.1080/13602381.2020.1759300

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2022). Partial Least Squares Structural Equation Modeling. In *Handbook of Market Research* (pp. 587–632). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-57413-4_15

Schallmo, D. R. A., & Williams, C. A. (2018). History of Digital Transformation. In *Digital Transformation Now!* (pp. 3–8). Springer, Cham. https://doi.org/10.1007/978-3-319-72844-5_2

Schultz, B. (2013). Big data in big companies. *Baylor Business Review*, *32*(1), 20–21.

Sekaran, U., & Bougie, R. (2016). *Research Methods for Business: A Skill-Building Approarch* (7th Editio). Hoboken, NJ: John Wiley & Sons, Inc.

Shahbaz, M., Gao, C., Zhai, L., Shahzad, F., & Arshad, M. R. (2020). Moderating Effects of Gender and Resistance to Change on the Adoption of Big Data Analytics in Healthcare. *Complexity*, *2020*, 1–13. https://doi.org/10.1155/2020/2173765

Shaikh, A. A., & Karjaluoto, H. (2014). Mobile banking adoption: A literature review. *Telematics and Informatics*, *32*(1), 129–142. https://doi.org/10.1016/j.tele.2014.05.003

Sheikh, A. A., Rana, N. A., Inam, A., Shahzad, A., & Awan, H. M. (2018). Is e-marketing a source of sustainable business performance? Predicting the role of top management support with various interaction factors. *Cogent Business & Management*, *5*(1), 1516487. https://doi.org/10.1080/23311975.2018.1516487

Shirazi, F., Tseng, H.-T., Adegbite, O., Hajli, N., & Rouhani, S. (2022). New product success through big data analytics: an empirical evidence from Iran. *Information Technology & People*, *35*(5), 1513–1539. https://doi.org/10.1108/ITP-03-2020-0105

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research, 70*, 263–286. https://doi.org/10.1016/j.jbusres.2016.08.001

Statista. (2023). Global big data analytics market size 2021-2029.

Stief, S. E., Eidhoff, A. T., & Voeth, M. (2016). Transform to Succeed: An Empirical Analysis of Digital Transformation in Firms. *World Academy of Science, Engineering and Technology, International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, *10*(6), 1833–1842.

Strawn, G. . (2012). Scientific research: How many paradigms? *Educause Review*, *47*(26).

Sun, S., Cegielski, C. G., Jia, L., & Hall, D. J. (2018). Understanding the Factors Affecting the Organizational Adoption of Big Data. *Journal of Computer Information Systems*, *58*(3), 193–203. https://doi.org/10.1080/08874417.2016.1222891

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic Capabilities and Strategic Management. *Strategic Management Journal*, *18*(7), 509–533. https://doi.org/10.1093/0199248540.003.0013

Tenenhaus, M., Vinzi, V. E., Chatelin, Y. M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics & Data Analysis*, *48*(1), 159–205.

Thirathon, U., Wieder, B., & Ossimitz, M.-L. (2022). Determinants of analytics-based managerial decisionmaking. *International Journal of Information Systems and Project Management*, *6*(1), 27–40. https://doi.org/10.12821/ijispm060102

Tornatzky, L. G., & Fleischer, M. (1990). *Processes of technological innovation*. Lexington books.

Tran, L. Q. T., Phan, D. T., Herdon, M., & Kovacs, L. (2022). Assessing the Digital Transformation in Two Banks: Case Study in Hungary. *Agris On-Line Papers in Economics and Informatics*, *14*(2), 121–134. https://doi.org/10.7160/aol.2022.140210

Tran, T. L. Q., Herdon, M., Phan, T. D., & Nguyen, T. M. (2023). Digital skill types and economic performance in the EU27 region, 2020–2021. *Regional Statistics, 13*(3), 536–558. https://doi.org/10.15196/RS130307

