

FACTORS LEADING TO ORGANIZATIONAL ADOPTION OF
DATA SCIENCE AND ADVANCED ANALYTICS TECHNOLOGIES

HÜSEYİN TOLGA DURDU

BOĞAZİÇİ UNIVERSITY

2022

FACTORS LEADING TO ORGANIZATIONAL ADOPTION OF
DATA SCIENCE AND ADVANCED ANALYTICS TECHNOLOGIES

Thesis submitted to the
Institute for Graduate Studies in Social Sciences
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Management

by
Hüseyin Tolga Durdu

Boğaziçi University
2022

DECLARATION OF ORIGINALITY

I, Hüseyin Tolga Durdu, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature

Date

ABSTRACT

Factors Leading to Organizational Adoption of Data Science and Advanced Analytics Technologies

The aim of this research was to propose and empirically test critical factors for Data Science and Advanced Analytics Technologies (DSAAT) adoption based on the prior organizational technology adoption literature. First, drawing on Technology-Organization-Environment (TOE) framework, we developed a survey instrument adapting validated measurement scales from prior literature. Then, we collected data from 140 respondents and empirically tested using PLS-SEM methodology the facilitating influence of technological, organizational, and environmental factors on the adoption intention, and inhibiting influence of lack of data driven decision making and internal expertise on the extent of actual adoption. Our results suggest that environmental factors i.e. mimetic pressure and the availability of external expertise play a significant role in influencing development of adoption intention, while there is no support for the influence of technological factors. The results of respondents from business functions suggest that the extent of data driven decision making has a positive effect on the extent of actual adoption, while for IT function respondents the lack of internal expertise moderates the effect of adoption intention on the extent of actual adoption. Considering the emergence of big data and hence Data Science and Advanced Analytics Technologies allowing extracting value from it, we feel that developing an adoption model and empirically testing it would be a stepping stone for further research on understanding the successful adoption and usage of these technologies in the future.

ÖZET

Veri Bilimi ve İleri Analitik Teknolojilerinin Kullanımına Etki Eden Organizasyonel Faktörler

Bu araştırmanın amacı, Veri Bilimi ve İleri Analiz Teknolojileri (VBİAT) nin organizasyonlar tarafından benimsenmesine etki eden faktörleri önermek ve bu faktörlerin etkisini veriye dayalı olarak test etmektir. Çalışmamız için öncelikle Teknoloji-Organizasyon-Çevre (TOÇ) çerçeve çalışmasından yararlanarak ve mevcut literatürdeki doğrulanmış ölçekleri uyarlayarak bir anket aracı geliştirdik. Daha sonra 140 katılımcıdan veri toplayarak, teknolojik, örgütsel ve çevresel faktörlerin VBİAT benimseme niyetine kolaylaştırıcı etkisi ile veriye dayalı karar verme sürecinin ve iç uzmanlık eksikliğinin bu teknolojilerin gerçek kullanımına etkisini veri dayalı olarak PLS-SEM yöntemi ile test ettik. Sonuçlarımız, çevresel faktörlerin, yani rekabet baskısının ve dış uzmanlığa erişimin, VBİAT benimsenmesinde önemli bir rol oynadığını, ancak teknolojik faktörlerin etkisini gösteren bir kanıt bulamadığımızı göstermektedir. İş birimleri katılımcılarının sonuçlarına göre veri odaklı karar verme eğiliminin olmaması gerçek kullanımı engellerken, iç uzmanlık eksikliği ise teknoloji birimleri katılımcılarının sonuçlarına göre benimsenme niyetinin gerçek kullanıma etkisini modere etmektedir. Büyük verinin kavramının ortaya çıkması ve büyük veriden değer çıkarmaya izin veren Veri Bilimi ve İleri Analitik Teknolojilerinin gelişimi göz önüne alındığında, bu konuda bir benimseme modeli geliştirmenin ve veriye dayalı olarak test etmenin, bu teknolojilerin başarılı bir şekilde benimsenmesini ve kullanımını anlamak için bir daha sonra yapılacak araştırmalar için yol gösterici olacağını düşünüyoruz.

CURRICULUM VITAE

NAME: Hüseyin Tolga Durdu

DEGREES AWARDED

PhD in Management, 2022, Boğaziçi University

MA in Business Information Systems, 2007, Boğaziçi University

BS in Industrial Engineering, 1996, Boğaziçi University

AREAS OF SPECIAL INTEREST

Information systems, management information systems, business analytics, data science, machine learning, big data, and technology management.

PROFESSIONAL EXPERIENCE

Instructor, Financial Engineering MS Program, Ozyegin University, 2020 - present

Executive Advisor to CEO, Data Management Office, Fibabanka, 2019 - present

Co-founder and Managing Partner, Accelera Yazılım, 2014 - present

Information Management, Senior Vice President, ING Bank, 2002 - 20014

Financial Planning, Assistant Manager, Borusan Holding, 2001

Investment Banking, Associate Manager, Ventura Partners, 2000-2001

Investment Banking, Associate Manager, Bank Kapital, 1996-2000

AWARDS AND HONORS

Highest Honors List, Boğaziçi University, 2007

Ranked 1st in Business Information Systems Class of 2007 (GPA:4.00)

Highest Honors List, Boğaziçi University, 1996

Ranked 2nd in Industrial Engineering Class of 1996 (GPA: 3.77)

SERVICES TO PROFESSION

Founding President of Veri Yönetimi Derneği, Data Management International Turkey Chapter, 2012-2018.

Instructor for Data Management and Data Analytics courses at Turkish Banking Association, 2014-2019.

Book Reviewer in Data Management Domain for Elsevier Publishing.

PUBLICATIONS

Conference Proceedings

Arsal, R.E, Durdu, H.T., and Tongarlak, M.H. (2022). Organizational Enablers of Data-Driven Digital Transformation: A Case Study from Banking Industry. IEEE Technology & Engineering Management Conference (TEMSCON), 2022.

Güntay, L., Bozan, E., Tığrak, Ü., Durdu, T., Özkahya, E.C., (2022). An Explainable Credit Scoring Framework: A Use Case of Addressing Challenges in Applied Machine Learning. IEEE Technology & Engineering Management Conference (TEMSCON), 2022.

ACKNOWLEDGEMENTS

Completion of this thesis would not have been possible without the guidance and support I received from my current thesis advisor Assistant Professor Dr. Hayri Tongarlak and my former thesis advisor Prof. Dr. Stefan Koch.

I would like to also express my deepest gratitude to Prof. Dr. Ayşegül Toker for giving me the opportunity to join the prestigious Boğaziçi University PhD program in Management. Looking back, I still remember her cautioning me about the difficulties of this endeavour while having other professional responsibilities in the industry. Yet, she also provided invaluable guidance and endless encouragement along the way.

I could not also have undertaken this journey without the camaraderie of Ergün Aarsal whom I spent countless hours working side by side. Alongside him, I had the opportunity to hone my skills and produce quality outcomes. I would also like to acknowledge Prof. Dr. Emin Anarım for encouraging me to start this endeavour in the first place.

Lastly, I could not have undertaken this journey without the understanding and support my wife and daughter extended to me. Their support kept my spirit and motivations high enabling me to open another chapter in my life.

This dissertation is dedicated to

my beloved wife, daughter, and parents

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION	1
1.1 Overview.....	1
1.2 Research objectives	8
1.3 Proposed model	10
1.4 Research design	10
1.5 Contributions	12
1.6 Outline of dissertation.....	13
CHAPTER 2: LITERATURE REVIEW	14
2.1 Data science and advanced analytics (DSAA) literature in IS	14
2.2 Technology adoption literature.....	33
2.3 Organizational technology adoption literature	36
CHAPTER 3: THEORETICAL DEVELOPMENT	43
3.1 Technology, organization, environment framework	43
3.2 Technology characteristics and relevant factors.....	44
CHAPTER 4: METHODOLOGY	51
4.1 Research design	51
4.2 Sample design.....	53
4.3 Research instrument.....	58
4.4 Data analysis	71

CHAPTER 5: RESULTS	75
5.1 Data examination and preprocessing	76
5.2 Sample characteristics	77
5.3 Measurement model evaluation	80
5.4 Structural model evaluation	93
CHAPTER 6: CONCLUSION	103
6.1 Summary of findings	103
6.2 Research contributions and managerial implications	104
6.3 Limitations and future research directions	105
APPENDIX A: ETHICS COMMITTEE APPROVAL	106
APPENDIX B: SURVEY INSTRUMENT INTRODUCTION	107
APPENDIX C: SURVEY INSTRUMENT	109
APPENDIX D: DESCRIPTIVE STATISTICS OF SURVEY RESPONSES	114
APPENDIX E: MEASUREMENT MODEL OUTER LOADINGS	116
APPENDIX F: CROSS LOADINGS	118
REFERENCES	120

LIST OF TABLES

Table 1. Success Factors for Effective Use of Big Data	23
Table 2. Potential Constructs of BDA Research	30
Table 3. Theoretical Models Used in Innovation Adoption Literature	34
Table 4. Organizational Factors	42
Table 5. Technological Factors	42
Table 6. Environmental Factors	42
Table 7. Organizational Adoption Constructs and Definitions	48
Table 8. Initial List of Survey Participants	55
Table 9. Sample Size Recommendation for PLS-SEM	57
Table 10. List of Constructs and Measurement	61
Table 11. Breakdown of Survey Respondents by Industry	77
Table 12. Breakdown of Survey Respondents by Managerial Responsibility	78
Table 13. Share of DSAA in Overall Responsibilities of Respondent	79
Table 14. Duration Since Initial Trial Use of DSAAT	79
Table 15. Descriptive Statistics about Survey Respondents	80
Table 16. Algorithmic Options and Parameter Settings for Measurement Model....	82
Table 17. Construct Reliability and Convergent Validity	85
Table 18. Discriminant Validity - Fornell-Larcker Criterion.....	88
Table 19. Heterotrait-Monotrait Ratio (HTMT)	90
Table 20. Measurement Model Evaluation Summary	91
Table 21. Measurement Model Collinearity Statistics (VIF).....	92
Table 22. Higher-Order Constructs Indicator Weights	92
Table 23. Parameter Settings for Bootstrapping	95

Table 24. Structural Model Collinearity Statistics (VIF).....	97
Table 25. Structural Model Path Coefficients.....	98
Table 26. Explained Variance of Endogenous Constructs.....	99
Table 27. Coefficient of Determination (R^2 Value)	100
Table 28. Multi-Group Analysis Structural Model Path Coefficients	101
Table 29. Effect Size f^2	102

LIST OF FIGURES

Figure 1. Nexus of forces	2
Figure 2. High level conceptual model	10
Figure 3. Big data “insight extraction” process	16
Figure 4. Big data research agenda for IS	20
Figure 5. Conceptual classification of big data and analytics challenges	23
Figure 6. Big data analytics CSFs classification	25
Figure 7. Integrated model of big data value realization	26
Figure 8. Resource supporting BDA capability	27
Figure 9. Big data analytics resource management framework	28
Figure 10. Big data analytics value creation framework.....	29
Figure 11. Factors effecting big data adoption intent	32
Figure 12. Diffusion process	38
Figure 13. Diffusion of innovations (DOI) theory at firm level	39
Figure 14. Technology, organization, environment (TOE) framework	41
Figure 15. Conceptual model	49
Figure 16. Research plan.....	51
Figure 17. Steps of the data analysis	75
Figure 18. Measurement model specification	83
Figure 19. Structural mode.....	98

ABBREVIATIONS

ACTADP: Actual Adoption

ADPINT: Adoption Intention

AVE: Average Variance Extracted

BDA: Big Data Analytics

BI: Business Intelligence

CB-SEM: Covariance Based Structured Equation Modeling

CI: Confidence Interval

CMPLX: Complexity

CMPTBL: Compatibility

CRM: Customer Relationship Management

CSTSAV: Cost Savings

DDD: Data Driven Decision Making

DOI: Diffusion of Innovation

DSAA: Data Science and Advanced Analytics

DSAAT: Data Science and Advanced Analytics Technologies

DSS: Decision Support Systems

ERP: Enterprise Resources Planning

EXTEXP: External Expertise

f²: Effect Size

GoF: Goodness-of-fit

HTMT: Heterotrait-Monotrait Ratio

IS: Information Systems

IT: Information Technology

LCKINTEXP: Lack of Internal Expertise

MANOVA: Multivariate Analysis of Variance

MIMPRS: Mimetic Pressure

MIS: Management Information Systems

OLS: Ordinary Least Squares

ORGSCLK: Organization Slack

PCA: Principal Component Analysis

PLS-SEM: Partial Least Squares Structured Equation Modeling

RELADV: Relative Advantage

R²: R Squared

SC: Supply Chain

TMBLF: Top Management Belief

TMPRT: Top Management Participation

TOE: Technology, Organization, Environment

VIF: Variance Inflation Factor

CHAPTER 1

INTRODUCTION

1.1 Overview

In this section, drawing on the practitioner literature, we first outline the developments that led to the emergence of the phenomenon called big data, and related concepts data science and advanced analytics. Then, we turn our attention to the underlying developments such as the advent of web scale companies, their use of data to create business value, and role of open source software in doing so. In light of these developments, we contemplate about the unique characteristics of big data, data science, and advanced analytics, while proposing potential difficulties organizations in traditional industries might be facing in adopting these transformative concepts and related technologies. Lastly, we close this section by providing common definitions of big data, data science, and advanced analytics.

1.1.1 Emergence of data science and advanced analytics

The last couple of decades brought about accelerated pace of technological advancement in information technologies. First, internet became ubiquitous paving way to other developments. This was followed by the advent of social media, mobile devices, and always on connectivity. The widespread availability of broadband internet also made cloud computing more viable. Gartner (2012) coined the term nexus of forces (Figure 1) to describe the convergence and mutual reinforcement of social, mobility, cloud and information patterns that drive new business models.

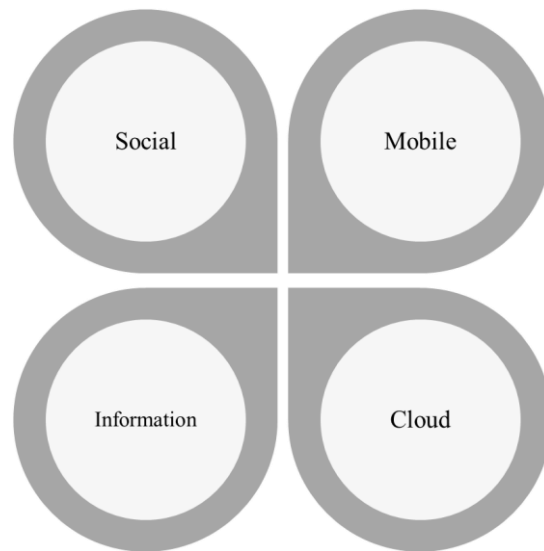


Figure 1. Nexus of forces
Source: Gartner, 2012

The increasing communication channels and variety of data produced by them gave rise to a new phenomenon called big data (Gartner, 2012).

Many in industry concluded that ability to capitalize on this convergence was depended on developing capabilities to use data in an intelligent way. The Economist (2010) claimed that data became the new raw material of business world, at least as valuable as capital and labour. On the other hand, McKinsey (2011), highlighting the difficult of developing the capabilities to manage and analyze big data, declared that it will be the next frontier for innovation, competition, and productivity. McKinsey (2011) further claimed that big data has the potential to produce 250 billion Euros potential value to Europe's public sector administration and more than 600 billion US Dollars annual surplus to consumers globally from using personal location data.

In October 2012, Harvard Business Review published a special issue bringing the term big data to the front and center. In this issue, McAfee et al. (2012) proclaimed that organizations are collecting so much data that they do not even know what to do with it. For turning all this data into competitive advantage, organizations

of all sizes would need new capabilities and a new management style (McAfee et al., 2012). In their much cited article, published in this Harvard Business Review special issue, McKinsey consultants Barton and Court (2012) argue that writing-off big data as another fad would be a costly mistake. They further stressed that data is useless unless used wisely by building analytics models focused on improving business results and transforming organization capabilities to implement analytical results (Barton & Court, 2012).

The interest in big data, advanced analytics and related concepts such as data science, machine learning and artificial intelligence has never waned ever since. Thomas Davenport argued that 2010s marked the start of a new era he dubbed Analytics 3.0, in which every sizable firm in every industry would be capitalizing on data (Davenport, 2013). Since then, the usage of the term data science became increasingly more common than the term big data. In a similar vein, big data became the ingredient for the emerging discipline data science which is regarded as the study of the extraction of knowledge from data (Dhar, 2013).

Fortune (2015) announced that McKinsey acquired a boutique data science consultancy firm Quantum Black, the company helping make sense of Formula 1 data. This was followed by a flurry of acquisitions by all the leading consulting companies such as McKinsey, Boston Consulting Group, and Bain in an effort to quickly shore up their ranks having capabilities in data science and advanced analytics (Rovit, 2020). This was a clear sign that organizations in various industries started to take data and advanced analytics more seriously, creating a shortage of talent foreseen earlier by Davenport and Patil (2012). In an article marking the level of interest in using data to create value, The Economist (2017) claimed that the data is becoming the new oil, replacing the role of oil as the most valuable resource.

Following the initial hype around amassing big data, businesses turned their attention to making sense of the data they possess to derive value using advanced analytics, data science, machine learning and artificial intelligence (AI). Hence, almost all the major consulting companies published reports on topics such as competing in a data-driven world (McKinsey, 2016), developing data science capabilities (Booz Allen, 2015; Booz Allen, 2017), and building a data-powered enterprise (Capgemini, 2019).

More recently, Davenport (2018) ushered in a new era, he called, Analytics 4.0 in which organizations embracing data science, machine learning, agile methodologies, cloud computing, and open source technologies gaining significant competitive advantages. In their latest report, titled “The Data-Driven Enterprise of 2025”, McKinsey (2022) outlined seven characteristics defining the new data-driven enterprise, and foresee that in the leading organizations data will be embedded in every decision, interaction, and process (McKinsey, 2022).

Even though the amount of data collected by companies keep growing exponentially, for organizations other than some innovators and early adopters, widespread adoption of big data, and data science and advanced analytics tools and techniques is taking longer than originally expected (McKinsey, 2016; McKinsey, 2022). Especially organizations in traditional industries are struggling to create strategies and mobilize resources to start on big data, and data science and advanced analytics investments (Iansiti & Nadella, 2022; McKinsey, 2016).

1.1.2 Characteristics of data science and advanced analytics

A closer look into the factors behind the emergence of big data, data science and advanced analytics reveal valuable insights. Along with the developments in multiple fronts as summarized by Gartner (2012) as the concept of nexus of forces new business models emerged in 2000s.

This led to the rise of companies like Google, Amazon, Facebook and Twitter. For these web scale consumer companies data was their lifeblood, and data science created a new data economy in which these leading companies invest in building infrastructure, tool, and talent. (Cao, 2017; Davenport, 2013) The sheer scale of these companies led them to force the boundaries of existing commercial technology platforms. This was both limiting their capabilities and becoming more and more costly to maintain and scale.

In order to overcome the limitations of existing technology platforms, they started developing and later open sourcing new technologies like Map Reduce, Hadoop, NoSQL databases (Davenport, 2013). First, they used these technologies for their operational processes. But since data was central to their business models later they started using these new technologies to extract meaning from data as well. Combining these new technology platforms with existing statistics, data mining, machine learning and operations research techniques companies started to use analytics at an unprecedented scale to create competitive advantage. (Davenport, 2018)

While innovator web scale consumer companies and some other early adopter pioneers are gaining competitive advantage using big data and advanced analytics, companies in more mainstream industries such as finance, telecom, retail, and manufacturing are struggling to benefit from these potentially disruptive technologies

(Iansiti & Nadella, 2022; McKinsey, 2016). Davenport (2018) expects that in the next decade the focus for most firms will shift from faster transactions and information delivery to finding ways to better analyse, monetize, and capitalize on all the available information. Today, for the regular organization data science and advanced analytics is comparable to what the World Wide Web was in the 1990s. Because of the fast pace of change organizations feels obliged to find ways to make use of data assets. (Davenport, 2018)

Compared to other enterprise technologies such as enterprise resources planning (ERP), customer relationship management (CRM), and even business intelligence (BI) and data warehouses, these new data science and advanced analytics technologies have some interesting contrasts making it difficult for traditional companies with limited means to start benefiting from them (Iansiti & Nadella, 2022).

First of all, these technologies are the result of user organizations' innovations. The lead user organizations developing these technologies have been focusing on use value rather than sales value of the resulting IT assets. In order to increase pace of development and hence the use value they expect to reap, they resorted to open innovation and started open sourcing these technologies. This resulted in a fragmented technology stack in which most of the software components were back-end server side stuff developed by developers for other developers to integrate into their systems (McAfee, Brynjolfsson, Davenport, Patil, & Barton, 2012).

This is in contrast to other enterprise technologies where there are a few large vendors with mature and well supported software platforms. Unlike other enterprise technologies, for these open-source tools and technologies upfront acquisition cost is not an entry barrier (McAfee et al., 2012). Yet, it is hard to find people with

knowledge and skills necessary to implement and maintain these solutions (Davenport & Patil, 2012; McAfee et al., 2012; McKinsey, 2011). According to the TDWI's Predictive Analytics Survey, companies cited getting started with the right project and lack of skills of staff along with complexity related issues as the top challenges. Cost again remained at the bottom of the list (Halper & Krishnan, 2013).

The last but maybe one of the most important differences is while enterprise technology implementation projects require engineering mindset, big data and especially data science and advanced analytics projects require an experimentation culture in which testing and learning is the norm. In order to evolve from technology focus to analytics focus organizations need to institutionalize this test and learn culture. Booz Allen (2015) suggests that willingness to fail and embracing experimentation is at the core of data science, and in addition to a willingness to fail organizations should strive to learn from these failures quickly.

On top of these difficulties, many companies are having difficulty finding the right use cases to start with. So, despite all the hype around big data, data science, and advanced analytics and low entry barriers, real adoption among the majority is lagging behind the expectations. In order to help alleviate this problem, McKinsey (2016) developed a use case driven methodology spanning more than 300 use cases across industries to help client organizations select the most promising problem areas to tackle first.

1.1.3 Defining big data, data science and advanced analytics

Gartner defines big data as high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making. On the other hand, Gartner defines advanced

analytics businesses use sophisticated tools and techniques to examine data for insights, make predictions, or generate recommendations. Another research company Forrester focuses on the capabilities in defining the big data as the frontier of a firm's ability to store, process, and access (SPA) all the data it needs to operate effectively, make decisions, reduce risks, and serve customers.

Edison Data Science Framework (2020) defines data science as the extraction of actionable knowledge from raw data. The entire data science life cycle incorporates principles, techniques, and methods from many disciplines including data engineering, advanced analytics, statistics, machine learning, visualization, and in the context of big data, also big data engineering (Cuadrado-Gallego & Demchenko, 2020).

As an indication of their ubiquity, big data and data science even found their way to Oxford and Merriam-Webster English dictionaries. According to Oxford English Dictionary, big data is defined as “data sets that are too large and complex to manipulate or interrogate with standard methods or tools”, while data science is defined as “the use of scientific methods to obtain useful information from computer data, especially large amounts of data”.