Vărzaru, A. A. (2022). Assessing Digital Transformation of Cost Accounting Tools in Healthcare. *International Journal of Environmental Research and Public Health*, *19*(23), 15572. https://doi.org/10.3390/ijerph192315572

Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*. https://doi.org/10.1287/mnsc.46.2.186.11926

Verma, S., & Chaurasia, S. (2019). Understanding the Determinants of Big Data Analytics Adoption. *Information Resources Management Journal*, *32*(3), 1–26. https://doi.org/10.4018/IRMJ.2019070101

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *Journal of Strategic Information Systems*, *28*(2), 118–144. https://doi.org/10.1016/j.jsis.2019.01.003

Wahyudi, A., Farhani, A., & Janssen, M. (2018). Relating Big Data and Data Quality in Financial Service Organizations. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 11195 LNCS, pp. 504–519). Springer International Publishing.

https://doi.org/10.1007/978-3-030-02131-3_45

Wang, C. L., & Ahmed, P. K. (2007). Dynamic capabilities: A review and research agenda. *International Journal of Management Reviews*, *9*(1), 31–51. https://doi.org/10.1111/j.1468-2370.2007.00201.x

Wang, F. Y. (2012). A big-data perspective on AI: Newton, Merton, and analytics intelligence. *IEEE Intelligent Systems*, *27*(5), 2–4.

Wang, Y. M., Wang, Y. S., & Yang, Y. F. (2010). Understanding the determinants of RFID adoption in the manufacturing industry. *Technological Forecasting and Social Change*, *77*(5), 803–815. https://doi.org/10.1016/j.techfore.2010.03.006

Wang, Y., Xiuping, S., & Zhang, Q. (2021). Can fintech improve the efficiency of commercial banks? —An analysis based on big data. *Research in International Business and Finance*, *55*(September 2020), 101338. https://doi.org/10.1016/j.ribaf.2020.101338

Wikibon. (2018). Big data and analytics software market share from 2015 to 2017, by vendor. Retrieved June 7, 2023, from https://www.statista.com/statistics/491542/big-data-software-by-leading-vendor-share/

Yadegaridehkordi, E., Nilashi, M., Shuib, L., Hairul Nizam Bin Md Nasir, M., Asadi, S., Samad, S., & Fatimah Awang, N. (2020). The impact of big data on firm performance in hotel industry. *Electronic Commerce Research and Applications*, *40*(December 2019), 100921. https://doi.org/10.1016/j.elerap.2019.100921

Youssef, M. A. E.-A., Eid, R., & Agag, G. (2022). Cross-national differences in big data analytics adoption in the retail industry. *Journal of Retailing and Consumer Services*, *64*(November 2021), 102827. https://doi.org/10.1016/j.jretconser.2021.102827

Yu, D., Tao, S., Hanan, A., Ong, T. S., Latif, B., & Ali, M. (2022). Fostering Green Innovation Adoption through Green Dynamic Capability: The Moderating Role of Environmental Dynamism and Big Data Analytic Capability. *International Journal of Environmental Research and Public Health*, *19*(16), 10336. https://doi.org/10.3390/ijerph191610336

Zelenka, M., & Podaras, A. (2021). Increasing the Effectivity of Business Intelligence Tools via Amplified Data Knowledge. *Studies in Informatics and Control*, *30*(2), 67–77. https://doi.org/10.24846/v30i2y202106

Zhan, Y., & Tan, K. H. (2020). An analytic infrastructure for harvesting big data to enhance supply chain performance. *European Journal of Operational Research*, *281*(3), 559–574. https://doi.org/10.1016/j.ejor.2018.09.018

Zhao, X., Lynch Jr, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, *37*(2), 197–206.

Zhou, L., Pan, S., Wang, J., & Vasilakos, A. V. (2017). Machine learning on big data: Opportunities and challenges. *Neurocomputing*, *237*(January), 350–361. https://doi.org/10.1016/j.neucom.2017.01.026

Zhou, Z. H., Chawla, N. V, Jin, Y., & Williams, G. J. (2014). Big data opportunities and challenges: Discussions from data analytics perspectives [discussion forum]. *IEEE Computational Intelligence Magazine*, *9*(4), 62–74.

Zhu, L., Yu, F. R., Wang, Y., Ning, B., & Tang, T. (2019). Big Data Analytics in Intelligent Transportation Systems: A Survey. *IEEE Transactions on Intelligent Transportation Systems*, *20*(1), 383–398. https://doi.org/10.1109/TITS.2018.2815678

Zikopoulos, P., Deroos, D., Parasuraman, K., Deutsch, T., Giles, J., & Corrigan, D. (2012). *Harness the power of big data The IBM big data platform*. *McGraw Hill Professional*. McGraw Hill Professional.

LIST OF PUBLICATIONS

Journal articles

Tran, L. Q. T., Herdon, M., **Phan, D. T**., & Nguyen, M. T. (2023). Digital skill types and economic performance in the EU-27 region, 2020-2021. *Regional Statistics, 3(3), 536-558.* *https://doi.org/10.15196/RS130307*

Setiawan, B., **Phan, T. D**., Medina, J., Wieriks, M., Nathan, R. J., & Fekete-Farkas, M. (2023). Quest for financial inclusion via digital financial services (Fintech) during COVID-19 pandemic: case study of women in Indonesia. *Journal of Financial Services Marketing*, 1-15. https://doi.org/10.1057/s41264-023-00217-9

Dai, T. P. (2022). Assessing London International Financial Center in Double Uncertain Scenarios. *Studies in Business and Economics*, 17(2), 270-281. https://doi.org/10.2478/sbe-2022-0038

Tran, L. Q. T., **Phan, D. T**., & Nguyen, M. T. (2022). Digital Economy: A Comparative Study in ASEAN. *Theory, Methodology, Practice - Review of Business and Management*, *18*(02), 83-92. https://doi.org/10.18096/TMP.2022.02.05

Phan, D. T., & Tran, L. Q. T. (2022). Building a Conceptual Framework for Using Big Data Analytics in the Banking Sector. *Intellectual Economics*, 1(16), 5-23. https://doi.org/10.13165/IE-22-16-1-01

Tran, L. Q. T., **Phan, D. T**., Herdon, M., & Kovacs, L. (2022). Assessing the Digital Transformation in Two Banks: Case Study in Hungary. *AGRIS on-line Papers in Economics and Informatics*, *14*(665-2022-769), 121-134. https://doi.org/10.7160/aol.2022.140210

Dai Thich, P. (2021). A Study on Behavior Intention to Adopt Mobile Banking Apps. *International Journal of E-Services and Mobile Applications (IJESMA)*, *13*(2), 60-72. https://doi.org/10.4018/IJESMA.2021040104

Phan, D. T. (2020). I have seen the future, and it rings-What we know about mobile banking research. Theory Methodology Practice: Club of Economics in Miskolc, 16(2), 69-80. https://doi.org/10.18096/TMP.2020.02.07

Conference papers

Thich D. Phan. (2022). An analysis of digital skills and big data analytics adoption in European SMEs: A correlation study. In: Veresne, Mariann Somori; Liptak, Katalin;

Harangozo, Zolt (ed.). “*Balance and Challenges – Sustainability*” XII. International Scientific Conference. Miskolc University, Hungary. p. 725-730.

Phan Dai Thich, Tran Lam Quynh Trang. (2021). Big data analytics usage and banking performance: a conceptual framework. In: Alina Badulescu (ed.). *Emerging markets economics and business. Contribution of young researchers.* Proceedings of the 12th international conference of doctoral students and young researchers. Oradea University, Oradea, Romania. p. 352-356.

Phan Dai Thich. (2020). Bibliographic literature review in big data adoption. In: Szucs, Peter (ed.) *Doctoral students' Forum*, Miskolc, November 19, 2020: Sectional publication of the Faculty of Economics. p. 135-141.