1.2 Research objectives

There is a surge of interest in big data and accompanying data science and advanced analytics technologies both in business and academia. Although the concepts and methods are not totally new having roots in decision support systems, business intelligence, data mining and machine learning, the scale of data sets that needs to be processed and the cross-functional nature of the methods used makes this new technology domain different. Considering the highly fragmented and open-source

ecosystem developing these new technologies, it is in stark contrast to other enterprise systems such as enterprise resources planning (ERP), customer relationship management (CRM), and business intelligence (BI) dominated by commercial IT vendors. Because of this, it might be expected that both the initial adoption and later infusion processes of these new technologies to be different as well.

Although we are still in the early stages of widespread adoption, because of the initial success of the innovator organizations and the heightened sense of urgency among the other organizations about the importance of these new technologies there is need for research to investigate which factors facilitates or inhibits the adoption of Data Science and Advanced Analytics Technologies. As the adoption of these technologies become more common this research should be complemented by further research in the later infusion stages.

The aim of this research is to first find out critical factors for the adoption of Data Science and Advanced Analytics Technologies (DSAAT) based on the existing organizational technology adoption literature. After the factors most relevant for DSAAT identified for technological, organizational, and environmental domains, relationship of these factors to the adoption intention and actual adoption empirically tested. Due to the level of publicity about the subject, it can be expected that majority of the organizations already developed an intent. So, we focused also on actual adoption in the form of production implementations rather than sole intention to adopt. Actual adoption was defined as the use of these technologies in a production environment operationalizing data driven decision making.

1.3 Proposed model

Based on Technology Organization Environment (TOE) Framework and examination of prior literature for potential factors influencing organizational adoption of technology, we propose the high level conceptual model in Figure 2. In our proposed model, technological and environmental factors directly affect adoption intention, while some organizational factors play facilitating role at the intention stage others play an inhibiting role at the actual adoption stage.

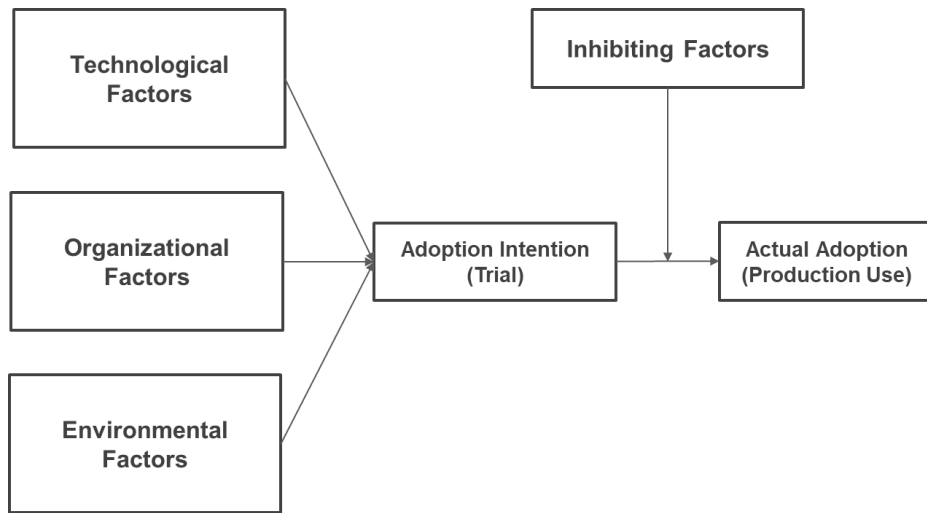


Figure 2. High level conceptual model

1.4 Research design

After conducting a thorough literature review of academic and practitioner journals, we created a list of potential factors to be tested empirically. Then, we formed a panel group of experts from academy and practitioners from industry whose organizations already had adopted or considered to adopt Data Science and Advanced Analytics Technologies in the past. We finalized our proposed model and associated hypotheses after having in-depth discussions with these experts and practitioners. We deemed this stage crucial to our research since very little research

has been conducted to examine Data Science and Advanced Analytics Technology (DSAAT) adoption specifically.

Once we identified all the relevant and critical factors for data science and advanced analytics technology adoption, we applied a quantitative field survey research methodology. Because we were both interested in the trial and production use stages of adoption, we designed a survey to measure both the adoption intent and the extent of actual adoption. This means that our sample consisted of IT and business professionals from organizations that might not have fully adopted these technologies, but had been considering the adoption of them. We used a purposive sampling strategy and in order to increase the sample size applied several rounds of snowball sampling. As part of the survey instrument, we asked questions to make sure that the participants satisfy our sample frame requirements. We used Technology Organization Environment (TOE) framework commonly used organization adoption studies and adapted validated measurement scales from prior literature. Finally, we analyzed the data to test the effect of hypothesized factors on both adoption intent and extent of actual adoption using Partial Least Squares Structured Equation Modelling (PLS-SEM) technique.

To summarize a step-by-step flow-chart:

- Completed literature review to identify constructs of the study.
- Formed a panel group from IT and business professionals to triangulate our literature review.
- Designed and finalized the survey instrument.
- Pre-tested the survey instrument with experts to ensure validity of the instrument.

- Created a list of potential industries and organizations in those industries.
- Created a list of potential contacts working in target organizations with the help of our professional contacts and publicly available information.
- Collected data from the initial sample of contacts and employed snowball sampling to increase the sample size.
- Analysed data using PLS-SEM technique to test the hypotheses in our model.
 - First analysed the measurement model to make sure that we have reliable and valid measurements.
 - Then analysed structural model where we examine the effects of technological, organization, and environmental factors on adoption intention as well as the effect of inhibiting factors on the relationship between adopt intention and extent of actual adoption.

1.5 Contributions

We are hoping that this study will strongly contribute to both academy and practice. First, to the best of our knowledge, there is very little work on Data Science and Advanced Analytics adoption at the organizational level. Considering the emergence of big data and hence Data Science and Advanced Analytics Technologies allowing extracting value from big data, we feel that developing an adoption model and empirically testing it would be a steppingstone for further research on understanding the successful adoption and usage of these technologies in the future.

Our analysis at both intention and adoption stages should help organizations avoid the pitfalls of ignoring critical organizational factors before jumping on the big data bandwagon. Hence, our findings will hopefully help organizations to successfully adopt these potentially transformative technologies rather than getting

stuck at the intent level and use the systems effectively to gain competitive advantage within their industries.

1.6 Outline of dissertation

This chapter summarized the context around this research study, and introduced the research objectives, research design, proposed model, and expected contributions.

Chapter 2 provides a review of the relevant literature covering MIS research streams related to data science and advanced analytics, and technology adoption literature.

Based on findings from literature review, in Chapter 3 we developed the proposed research model and explain the hypothesized relationships. Chapter 4 describes in detail the methodologies used for research design, sample design, development of the research instrument, and data analysis. Chapter 5 explains the application of PLS-SEM methodology and discusses the results of the study. Finally, in Chapter 6 we conclude the dissertation by providing a summary of contributions, limitations, and future research directions.

CHAPTER 2

LITERATURE REVIEW

2.1 Data science and advanced analytics (DSAA) literature in IS

In this section, we start the review of the literature by looking into various research streams in MIS literature covering big data, advanced analytics, and data science.

We searched “Web of Science” database by “Topic” with keywords “big data analytics” and “data science” separately, and along with “adoption” keywords. We first grouped the results by academic disciplines and then further examined the abstracts of the results to filter out the literature that only covers the technical aspects of the topic, such as academic papers written in computer science and engineering domains.

We then group the results by their theme and ended up with the following categories:

- Conceptual (e.g., Chen, Chiang, & Storey 2012; Corea, 2016; Donoho, 2017; Gandomi & Haider, 2015; Provost & Fawcett, 2013;)
- Literature Review and Research Directions (e.g., Abbasi, Sarker & Chiang, 2016; Agrawal et al., 2014; Cao, 2017; De Mauro, Greco, & Grimaldi, 2015; Phillips-Wren, Iyer, Kulkarni, & Ariyachandra, 2015)
- Critical Success Factors (e.g. Al-Sai, Abdullah, & Husin, 2020; Ajah et al., 2019; Baesens, Bapna, Marsden, Vanthienen, & Zhao, 2016; Sivarajah, Kamal, Irani, & Weerakkody, 2017; Surbakti, Wang, Indulska, & Sadiq, 2020)

- Capabilities and Value (e.g. Brynjolfsson, Hitt, & Kim, 2011; Grover, Chiang, Liang, & Zhang, 2018; Gupta & George, 2016; Günther, Mehrizi, Huysman, & Feldberg, 2017; Mikalef, Pappas, Krogstie, & Pavlou, 2020)
- Adoption (e.g. Agrawal, 2015; Inamdar et al., 2020; Lai, Sun, & Ren, 2018; Maroufkhani, Ismail, & Ghobakhloo, 2020; Sun, Cegielski, Jia, & Hall, 2018; Wimmer & Aasheim, 2019)

In the remaining part of this section, we summarized some of the prominent themes from each of these research streams.

2.1.1 Conceptual research in DSAA

Articles in the conceptual research stream defines main concepts such as big data, data science, and advanced analytics. Chen et al. (2012) describes big data and big data analytics as the data sets and technologies that are so large and complex that they require advanced and specialized storage, management, analysis and visualization technologies. The roots of big data analytics research can be traced back to research studies on decision support systems (DSS), business intelligence (BI) and business analytics (BA). Big data analytics refers to the advanced business intelligence & analytics (BI&A) technologies that are grounded mostly in data mining, machine learning and advanced statistical analysis (Chen et al. 2012).

Provost and Fawcett (2013) define a related concept data science as principles, processes and techniques for understanding phenomena via the automated analysis of data. Considering the ubiquity of data and information in today's world data science enables the organizations to base decisions on data. Data-driven decision making (DDD) can be defined as the practice of using results of data

analysis as the basis of decisions rather than relying on pure intuition. (Provost & Fawcett, 2013).

Gandomi and Haider (2015) posits that big data is useless in a vacuum and its real value is realized when it enables data based decision making. They break down the process of extracting valuable information from high volumes of fast moving data into 5 stages shown in Figure 3. Overall big data management process consists of two sub-processes, data management and analytics. Data management in turn consists of processes and supporting tools to prepare the data for analysis. Analytics, on the other hand, refers to techniques used to analyze and extract insights from big data (Gandomi & Haider, 2015)

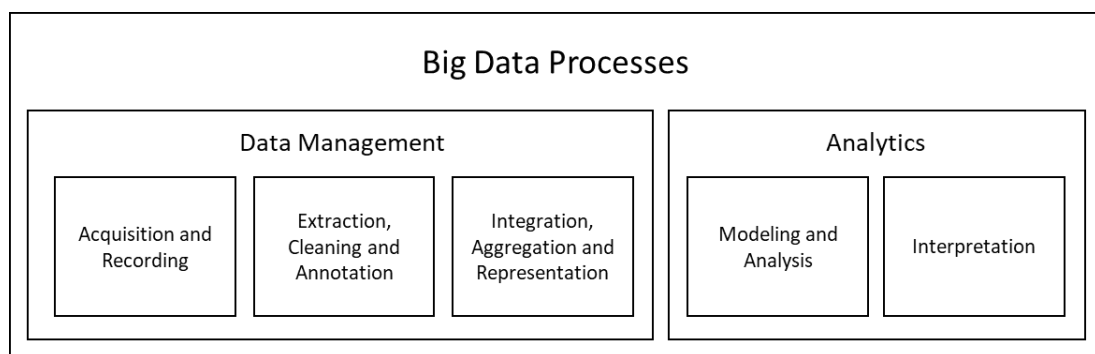


Figure 3. Big data “insight extraction” process
Source: Gandomi and Haider (2015)

Donoho (2017) traces the roots of data science back to John Tukey’s famous “The Future of Data Analysis” article. He relates various disciplines such as data science, big data, and statistics, and compare their philosophical differences. Next, he moves on to define Greater Data Science (GDS) as an all-encompassing concept comprised of activities classified into 6 divisions.

- Data Gathering, Preparation, and Exploration
- Data Representation and Transformation

- Computing with Data
- Data Modeling
- Data Visualization and Presentation
- Science about Data Science

GDS defines data science as the science of learning from data. In essence, it scientifically studies the methods involved in the whole lifecycle of analysis and processing of data, proposes technology solutions to improve these methods in an evidence-based manner. It can be expected that, this science will grow both in scope and impact in the future as both scientific data and data about science become ubiquitous (Donoho, 2017).

Corea (2016) observes that due to the hype around them phrases such as big data and data science are misused, and by convention many people often use them interchangeably. Data science can be differentiated from earlier concepts by its emphasis on innovative new approaches and technologies to extract valuable insights even from low-value big data that fails to fit conventional database systems (Corea, 2016). There are also other common misconceptions that needs to be addressed to realize the full potential of big data and data science such as (Corea, 2016): “More data means higher accuracy”, “If you want to do big data, you have to start big”, “Data equals objectivity”, “Your data will reveal you all the truth”.

Although the concept of artificial intelligence predates data science, the latter can also be thought of representing the infancy stage of artificial intelligence (AI). As the field that studies the methods to create computers displaying intelligent behaviour normally associated by humans, AI can be thought to use big data as the fuel that powers it (Corea, 2016).

2.1.2 Literature review research in DSAA

Articles in the literature review and research directions research stream gives an overview of existing research and provides suggestions for future research directions. Agarwal and Dhar (2014) posits that the tsunami of information caused by the availability of big data pose new research questions. They invite the other scholars to ponder about the following five questions to better frame the enormous research challenges:

- Are the concepts of big data, analytics, and data science something new?
- What are the strengths that the IS community brings to research in these concepts?
- What are relevant and interesting research questions, and new areas of inquiry?
- How can we mend the philosophical differences of predictive vs explanatory approaches?
- How should editors and reviewers evaluate research in this domain?

The foundation of data science and business analytics have been around for a long time in different disguises, but the recent developments about the availability of big data create significant new questions and opportunities for research. The IS discipline is well positioned to contribute to this area with its long tradition of conducting research at the intersection of technology, data, business, and society (Agarwal & Dhar, 2014).

De Mauro et al. (2015) considers big data as a trending buzzword clouded by a lot of conceptual vagueness in both academia and the industry. The term is used to describe a variety of concepts from the data processing technologies on one hand to

the transformational impact on business and society on the other. This lack of formal definition led researchers to diverge into multiple research directions.

An analysis of frequent keywords reveals four general research themes (i.e., Information, Technology, Methods, Impact). Consequently, De Mauro et al. (2015) based on their literature review proposes the following definition of big data to further a joint research agenda: “big data represents the information assets characterized by such a High Volume, Velocity and Variety to require specific Technology and Analytical Methods for its transformation into Value.” (De Mauro et al., 2015)

Motivated by the recent interest in big data from the IS discipline, Abbasi et al. (2016) summarized varying perspectives on promising big data research topics and the challenges posed by big data. They propose that there is a need for an inclusive research agenda focusing on the intersection of characteristics of big data, necessary resources, and capabilities. As can be seen in Figure 4, they argue that behavioural, design science, and economics of IS research streams have much to contribute by the following:

- Critically engaging with fundamental issues, such as epistemology, methodology, ethics, and the design of novel artifacts
- Rethinking decision models of small data era and adapting them for use in the era of big data
- Assessing economic and societal outcomes of big data via multi-paradigmatic research initiatives

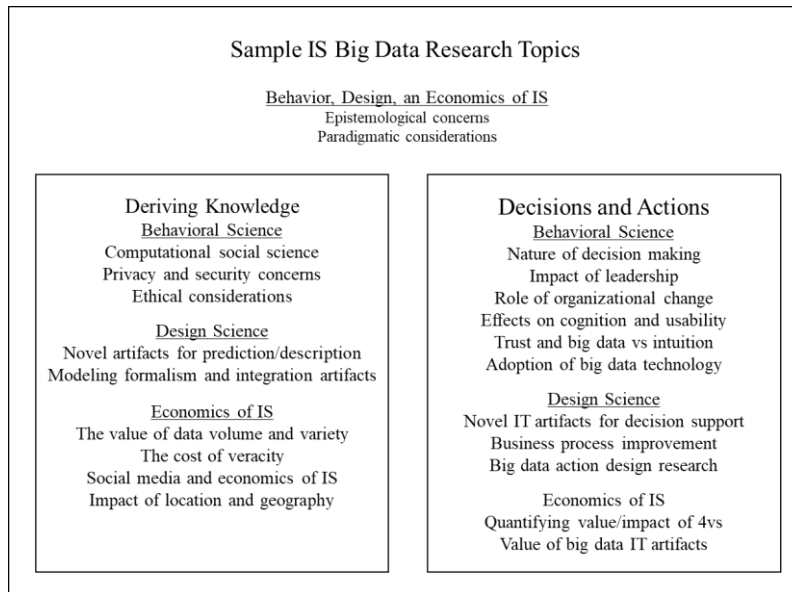


Figure 4. Big data research agenda for IS
 Source: Abbasi et al., 2016

Cao (2017) defines data science as a new trans-disciplinary field that builds on and integrates various relevant disciplines. Prominent ones among these disciplines are statistics, informatics, computing, communication, management and sociology.

Cao (2017) further argues that research challenges and opportunities mainly stem from the inadequacies of existing methodologies, theories, and systems and can be categorized into six major categories: i.e., data/business understanding challenges, mathematical and statistical foundation challenges, data/knowledge engineering and analytics challenges, quality and social issues challenges, data value/impact and utility challenges, data-to-decision and action-taking challenges. Addressing these challenges requires a multidisciplinary effort and the development of an approach from complex system perspective. If this is achieved there is unique opportunities for breakthrough research and cutting-edge technological innovation (Cao, 2017).

Phillips-Wren et al. (2015) triangulated the results of a thorough academic literature review with a review of practitioner literature and interviews with

practitioners. They claim that further academic research in the areas they propose is likely to increase the relevance of academic research to practice.

Analysis area stand out as having the most interesting unanswered questions due to the need for massive scale and the difficulty of using the results for decision support. On the usage side, there are plethora of behavioural research questions around topics such as the ideal composition of data science teams, user engagement for big data capabilities, and integration of big data analytics in the processes (Phillips-Wren et al., 2015).

2.1.3 Critical success factors research in DSAA

Articles in the critical success factors research stream highlights the challenges posed by big data and identifies critical success factors for applying data science and advanced analytics to address these challenges. Baesens et al. (2016) argue that big data and data science is transforming the world and this emerging landscape creates many research opportunities. They propose various critical success factors such as high quality data, trust of decision maker in both the data and the methods, skill set of the business analysts, ability to integrate systems and integration environments, breaking down the silos imposed by business boundaries,

Ajah et al. (2019) reviewed the recently developing big data and data science literature and highlighted a set of critical success factors commonly proposed for achieving success with big data analytics implementations. These factors include having an expertise both at the business and analytics teams, knowledgeable management teams, ability to architect complex systems. There are also pitfalls to avoid such as a false sense of security, inclination to store data without apparent business benefit, straining IT resources, inability to navigate security, privacy and compliance challenges (Ajah et al., 2019).

Based on review of the extant literature Surbakti et al. (2020) posits that although organizations' interest in big data and analytics is growing, and there is increasing evidence that investments in big data solutions may fail to deliver intended value. Hence, they set out to suggest a rigorous academic guidance about the critical success factors enabling effective use of big data. They categorized 41 factors they identified under 7 themes, i.e., data quality, data privacy, security and governance, perceived organizational benefit, process management, people aspects, systems, tools, and technologies, and organizational aspects. Table 1 summarizes these 41 critical success factors by 7 main research themes.

Sivarajah et al. (2016) conducted a systematic literature review analysis to identify big data and big data analytics challenges confronted by the organizations, and methods proposed to overcome those challenges. Considering the potential impact of big data and advanced analytics, they argue that this understanding is crucial for making robust investment decisions. They categorized these challenges under three categories as seen at Figure 5: data challenges, process challenges, and management challenges.

Table 1. Success Factors for Effective Use of Big Data

Research Theme	Factors
Organizational Aspects	organizational cultural competence, talent management, change management program, strategic alignment, project management, performance management, organizational structure and size, interdepartmental collaboration, communication, top management support, environmental effect, clear goals, focus on innovation
Systems, Tools, and Technologies	system quality, IT infrastructure, vendor support
People Aspects	people's knowledge and skills, trust, champions, employee engagement, user participation, individual characteristic
Data Privacy, Security, and Governance	data privacy and security, data governance
Data Quality	data quality and information quality, data completeness, data currency, data access, data relevance, data accuracy, data consistency
Process Management	process orientation, IT business process integration, data management
Perceived Organizational Benefit	perceived value, perceived risk, perceived usefulness, perceived ease of use, intention to use, perceived observability, cost–benefit analysis

Source: Surbakti et al., 2020

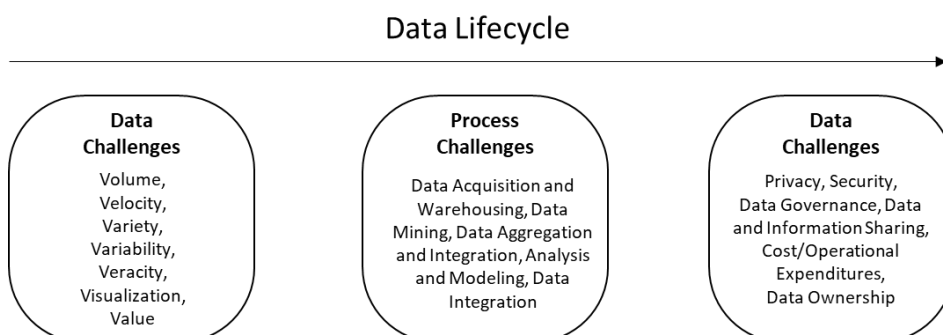


Figure 5. Conceptual classification of big data and analytics challenges

Source: Sivarajah et al., 2016

In their systematic literature review for identifying critical success factors for big data and big data analytics, Al-Sai et al. (2020) created a classification framework to provide an in-dept understanding of factors contribution to successful implementation of big data analytics. In doing so, they aimed to summarize recent research into critical factors for successful use of big data and guide further research in those areas deemed important in existing literature. In their framework, critical factors are classified into 5 categories, namely organization, people, technology, data management, and governance. Figure 6 summarizes the results of their classification framework.

2.1.4 Capabilities and business value research in DSAA

Articles in capabilities and business value research stream summarized the important capabilities and their effect on achieving business value from big data, analytics and data science. In their research to investigate the effect of data-driven decision (DDD) making on firm performance Brynjolfsson et al. (2011) showed that for firms being one standard deviation higher on the scale they developed for data-driven decision making translates to 4%-6% higher productivity. The relationship between data-driven decision making and performance also can be seen in other various important performance measures such as asset utilization, return on equity and market value. Their research also distinguished the impact of investments in IT and investments in capabilities supporting data driven decision making.