Phan Dai Thich. (2019). Behavior intention toward adopting mobile banking applications: a case study of young customers in Vietnam. In: Veresne, Mariann Somori (ed.). “*Balance sheet and challenges*” XI. International scientific conference. University of Miskolc, Hungary. p. 172-179.

APPENDICES

Appendix A: Survey Questionnaire

Part I: Information of Respondents

Your age:

☐ 20-29 years old ☐ 30-49 years old ☐ 50 years old or older

Your experience in digital transformation and big data projects:

☐ less than 1 years ☐ 1-3 years ☐ 3-5 years ☐ more than 5 years

Which type of organization is best to describe your employer?

☐ Commercial bank

☐ Saving banks

☐ Cooperative banks

☐ Other:

Which area/ department are you working in?

☐ General management

☐ Information Technology

☐ Data management

☐ Risk management

☐ Marketing / Customer services

☐ Others:

Your organization Size (Total Assets):

☐ Less than EUR 50 billion

☐ Between EUR 50-100 billion

☐ More than EUR 100 billion

Part II: Questions about using Big Data Analytics in your organization.

At my organization, Big Data Analytics is used to:

A low score (1) represents that the organization never uses BDA in this manner, a high score (5) represents that the organization frequently uses BDA in this manner. Use of BDA: Consider the general use of BDA in your organization in the last three years.

	1	2	3	4	5
Develop new products or services.	☐	☐	☐	☐	☐
Implement new business processes.	☐	☐	☐	☐	☐
Create new customer relationships.	☐	☐	☐	☐	☐
Change the way of doing business.	☐	☐	☐	☐	☐
Create/ build a new method for risk management / to detect fraud or financial crime prevention.	☐	☐	☐	☐	☐

BDA infrastructure.

Please indicate your level of agreement with the following statement (1 represents strongly disagree, 5 represents strongly agree).

	1	2	3	4	5
The IT infrastructure within our organization is sufficient for BDA.	☐	☐	☐	☐	☐
Databases within our organization are available for our company.	☐	☐	☐	☐	☐
The organization facilitates the use of BDA by analytical professionals.	☐	☐	☐	☐	☐
When implementing BDA, the organization will have sufficient supporting staff.	☐	☐	☐	☐	☐
When implementing BDA, the organization will have sufficient support for training.	☐	☐	☐	☐	☐

The support from top management in using BDA.

Please indicate your level of agreement with the following statement (1 represents strongly disagree, 5 represents strongly agree).

	1	2	3	4	5
Our top management promotes the use of BDA in the organization.	☐	☐	☐	☐	☐
Our top management creates support for BDA initiatives within the organization.	☐	☐	☐	☐	☐
Our top management promotes BDA as a strategic priority within the organization.	☐	☐	☐	☐	☐
Our top management is interested in the news about using BDA.	☐	☐	☐	☐	☐

BDA skills at your organization

Please indicate your level of agreement with the following statement. (1 represents strongly disagree, 5 represents strongly agree).

	1	2	3	4	5
Our organization provides BDA training to our employees.	☐	☐	☐	☐	☐
Our organization hires new employees that already have the BDA skills.	☐	☐	☐	☐	☐
Our BDA staff has the right skills to accomplish their jobs.	☐	☐	☐	☐	☐
Our BDA staff has the proper education to fulfill their jobs.	☐	☐	☐	☐	☐

Data Quality at your organization

Please indicate your level of agreement with the following statement (1 represents strongly disagree, 5 represents strongly agree).