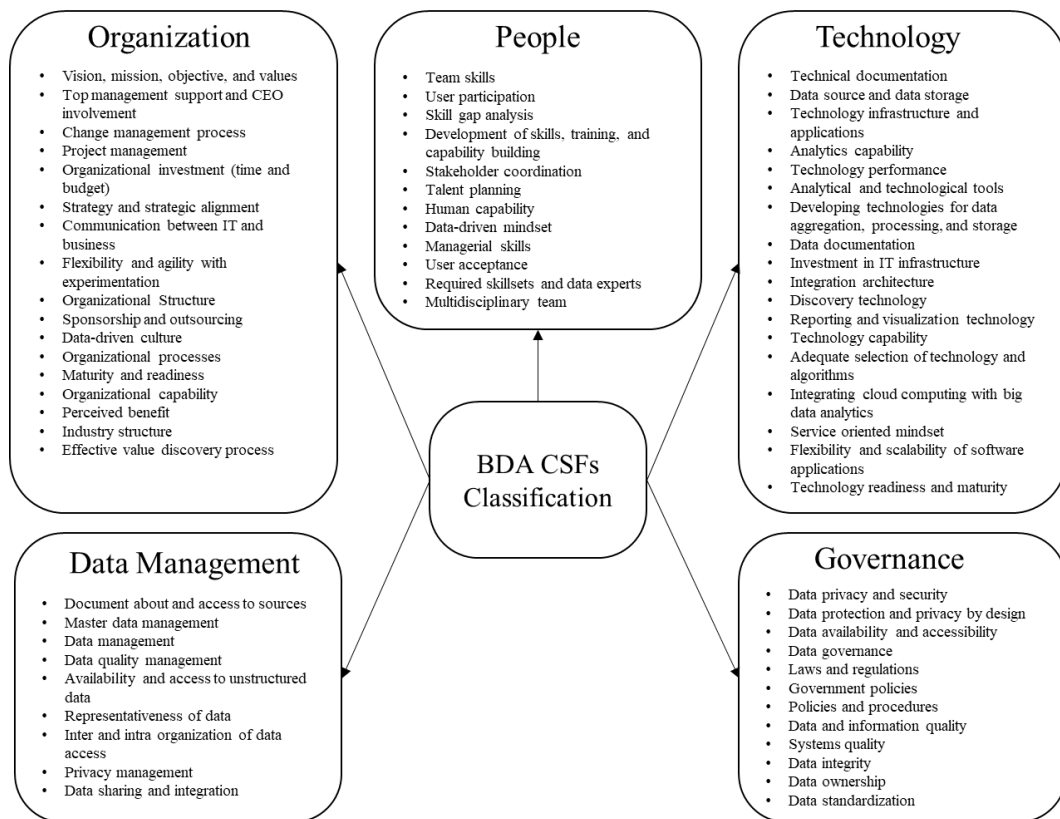


Figure 6. Big data analytics CSFs classification
Source: Al-Sai et al., 2020

Günther et al. (2017) posits that despite big data having been considered a transformational technological development, we have a very limited understanding of how and if organizations translate its potential into real value. After an in-depth systematic literature review, they identified six areas of inquiry about how organizations realize value from big data. They further identify portability and interconnectivity as two socio-technical features of big data that influence value realization: They stressed that organizations need to continuously realign work practices, organizational models, and stakeholder interests to reap the real value from big data. Their findings are synthesized in the form of integrated model that can be seen at Figure 7.

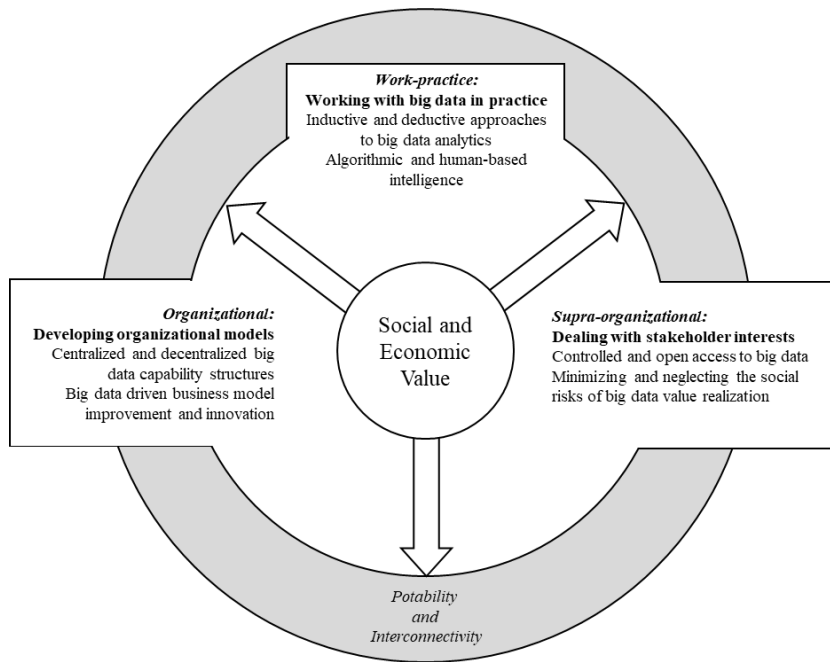


Figure 7. Integrated model of big data value realization
 Source: Günther et al., 2017

Following a review of extant empirical research, Gupta and George (2016) observed that investments in big data and data science technology alone proved insufficient to produce competitive advantage. They propose that organizations need to support these investments by creating capabilities competing organizations find hard to achieve. Based on the resource-based theory and current research in big data and advanced analytics, they identified the set of resources that allow building a big data analytics capability. They defined big data analytics (BDA) capability as “a firm’s ability to assemble, integrate, and deploy its big data-specific resources” and categorized the resources needed to develop it into tangible, human skills, and intangible types as seen in Figure 8. They created a survey instrument and collected data to measure the big data capability of the organizations. Upon testing their proposed framework, they found evidence that big data analytics capability indeed leads to superior firm performance.

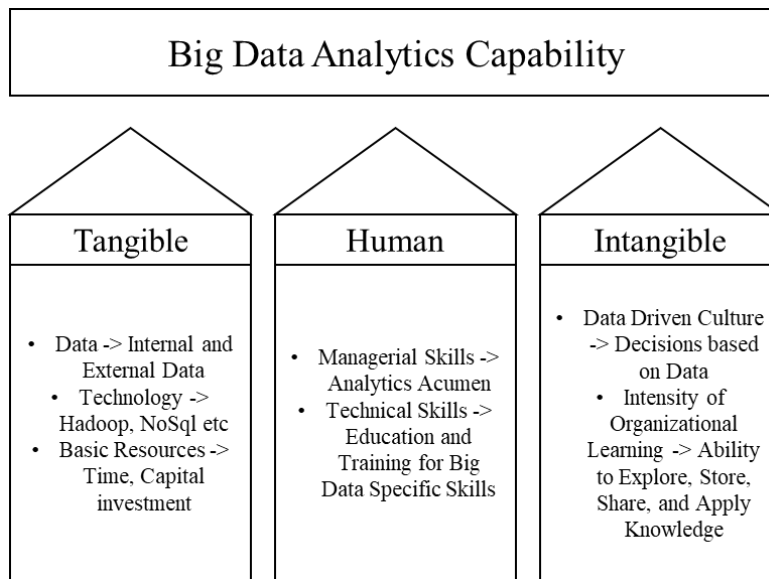


Figure 8. Resource supporting BDA capability

Source: Gupta & George, 2016

In the editorial of Special Issues of Information & Management titled “Big data and business analytics: A research agenda for realizing business value”, Mikalef et al. (2020) posits that number of companies adopting big data and business analytics increased considerably, requiring a shift in the focus to challenges in routinization of this transformative technology. They argued that this would create renewed opportunities for breakthrough research spanning different domains at the intersection of multiple research traditions. They presented a framework as seen in Figure 9 outlining components of big data and advanced analytics resources and highlighted interesting new research directions. They further stressed that widespread adoption of big data and advanced analytics as core components of contemporary organizations require an examination of the best ways to generate business value by using these core components.

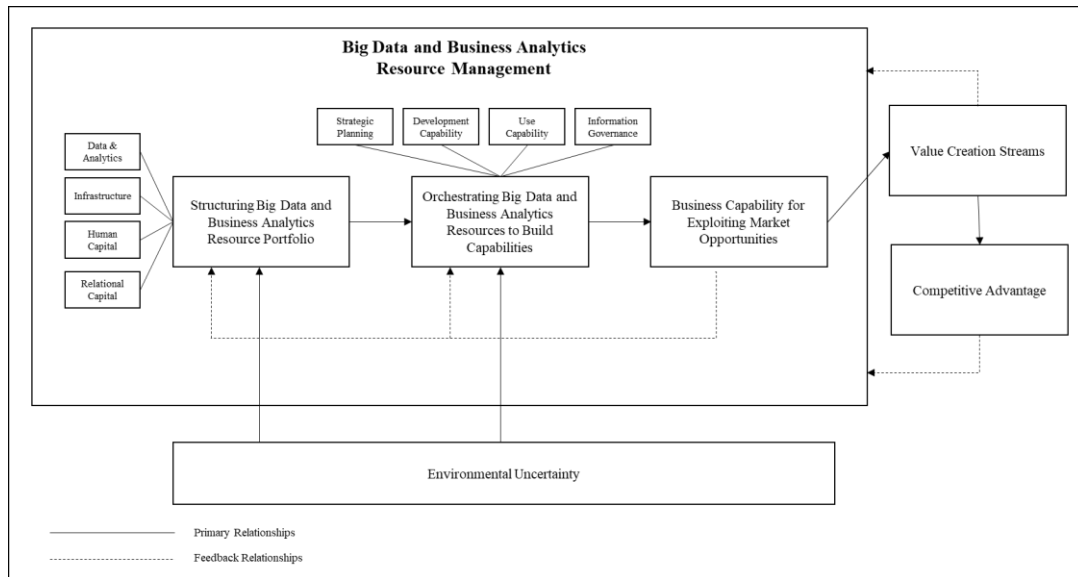


Figure 9. Big data analytics resource management framework
Source: Mikalef et al., 2020

Based on this resource management framework at Figure 9, Mikalef et al.

(2020) proposed following research questions:

- How can organization utilize big data and analytics methods to provide actionable insight?
- How can organizations bundle and orchestrate resources to increase their big data and business analytics leveraging capacity?
- Which organizational capabilities can big data analytics facilitate? What is the business value of doing so?

Grover et al. (2018) argued that despite the hype surrounding big data analytics (BDA) projects, it is unclear whether these initiatives provide any strategic value. They pointed to the lack of research regarding the ability of BDA on creating organizational value. They define the ultimate success of any BDA project as its ability to realize strategic business value. Then, they illustrated the framework through BDA applications in practice. Based on earlier IT value creation frameworks

they proposed a value manifestation framework (Figure 10) to help firms in formulating their BDA strategy.

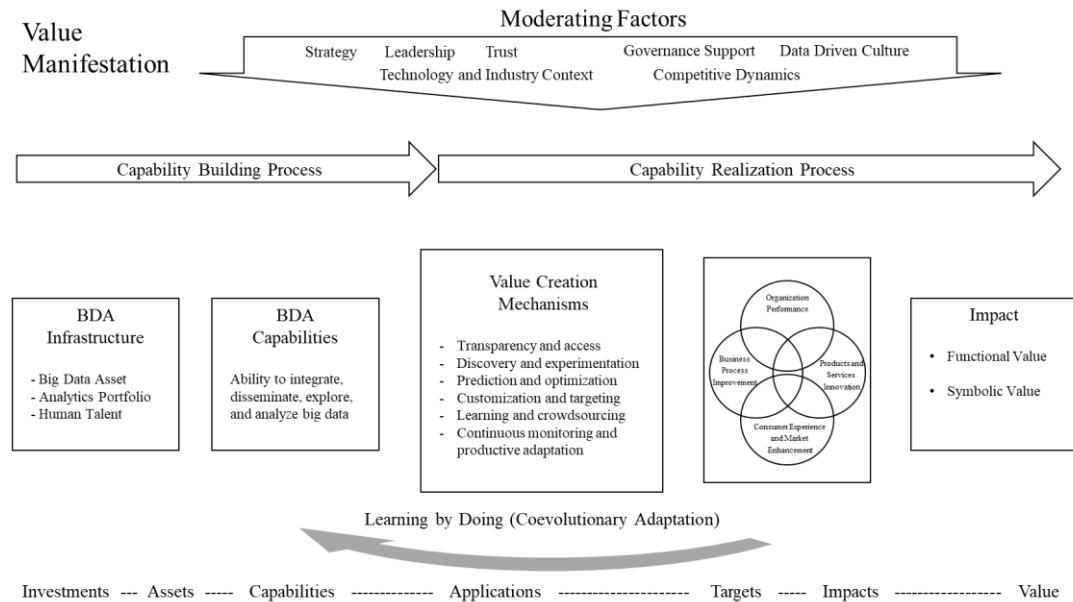


Figure 10. Big data analytics value creation framework
Source: Grover et al., 2018

Adapting VRIO (valuable, rare, costly to imitate, organizationally embedded) developed based on resource-based view (RBF) of firms, they proposed a firm developing a BDA strategy to ask the following questions to assess whether BDA has potential to generate strategic business value:

- Valuable: Does the BDA has the potential to provide valuable insights to create new business opportunities and/or counteract competitive threat?
- Rare: Are your big data content and/or analytics capability rare?
- Costly to imitate (imitable): Is it difficult or very costly for your competitors to imitate your BDA capabilities?
- Organizationally embedded: Is it in your organizations DNA to exploit these valuable, rare, and costly-to-imitate BDA capabilities?

Grover et al. (2018) also provided an extended list of the potential constructs at Table 2 for future research by categorizing them under 6 main concepts.

Table 2. Potential Constructs of BDA Research

Concept	Construct
BDA infrastructure	Dollar investment in BDA assets, Diversity of data sources, Volume, variety, and velocity of data, Quality of data analysis skills, Type of analytical toolset
BDA capabilities	Ability to integrate databases, Ability to disseminate structured and unstructured data, Ability to handle data volume, Ability to handle data variety, Ability to integrate external data Ability to design data analysis, Ability to interpret data outputs
Value targets	Quality of decisions, Business process efficiency, Product innovation, Service innovation, Customer experience, Market expansion
Value creation Mechanisms	Transparency, Access, Discovery, Experimentation, Prediction, Optimization, Customization, Targeting, Learning, Crowdsourcing, Monitoring, Proactive adaptation, Number of mechanisms triggered
Contextual	BDA strategy, Data-driven culture, Leadership, Data governance, Competitiveness, Environmental, Uncertainty
Impact	Profitability, Return on assets, Equity price, Symbolic value (image, satisfaction, industry leadership), Fraud detection, Market forecasting

Source: Grover et al., 2018

2.1.5 Adoption research in DSAA

Due to novelty of the big data and data science, there is not much published empirical IS research about the managerial aspects of big data and advanced analytics. Published research in big data and data science adoption is still quite limited and empirical published research is limited to only a handful of articles (Inamdar et al., 2020). A closer look into these articles reveal that there is almost none published research about big data and data science adoption in the leading journals. We also reviewed the results of our earlier search of “Web of Science” database for “big data analytics” and “data science” to check specifically for any published adoption related article in the AIS designated basket-of-eight leading journals. Due to the long publishing lifecycle of these journals, we can assume that there might be empirical research in the publication pipeline. In their big data

analytics value framework article at MIS Quarterly, leading researchers Grover et al. (2018) emphasized the importance of big data analytics for organizations and called other scholars to contribute to this research stream. This further validates the timeliness of our research. When we further evaluated the quality of the remaining published articles, we found only three worth mentioning here.

Sun et al. (2018) proposed a big data adoption research model based on diffusion of innovations (DOI) theory (Rogers, 1995) and specifically TOE framework (Tornatzky & Fleischer, 1990). In their conceptual model seen at Figure 11, they proposed 26 factors as having potential to influence intention to adopt big data technologies.

Maroufkhani et al. (2020) studied the effects of technological, organizational, and environmental factors on big data analytics (BDA) adoption and firm performance in the context of Iranian SMEs. They also used TOE framework (Tornatzky & Fleischer, 1990) to frame their research model and collected data via a quantitative field survey instrument. Their findings indicate that TOE factors have significant positive effects on BDA adoption, and while BDA adoption in turn improves performance of firms.

In another empirical study, Agrawal et al. (2015) developed and validated a research model to investigate the contextual factors effecting BDA adoption in emerging economies context. They also developed their model based on diffusion of innovation (DOI) theory (Rogers, 1995) and Technology-Organization-Environment (TOE) framework (Tornatzky & Fleischer, 1990).

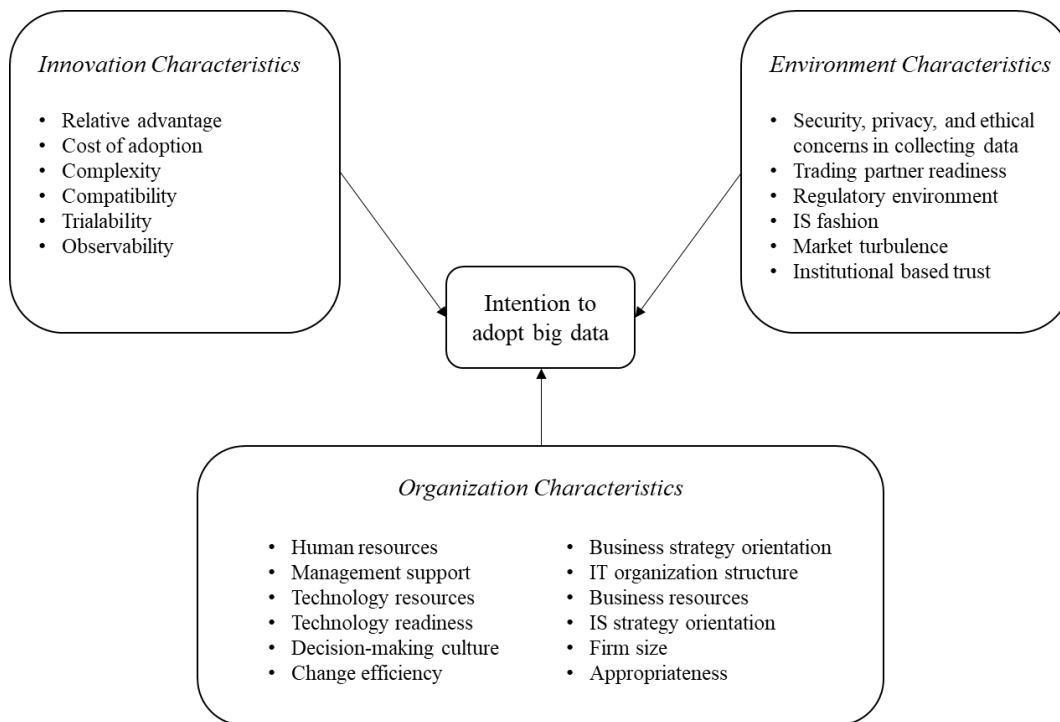


Figure 11. Factors effecting big data adoption intent
Source: Sun et al., 2018

They found six factors i.e. complexity, compatibility, regulatory support, organizational size, competition intensity, and environmental uncertainty to be significant determinants of BDA adoption, and three factors i.e., relative advantage, absorptive capacity, and technological resource competence to be non-significant. Two significant determinants, regulatory support and complexity were found to be inhibitors, while other significant determinants were facilitators of BDA adoption (Agrawal et al., 2015).

Lai et al. (2018), in another empirical survey research, based on previous literature on big data Analytics (BDA) and supply chain (SC) management, investigated the factors influencing organizations' intention to adopt BDA. In their proposed model, they categorized potential influences into four categories i.e., technological factors, organizational factors, environmental factors, and SC

characteristics. The empirical results indicate that perceived benefits and top management support are the strongest factors that influence the adoption intention. Furthermore, they found SC connectivity and other environmental factors i.e., competitors' adoption and government policy to moderate the driving forces and adoption intention.

2.2 Technology adoption literature

The last few decades have witnessed information technologies' gradual evolution into one of the most important sources of competitive advantage. Firms invest in IT to improve work efficiency, and ultimately to boost their organizational effectiveness. In order to meet these goals, a successful adoption of these information technologies (IT) is a crucial first step for organizations (Agarwal, 2000).

IT adoption occurs at two different but interrelated levels: First level is the "organizational level". IT can be adopted or diffused within the whole organization or at sub-level units such as the IT department. Hence, organizational-level adoption research can be defined as "research that examines adoption and diffusion of IT-based innovations by organizations or organizational units" (Jeyaraj et al., 2006). Second level where the adoption may take place is the "individual level". The acceptance of the technology by the users is critical since IT must be utilized by the end users to take advantage of the technologies that are adopted by the organizations. We define individual-level adoption research as "research that examines adoption and diffusion of IT-based innovation by the individuals" (Jeyaraj et al., 2006).

Because of the importance of adoption at both levels, IS research has focused on both of them extensively in the last 30 years or so. Although researchers mostly investigate these levels separately, both of them aim to find answers to the same

question: What factors facilitate or hinder the adoption and diffusion of IT-based innovations? (Fichman, 2004).

In technology adoption literature there are several theories that are used the most. These are the technology acceptance model (TAM) (Davis, 1986; Davis, 1989; Davis et al., 1989), theory of planned behaviour (TPB) (Ajzen, 1985; Ajzen, 1991), unified theory of acceptance and use of technology (UTAUT) (Venkatesh et al., 2003), DOI (Rogers, 1995), and the TOE framework (Tornatzky & Fleischer, 1990). Hameed, Counsell, & Swift (2012) in an effort to create a unified conceptual model of innovation adoption categorized the existing research according to the theory and models used in them. As can be seen in Table 3, TAM, TPB, and UTAUT are individual level acceptance models while DOI and TOE tries to explain adoption at the organization level.

Individual-level researchers examine the factors that lead to the “adoption behavior” of the users. While doing so, they frame their studies around several theories. These theories are largely rooted in the social psychology research and rely on cognitive perspectives that assume adopters as rational decision makers.

Table 3. Theoretical Models Used in Innovation Adoption Literature

Innovation Theories/Frameworks	Number of Studies	
	Organizational Level Analysis	Individual Level Analysis
Diffusion of Innovation (DOI)	28	3
Perceived Characteristics of Innovation (PCI)	1	0
Technology Acceptance Model (TAM)	11	26
Theory of Planned Behaviour (TPB)	4	12
Theory of Reasoned Action (TRA)	5	14
Technology Acceptance Model 2 (TAM2)	0	2
Technology Organization Environment (TOE) Model	35	0
Tricore	2	0
Task Technology Fit (TTF)	0	1
Unified Thoery of Acceptance and Use of Technology (UTAUT)	0	1
None/Others	81	12

Source: Hameed et al., 2012

The cognitive-based social psychology theories that are used the most are theory of reasoned action (TRA; Ajzen & Fishbein, 1980) and theory of planned behavior (TPB; Ajzen, 1985; Ajzen, 1991). These two theories also serve as the basis for technology acceptance model (TAM; Davis, 1986; Davis et al., 1989) as well as unified theory of acceptance and use of technology (UTAUT; Venkatesh et al., 2003), the two dominant theoretical lenses for the individual-level adoption studies.

Briefly, these studies suggest that an individual's adoption behavior is most significantly influenced by his/her intention to adopt, which is influenced by their salient beliefs about the technology in question. These studies identify "perceived usefulness/expected benefits" and "ease of use" of technologies as the two most salient beliefs. Hence, researchers have focused on understanding the personal (e.g., self-efficacy, personal innovativeness, computer anxiety, motivation) and organizational factors (e.g., top management support, facilitating conditions, IT support staff) that would improve the usefulness and ease of use perceptions of the users.

On the other hand, organizational-level research generally adopts two theories that complement each other: Rogers' Diffusion of Innovation Theory (DOI, 1995) and Tornatzky and Fleischer's Technology Organization Environment model (TOE, 1990). While it is not a theory on its own, the TOE framework mainly serves as a framework that suggest researchers to consider all of the three higher level factors while investigating organizational adoption or diffusion. The technological, organizational, and environmental factors that are selected for each study differ depending on the type of IT, organizational context and the industry where the organization competes. In terms of technological factors, DOI sheds light to IT-

related determinants of adoption, i.e., “relative advantage”, “complexity”, “compatibility”, “observability” and “trialability”.

Traditional innovation adoption frameworks were initially developed to explain individual behavior. So, their relevance and applicability to technology adoption in an organizational setting when there is an organizational mandate became a focus of many researchers (Gallivan, 2001). Process and research models emerged to explain organizational implementation of innovations. Rogers (1995) in addition to his initial individual innovation adoption model proposed a different five-stage organizational innovation adoption model that can be classified as initiation (stage 1-2) and implementation stages (3-5).

Although Roger’s model was the first process model that tries to explain organizational adoption and implementation others followed. The most well-known model of technology implementations in organizations is the six-stage model proposed by Zmud and his colleagues. The stages of this model are initiation, adoption, adaptation, acceptance, routinization, and infusion. (Cooper & Zmud, 1990).

2.3 Organizational technology adoption literature

In this section, we will explain the most prominent organizational technology adoption theories Diffusion of Innovations (DOI) and Technology Organization Environment (TOE) in detail.

2.3.1 Diffusion of innovations (DOI) theory

DOI theory tries to explain the factors and mechanisms that leads to the adoption of innovations by individuals and organizations. First proposed at the individual level it is the most widely used and broadly accepted theory explaining technology innovation adoption. DOI theory sees innovations as being communicated through certain channels over time and within a particular social system (Rogers, 1995).

Rogers proposed five attributes of innovations playing a role in the ease of adoption by individuals. These are relative advantage, compatibility, trialability, observability and complexity of the innovation. Innovations that have these five attributes have a smooth and faster adoption. The first four factors are expected to be positively correlated with rate of adoption while the last factor, complexity, is expected to be negatively correlated with rate of adoption (Rogers, 1995)

Individuals seem to have different degrees of willingness to adopt innovations. Because of this the population adopting an innovation is generally observed to be normally distributed over time (Rogers, 1995). This normal distribution can be broken down into segments leading to the classification of individuals into the following five categories in terms of individual innovativeness: innovators, early adopters, early majority, late majority, laggards. (Rogers, 1995).

The characteristics of the members of each category can be summarized below:

- Innovators - venturesome, educated, multiple info sources
- Early adopters - social leaders, popular, educated
- Early majority - deliberate, many informal social contacts
- Late majority - skeptical, traditional, lower socio-economic status
- Laggards - neighbors and friends are main info sources, fear of debt

The bell shaped adoption curve can be converted to a cumulative percentage curve (Figure 12). This characteristic S curve shows both the initial increasing rate and then levelling-off of the innovation adoption within the population (Rogers, 1995).

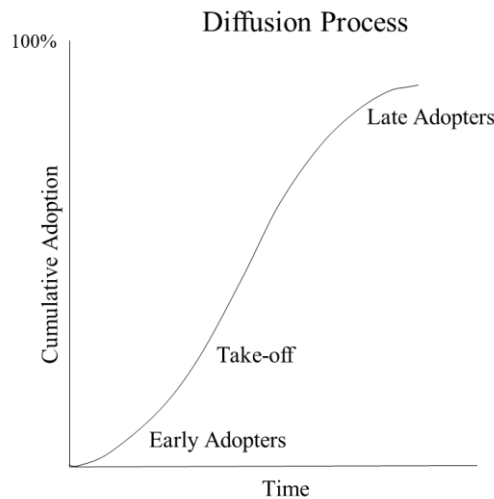


Figure 12. Diffusion process
Source: Rogers, 2005

Based on DOI theory, Moore and Benbasat (1991) extended the five factors impacting the adoption of innovations proposed by Rogers. They proposed that in IS context eight factors impact the adoption of IT. These are voluntariness, relative advantage, compatibility, image, ease of use, result demonstrability, visibility, and trialability. They also developed and validated scales used to operationalize these factors in their study.

At the firm-level, Rogers (1995) propose that innovativeness is related to three group of independent variables as can be seem in Figure 13. These are individual (leader) characteristics, internal organizational structural characteristics, and external characteristics of the organization.

In this model Rogers (1995) defines individual characteristics as the leader attitude toward change. Internal characteristics of organizational structure includes centralization, complexity, formalization, interconnectedness, organization slack, and size.

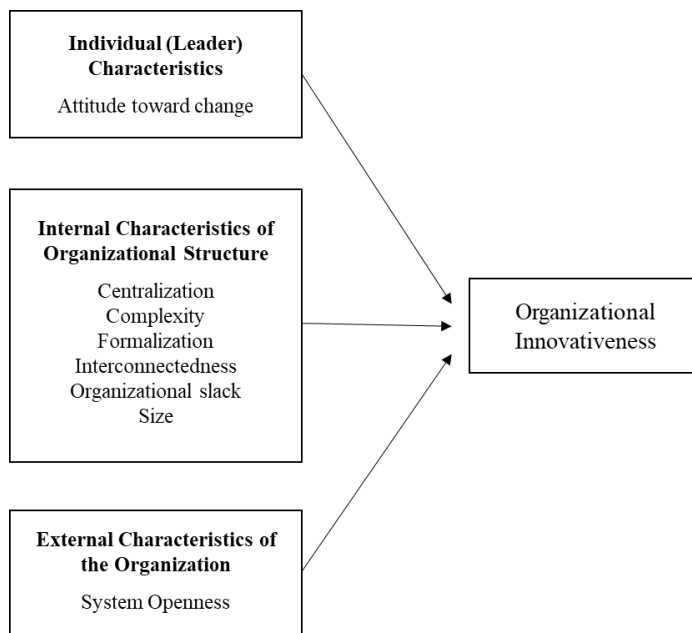


Figure 13. Diffusion of innovations (DOI) theory at firm level
Source: Rogers, 2005

Each of these characteristics are further defined by Rogers (1995) as:

“centralization is the degree to which power and control in a system are concentrated in the hands of a relatively few individuals”; “complexity is the degree to which an organization’s members possess a relatively high level of knowledge and expertise”; “formalization is the degree to which an organization emphasizes its members’ following rules and procedures”; “interconnectedness is the degree to which the units in a social system are linked by interpersonal networks”; “organizational slack is the degree to which uncommitted resources are available to an organization”; “size is the

number of employees of the organization”. The last group of external characteristics of organizational structure refers to system openness.

DOI theory has been adapted and applied in IS literature many times in different technology contexts. In one of them Bradford and Florin (2003) examined the role of innovation diffusion factors on the implementation success of ERP systems. They found technical compatibility, perceived technical complexity, and competitive pressure as important antecedents to the adoption of innovations.

2.3.2 Technology, organization, environment (TOE) framework

Tornatzky and Fleisher (1990) defines the process by which a firm adopts and implements technological innovations to be influenced by technological, organizational, and environmental contexts (Figure 14).

The technological context is comprised of all the relevant internal and external technologies. The organizational context includes characteristics and resources of the firm: the firm’s size, degree of centralization, degree of formalization, managerial structure, human resources, amount of slack resources, and linkages among employees. The environmental context refers to the size and structure of the industry, the firm’s competitors, the macroeconomic context, and the regulatory environment (Tornatzky & Fleisher, 1990).

Tornatzky and Fleisher (1990) propose that these three group of factors present both constraints and opportunities for technological innovation. So, these three group of factors influence both each other and the way a firm structures the technological innovation decision making.

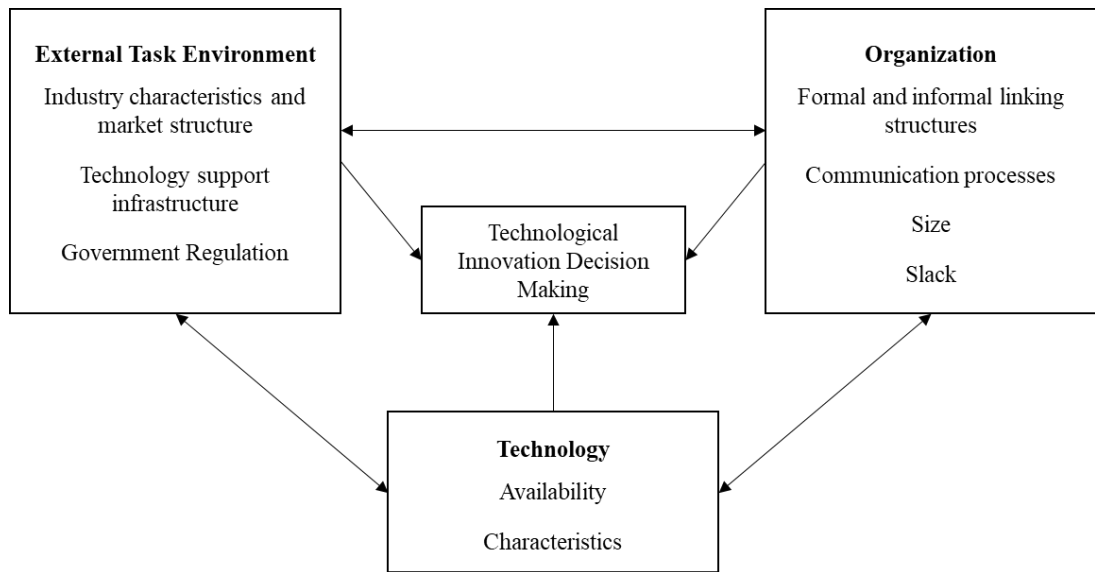


Figure 14. Technology, organization, environment (TOE) framework
Source: Tornatzky & Fleisher, 1990

Basole, Seuss, and Rouse (2013) analyzed existing multi-disciplinary research of the enterprise-level IT innovation literature by means of text analytics methods. They first identified 472 articles and classified these according to functional discipline, publication, research methodology, and information technology type. Then they used text analytics methods to identify adoption determinants and their relationships in existing literature. Their results create an extensive list of potential critical factors for adoption for future studies.

Table 4, 5 and 6 summarizes their findings.

Table 4. Organizational Factors

1977-1979 (n = 4)	1980-80 (n = 32)	1990-99 (n = 134)	2000-08 (n = 290)
Strategy (100%)	Experience (53%)	Experience (69%)	Experience (72%)
Organization structure (75%)	Integration (44%)	Integration (54%)	Infrastructure (58%)
Experience (50%)	Strategy (44%)	Strategy (42%)	Integration (51%)
Performance gap (50%)	Organizational structure (41%)	Infrastructure (40%)	Strategy (47%)
Slack resources (25%)	Organizational size (19%)	Organizational structure (31%)	Trust (37%)
User involvement (25%)	Infrastructure (16%)	Top management support (21%)	Top management support (24%)
Top management support (25%)	Top management support (12%)	IS planning (13%)	Financial resources (22%)
	Maturity (12%)	Human resources (12%)	Human resources (18%)
	Organization culture (12%)	Organizational size (11%)	Organizational structure (17%)
	Financial resources (9%)	Trust (10%)	Maturity (16%)
	IS planning (9%)	Financial resources (10%)	Organizational readiness (9%)
	User involvement (9%)	Maturity (10%)	Organizational size (9%)
	Communicability (6%)	User involvement (10%)	User involvement (8%)
	Technology policy (6%)	Organizational culture (9%)	IT resources (7%)
	Trust (6%)	Communicability (4%)	IS planning (6%)

Source: Basole et al., 2013

Table 5. Technological Factors

1977-1979 (n = 4)	1980-80 (n = 32)	1990-99 (n = 134)	2000-08 (n = 290)
Complexity (100%)	Complexity (56%)	Cost (62%)	Cost (62%)
Cost (50%)	Cost (47%)	Complexity (46%)	Complexity (57%)
	Compatibility (28%)	Compatibility (31%)	Compatibility (41%)
	Image (16%)	Relative advantage (22%)	Relative advantage (26%)
	Relative advantage (9%)	Image (15%)	Image (22%)
	Triability (6%)	Triability (11%)	Perceived benefits (18%)
	Ease of use (3%)	Ease of use (19%)	Ease of use (14%)
	Perceived benefits (3%)	Perceived benefits (6%)	Triability (13%)
		Perceived usefulness (6%)	Perceived usefulness (11%)
		System quality (2%)	Result demonstrability (1%)
			System quality (1%)

Source: Basole et al., 2013

Table 6. Environmental Factors

1977-1979 (n = 4)	1980-80 (n = 32)	1990-99 (n = 134)	2000-08 (n = 290)
Government (50%)	Power (53%)	Power (57%)	Dependency (83%)
Power (50%)	Competitive pressure (41%)	Uncertainty (45%)	Power (68%)
Uncertainty (50%)	Government (41%)	Competitive pressure (37%)	Competitive pressure (56%)
	Uncertainty (37%)	Government (36%)	Government (52%)
	Dependency (3%)	Dependency (7%)	Uncertainty (39%)
	Environmental complexity (3%)	Social network (3%)	External pressure (15%)
		Environmental complexity (2%)	Vendor support (8%)
		Vendor support (2%)	Social network (6%)
		Customer interaction (1%)	Interconnectedness (5%)
			Coercive pressure (3%)
			Competitive intensity (3%)
			Normative pressure (3%)
			Customer interaction (2%)

Source: Basole et al., 2013

CHAPTER 3

THEORETICAL DEVELOPMENT

3.1 Technology, organization, environment framework

We developed our conceptual model from the theoretical lens of the TOE framework. TOE framework has been widely used for technology adoption studies since its introduction by Tornatzky and Fleisher (1990). It consists of three elements, technological, organizational, and environmental contexts that present “both constraints and opportunities for technological innovation” (Tornatzky & Fleisher, 1990).

Earlier empirical research has based its theoretical roots in this framework both for initial adoption decisions as well as post-adoption usage of technologies. For example, at the initial adoption stage, Chau and Tam (1997) investigated the factors that influence the adoption of open systems, while Grover (1993) developed a model for the adoption of interorganizational systems. At the post-adoption stage, Zhu and Kramer (2005) examined the retail industry, and empirically tested the variations in the usage and value of e-business by organizations.

Overall, TOE framework sets the stage by providing a basic ground for the researchers to build their study on, then the researchers identify appropriate factors depending on the context and the technology under investigation (Baker, 2012). That is why it is generally accepted as a generic theory or research framework rather than a theory on its own and research studies using the framework thought to have an exploratory research paradigm (Baker, 2012; Zhu & Kreamer, 2005). Due to the novel nature of Data Science and Advanced Analytics Technologies, we also adopted

this stance and carefully reviewed existing literature to identify salient factors that can be used as part of TOE framework.

In our research, we investigate the Technological, Organizational, and Environmental factors that would influence the adoption process of Data Science and Advanced Analytics Technologies as a whole. We argue that, due to its novel nature, the adoption of these technologies needs to be studied by identifying and testing its antecedent factors found in existing literature, while also ensuring content validity by consulting experts from industry and academia regarding the relevancy of these factors.

3.2 Technology characteristics and relevant factors

Following our justification for the TOE framework and the nature of Data Science and Advanced Analytics characteristics, the set of factors and hypotheses in our research model has been finalized via a qualitative study of in-depth interviews with a select group of practitioners from the industry and academia. During this assessment we drew upon the factors outlined in literature review and meta-analysis studies in extant literature such as Basole et al. (2013), Hameed et al. (2012), and Oliveira, Thomas, and Espadanal (2014) and evaluated their suitability for inclusion in our model with experts from the field. We followed the approach suggested by methodological researchers to ensure content validity of instrument. Specifically, we interviewed practitioner and academic experts to evaluate the relevance and validity of both the individual items and the overall instrument (Tojib & Sugianto, 2006). The aim of this process was to introduce the experts the broader set of factors in extant technology adoption literature and to retain the best items which are believed to adequately measure a desired content domain. In this section, we summarize the

results of this evaluation process and explain the rationale for inclusion of the factors in the final proposed research model.

Data Science and Advanced Analytics Technologies differ from other enterprise information technologies by certain characteristics. From a technological perspective, its “low-cost” nature means that managers may find it too attractive to develop adoption intentions, and they can be mistaken for expecting a relatively easy infusion and an immediate return.

However, using Data Science and Advanced Analytics Technologies needs top-notch skillsets that include both engineering precision and analytics experimentation, which may play an inhibiting role in the effective use of big data. It means that Data Science and Advanced Analytics Technology adoption will require a highly-fragmented technology stack that can be found only in organizations that boast very strong IT infrastructure with skilled IT staff. We also expect an organization’s decision-making culture as an important factor, since there is a high degree of difficulty in finding the right use cases. The system will likely provide the most satisfactory returns only when the decision making culture supports operationalizing the results of the Data Science and Advanced Analytics.

While these characteristics of Data Science and Advanced Analytics help us identify the technological and organizational antecedents of its adoption, we turn to Institutional Theory to distinguish the environmental factors that are relevant to our research. Institutional Theory perspective argues that “organizations respond to an environment, which consists of organizations responding to an environment of organizations’ responses” (Schelling, 1978). Burt (1987) and Di Maggio and Powell (1983) emphasize the interconnected nature of similar organizations within the same environment, and how they are subject to isomorphic pressures. While there are three

types of isomorphic pressures (coercive, mimetic, and normative), we will focus only on mimetic pressure, since the other two operate through interconnected or interorganizational relations. Isomorphic pressures have also been used in technology adoption studies. For example, Teo, Wei, and Benbasat (2003) examined the adoption of financial electronic data interchange (FEDI) at the firm-level, and found that all three types of pressures have significantly and positively influenced the FEDI adoption.

According to the institutional theory, mimetic pressures are manifested either in “the prevalence of a practice” or “the perceived success of organizations within the focal organization’s industry that have adopted the practice” (Haveman, 1993). DiMaggio and Powell (1983) also suggest that organizations tend to model themselves after other organizations to acquire what is called a status-conferring legitimacy, regardless of the value of the technology or the practice.

Empirical findings support this viewpoint. Banker and Kaufman (1988) found that ATM technology in banking industry was adopted due to mimetic pressures although its business value was yet to be established. Considering the recent emergence of big data, data science and the hype surrounding it, we expect mimetic pressures to play an important role in the development of adoption intentions. Similar to previous work by Teo et al. (2003), we can operationalize these pressures as extent of adoption among competitors and perceived success of competitor adopters.

After the content validity assessment, we consider the following as our final set of hypothesized factors:

- Technological factors: Relative Advantage, Perceived Complexity, Technical Compatibility, Cost

- Organizational factors: Top Management Belief, Top Management Participation, Organizational Slack
- Environmental factors: Mimetic Pressure, Availability of External Expertise
- Inhibiting factors: Extent of Data Driven Decision-making and Lack of Internal Expertise

Control variables: Respondent Role (Business vs IT)

Table 7 lists the definitions of the constructs that represent our earlier set of hypothesized factors.

Our literature review reveals that one of the caveats of both the individual and organizational-level technology adoption studies is the sole measurement of the intention without moving forward to the actual behaviour stage, which refers to the full implementation and actual adoption of the systems. These studies are built on the theoretical principles of either/both Theory of Reasoned Action (TRA, Ajzen and Fisbein 1973) or/and Theory of Planned Behavior (TPB, Ajzen,1991), which assume that behavior should follow intention. While the correlation between the two is generally found to be high in many contexts, researchers often treat intention as a proxy to behavior, mainly because of convenience. While we do not counter-argue against the strong link between intention and behaviour, we argue that intent is provisional, and there may be certain conditions in which the actual behaviour (i.e. adoption) may not occur (e.g. Zimmer, Aarsal, Al-Marzouq, Moore, & Grover, 2010). In the case of Data Science and Advanced Analytics adoption, we posit that for majority of organizations adoption intention is already high due to lower upfront costs, high trialability, and perceived competitive pressure. Hence, trying to measure

Table 7. Organizational Adoption Constructs and Definitions

Construct	Definition
Relative Advantage	“the degree to which use of an innovation is perceived as being better than its precursor.” (Moore and Benbasat, 1991)
Complexity	“the degree to which a system is perceived as relatively difficult to understand and use.” (Rogers, 1995; Moore and Benbasat, 1991)
Compatibility	“the degree to which an innovation is perceived as being consistent with existing values, needs and past experiences of potential adopters.” (Moore and Benbasat 1991; Rogers, 1995)
Cost Savings	the degree to which a technology is perceived as having benefits outweighing the set-up costs, as reducing existing costs, and as having low ongoing maintenance costs (Oliviera et al., 2011; Sangle, 2011)
Top Management Belief	a subjective psychological state regarding the potential of the technology under consideration (Liang, Saraf, Hu, & Xue, 2007)
Top Management Participation	the behavior and actions performed to facilitate assimilation of the technology under consideration (Liang et al., 2007)
Organizational Slack	“the pool of resources in an organization that is in excess of the minimum necessary to produce a given level of organizational output” (Nohria & Gulati, 1996)
Mimetic Pressure	the degree of perceived uncertainty that encourages organizational imitation (DiMaggio et al., 1983; Liang et al., 2007)
External IT Expertise	“the effectiveness of consultants and the quality of vendor support” (Thong, Yap, & Raman, 1996)
Data-Driven Decision-making	the extent of usage and existence of data for the creation of a new product or service and business decision making in the entire company (Brynjolfsson et al., 2011)
Lack of Internal Expertise	The degree to which the internal experts is unable to meet the technical requirements to successfully complete DSAAT projects (Ewusi-Mensah & Przasnyski, 1991)
Adoption Intention	to what extent is DSAAT expected to be used to support key business activities in business processes (Tallon & Pinsonneault, 2011)
Actual Adoption / Usage	to what extent is DSAAT used to support key business activities in business processes (Tallon & Pinsonneault, 2011)

only the intention stage may be misleading in understanding the actual adoption/usage of Data Science and Advanced Analytics Technologies by the organizations.

Therefore, we posit that in order to better understand the factors leading to real adoption we need to directly focus on organizations that already piloted use of Data Science and Advanced Analytics Technologies. We call this stage the “Trial” stage in which the companies try to evaluate whether full adoption is beneficial for the organization.

In order to understand whether the firms are actually using the technology we define actual adoption as “Production Use” stage in which after the pilot period the organization automate and operationalize the technology by embedding the results in business processes. This will help us understand the link between manifested intention (“Trial”) and actual adoption (“Production Use”).

Figure 15 summarizes our conceptual model in which technological and environmental factors directly affect adoption intention, while some organizational factors play a facilitating role at the intention stage others play an inhibiting role at the actual adoption stage.