	1	2	3	4	5
We have access to very large, unstructured, or fast-moving data for analysis.	☐	☐	☐	☐	☐
We integrate data from multiple internal sources into a data warehouse for easy access	☐	☐	☐	☐	☐
We integrate external data with internal to facilitate a high-value analysis of our business environment.	☐	☐	☐	☐	☐
We use real-time data.	☐	☐	☐	☐	☐

Part III: Questions about the performance of your organization

Market performance

Compare with your major competitors, how do you rate your organization's performance in the following areas over the past 3 years. *(1 represents strongly disagree, 5 represents strongly agree).*

	1	2	3	4	5
We are exploring new markets more quickly than competitors.	☐	☐	☐	☐	☐
We are introducing new products or services into the market faster than competitors.	☐	☐	☐	☐	☐

Our success rate of new products or services has been higher than competitors.	☐	☐	☐	☐	☐
Our market share has exceeded that of competitors.	☐	☐	☐	☐	☐

Risk management performance

Compare with your major competitors, how do you rate your organization's performance in the following areas over the past 3 years. *(1 represents strongly disagree, 5 represents strongly agree).*

	1	**2**	**3**	**4**	**5**
Our banks perform better in identifying risks compared to competitors.	☐	☐	☐	☐	☐
Our banks perform better in calculating and simulating risk indicators in real-time.	☐	☐	☐	☐	☐
Our banks perform better in controlling and managing reputational risk, operational risk, and compliance.	☐	☐	☐	☐	☐
Our banks perform better in reporting, calculating risk exposure, and performing stress tests in real-time.	☐	☐	☐	☐	☐

Appendix B: The results of the Bootstrapped HTMT values

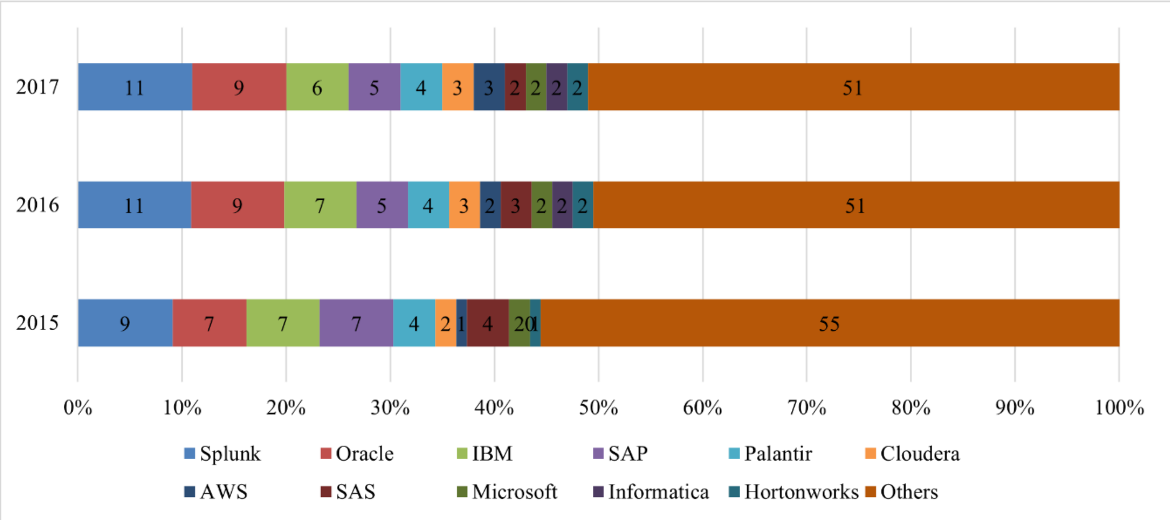
	Original Est.	Bootstrap Mean	Bootstrap SD	T Stat.	2.5% CI	97.5% CI
BDA infrastructure → BDA skills	0.868	0.868	0.033	26.349	0.798	0.927
BDA infrastructure → Top management support	0.736	0.734	0.060	12.165	0.605	0.842
BDA infrastructure → Data quality	0.697	0.698	0.072	9.679	0.548	0.832
BDA infrastructure → USAGE	0.663	0.663	0.058	11.440	0.543	0.769
BDA infrastructure → Size	0.086	0.121	0.064	1.357	0.035	0.278
BDA infrastructure → Market performance	0.532	0.530	0.082	6.525	0.359	0.679
BDA infrastructure → Risk management	0.545	0.542	0.077	7.042	0.379	0.681
BDA skills → Top management support	0.798	0.797	0.047	16.875	0.694	0.878
BDA skills → Data quality	0.664	0.665	0.077	8.601	0.508	0.810
BDA skills → USAGE	0.694	0.692	0.056	12.309	0.573	0.794
BDA skills → Size	0.180	0.188	0.089	2.020	0.048	0.377
BDA skills → Market performance	0.591	0.589	0.078	7.530	0.424	0.731
BDA skills → Risk management	0.470	0.469	0.079	5.936	0.305	0.617
Top management support → Data quality	0.639	0.641	0.079	8.109	0.479	0.789
Top management support → USAGE	0.543	0.541	0.085	6.352	0.361	0.697
Top management support → Size	0.124	0.146	0.040	3.112	0.079	0.234
Top management support → Market performance	0.500	0.497	0.081	6.185	0.326	0.645
Top management support → Risk management	0.379	0.377	0.089	4.282	0.198	0.546
Data quality → USAGE	0.552	0.553	0.088	6.273	0.375	0.718
Data quality → Size	0.091	0.135	0.055	1.644	0.044	0.264
Data quality → Market performance	0.426	0.434	0.088	4.836	0.271	0.608
Data quality → Risk management	0.428	0.433	0.084	5.119	0.276	0.593
USAGE → Size	0.232	0.233	0.085	2.719	0.079	0.405
USAGE → Market performance	0.599	0.597	0.077	7.774	0.434	0.739
USAGE → Risk management	0.512	0.509	0.081	6.292	0.335	0.658
Size → Market performance	0.200	0.202	0.085	2.343	0.058	0.374
Size → Risk management	0.155	0.166	0.085	1.824	0.039	0.345
Market performance → Risk management	0.694	0.693	0.064	10.833	0.556	0.805