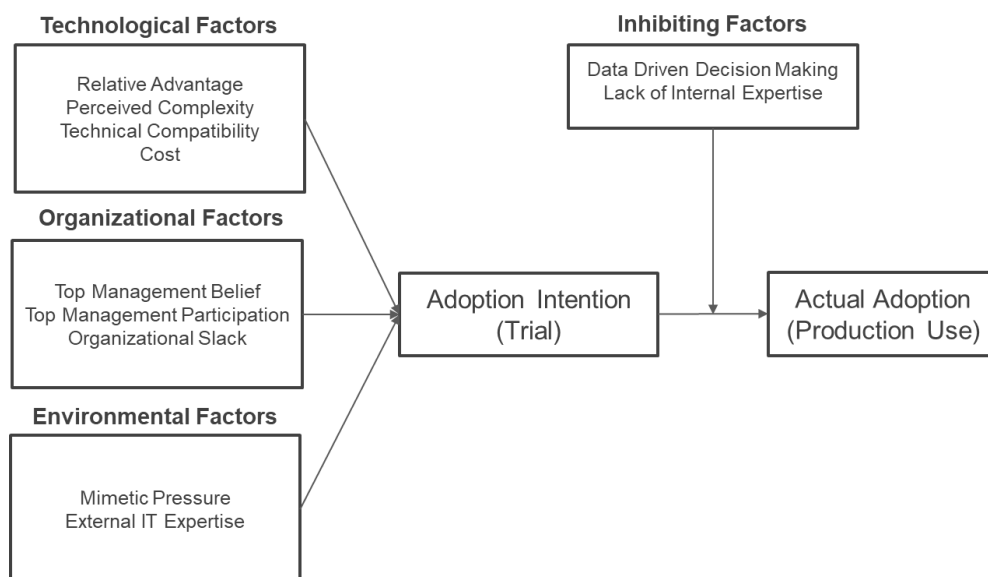


Figure 15. Conceptual model

Based on our proposed model, we posit the following propositions:

- P1. The technological factors (especially cost and relative advantage) will influence the adoption intention significantly. → Low cost makes Data Science and Advanced Analytics adoption look attractive and feasible, and their benefits will be perceived high.
- P2. The environmental factors (mimetic pressures and availability of external expertise) will influence the adoption intention significantly → Data Science and Advanced Analytics is one of the hottest recent topics, many successful firms are using it, and others may feel the need to keep up with the trend to gain competitive advantage.
- P3. The organizational factors (top management support in the form of both belief and participation, and organizational slack) will have significant influence on adoption intention → These factors are likely to facilitate the intention formation while the other organizational factors, availability of expertise and the extent of data driven decision making will have an influence on the actual adoption stage.
- P4. The inhibiting factors, i.e., the extent of data driven decision-making and availability of internal expertise will negatively moderate the relationship between intention to adopt and actual adoption.

CHAPTER 4

METHODOLOGY

4.1 Research design

The goal of this study is to develop and quantitatively test a model that explains the factors leading to successful organizational adoption of Data Science and Advanced Analytics Technologies (DSAAT).

This chapter describes the research paradigm and methodology used in this study detailing unit of analysis, sample design, instrument design, data collection, and data analysis procedures and methods. Based on our positivist research paradigm, quantitative survey methodology is adopted as it is also the prominent research methodology approach in IS research (Galliers et al., 1987; Palvia, Daneshvar, Ghoshal, Uppala, & Wang, 2015; Palvia, Mao, Salam, & Soliman, 2003) in general and also the preferred approach for rigor in organization technology adoption literature (Choudrie et al., 2005; Riedl et al., 2011; Williams, Dwivedi, Lal, & Schwarz, 2009).

Constructs operationalized via adopting existing measures in the prior literature. A field survey was used to collect data from carefully selected sample. Later, PLS-SEM has been used to test the causal relationships in the proposed model. Figure 16 outlines the overall research plan of our study.

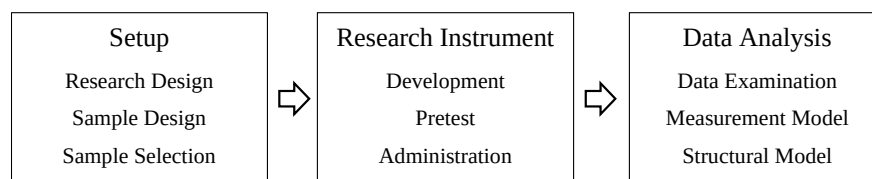


Figure 16. Research plan

4.1.1 Unit of analysis

The current study investigates the factors leading to successful organizational adoption of an emerging technology. So, the unit of analysis in our research is organization.

Although the most widely used method in organizational adoption literature has been survey research (Choudrie et al., 2005; Karanja & Zaveri, 2013; Palvia et al. 2015), researchers prefer to use case research method to investigate factors in organizational adoption of technology especially when the unit of analysis is the organization (Choudrie et al., 2005). This can be mostly attributed to the difficulty in reaching out to large enough sample of quality participants as well as the potential biases resulting from aggregating data coming from individual participants. There is considerable number of research using senior professionals as organizational representatives (Karanja & Zaveri, 2013) to investigate organizational phenomena when it is feasible to do so. So, considering the availability of potential participants as organizational representatives and the potential for rich quantitative analysis, we decided to employ a field survey methodology targeting senior professionals in a related role.

4.1.2 Control variables

The use of control variables allows researchers to eliminate alternative explanations and confounders, while also making it possible to test causal relationships in subgroups. (Jenkins, 1985; Lucas, 1989). In our study, we included the fact that whether a participant is part of a business or IT function of their organizations as a control variable. PLS-SEM methodology allows the use of categorical variables as

controls either as dummy encoded variables or estimating a model per subgroup (Hair et al., 2017).

4.2 Sample design

Sample design consists of selection of sample frame, sample selection process, and determination of the appropriate sample size (Churchill, 1991). Selection of the sample frame is the most important element of the sample design, as representativeness of the sample frame impacts the credibility of the results of the study. (Lucas, 1989; Pinsonneault et al., 1993)

Survey research in which the unit of analysis is organization, inevitably target questions to individual members of the group under study. When multiple members of the same group answer survey instruments, the inherent covariation within the group increase the probability of detecting significant findings (Barcikowski, 1981; Baroudi, 1989). In order to mitigate this problem, while selecting the survey participants we aimed to approach one senior professional from business and IT functions of a specific organization each.

The population for our study consists of companies that have the potential to adopt Data Science and Advanced Analytics Technologies. In order to select a representative sample, we first identified criteria to select industries and companies to include in our sample frame. Industries and companies that satisfy one or more of the following criteria are included in the sample frame:

- Industries in which companies serve large number of customers
- Industries in which there is high information intensity
- Industries known to be as early adopters of data related technologies

4.2.1 Sample selection process

Considering the difficulties of employing such a study in multiple languages in a cross cultural context, we limited our sample to organizations in Turkey. Banking, Telecommunications, Insurance, Retail, and E-Commerce industries are characterized by having retail customer-focused companies, high information intensity, and being early adopters of data related technologies. So, we selected companies in these industries in Turkey as our target based on the criteria defined in section 4.2.

Although organizational adoption theories mainly focus on a single country context (Williams et al., 2009), limiting our sample frame to a handful of selected industries further had the potential to reduce generalizability of our study. Hence, we decided to include other companies that satisfy similar criteria but operate in different industries in Turkey.

Considering the difficulty of reaching out to qualified participants in the final list of target organizations in the sample frame, a list potential survey participants was created using LinkedIn search, web search, and researcher's contacts in the field. Only senior professionals who has the potential to represent their organizations views are included in the final list after cross-checking their final role via references and online information. The final list included 158 people. Composition of the final list of potential survey participants to be contacted is summarized in Table 8.

4.2.2 Sample size

Minimum sample size requirement in quantitative survey research based MIS literature has been a contentious issue (Hair et al., 2017; Westland, 2010). Especially

Table 8. Initial List of Survey Participants

Industry	Business	IT	Other	Total
Banking	27	17	6	50
Telecom	5	9	6	20
E-commerce	3	3	8	14
Retail	4	8	2	14
Holding	4	8		12
Insurance	1	4	5	10
Manufacturing	1	5	4	10
Services	4	3	2	9
Technology		9		9
Energy		4		4
Airline	1	1	1	3
Media	2	1		3
Total	52	73	34	158

when the unit of analysis is the organization, the size of the population becomes inevitably low, making it harder for researchers to reach meaningful sample sizes (Baroudi et al., 1989; Pinsonneault et al., 1993).

The issue of small sample sizes became one of the most cited reasons for the use of PLS-SEM as the preferred approach due to relatively low sample size requirements even in the case of complex model structures (Hair, Matthews, Matthews, & Sarstedt, 2017; Reinartz, Haenlein, & Henseler, 2009). Thus, researchers in search of a way around minimum sample size requirements often revert to the rough the guideline known as 10-times rule (Goodhue, Lewis, & Thompson, 2012; Marcoulides & Saunders, 2006).

As a heuristic, 10-times rule suggest that the minimum sample size in a PLS-SEM analysis should be equal to the larger of the following (Barclay, Higgins, & Thompson, 1995; Hair et al., 2017; Kock & Hadaya, 2018):

- 10 times the maximum number of formative indicators used to measure an exogenous latent variable
- 10 times the maximum number of structural paths directed at a particular endogenous latent variable

Although use of the 10 times rule heuristic is quite common in PLS-SEM based MIS literature, simulation studies showed its inadequacy in providing the necessary guardrails for catching significant effects when effect sizes are relatively small (Goodhue et al., 2012; Marcoulides & Saunders, 2006; Westland, 2010). In order to ensure survey based studies provide sufficient evidence for the hypothesized relationships in a PLS-SEM model, sample size recommendations provided by Cohen (1992) should be followed (Akter, D'ambra & Ray, 2011; Hair et al., 2017). The guidelines provided by Cohen (1992) consider significance levels, statistical power, and effect sizes while recommending minimum sample sizes.

Due to the inherent nature of having small population sizes leading to sample sizes in organizational adoption studies, we were expecting to face the same problems while trying to reach sufficient sample size. Accommodating relatively smaller sample sizes, PLS-SEM was a good fit for us as it estimates path coefficients in a sequential manner via multiple OLS regression runs. (Hair et al. 2017; Reinartz et al., 2009;).

Our research model has 11 exogenous latent variables, 3 higher-order reflective-formative latent variables, 2 moderating variables, and 2 endogenous latent variables. The maximum number of rows pointing toward a formatively measured

variable was 5, while the final exogenous variable also had 5 rows pointing to it.

Thus, the rough guideline provided by 10-times rule suggests that we have at least 50 observations.

On the other hand, Hair et al. (2017) propose that the part of the model with the biggest number of predictors should determine the minimum sample size via a proper power analysis. Table 9 shows the sample sizes necessary to identify minimum R² values in any of the endogenous latent variables in the structural model for significance levels of 5% or 10%, and a statistical power of 80% (Cohen, 1992; Hair et al. 2017).

Table 9. Sample Size Recommendation for PLS-SEM for a Power of 80%

Maximum Number of Arrows Pointing at a Construct	Significance Level							
	%5				%10			
Effect Size	Minimum R ²				Minimum R ²			
	0.10	0.25	0.50	0.75	0.10	0.25	0.50	0.75
2	90	33	14	8	72	26	11	7
5	122	45	20	12	99	37	17	10
10	156	59	27	18	129	49	24	16

Source: Cohen (1992): A power primer. Psychological Bulletin 112: 155–159.

Since the maximum number of arrows pointing to a single latent variable in either formative measurement model or the structural model is 5 in our proposed model, Table 9 suggests that we will need up to 156 observations to reach a statistical power of 80% for identifying relatively weak R² values of 10% with a

Type I error of 5%. So, we decided to target collecting survey responses from at least 150 participants.

Hair et al. (2017) also propose running statistical power analyses using programs such as G*Power. There are also alternative methods proposed such as inverse square root and gamma-exponential methods (Kock & Hadaya, 2018). Post-hoc calculation of inverse square root method which takes into account the minimum path coefficients in our model also yields a sample size estimation of around 200 observations.

4.3 Research instrument

In this section, we first describe how the survey instrument was developed based on existing measures in the literature. The section proceeds by describing how we pretested, and later administered the survey instrument taking into account recommendations in the literature.

Survey research is the most common methodology in MIS research (Palvia et al., 2004; Palvia et al., 2015; Riedl & Rueckel, 2011). Since survey research is about trying to measure latent constructs that cannot be directly observed, quality of the instrumentation is of utmost importance. (Lewis, Templeton, & Byrd, 2005; MacKenzie, Podsakoff, & Podsakoff, 2011; Newsted, Huff, & Munro, 1998; Straub, 1989). Various studies consider development of MIS instruments as an MIS research category on its own, and there is also a separate research stream proposing guidelines for MIS survey instrument development and validation (Chen & Hirschheim, 2004; Orlikowski & Baroudi, 1991).

Developing and validating measurement scales for survey instruments is a costly and challenging activity. Moreover, using previously developed and validated

survey instruments both brings rigor into the research study under consideration and contributes to the cumulative tradition of research. (Lewis et al., 2005; MacKenzie et al., 2011).

4.3.1 Survey instrument development

Following a comprehensive literature review about organizational adoption in MIS literature, we identified factors that has the potential to influence organizational adoption of Data Science and Advanced Analytics Technology (DSAAT).

After finalizing the research model and the related constructs, we followed the guidelines in literature for construct operationalization, scale adoption, and validation proposed by Lewis et al. (2005) and MacKenzie et al. (2011).

We first examined prior literature to identify how the constructs in our research model had been previously operationalized. As suggested by Mackenzie et al. (2011) we adopted existing measurement scales to generate items representing the construct. To assess the content validity of the items, we consulted domain experts to evaluate whether individual items represent an aspect of the content domain of the construct and all items collectively represent of the entire content domain of the construct. (Mackenzie et al. 2011). After incorporating the feedback from the domain experts, we translated the survey items to Turkish and collected a second round of feedback from domain experts about whether translation affected the meaning of the items.

Once the content validity assessments were done, we proceeded by formally specifying the measurement model as suggested by Mackenzie et al. (2011). Table 10 contains the final list of constructs in the survey instrument grouped along with logical grouping of the constructs and type of the measurement models.

To be communicated while introducing the survey instrument, we also defined our focal technologies and validated this definition with domain experts. Consequently, Data Science Advanced Analytics Technologies (DSAAT) are defined as:

- Advanced data management technologies for building highly scalable data systems (i.e., Hadoop, NoSQL Database Systems, Stream Processing Engines)
- New generation mathematical modeling and machine learning platforms and programming languages that enable end-to-end automation of analytics activities (i.e., R, Python, Spark/Scala).

In the remaining part of this section, we will summarize the definitions of the constructs and as well as the adopted scales to operationalize them from the prior literature.

Constructs about technological factors in our proposed model, Relative Advantage, Complexity, Compatibility, and Cost Savings are operationalized by adapting the scales used in Oliveira et al. (2014).

Table 10. List of Constructs and Measurement

Construct	Construct Group	Measurement Type
Relative Advantage	Technological Factors	Reflective
Complexity	Technological Factors	Reflective
Compatibility	Technological Factors	Reflective
Cost Savings	Technological Factors	Reflective
Top Management Belief	Organizational Factors	Reflective
Top Management Participation	Organizational Factors	Reflective
Organizational Slack	Organizational Factors	Reflective
Mimetic Pressure	Environmental Factors	Reflective
External IT Expertise	Environmental Factors	Reflective
Data-Driven Decision-making (DDD)	Inhibiting Factors	Reflective
Lack of Internal Expertise	Inhibiting Factors	Reflective
Adoption Intention		Formative
Actual Adoption / Usage		Formative

Relative Advantage is operationalized as:

- Stem: Please rate your level of agreement for the following when you first considered (“Trial”) using DSAAT

- Scale: 1= *Strongly disagree*, 4=*Neutral*, 7 = *Strongly agree*

- Items

RELADV1: DSAAT allows you to manage business operations in an efficient way.

RELADV2: The use of DSAAT improves the quality of operations.

RELADV3: Using DSAAT allows you to perform specific tasks more quickly.

RELADV4: The use of DSAAT offers new opportunities

RELADV5: Using DSAAT allows you to increase business productivity

Complexity is operationalized as:

- Stem: Please rate your level of agreement for the following when you first considered (“Trial”) using DSAAT
- Scale: 1= *Strongly disagree*, 4=*Neutral*, 7 = *Strongly agree*
- Items

CMPLX1: The use of DSAAT would require a lot of mental effort.

CMPLX2: The use of DSAAT would be frustrating.

CMPLX3: The use of DSAAT would be too complex for business operations.

CMPLX4: The skills needed to adopt DSAAT would be too complex for employees.

Compatibility is operationalized as:

- Stem: Please rate your level of agreement for the following when you first considered (“Trial”) using DSAAT
- Scale: 1= *Strongly disagree*, 4=*Neutral*, 7 = *Strongly agree*
- Items

CMPTBL1: The use of DSAAT would fit the work style of the company.

CMPTBL2: The use of DSAAT would be fully compatible with current business operations.

CMPTBL3: Using DSAAT would be compatible with your company’s corporate culture and value system.

CMPTBL4: The use of DSAAT would be compatible with existing hardware and software in the company.

Cost Savings is operationalized as:

- Stem: Please rate your level of agreement for the following when you first considered (“Trial”) using DSAAT
- Scale: 1= *Strongly disagree*, 4=*Neutral*, 7 = *Strongly agree*
- Items

CSTSAV1: The benefits of DSAAT would be greater than the costs of this adoption.

CSTSAV2: With DSAAT there would be a reduction of energy costs and environmental costs.

CSTSAV3: Maintenance costs of DSAAT would be very low.

Constructs about organizational factors in our proposed model, Top Management Belief and Top Management Participation are operationalized by adapting the scales used in Liang et al. (2007). The other organizational factor Organizational Slack is operationalized by adapting slack scale used in Nohria and Gulati (1996).

Top Management Belief is operationalized as:

- Stem: The senior management of our firm was believing that...
- Scale: 1= *Strongly disagree*, 4=*Neutral*, 7 = *Strongly agree*
- Items

TMBLF1 DSAAT had the potential to provide significant business benefits to the firm.

TMBLF2 DSAAT would create a significant competitive arena for firms.

TMBLF3 it was NOT necessary to use DSAAT to conduct business activities.

Top Management Participation is operationalized as:

- Stem: The senior management of our firm actively...
- Scale: 1= *Strongly disagree*, 4=*Neutral*, 7 = *Strongly agree*

TMPRT1: ...articulated a vision for the organizational use of DSAAT.

TMPRT2: ...formulated a strategy for the organizational use of DSAAT.

TMPRT3: ...established goals and standards to monitor the DSAAT projects.

Organizational Slack is operationalized as:

- Stem: Assume that due to some sudden development, 10% of the time of all people working in your department has to be spent on work totally unconnected with the tasks and responsibilities of your department.
- Scale: 1= *Will not be affected*, 4= *Would fall by about 10 percent, the same as the proposed reduction in resources*, 7= *Will fall by 20% or more*
- Items

ORGSLCK1: How seriously will your output be affected over the next year?

ORGSLCK2: How significantly will your work be affected over the next year?

One of the environmental factors in our proposed model, Mimetic Pressure is operationalized by adapting the scale used in Liang et al. (2007). The other environmental factor Availability of External Expertise is operationalized by

adapting items from Consultant Effectiveness and Vendor Support scales used in Thong et al. (1996).

Mimetic Pressure is operationalized as:

- Stem: Our main competitors who have adopted DSAAT ...
- Scale: 1= *Strongly disagree*, 4= *Neutral*, 7 = *Strongly agree*

MIMPRS1: ...have greatly benefited

MIMPRS2: ...are favorably perceived by others in the same industry

MIMPRS3: ...are favorably perceived by their suppliers and customers

Availability of External IT Expertise is operationalized as:

- Stem: External expertise about DSAAT was available in the market
- Scale: 1= *Strongly disagree*, 4= *Neutral*, 7 = *Strongly agree*

EXTEXP1: For recommending suitable DSAAT solutions

EXTEXP2: For managing implementation of DSAAT solutions

EXTEXP3: For providing technical support for DSAAT solutions

EXTEXP4: For providing training about DSAAT solutions

EXTEXP5: For managing the relationship with other parties in DSAAT solution implementations

One of the inhibiting factors in our proposed model, Data Driven Decision-making is operationalized by adapting the scale used in Brynjolfsson et al. (2011).

The other inhibiting factor Lack of Internal Expertise is operationalized by adapting several items each from Technological Factors and IS Personnel Issues scales used in Ewusi-Mensah and Przasnyski (1991).

Data Driven Decision-making is operationalized as:

- Stem: Please rate your level of agreement for the following arguments
- Scale: 1= *Strongly disagree*, 4= *Neutral*, 7 = *Strongly agree*
- Items

DDD1: We use data for the creation of a new service and/or product.

DDD2: We depend on data to support our decision making (the work practices and environment of the entire company).

DDD3: We have the data we need to make decisions.

Lack of Internal Expertise is operationalized as:

- Stem: Please rate your level of agreement for the following arguments
- Scale: 1= *Strongly disagree*, 4= *Neutral*, 7 = *Strongly agree*
- Items

LCKINTEXP1: DSAAT implementation depended on level of technical expertise lacking in organization

LCKINTEXP2: Internal staff turnover contributed to difficulties in DSAAT implementation

LCKINTEXP3: DSAAT implementation was too ambitious or technically difficult to accomplish

LCKINTEXP4: DSAAT implementation was the first one of its kind or size attempted by the IS/user departments

LCKINTEXP5: Internal staff underestimated the DSAAT information systems requirements

Considering the need for timely insight for the adoption of Data Science and Advanced Analytics Technologies (DSAAT), and complexities of carrying out a longitudinal study in organizational adoption context, we decided to ask both the original intention and the extent of ensuing usage in a cross-sectional research design. For this purpose, we asked participants to rate both their initial intention during the Trial stage and their later Production Use in critical business processes.

We defined Trial and Production Use as:

- Trial: Ad-hoc use, testing or piloting of DSAAT.
- Production Use: Stage in which the DSAAT is operationalized by automating and embedding the results in business processes.

In technology adoption literature, Adoption Intention is used more frequently as the dependent variable rather than Actual Use. Although adoption intention is determined by a richer set of indirect influences including normative pressures and innovation characteristics, actual usage is influenced more by direct experience (Karahanna, Straub, & Chervany 1999). Therefore, taking into account the hype around Data Science and Advanced Analytics Technologies using Adoption Intention as the final target variable would limit the practical value of our research. To mitigate this, in addition to a item to measure adoption intention, we included an item to measure the actual extent of adoption by asking how much DSAAT support critical business processes. To do this we adapted Process Level IT use scale used in Tallon and Pinsonneault (2011) to develop both Adoption Intention and Actual Use items.

Adoption Intention construct is operationalized as:

- Stem: When you first considered using DSAAT, to what extent you expected DSAAT to be used to support key business activities in each of the following business processes?
- Scale: 1= *Low DSAAT Use*, 4= *Medium DSSAAT Use*, 7= *High DSAAT Use*
- Processes: Supplier Relations / Product & Service Enhancement / Productions & Operations / Marketing & Sales / Customer Relations.

Actual Adoption/Usage is operationalized as:

- Stem: Now, to what extent DSAAT is used to support key business activities in each of the following business processes?
- Scale: 1= *Low DSAAT Use*, 4= *Medium DSSAAT Use*, 7= *High DSAAT Use*
- Processes: Supplier Relations / Product & Service Enhancement / Productions & Operations / Marketing & Sales / Customer Relations.

4.3.2 Survey instrument pretesting

Using more rigorous scientific standards especially for data collection procedures have been encouraged by leading researchers (Mitchell, 1985, Straub, 1989).

Although developing and validating a high quality instrument is costly and requires hard work, it is a crucial step in obtaining accurate results, and reliable and valid conclusions. (Boudreau, Gefen, & Straub, 2001; Ju, Chen, Sun, & Wu., 2006).

Various researchers such as Pinsonneault and Kraemer (1993), Grover et al. (1993), and Malhotra and Grover (1998) have proposed criteria for the assessment of instrument quality in survey research. Pretesting in which the draft instrument's validity is qualitatively tested by experts, has been proposed as an important step by

these researchers. Boudreau et al. (2001) propose that regardless of the skill of the researchers every instrument should be pretested before data collection. Alreck, Settle, and Robert (1995) defines a pretest as a preliminary trial of some or all aspects of the instrument to ensure that there are no unanticipated difficulties. In addition, Straub (1989) recommends researchers to assess as many validities as possible during the pretest process.

In some studies, pretested survey instruments from previous studies are used as-is in a similar context with similar respondent characteristics. In this case there is no need to pretest the survey instrument again (Pinsonneault & Kraemer, 1993). In our case, we were adapting measurement scales from a variety of sources and using them together after translating them into another language. So, as suggested by prior literature, we ran several rounds of pretests with interviewees from various profiles such as academic experts in methodology, academic experts in information systems, and practitioners from the industry. After personal interviews with 11 experts from both academia and industry, we modified some of the items in the instrument to clarify meaning and correct identified weaknesses.

We have received feedback about the length of the instrument, the format of the scales, and question ambiguity from our respondents while pretesting the survey. In addition to the feedback regarding measurement scale items, the discussions led to the addition of some more descriptive questions to better understand the responder characteristics and if needed to use these as control variables.

4.3.3 Survey administration

Due to the difficulty of administering pen and paper version of the survey instrument, we decided to administer its digital version via an online survey tool. Recent research documents the prevalence of the use of digital instruments, since they are easier to administer, less costly, and preferred by research participants, while producing similar results (Buchanan & Hvizdak, 2009; Evans & Mathur, 2018).

In order to make sure that the respondents belong to the sample frame, the online survey instrument started with an introduction explaining the purpose of the research as well as qualified respondents. Introduction section of the survey instrument can be found at Appendix B. As a precaution, below set of descriptive questions was added to the survey instrument to be later used for screening:

- Industry
- Number of Employees
- IT Team Size
- DSAA Team Size
- Whether respondent part of IT or Business function
- Title/Level of Respondent
- Title/Level of Respondent's Immediate Manager
- Total number of people reporting to Respondent
- Percentage of DSAA in Overall Responsibilities of Respondent
- How long since the Respondent's Organization initiated "Trial Use" of DSAA

Since the initial list of potential participants would not be enough to reach the required sample size of around 150, while administering the survey we also tried to

implement snowball sampling. We did this by reaching out to each participant separately via phone and personalized electronic messages and asking them to refer us to qualified participants. This way we managed to keep the response rate high and ensure that the additional participants via snowball sampling met the sample frame requirements. Initially, all the participants in the initial potential list of 158 people contacted personally via available contact details and invited to participate in the research survey. In addition to the snowball sampling, a second round of in-depth LinkedIn search carried out. These resulted in additional potential survey respondents of 113 people. These people were also invited to participate in the online survey. Out of the total number of 271 potential contacted participants, 161 responses have been collected with response rate of around 60%. After careful examination and screening out of the problematic responses, we ended up with 140 observations, not too far away from our target sample size of 150.

4.4 Data analysis

Structural equation modeling (SEM) or Path Modeling is a family of statistical techniques that have grown in popularity both in management and social sciences. Although more traditional techniques such as MANOVA and OLS Regression has still their place, these newer techniques allow more flexibility in modelling and analyzing complex phenomena by testing entire theories. In the last two decades, it has become the standard for analyzing complex cause and effect relationships between multiple latent constructs in management research (Hair et al. 2017; Hair, Ringle, & Sarstedt, 2011; Henseler, Hubona, & Ray, 2016; Marcoulides & Saunders, 2006).

For many years, SEM was synonymous with covariance-based SEM (CB-SEM) and until recently the alternative PLS-SEM was a methodology many researchers barely familiar with (Henseler et al., 2016). As the interest in it in management research in general and MIS research in particular increased, it attracted growing criticism for being as a less rigorous method for path modelling in comparison to CB-SEM (Peter, 2018; Peter & Hadavi, 2021). As a parametric technique, CB-SEM requires that its assumptions such as multivariate normality, minimum sample size, and homoscedasticity be met so that it can be used for model estimation (Diamantopoulos & Siguaw 2000; Hair et al., 2011). PLS-SEM, on the other hand is a non-parametric technique with many advantages not offered by CB-SEM (Hair et al. 2011; Henseler et al., 2016; Sarstedt et al., 2022). When it is not possible to satisfy the assumptions of CB-SEM or the research focus is more exploratory in nature, PLS-SEM became the preferred method since it provides more robust estimations of the structural model (Hair et al., 2011; Hair, Hollingsworth, Randolph, & Chong, 2017; Henseler et al., 2016;).

In contrast to covariance based CB-SEM in which you test whether your data fits your whole proposed model in a single estimation procedure, variance-based PLS-SEM works iteratively first creating proxies of latent variables and then estimating the structural model parameters by using partial least squares regressions (Hair et al., 2017; Henseler et al., 2016). The method was originally invented by Wold (1974) for the analysis of high dimensional data. Over the years, it has been extended and underwent various modifications to become a fully-fledged SEM method that can handle construct measurement and path modeling simultaneously (Dijkstra & Henseler, 2016; Hair et al., 2017).

Due to its non-parametric nature and especially its ability to achieve high statistical power even in the case of small sample sizes PLS-SEM started to become the dominant SEM methodology in MIS research. (Henseler et al., 2016; Peter, 2018; Peter & Hadavi 2021). This led leading researchers such as Reinartz et al. (2009) to question dubious justifications of researchers to prefer PLS-SEM over CB-SEM.

SEM methodology researchers such as Reinartz et al. (2009) and Hair et al. (2017) recommends the use of PLS-SEM when the following conditions are encountered:

- Parametric assumptions of CB-SEM are not met due to non-normality or heteroscedasticity
- It is hard to obtain large enough sample sizes due to the nature of sample frame as in the case of organizational technology adoption
- There is complex theoretical model with large number of indicators measuring numerous endogenous and exogenous constructs
- Main objective is to do exploratory analysis for theory building or prediction

Based on these criteria, we decided to use PLS-SEM for the assessment of our measurement models and examining the relationship between latent constructs in our proposed model. This was mainly due to the novelty of many of the characteristics of DSAAT, inevitably requiring an exploratory focus in our research. The other factor that influenced our decision was the ability of PLS-SEM to handle non-normal data and achieve high power even in the case of small sample sizes.

Yet, we are aware of the inherent risks of falling into the trap of ending up with a research model generated through torturing the data, hence fish a model fitting the data. (Goodhue et al., 2013; Peter & Hadavi, 2021). We heeded the precautions of these leading researchers and tried to report all the modifications to the model

with a discipline proposed by Hair et al. (2017). Thus, we claim that we embraced the responsible use of PLS-SEM and are confident that the estimated model adequately maps reality (Chin et al., 2020).

We used SmartPLS to run PLS-SEM analysis and Chapter 5 summarizes the results of the data analysis.

CHAPTER 5

RESULTS

After administering the survey instrument via an online survey tool, we collected the data as outlined in Figure 16. Research Plan. We examined the data for problems and preprocessed it as suggested by the guidelines in Hair et al. (2017). Then, we analyzed the sample characteristics to make sure that our purposive sampling strategy combined with a carefully executed snowball sampling produced a representative subset of the sample frame. Then, for the analysis of whether the data supports our proposed model, we ran PLS-SEM analysis via SmartPLS version 3.3.9. We first evaluated the measurement model for reliability and validity, and after finalizing the measurement model we evaluated the structural model to test the hypotheses in our proposed model. Figure 17 summarizes the steps of the data analysis.

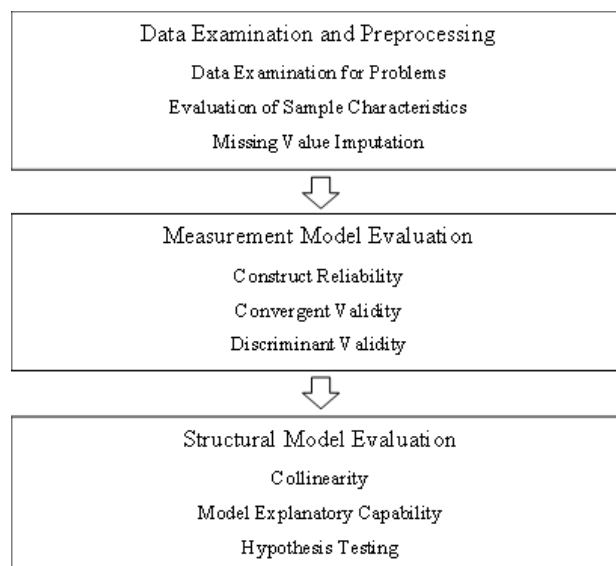


Figure 17. Steps of the data analysis

5.1 Data examination and preprocessing

We started the analysis with 161 observations collected via online survey administration tool. A preliminary analysis of the collected data revealed that data is mostly clean with relatively low missing values. There was however an issue of duplicate observations apparently caused by the online survey administration. After eliminating these redundant responses coming from the same survey respondents, we checked the data for issues such as suspicious response patterns (e.g. straight lining), missingness, and outliers. We first eliminated data with very low variance and more than 6 missing answers case-wise to end up with 140 observations. We imputed the remaining missing values by average of the non-missing indicators belonging to same construct.

Then, we evaluated univariate outliers with a cut-off value of 3.5 standard deviations (Tabachnick, Fidell, & Ulmann 2007). Next, we tested skewness and kurtosis of the data as suggested by Cohen (1960), and checked whether there are problematic values of ± 2 (Tabachnick et al., 1996). Although, there were some items with high skewness or kurtosis, we decided to keep these items also taking into account ability of PLS-SEM to handle non-normal data (Hair et al., 2017). Finally, we checked for multivariate outliers via calculating Mahalanobis distances (Tabachnick & Fidell, 1996). While evaluating Mahalanobis distances, we identified sub-groups with different response patterns, but we did not deem it necessary to eliminate any sub-group. This led us to plan a group moderation analysis for subgroups of business vs IT responses.

In our final sample of 140 observations, 64 comes from respondents self-describing as coming from a business function while the remaining 76 comes from

respondents self-describing as coming from an IT function. Appendix D summarize the descriptive statistics of the items in the survey.

5.2 Sample characteristics

After finalizing the dataset via data examination and pre-processing, we evaluated the sample characteristics to make sure that we have a representative subset of the sample frame. We looked into the breakdown of number of respondents by industry, managerial responsibility, duration since start of trial use, and share of data science and advanced analytics (DSAA) in respondent's total responsibilities. Table 11 lists the number of respondents by industry. In parallel with the initial sample of potential respondents, the final sample included the largest number of respondents coming from banking, telecom, e-commerce, retail, and insurance industries respectively.

Table 11. Breakdown of Survey Respondents by Industry

Industry	Business	IT	Total
Banking	29	47	76
Telecom	6	4	10
E-Commerce	7	2	9
Retail	4	3	7
Insurance	5	1	6
Manufacturing	5		5
Energy	2	1	3
Conglomerate	1	1	2
Other	17	5	22
Total	77	64	140

Although the share of other industries in the final sample was similar to our initial potential participant list, the share of banking industry was considerably

higher. This can be attributable to the higher number of organizations in the banking industry that can afford to adopt new technologies.

Since our unit of analysis is organization, it is also important to evaluate whether the respondents have sufficient seniority and knowledge to answer questions on behalf of their organizations. A review of the breakdown of respondents by their managerial responsibilities in Table 12 indicates that more than 90% of respondents have a managerial responsibility. Table 13 shows the breakdown of respondents by the share of DSAA in overall responsibilities of respondents. More than two thirds of the respondents report that data science and advanced analytics occupy more than 20% of their responsibilities. These two figures suggest that the sample is comprised of individuals who can answer questions on behalf of their organizations, thus satisfying our sample frame requirements.

Table 12. Breakdown of Survey Respondents by Managerial Responsibility

Title	Count
C-Level	40
Director	6
Manager	53
Team Lead	30
Individual Contributor	11
Total	140

As a final check, we look into the breakdown of respondents by the duration since their organization's initial trial use of data science and advanced analytics technologies (DSAAT).

Table 13. Share of DSAA in Overall Responsibilities of Respondent

Percentage of DSAA in Overall Responsibilities	Count
More than 50%	75
20% to 50%	31
Less than 20%	32
Missing	2
Total	140

Table 14 indicates that for more than 90% of the respondents the duration since their initial trial use of DSAAT is more than a year, while for only 2 respondents this duration is less than 6 months. This further suggest that the respondents are coming from organizations with sufficient experience to evaluate actual adoption. Table 15 summarizes descriptive statistics of the companies in our final sample. The median total headcount and IT headcount are 5000 and 350 respectively, while the median DSAA headcount is 30. These numbers also indicate that the sample is comprised of companies coming from information intensive industries.

Table 14. Duration Since Initial Trial Use of DSAAT

Duration	Count
Less than 6 Months	2
6 Months to 1 Year	13
1 to 3 Years	37
3 to 5 Years	30
More than 5 Years	57
Missing	1
Total	140

Table 15. Descriptive Statistics about Survey Respondents

	Count	Mean	Stdev	Min	Q1	Median	Q3	Max
Total Headcount	139	9111	10499	21	1200	5000	13500	60000
IT Headcount	137	632	687	0	100	350	1000	3500
DSAA Headcount	134	57	77	1	10	30	70	500
Number of People Reporting to Respondent	122	28	78	1	5	10	23	700

5.3 Measurement model evaluation

After examining the data for issues such as duplicate observations, straight-lining, low variance, missing values, and outliers, we screened out invalid observations and imputed missing values as suggested by Hair et al. (2017). Next, we evaluated the sample characteristics to make sure that observations satisfy sample frame requirements.

Then, we loaded the data to SmartPLS software to run the PLS-SEM analysis. PLS-SEM analysis allows researchers to complete models (Hair et al., 2011) and requires the specification of both the measurement (also known as outer model) and structural models (also known as inner model) before running the analysis. The common implementation of algorithm then estimates all the model parameters sequentially following the approach suggested by Lohmöller (1989). Iterative estimation of latent construct scores in stage one is followed by calculation of the final parameter estimates via partial OLS regression method.

The initial step of the PLS-SEM analysis is to specify the measurement model. We followed the guidelines suggested by Hair et al. (2017) to specify the measurement model. Although TOE framework does not require specification of higher order constructs, we decided on specifying the measurement model by combining technological, organizational, and environmental factors under three

higher order constructs (Oliviera et al. 2014; Sarstedt, Hair, Cheah, Becker, & Ringle, 2019). This decision was mainly due to higher order constructs' ability to handle multicollinearity and reducing model complexity to ease sample size requirements (Ringle, Sarstedt, & Straub, 2012). After reviewing the adapted measurement scales' original specification in the prior literature, we decided to model them using a reflective-formative measurement perspective using the guidelines provided by Hair et al. (2017).

For the second order constructs, we follow the guidelines provided by Ringle et al. (2012) and Becker, Klein, and Wetzels (2012). It is suggested that when the lower-order constructs are reflectively measured and the higher order constructs do not share common causes but formed to represent general concepts fully, reflective-formative Type II model should be the preferred measurement model specification for higher order constructs. (Becker et al., 2012; Crocetta et al., 2021)

Since inhibiting factors Data Driven Decision Making and Lack of Internal Expertise do not satisfy the usual requirements for being part of a higher order construct, we specified them separately as individual latent constructs having reflective measurement model (Becker et al., 2012; Crocetta et al., 2021). As suggested by prior organization technology adoption literature (Oliviera et al., 2014), one would expect inhibiting factors moderate the effect of adoption intention on the extent of actual adoption. So, we included moderating effects of both Data Driven Decision Making and Lack of Internal Expertise in our model.

As suggested by latent construct measurement literature, we avoided single-item measures even in the case of adoption intention and extent of actual adoption. (Hair et al., 2017). So, due to their multidimensional conceptualization, we specified them as formative endogenous latent variables rather than using sum-scores as latent

exogenous construct scores (Hair et. al., 2017). This decision was also justified as a result of the pretest discussions with experts coming from the industry.

Then, we ran PLS-SEM algorithm using SmartPLS software. We used the algorithmic options and parameter settings at Table 16 such as structural model path weighting method, initial parameter values to start the PLS-SEM algorithm, stop criterion, and maximum number of iterations as suggested by SmartPLS guidelines and Hair et al. (2017).

Table 16. Algorithmic Options and Parameter Settings for Measurement Model

Option / Parameter	Setting
Weighting Method	Path Weighting Scheme
Initial Outer Weights	+1
Stop Criterion	10^{-7}
Maximum Number of Iterations	300
Higher-Order Model Specification	Repeated indicators approach

Figure 18 shows the final measurement model specification. In our final proposed model, three exogenous higher order latent constructs technological factors, organizational factors, and environmental factors are hypothesized as having a positive effect on endogenous latent construct adoption intention, while the positive hypothesized relationship between adoption intent and the other latent endogenous construct adoption intention is hypothesized to be positively moderated by the inhibiting factors extent of data driven decision making and lack of internal expertise.

5.3.1 Construct reliability and convergent validity

PLS-SEM assessment follows a two-step process, starting with an of evaluation of the measurement model followed evaluation of the structural model (Hair et al., 2011). Evaluation of measurement model starts with an assessment of the quality of measurement model via reliability and validity assessments. Any quality issues of the measurement model is resolved before moving on to structural model assessment. Then, if the characteristics of the measurement models are satisfactory the analysis continues with the assessment of structural model results. (Hair et al., 2017).

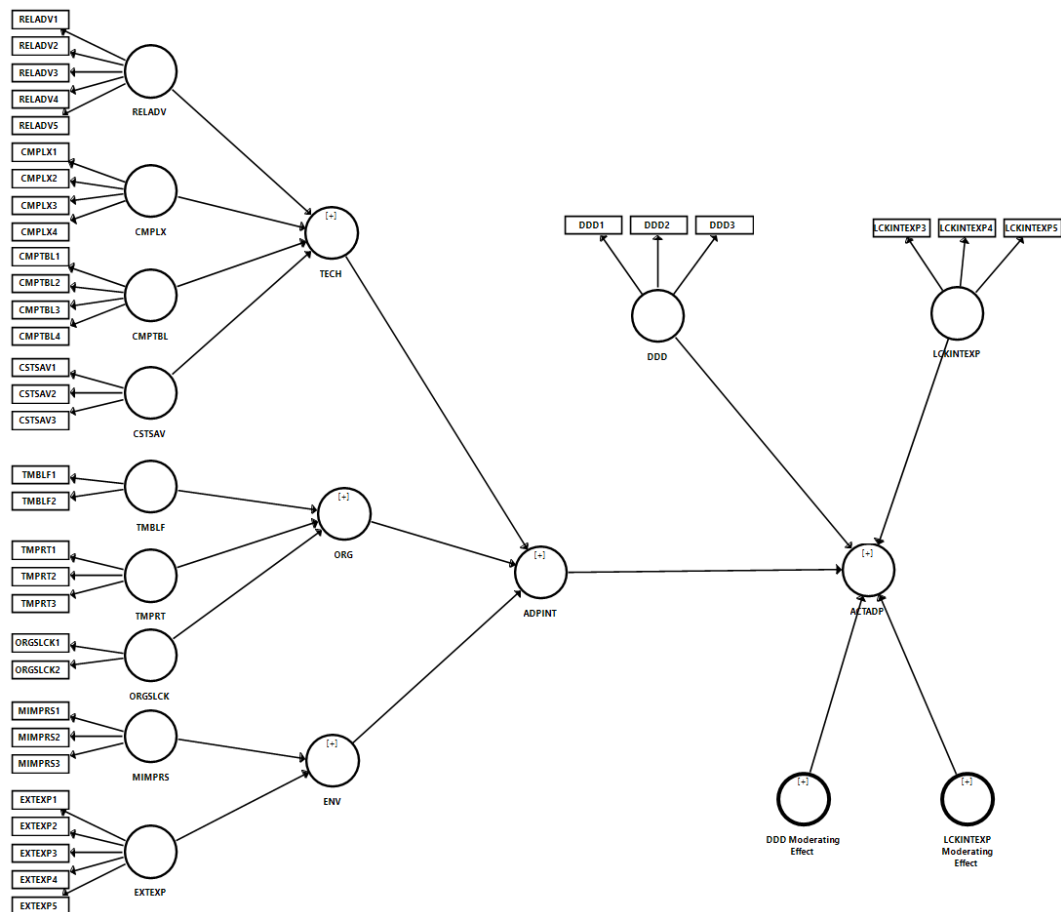


Figure 18. Measurement model specification

Hair et al. (2017) suggests the following rules of thumb for evaluating construct reliability and convergent validity of reflective measurement models.

- Evaluate internal consistency reliability by composite reliability and Cronbach's alpha.
 - Regard Cronbach's alpha as the lower bound and composite reliability as the upper bound of internal consistency reliability.
 - Internal consistency reliability should be higher than 0.70 in confirmatory research. In exploratory research, a value of above 0.60 is acceptable.
- Evaluate indicator reliability by the indicator's outer loadings.
 - Outer loadings should be higher than 0.70. Indicators with outer loadings between 0.40 and 0.70 should be considered for removal.
 - Removal is justified only if it leads to an increase in composite reliability and AVE, and a positive qualitative content validity assessment.
- Evaluate convergent validity by examining the AVE values.
 - AVE should be higher than 0.50.

After an initial evaluation of outer loading values of the reflective measurement model, we identified that all the items were loading to their respective constructs with a loading of at least around 0.6. Only the reverse coded 3rd item in top management team belief was loading poorly with a value of close to 0.2. So, before moving onto further evaluation, we removed this item also taking into account that the other two items' coverage would still preserve content validity of the construct.

Table 17 summarises results of the construct reliability and convergent validity assessments. All the exogenous latent constructs show a Cronbach's Alpha

value of at least 0.6 as suggested by Hair et al. (2017). Only three constructs, have relatively lower Cronbach's Alpha values of less than 0.7: i.e., CSTSAV (0.611), ORGSLCK (0.629) and LCKINTEXP (0.669). Yet, a further look into their composite reliability values of CSTSAV (0.789), ORGSLCK (0.799) and LCKINTEXP (0.819) suggest that these items also exhibit sufficient internal consistency reliability.

Since all the AVE values in Table 17 have values above 0.5 as suggested by Hair et al. (2017), we can conclude that items in our measurement model show sufficient convergent validity. Thus, our measurement model exhibits acceptable levels of internal consistency reliability and convergent validity, allowing us to proceed with discriminant validity assessment.

Table 17. Construct Reliability and Convergent Validity

	Cronbach's Alpha	Composite Reliability	AVE
RELADV	0.942	0.956	0.813
CMPLX	0.809	0.830	0.559
CMPTBL	0.861	0.906	0.708
CSTSAV	0.611	0.789	0.560
TMBLF	0.933	0.967	0.937
TMPRT	0.927	0.953	0.872
ORGSLCK	0.629	0.799	0.676
MIMPRS	0.875	0.923	0.801
EXTEXP	0.933	0.949	0.789
DDD	0.773	0.866	0.683
LCKINTEXP	0.669	0.818	0.603

5.3.2 Discriminant validity

Discriminant validity is defined as the extent to which a construct is truly distinct from other constructs. Thus, establishing discriminant validity of the construct indicates that it is unique and has the ability to capture phenomena not represented by other constructs in the model (Campbell & Fiske, 1959; Hair et al., 2017).