Appendix C: Significant factors toward adoption/usage BDA

Authors	Theories	Significant factors toward adoption/usage BDA
Verma & Chaurasia (2019)	TOE & DOI	RA; CX; CP; CPA; TOP; Technology readiness; Organizational data environment
Baig et al. (2021)	TOE & DOI	RA; CX; CP; CPA; TOP; Financial resources; Human expertise & skills; Security & privacy; GOV
Chaurasia & Verma (2020)	TOE & DOI	Perceived strategic value
Maroufkhani, Wan Ismail, et al. (2020)	TOE & DOI	RA; CX; CPA; Risk and insecurity; Trialability; Observability; TOP; External support; GOV; Organizational readiness; CP
Park & Kim (2019)	TOE	Perceived benefits; Technological capability; Financial investment competence; Data quality & integration; TOP; Security and privacy; GOV
Youssef et al. (2022)	TOE	Security concern; EXS; Competitive intensity; Firm size; TOP; Competence of IS staff and staffs' IS knowledge
Shahbaz et al. (2020)	TAM	PU; PEOU; Self-efficacy
Cabrera-Sánchez & Villarejo-Ramos (2020)	UTAUT	Performance expectancy; Social influence; Facilitating condition; Resistance use; Opportunity cost
El-Haddadeh et al. (2021)	TOE	Perceived benefits; CX; TOP
Lutfi et al. (2022)	TOE	RA; TOP; Organizational readiness; GOV
Maroufkhani, Tseng et al. (2020)	TOE	CX; Uncertainty and Insecurity; Trialability; Observability; TOP; Organizational resources; EXS
Yadegaridehkordi et al. (2020)	TOE	CPA; RA; TOP; Organizational resource; Firm size; External pressure; EXS; Security & privacy concern; IT expertise

Note: RA: Relative advantage; CX: Complexity; CPA: Compatibility; CP: Competitive pressure; TOP: Top manager support; GOV: Government policy, regulation, support; EXS: external support

Appendix D: The market share of BDA software provider in 2019



Source: (Wikibon, 2018)