Hair et al. (2017) suggests the following rules of thumb for evaluating discriminant validity of constructs in reflective measurement models.

- Evaluate discriminant validity by the HTMT criterion.
 - The confidence interval of the HTMT statistic should not include the value of 1 for all combinations of constructs.
- Also check the traditional discriminant validity assessment methods.
 - An indicator's outer loadings on its original construct should be higher than all its cross-loadings with other constructs.
 - Square root of the AVE of each construct should be higher than its highest correlation with any other construct (Fornell-Larcker criterion).

Researchers used to rely on the last two traditional measures of discriminant validity: checking whether any of the cross-loadings are higher than loading of associated construct and Fornell-Larcker criterion. So, these are typically the first way to assess the discriminant validity of the indicators. Appendix F clearly shows that none of the items have cross-loading higher than their loading with their associated construct. So, we proceeded to checking whether Fornell-Larcker criterion is satisfied. Table 18 tabulates inter-construct correlations with square root of AVE values on the diagonal. Since diagonal values of the table is significantly higher than

the inter-construct correlations, we conclude that Fornell-Larcker criterion is also satisfied.

Recent research on discriminant validity assessment found that cross-loadings and the Fornell-Larcker criterion might fail to detect discriminant validity issues under certain conditions such as when there is high inter construct correlations and when indicator loadings of the constructs vary only slightly (e.g., all indicator loadings vary between 0.60 and 0.80) (Henseler et al., 2016). Therefore, Henseler et al. (2016) suggest using HTMT as the primary method for discriminant validity assessments. So, to better evaluate discriminant validity, we next checked assessed HTMT ratio of correlations. HTMT is defined as the ratio of the between-trait correlations to the within trait correlations. More specifically, HTMT is the mean of all cross correlations of indicators across different constructs divided by the geometric mean of the average correlations of indicators measuring the same construct (Hair et al., 2017; Henseler et al., 2016). So, higher HTMT values especially above 0.90 indicate a lack of discriminant validity. Table 19 tabulates HTMT values of our measurement model. Apart from a couple of values above 0.5 all the values are well below the recommended threshold of 0.9.

However, since HTMT is a non-parametric statistic, it recommended to run a bootstrap analysis to check significance of HTMT values (Hair et al., 2017; Henseler et al., 2016). We ran a bootstrap analysis to see whether 95% confidence interval for HTMT values include the value of 1. We found that there were no constructs for which HTMT confidence interval includes the value of 1, further suggesting that our measurement model satisfies all discriminant validity criteria suggested by literature.

Table 18. Discriminant Validity - Fornell-Larcker Criterion

	RELADV	CMPLX	CMPTBL	CSTSAV	TMBLF	TMPRT	ORGLCK	MIMPRS	EXTEXP	DDD	LCKINTEXP
RELADV	0.901										
CMPLX	0.333	0.748									
CMPTBL	0.512	0.247	0.841								
CSTSAV	0.543	0.161	0.365	0.748							
TMBLF	0.386	0.057	0.462	0.259	0.968						
TMPRT	0.291	0.061	0.433	0.165	0.620	0.934					
ORGLCK	0.009	-0.077	0.039	0.197	0.092	0.038	0.822				
MIMPRS	0.456	0.169	0.335	0.330	0.242	0.217	-0.217	0.895			
EXTEXP	0.279	0.069	0.328	0.150	0.381	0.319	0.104	0.238	0.888		
DDD	0.315	0.020	0.457	0.235	0.336	0.431	0.063	0.213	0.299	0.827	
LCKINTEXP	0.159	0.369	0.059	0.225	-0.059	0.074	-0.031	0.257	-0.060	0.076	0.777

Since there are higher order reflective-formative latent constructs, an evaluation of collinearity of the first order constructs within the second order constructs would be necessary to make sure that the model would produce reliable parameter estimates. Hair et al. (2017) suggests that VIF values should be less than 5 among indicators, and if that is not the case eliminating indicators. In addition, in the case of higher order formative constructs one should consider whether a reflective-reflective measurement model would be a better alternative to mitigate the collinearity problems. Table 20 summarize the measurement model assessment criteria, while Table 21 shows multicollinearity (VIF) statistics of the formative second order latent constructs. Finally, you can see relative contribution of first order constructs in the higher order latent constructs at Table 22. These results suggest that we have a measurement model that satisfy quality criteria suggested in literature, and we can proceed with structural model evaluation.

Table 19. Heterotrait-Monotrait Ratio (HTMT)

	RELADV	CMPLX	CMPTBL	CSTSAV	TMBLF	TMPRT	ORGLCK	MIMPRS	EXTEXP	DDD
RELADV										
CMPLX	0.247									
CMPTBL	0.567	0.223								
CSTSAV	0.668	0.200	0.478							
TMBLF	0.412	0.090	0.514	0.313						
TMPRT	0.312	0.097	0.483	0.229	0.664					
ORGLCK	0.133	0.142	0.136	0.299	0.112	0.058				
MIMPRS	0.499	0.166	0.381	0.402	0.264	0.239	0.334			
EXTEXP	0.298	0.211	0.363	0.190	0.409	0.343	0.112	0.261		
DDD	0.368	0.076	0.558	0.313	0.401	0.516	0.100	0.249	0.346	
LCKINTEXP	0.200	0.550	0.134	0.352	0.093	0.129	0.139	0.353	0.174	0.154

Table 20. Measurement Model Evaluation Summary

Latent Variable	Indicators	Convergent Validity			Internal Consistency Reliability		Discriminant Validity
		Loadings	Indicator Reliability	AVE	Composite Reliability	Cronbach's Alpha	
		>0.70	>0.50	>0.50	0.60-0.90	0.60-0.90	HTMT CI Not Incl. 1
RELADV	RELADV1	0.907	0.822	0.813	0.956	0.942	YES
	RELADV2	0.915	0.838				
	RELADV3	0.914	0.836				
	RELADV4	0.841	0.707				
	RELADV5	0.927	0.860				
CMPLX	CMPLX1	0.915	0.837	0.559	0.830	0.809	YES
	CMPLX2	0.818	0.669				
	CMPLX3	0.563	0.317				
	CMPLX4	0.642	0.412				
CMPTBL	CMPTBL1	0.844	0.712	0.559	0.830	0.809	YES
	CMPTBL2	0.897	0.805				
	CMPTBL3	0.848	0.720				
	CMPTBL4	0.771	0.594				
CSTSAV	CSTSAV1	0.643	0.413	0.560	0.789	0.611	YES
	CSTSAV2	0.872	0.760				
	CSTSAV3	0.711	0.506				
TMBLF	TMBLF1	0.967	0.935	0.937	0.967	0.933	YES
	TMBLF2	0.969	0.939				
TMPRT	TMPRT1	0.921	0.848	0.872	0.953	0.927	YES
	TMPRT2	0.953	0.908				
	TMPRT3	0.927	0.860				
ORGSLOCK	ORGSLOCK1	0.980	0.961	0.676	0.799	0.629	YES
	ORGSLOCK2	0.625	0.391				
MIMPRS	MIMPRS1	0.853	0.727	0.801	0.923	0.875	YES
	MIMPRS2	0.907	0.823				
	MIMPRS3	0.923	0.852				
EXTEXP	EXTEXP1	0.885	0.783	0.789	0.949	0.933	YES
	EXTEXP2	0.937	0.878				
	EXTEXP3	0.923	0.852				
	EXTEXP4	0.843	0.711				
	EXTEXP5	0.848	0.720				
DDD	DDD1	0.787	0.619	0.683	0.866	0.773	YES
	DDD2	0.878	0.771				
	DDD3	0.812	0.659				
LCKINTEXP	LCKINTEXP3	0.811	0.657	0.603	0.818	0.669	YES
	LCKINTEXP4	0.650	0.422				
	LCKINTEXP5	0.855	0.730				

Table 21. Measurement Model Collinearity Statistics (VIF)

	TECH	ORG	ENV
RELADV	1.794		
CMPLX	1.136		
CMPTBL	1.389		
CSTSAV	1.441		
TMBLF		1.635	
TMPRT		1.624	
ORGLCK		1.009	
MIMPRS			1.060
EXTEXP			1.060

Table 22. Higher-Order Constructs Indicator Weights

	TECH	ORG	ENV
RELADV	0.615		
CMPLX	0.113		
CMPTBL	0.354		
CSTSAV	0.172		
TMBLF		0.445	
TMPRT		0.658	
ORGLCK		0.033	
MIMPRS			0.283
EXTEXP			0.894

5.4 Structural model evaluation

PLS-SEM is known as variance based path modeling technique in contrast to more traditional covariance based CB-SEM. Under the hood the technique essentially uses a sequence of principal component analysis (PCA) and ordinary least squares (OLS) regression runs to estimate both the measurement and structural models in an iterative fashion as part of a larger estimation framework. So, these underlying techniques determine its statistical properties (Hair et al., 2011; Henseler et al., 2016). In essence, the method after estimating the parameters of the reflective and formative measurement model, tries to maximize the explained variance in the endogenous latent variables by running partial OLS regressions. So, the focus is on the prediction of a specific set of hypothesized relationships rather than on explanations. (Hair et al., 2011; Henseler et al., 2016).

On the other hand, unlike CB-SEM, PLS-SEM is non-parametric method in nature. Although being an extension of OLS regression analysis for complex path models, it does not make any parametric assumptions regarding both data and the residuals, which is the case in regression analysis (Hair et al., 2017; Sarstedt & Mooi, 2014).

Due to this non-parametric nature, it relies on bootstrapping to test the significances of the model path coefficients as is the case for other non-parametric techniques. Since, PLS-SEM does not assume an underlying distribution, it produces a distribution from the data using bootstrapping, and then use it for significance testing (Hair et al., 2017).

Moreover, since it runs a series of partial OLS regressions for smaller subsets of the larger complex model, PLS-SEM generally achieve higher statistical power levels than CB-SEM. (Hair et al., 2011; Reinartz et al., 2009). These less restrictive

requirements about data characteristics and model specification make PLS-SEM an ideal method for testing path models in a less researched exploratory setting (Hair et al., 2017).

The concept of model fit, which is the first criterion to evaluate the quality of covariance based CB-SEM model, is not directly transferable to variance based PLS-SEM. In CB-SEM, there are numerous goodness-of-fit measures based on chi-squared based goodness-of-fit measures and their extensions. Although there are some researchers such as Henseler et al. (2016) endorsing the use of goodness-of-fit measures for PLS-SEM, there is lack of comprehensive assessment of their applicability for PLS-SEM. Other researchers such as Hair, Risher, and Sarstedt (2019) caution about the tentative nature of the recent literature and recommends not using even the bootstrap based model fit assessments such as SRMR suggested by Henseler et al. (2016). Methodology researchers also question whether the concept of model fit is even valuable in the context of PLS-SEM applications (Hair et al., 2017; Lohmöller, 1989). Hence, Shmueli, Ray, Estrada, and Chatla (2016) recently proposed new set of evaluation procedures that are more in line with prediction-oriented nature of PLS-SEM.

So, after finalizing measurement model evaluation and we proceeded to structural model evaluation to test the hypotheses in our proposed model. Our proposed model includes adoption intention and actual adoption as two endogenous latent variables. Three reflective-formative high-order latent variables representing technological, organizational, and environmental factors are hypothesized positively to influence organizational adoption intention. Adoption intention is in turn hypothesized to positively influence the extent of actual adoption.

There are also two hypothesized inhibiting factors, namely not having a data driven decision making culture and lack of internal expertise. Since data driven decision making is used in affirmative form in prior literature, we decided to follow the same approach and keep the measurement scale as affirmative. Hence, we hypothesized that having data driven decision making positively moderates the effect of adoption intention on the extent of actual adoption, while lack of internal expertise negatively moderates the effect of adoption intention on the extent of actual adoption.

Since, structural model evaluation is dependent on the quality of the bootstrapping procedure, we carefully selected all the options as suggested by recent literature (Hair et al., 2017; Hair et al., 2019). We then ran bootstrap analysis for the complete dataset as well as sub-groups of business and IT respondents.

Table 23. Parameter Settings for Bootstrapping

Option / Parameter	Setting
Number of Bootstrap Samples	5000
Sign Change Option	No (gives most conservative results)
Confidence Interval Method	Bias-corrected and Accelerated (BCa)
Test type	Two-tailed
Significance level	0.05

Hair et al. (2017) suggests the following rules of thumb for evaluating discriminant validity of constructs in reflective measurement models.

- Do not start your evaluation with traditional model fit assessments such as GoF of CB-SEM

- Proceed with structural model evaluation only after assessing the quality of the reflective and formative measurement models
- Examine all endogenous and exogenous latent constructs in the structural model for multicollinearity.
 - Each latent construct's variance inflation factor (VIF) value should be higher lower than 5.
 - If not consider eliminating constructs or creating higher-order constructs to treat collinearity problems.
- Evaluate the significance of path estimates and explained variance of endogenous constructs
 - Use bootstrapping results to assess the significance of path coefficients
 - Use R2 values to assess explained variance
- Examine bootstrap confidence intervals to further evaluate stability of path coefficient estimates
 - Use the percentile method for constructing confidence intervals
- Use adjusted R2 while comparing models specified with different set of constructs or observations
- Evaluate effect of exogenous constructs' contribution to R2 values endogenous constructs
 - Effect size (f^2) values of 0.02, 0.15, and 0.35 indicate a small, medium, or large effect, respectively
- Optionally, evaluate predictive relevance of constructs via cross-validated redundancy measures for each endogenous construct.
 - Use blindfolding procedure to create cross-validated Q2 values

- Check whether exogenous constructs have predictive relevance for the endogenous constructs
- Optionally, evaluate effect of exogenous constructs' contribution to Q^2 values endogenous constructs
 - Effect size (q^2) values of 0.02, 0.15, and 0.35 indicate a small, medium, or large predictive relevance, respectively

After running the bootstrap analysis, we started our evaluation by collinearity assessment of the inner model. Table 24 shows that all the variance inflation factor (VIF) values are well below the suggested upper value of 5. This suggests that there is no collinearity issue that would threaten the stability of path estimates.

Table 24. Structural Model Collinearity Statistics (VIF)

	ADPINT	ACTADP
ORG	1.350	
TECH	1.365	
ENV	1.318	
ADPINT		1.169
DDD		1.084
DDD Moderating Effect		1.180
LCKINTEXP		1.147
LCKINTEXP Moderating Effect		1.372

Then, we proceeded our analysis by examining the structural model path coefficients. Figure 19 shows the graphical representation of our structural model with the estimated path coefficients. Table 25 presents summary statistics about the structural model path coefficients such as t-values, p-values, and bias corrected

confidence intervals in addition to the original path coefficient estimates for the complete set of 140 observations. In addition, Table 26 shows the R2 values of two endogenous latent constructs adoption intention and extent of actual adoption.

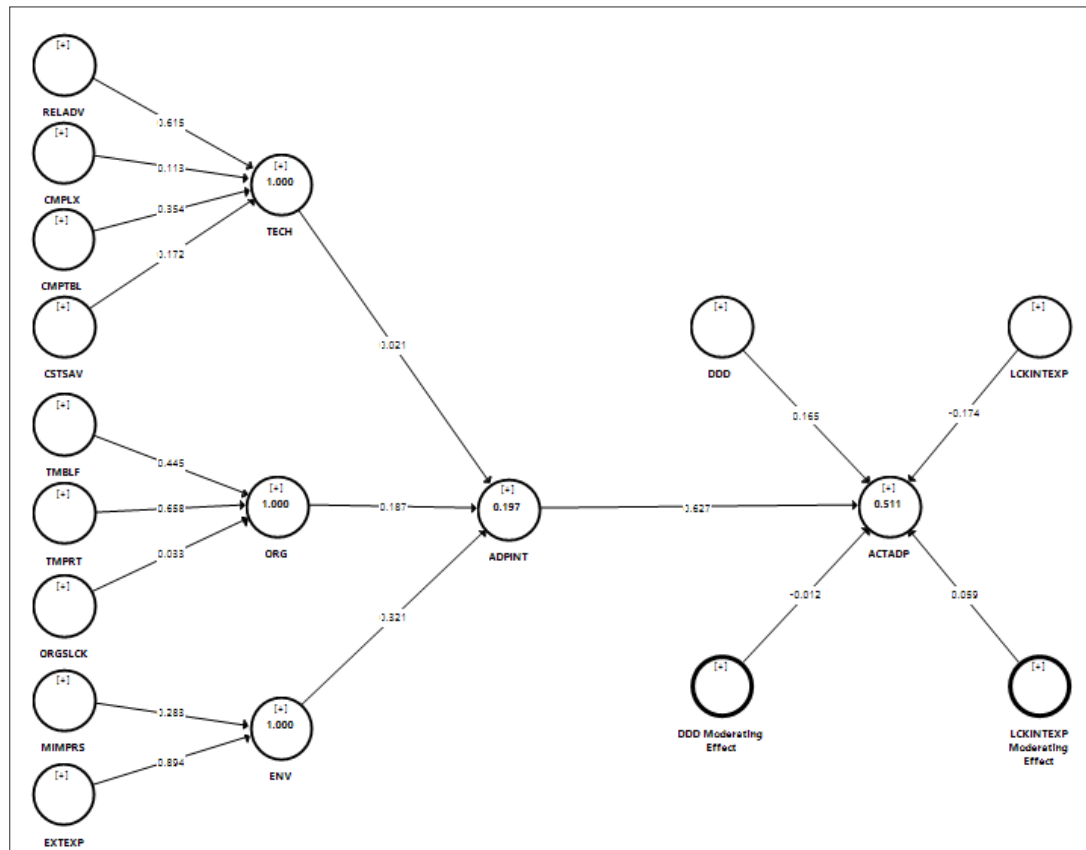


Figure 19. Structural model (n=140, complete observations)

Table 25. Structural Model Path Coefficients

	Path Coefficients	t values	p values	95% CI	Support
TECH -> ADPINT	0.021	0.134	0.894	[-0.288, 0.309]	No
ORG -> ADPINT	0.187	1.213	0.225	[-0.171, 0.45]	No
ENV -> ADPINT	0.321	2.288	0.022	[0.016, 0.523]	Yes
DDD -> ACTADP	0.627	4.672	0.000	[0.324, 0.804]	Yes
DDD Mod. -> ACTADP	0.165	1.124	0.261	[-0.075, 0.449]	No
LCKINTEXP -> ACTADP	-0.174	1.475	0.140	[-0.357, 0.059]	No
LCKINTEXP Mod. -> ACTADP	0.059	0.712	0.477	[-0.123, 0.203]	No
ADPINT -> ACTADP	0.894	16.869	0.000	[0.788, 0.985]	Yes

Table 26. Explained Variance of Endogenous Constructs

	R Square	R Square Adjusted
ACTADP	0.511	0.493
ADPINT	0.197	0.179

Before moving onto testing the hypothesis via evaluation of path coefficients, we evaluated R² values to assess how much exogenous variables collectively explain the variance of endogenous variables. Falk and Miller (1992) suggested that R² values should be at least greater than 0.10, While Hair et al. (2019) recommends R² values of 0.75, 0.50, and 0.25 as strong, moderate, and weak respectively, Chin (1998) earlier considered R² values of 0.67, 0.33 and 0.19 as substantial, moderate and weak, respectively (Roldan & Sanchez-Franco, 2012). Other scholars also caution against putting too much emphasis on R² values to evaluate theoretical adequacy of a model (Hair et al., 2019). Our model has R² values of 0.197 and 0.511 for adoption intention and extent of actual adoption respectively. We can consider our model as explaining the variance of adoption intention as relatively weakly. however, the extent of adoption is explained relatively strongly. These R² values suggest that our theoretical model has sufficient theoretical adequacy.

After evaluating the theoretical adequacy of our model via explained variances of the exogenous latent constructs, we moved onto testing of the hypothesized relationships in our structural model. An examination of bootstrap statistics of the path coefficient estimates at Table 25 suggest that only effect of environmental factors is significant with a p-value of 0.022, while the effect of technological factors and organizational factors are not significant. This finding further motivated us to estimate models for sub-groups of business and IT respondents. We evaluated the effect of categorical control variable (the respondent

being a member of business or IT function of their organizations) by estimating separate PLS-SEM models with the same measurement and structural path model. Table 27 summarizes explained variance values of endogenous latent constructs for business and IT sub-groups. Both models indicate theoretical adequacy similar to original path model estimated for the whole set of observations.

Table 27. Coefficient of Determination (R^2 Value)

Business	R Square	R Square Adjusted
ACTADP	0.548	0.509
ADPINT	0.352	0.319

IT	R Square	R Square Adjusted
ACTADP	0.726	0.707
ADPINT	0.187	0.153

An examination of the path coefficients of separate PLS-SEM model runs for business and IT respondents at Table 28 suggests that there is considerable variation in terms of the responses of these two subgroups. While path estimate values for technological factors is still insignificant for both business and IT respondents, organizational factors became significant at alpha value of 10% with a p value of 0.099.

On the other hand, examination of inhibiting factors' path coefficients reveals that on aggregate data driven decision making has a direct strongly significant effect at an alpha of 1%. When examined at the subgroup level, this direct effect of data driven decision making culture on extent of adoption intention has a partially significant direct effect for business respondents with a p-value of 0.103. However, for IT respondents moderating effect of lack of internal expertise became significant at an alpha of 5% with a p-value of 0.045. However, it is important to note that since

these figures are achieved with smaller sample sizes of 64 for business respondents and 76 for IT respondents, these effects have less statistical power.

Table 28. Multi-Group Analysis Structural Model Path Coefficients

Business (n=64)					
	Path Coefficients	t values	p values	95% CI	Support
TECH -> ADPINT	-0.033	0.138	0.890	[-0.405, 0.622]	No
ORG -> ADPINT	0.383	1.649	0.099	[-0.167, 0.743]	Yes
ENV -> ADPINT	0.310	1.839	0.066	[-0.057, 0.618]	Yes
DDD -> ACTADP	0.323	1.632	0.103	[-0.098, 0.659]	No
DDD Mod. -> ACTADP	0.002	0.024	0.980	[-0.169, 0.236]	No
LCKINTEXP -> ACTADP	-0.200	1.261	0.207	[-0.409, 0.208]	No
LCKINTEXP Mod. -> ACTADP	0.048	0.458	0.647	[-0.211, 0.207]	No
ADPINT -> ACTADP	0.491	2.668	0.008	[0.193, 0.881]	Yes
IT (n=76)					
	Path Coefficients	t values	p values	95% CI	Support
TECH -> ADPINT	0.020	0.080	0.936	[-0.467, 0.484]	No
ORG -> ADPINT	0.194	0.679	0.497	[-0.599, 0.545]	No
ENV -> ADPINT	0.330	1.184	0.237	[-0.173, 0.69]	No
DDD -> ACTADP	0.154	0.967	0.334	[-0.087, 0.554]	No
DDD Mod. -> ACTADP	0.100	0.851	0.395	[-0.024, 0.513]	No
LCKINTEXP -> ACTADP	-0.142	0.948	0.343	[-0.399, 0.178]	No
LCKINTEXP Mod. -> ACTADP	0.267	1.953	0.051	[0.219, 0.485]	Yes
ADPINT -> ACTADP	0.472	2.005	0.045	[-0.923, 0.65]	Yes

As suggested by Hair et al. (2019), we also examined the effect size statistics to evaluate effect of exogenous constructs' contribution to R2 values endogenous constructs. Table 29 show the effect size (f2) values for the complete set of observations as well as for business and IT respondent sub-groups. As suggested by prior literature, adoption intention has quite a strong effect on extent of actual adoption for all three models. While environmental factors have a moderate effect for

adoption intention in all three models, data driven decision making has a stronger effect in business respondent model. Moderating effect of lack of internal expertise is quite strong with an effect size value of 0.564 in IT respondent model. This can be attributable to the fact that difficulty of finding experts with right skill sets is mainly felt by senior IT professionals in our sample.

Table 29. Effect Size f^2

	Complete		Business		IT	
	ADPINT	ACTADP	ADPINT	ACTADP	ADPINT	ACTADP
TECH	0.000		0.001		0.000	
ORG	0.032		0.124		0.042	
ENV	0.097		0.095		0.113	
DDD		0.052		0.168		0.082
DDD Moderating Effect		0.000		0.000		0.017
LCKINTEXP		0.054		0.081		0.065
LCKINTEXP Moderating Effect		0.006		0.006		0.564
ADPINT		0.687		0.409		0.621

CHAPTER 6

CONCLUSION

6.1 Summary of findings

The aim of this research was to propose critical factors for Data Science and Advanced Analytics Technologies (DSAAT) adoption based on the prior organizational technology adoption literature, and to empirically test the influence of these factors to find most relevant ones.

Based on the data collected from 140 organizational respondents, we used PLS-SEM methodology to test the effect of the facilitating influence of technological, organizational, and environmental factors on the adoption intention, and inhibiting influence of the extent of data driven decision making and the lack of internal expertise on the extent of actual adoption.

Our results suggest that environmental factors i.e., mimetic pressure and the extent of external expertise play a significant role in influencing development of adoption intention, while there is no support for the influence technological factors. Technology characteristics are known to be more important in early stages of innovation cycle of a specific organization (Hameed et al., 2012). The lack of support for the influence of technological factors might be due to the rapid adoption of these new technologies leading to what Rogers (1995) call early majority phase. If this is the case role of technology characteristics can be less important than organizational and environmental characteristics as seen in earlier technology adoption studies. Rogers (1995) suggests that for some technologies it is harder to show the results leading to low observability. Grover et al. (2018) argues that there is growing scepticism whether big data analytics provide value. So, after the initial

difficulties encountered, organizations may have already understood that one needs much more than technology itself. As suggested by Grover et al. (2018), there is a need for organizations to employ VRIO analysis to better understand the dimensions of this capability.

There is also considerable variation among the perceptions of respondents coming from business and IT functions. The results of respondents coming from business functions suggest that the extent of data driven decision making has a positive effect on the extent of actual adoption, while the lack of internal expertise moderates the effect of adoption intention on the extent of actual adoption for respondents coming from IT functions

6.2 Research contributions and managerial implications

Considering the emergence of big data and hence Data Science and Advanced Analytics Technologies allowing extracting value from big data, we feel that developing an adoption model and empirically testing it would be a steppingstone for further research on understanding the successful adoption and usage of these technologies in the future. Despite all the publicity around big data and heightened expectations about the potential value of the application of Data Science and Advanced Analytics Technologies, practitioners in industry have been finding it difficult to realize the value of big data.

This research indicates that there is a significant difference in the perceptions of individuals coming from business and IT functions. While the lack data driven decision making culture is seen as an inhibiting factor by respondents coming from business functions, respondents coming from IT functions see the effect of difficulty of finding appropriate talent as a roadblock. So, for top management teams

reconciling these differing views among their members could help facilitate value creation using big data.

6.3 Limitations and future research directions

This research examined the factors leading to organizational adoption of Data Science and Advance Analytics Technologies in a single developing country setting in Turkey. So, the results may not hold true for the context of other countries.

As is the case in organizational adoption field survey based research, we approached senior representatives of the organizations who might represent the organizations' views about the use of these technologies. Yet, there is the potential for bias as these people's views might not accurately represent their organizations' view as a whole.

The lack of support for the influence of technological factors can also be due to the cross-sectional nature of the data collection process. The resulting variance based research model might not be able to capture the influence of the technological factors that may develop over time. So, a longitudinal study adopting a process based approach would have captured the influence of technological factors.

Extent of actual adoption is also collected as part of the survey instrument as a perceived self-reported measure. There is a need for a proxy measure collected from secondary data to mitigate the self-reporting bias.

So, we believe that findings of this research should be replicated in other context and triangulated by other qualitative research methods such as a multi-site positivist case study.

APPENDIX A

ETHICS COMMITTEE APPROVAL

T.C.

BOĞAZİÇİ ÜNİVERSİTESİ

Sosyal ve Beşeri Bilimler Yüksek Lisans ve Doktora Tezleri Etik İnceleme Komisyonu

Sayı: 2019-34

28 Mart 2019

Hüseyin Tolga Durdu
İşletme

Sayın Araştırmacı,

"Veri Bilimi ve İleri Analitik Teknolojilerinin Organizasyonel Kullanımına Etki Eden Faktörler" başlıklı projeniz ile ilgili olarak yaptığımız SSB-EAK 2019/05 sayılı başvuru komisyonumuz tarafından 28 Mart 2019 tarihli toplantıda incelenmiş ve uygun bulunmuştur.

.....
Dr. Öğr. Üyesi İnci Aşkar

.....
Prof. Dr. Feyza Çorapçı

.....
Doç. Dr. Mehmet Yiğit Güdral

.....
Doç. Dr. Ebru Kaya

.....
Dr. Öğr. Üyesi Şebnem Yalçın

APPENDIX B

SURVEY INSTRUMENT INTRODUCTION

Items	Description
Research Summary	Kurumlar tarafından toplanan veri miktarı giderek artarken Veri Bilimi ve İleri Analitik teknolojilerinin kullanımı konusunda da artan bir ilgi olduğunu görüyoruz. Bu ilgiye rağmen bazı yenilikçi ve öncü kurumlar dışında bu teknolojilerin ve ilgili tekniklerin yaygın bir şekilde kullanılmaya başlamalarının beklenenden uzun sürdüğünü gözlemliyoruz. Bu araştırmanın amacı Veri Bilimi ve İleri Analitik Teknolojilerinin organizasyonlar tarafından benimsenmesi ve etkin olarak kullanılmasına etki eden kritik faktörlerin belirlenmesidir. Bu nedenle teknolojik, organizasyonel ve çevresel faktörler ile benimsenme ölçütlerinin arasında ilişkinin inceleneceği bir anket yapılmaktadır.
Definitions	Bu araştırma kapsamında Veri Bilimi ve İleri Analitik Teknolojileri (VBİAT) aşağıdaki şekilde tanımlanmaktadır: - İleri Veri Yönetimi Teknolojileri: Yüksek oranda ölçeklenebilir veri sistemleri kurmak için kullanılan. (Hadoop, NoSQL Database, Stream Processing Engines vb) - Yeni jenerasyon Matematiksel Modelleme ve Makine Öğrenmesi Platformları ve Programlama Dilleri: Tüm analitik aktivitelerinin uçtan uca otomasyonunu mümkün kılan. (R, Python, Spark/Scala)
Target Industries	Bu araştırma bilgi yoğun sektörlerdeki bilgi teknolojileri ve analitik uygulamaları müşterilerine ürün ve hizmet sunmak için kullanan şirketlere yöneliktir. Bu nedenle bu teknolojileri kendileri geliştiren, satan, destekleyen veya müşterilerinin bu teknolojileri uygulamaya alması için yardımcı olan teknoloji şirketleri araştırma kapsamı dışında tutulmuştur. Araştırmanın içermesi hedeflenen ana sektörler bankacılık, sigorta, telekom, perakende zincirleri, elektronik ticaret, enerji ve üretimdir. Yine de bu teknolojileri kullandığı bilinen başka sektörlerden bazı firmalar da örnekleme dahil edilmiştir.

Items	Description
Eligible Respondents	Uygun katılımcılar iş ve bilgi teknolojileri ekiplerinden, bölümlerinin ana sorumluluklarından birisi Veri Bilimi ve İleri Analitik Teknolojilerinin direkt olarak kullanımı olan orta ve üst düzey yöneticilerdir.
Research Policy	Araştırma süreci boyunca bütün cevaplarınız kesin gizlilik prensibi ile saklanacaktır. Tüm cevaplarınız anonim olarak değerlendirilecek ve sadece toplu olarak analiz edilecektir. İletişim bilgilerinizi paylaşmanız durumunda özet araştırma sonuçları tarafınız ile paylaşılacaktır.
How to Answer	Lütfen aşağıdaki soruları organizasyonunuzun Veri Bilimi ve İleri Analitik Teknolojilerini (VBİAT) kullanımını dikkate alarak cevaplayınız. Veri Bilimi ve İleri Analitik Teknolojilerinin benimsenmesi ve yaygın kullanımına etki eden faktörleri belirlemek için kurumunuzun bu teknolojileri kullanmayı değerlendirme ("Deneme") aşamasındaki durumu ile bu teknolojilerin halihazırda kritik iş süreçlerinizde kullanım seviyesi ("Gerçek Kullanım") hakkında bilgiye ihtiyacımız var. Bu anket sorularını cevaplarken aşağıdaki tanımları dikkate alınız. • Deneme: VBİAT'nin düzensiz ("ad hoc"), test veya pilot proje amacı ile kullanımı • Gerçek Kullanım: VBİAT'nin otomatize edilmesi ve iş süreçleri ile entegre edilerek kullanımı.

APPENDIX C

SURVEY INSTRUMENT

Constructs		Survey Item	Scale
Technological Factors		Aşağıdaki ifadelere katılma durumunuzu VBİAT'lerini kullanmayı ilk değerlendirdiğiniz ("Deneme") dönemi düşünerek belirtiniz	
Relative Advantage	RELADV1	VBİAT'nin iş operasyonlarınızı etkin bir şekilde yönetmenize imkan sağlayacağını düşünüyorsunuz.	(1= kesinlikle katılmıyorum; 7 = kesinlikle katılıyorum)
	RELADV2	VBİAT kullanımının operasyonlarınızın kalitesini artıracağını düşünüyorsunuz.	
	RELADV3	VBİAT'nin belirli işlerinizi hızlandıracağını düşünüyorsunuz.	
	RELADV4	VBİAT'nin yeni iş fırsatları yaratacağını düşünüyorsunuz.	
	RELADV5	VBİAT'nin üretkenliğinizi artıracağını düşünüyorsunuz.	
Complexity	CMPLX1	VBİAT kullanımının çok fazla mental efor gerektireceğini düşünüyorsunuz.	
	CMPLX2	VBİAT kullanımının beklenenden zor olabileceğini düşünüyorsunuz.	
	CMPLX3	VBİAT'nin iş operasyonlarınız için çok karmaşık olabileceğini düşünüyorsunuz.	

	CMPLX4	VBİAT kullanmak için gerekli olan yetkinliklerin çalışanlarınız için çok karmaşık olabileceğini düşünüyordunuz.	
Compatibility	CMPTBL1	VBİAT kullanmanın kurumunuzun çalışma şekline uyacağını düşünüyordunuz.	
	CMPTBL2	VBİAT'nin mevcut iş süreçleriniz ile tamamiyle uyumlu olacağını düşünüyordunuz.	
	CMPTBL3	VBİAT kullanmanın kurum kültürünüz ve değerlerinizle uyumlu olacağını düşünüyordunuz.	
	CMPTBL4	VBİAT'nin kurumunuzun mevcut yazılım ve donanımları ile uyumlu olacağını düşünüyordunuz.	
Cost Savings	CSTSAV1	VBİAT'nin faydalarının kullanmaya başlama maliyetlerinden fazla olacağını düşünüyordunuz.	
	CSTSAV2	VBİAT ile maliyetlerde düşüş olacağını bekliyordunuz.	
	CSTSAV3	VBİAT'nin bakım ve idame maliyetlerinin çok düşük olacağını bekliyordunuz.	
Organizational Factors		Aşağıdaki ifadelere katılma durumunuzu VBİAT'lerini kullanmayı ilk değerlendirdiğiniz ("Deneme") dönemi düşünerek belirtiniz	
Top Management Belief		Kurumumuzun üst yönetimi VBİAT'nin	(1= kesinlikle katılmıyorum; 7 = kesinlikle katılıyorum)
	TMBLF1	...kurumumuz için kaydadeğer bir iş değeri yaratma potansiyeline sahip olduğuna inanıyordu.	

	TMBLF2	...kurumlar için yeni bir rekabet alanı yaratacağını düşünüyordu.	
	TMBLF3	...iş faaliyetlerini sürdürmek için VBİAT'nin gerekli olmadığı kanaatindeydi.	
Top Management Participation		Kurumumuzun üst yönetimi aktif bir şekilde...	(1= kesinlikle katılmıyorum; 7 = kesinlikle katılıyorum)
	TMPRT1	...VBİAT'nin kullanımı için bir organizasyonel vizyona sahipti.	
	TMPRT2	...VBİAT'nin organizasyonel kullanımı için bir strateji geliştirmişti.	
	TMPRT3	...VBİAT projelerinin değerlendirilmesi için hedefler ve standartlar belirlemişti.	
Organizational Slack	ORGLCK1	Beklenmedik bir gelişme nedeni ile bölümünüzde çalışan kişilerin zamanlarının %10'unu bölümünüz görev tanımı ile ilgili olmayan işlere harcamak zorunda kaldıklarını varsayın. Bu durumda bölümünüzün iş çıktıları önümüzdeki yıl boyunca nasıl etkilenirdi?	1, "çıktı etkilenmez" 4, "çıktı da kaynaklardaki düşüş gibi %10 düşer" 7, "çıktı %20 veya daha fazla düşer"
	ORGLCK2	Yine benzer bir şekilde bölümünüzün faaliyet bütçesinin %10'u oranında azaltıldığını varsayın. Bu durumda bölümünüz işleri ne oranda etkilenirdi?	1, "etkilenmez" 4, "kaynaklardaki düşüş gibi %10 etkilenir" 7, "%20 veya daha fazla etkilenir"
Environmental Factors		Aşağıdaki ifadelere katılma durumunuzu VBİAT'lerini kullanmayı ilk değerlendirdiğiniz ("Deneme") dönemi düşünerek belirtiniz	
Mimetic Pressure		VBİAT'yi kullanmaya başlayan ana rakiplerimiz...	(1= kesinlikle katılmıyorum; 7 = kesinlikle katılıyorum)
	MIMPRS1	...kayda değer fayda sağladılar.	

	MIMPRS2	...sektördeki diğer kurumlar tarafında daha olumlu algılanmaya başladılar.	
	MIMPRS3	...tedarikçi ve müşterileri tarafından daha olumlu algılanmaya başladılar.	
		Piyasada VBİAT ile ilgili aşağıdaki konulardaki dış uzmanlığa erişebiliyoruz:	
External IT Expertise	EXTEXP1	Doğru VBİAT çözümlerinin önerilmesi için.	(1= kesinlikle katılmıyorum; 7 = kesinlikle katılıyorum)
	EXTEXP2	VBİAT çözümlerinin geliştirilmesi ve uygulamaya alınması için.	
	EXTEXP3	VBİAT çözümleri ile ilgili teknik destek için.	
	EXTEXP4	VBİAT ile ilgili eğitim çözümleri için.	
	EXTEXP5	VBİAT çözümleri uygulama projelerinde paydaşların yönetimi için.	
Inhibiting Factors		Aşağıdaki ifadelere katılma durumunuzu belirtiniz	
Lack of Internal Expertise	LCKINTEXP1	VBİAT'nin geliştirilme ve uygulamaya alınması çalışmaları kurumda bulunmayan bir teknik deneyim gerektiriyordu.	(1= kesinlikle katılmıyorum; 7 = kesinlikle katılıyorum)
	LCKINTEXP2	Çalışanların işten ayrılmaları VBİAT geliştirilmesi ve uygulamaya alınması konusunda zorluklara sebep oldu.	
	LCKINTEXP3	VBİAT geliştirilme ve uygulama alınması çalışmaları çok iddialıydı veya teknik olarak çok zordu.	
	LCKINTEXP4	VBİAT geliştirilme ve uygulamaya alınması çalışmaları kurum açısından boyut ve zorluk olarak bir ilk olma özelliği taşıyordu.	

	LCKINTEXP5	Çalışanlar için VBİAT geliştirilme ve uygulama alınması çok iddialıydı veya teknik olarak çok zordu.	
Data-Driven Decision-making (DDD)	DDD1	Yeni bir ürün ya da servis yaratmak için veriyi analiz ederiz.	(1= kesinlikle katılmıyorum; 7 = kesinlikle katılıyorum)
	DDD2	Kurumumuzun tüm faaliyetlerinde kararlarımızı veriye dayalı olarak veririz.	
	DDD3	Karar vermek için ihtiyacımız olan veriye sahibiz.	
Adoption Measures			
Adoption Intention		VBİAT'ni kullanmayı ilk düşündüğünüzde VBİAT'nin aşağıdaki süreçlerinizi ne ölçüde destekleyeceğini düşünüyordunuz.	(1: Çok Az; 7: Çok Fazla)
	ADPINT	Tedarikçi Yönetimi / Ürün ve Hizmet Geliştirme / Üretim ve Operasyon / Pazarlama ve Satış / Müşteri İlişkileri Yönetimi / Finans ve Risk	
Actual Adoption / Usage		Halihazırda VBİAT'nin aşağıdaki süreçlerinizi ne ölçüde destekleyediğini düşünüyorsunuz.	
	ACTADP	Tedarikçi Yönetimi / Ürün ve Hizmet Geliştirme / Üretim ve Operasyon / Pazarlama ve Satış / Müşteri İlişkileri Yönetimi / Finans ve Risk	

APPENDIX D

DESCRIPTIVE STATISTICS OF SURVEY RESPONSES

Construct	Indicator	Mean	Stdev	Skewness	Kurtosis
Relative Advantage	RELADV1	6.16	1.14	-1.83	3.95
Relative Advantage	RELADV2	6.18	1.10	-1.88	4.22
Relative Advantage	RELADV3	6.22	1.08	-1.68	2.75
Relative Advantage	RELADV4	6.06	1.35	-1.73	2.70
Relative Advantage	RELADV5	6.30	1.07	-2.30	6.58
Complexity	CMPLX1	5.09	1.51	-0.69	-0.09
Complexity	CMPLX2	4.75	1.61	-0.66	-0.11
Complexity	CMPLX3	3.78	1.79	0.25	-1.10
Complexity	CMPLX4	4.41	1.62	-0.26	-0.91
Compatibility	CMPTBL1	5.27	1.42	-0.84	-0.02
Compatibility	CMPTBL2	4.65	1.56	-0.25	-0.96
Compatibility	CMPTBL3	5.15	1.41	-0.70	-0.07
Compatibility	CMPTBL4	4.37	1.58	-0.23	-0.70
Cost Savings	CSTSAV1	5.22	1.72	-0.81	-0.41
Cost Savings	CSTSAV2	5.30	1.46	-0.76	-0.06
Cost Savings	CSTSAV3	4.08	1.60	0.18	-0.90
Top Management Belief	TMBLF1	5.69	1.43	-1.21	0.86
Top Management Belief	TMBLF2	5.79	1.43	-1.43	1.63
Top Management Belief	TMBLF3	5.07	1.82	-0.85	-0.33
Top Management Participation	TMPRT1	4.89	1.78	-0.68	-0.60
Top Management Participation	TMPRT2	4.41	1.82	-0.31	-1.02
Top Management Participation	TMPRT3	4.03	1.78	-0.09	-0.91
Organizational Slack	ORGSLCK1	3.86	1.93	-0.03	-1.15
Organizational Slack	ORGSLCK2	4.32	1.75	-0.42	-0.60
Mimetic Pressure	MIMPRS1	5.42	1.37	-0.81	0.43
Mimetic Pressure	MIMPRS2	5.55	1.34	-1.18	1.81

Construct	Indicator	Mean	Stdev	Skewness	Kurtosis
Mimetic Pressure	MIMPRS3	5.39	1.36	-0.95	1.12
External IT Expertise	EXTEXP1	5.06	1.40	-0.46	-0.52
External IT Expertise	EXTEXP2	5.14	1.33	-0.60	-0.29
External IT Expertise	EXTEXP3	5.10	1.40	-0.77	0.27
External IT Expertise	EXTEXP4	5.35	1.31	-0.83	0.34
External IT Expertise	EXTEXP5	4.97	1.38	-0.46	-0.57
Data-Driven Decision-making	DDD1	5.99	1.20	-1.45	2.22
Data-Driven Decision-making	DDD2	5.32	1.44	-0.98	0.54
Data-Driven Decision-making	DDD3	5.48	1.33	-1.02	0.85
Lack of Internal Expertise	LCKINTEXP1	5.49	1.43	-1.20	1.16
Lack of Internal Expertise	LCKINTEXP2	4.62	1.68	-0.53	-0.57
Lack of Internal Expertise	LCKINTEXP3	4.35	1.46	-0.29	-0.70
Lack of Internal Expertise	LCKINTEXP4	4.82	1.64	-0.47	-0.58
Lack of Internal Expertise	LCKINTEXP5	4.13	1.45	-0.17	-0.81
Adoption Intention	ADPINT1	4.54	1.42	-0.28	0.18
Adoption Intention	ADPINT2	5.92	1.04	-0.87	0.67
Adoption Intention	ADPINT3	5.74	1.13	-1.26	2.86
Adoption Intention	ADPINT4	6.36	0.94	-1.67	3.10
Adoption Intention	ADPINT5	6.13	0.97	-1.12	1.33
Adoption Intention	ADPINT6	5.96	1.17	-1.32	2.11
Actual Adoption	ACTADP1	3.96	1.58	-0.36	-0.17
Actual Adoption	ACTADP2	5.30	1.29	-0.78	0.86
Actual Adoption	ACTADP3	5.25	1.40	-0.71	0.49
Actual Adoption	ACTADP4	5.89	1.22	-1.28	1.85
Actual Adoption	ACTADP5	5.46	1.39	-0.60	-0.33
Actual Adoption	ACTADP6	5.54	1.49	-1.20	1.31

APPENDIX E

MEASUREMENT MODEL OUTER LOADINGS

	RELADV	CMPLX	CMPTBL	CSTSAV	TMBLF	TMPRT	ORGLCK	MIMPRS	EXTEXP	DDD	LCKINTEXP
RELADV1	0.907										
RELADV2	0.915										
RELADV3	0.914										
RELADV4	0.841										
RELADV5	0.927										
CMPLX1		0.915									
CMPLX2		0.818									
CMPLX3		0.563									
CMPLX4		0.642									
CMPTBL1			0.844								
CMPTBL2			0.897								
CMPTBL3			0.848								
CMPTBL4			0.771								
CSTSAV1				0.643							
CSTSAV2				0.872							
CSTSAV3				0.711							
TMBLF1					0.967						
TMBLF2					0.969						
TMPRT1						0.921					

	RELADV	CMPLX	CMPTBL	CSTSAV	TMBLF	TMPRT	ORGLCK	MIMPRS	EXTEXP	DDD	LCKINTEXP
TMPRT2						0.953					
TMPRT3						0.927					
ORGLCK1							0.980				
ORGLCK2							0.625				
MIMPRS1								0.853			
MIMPRS2								0.907			
MIMPRS3								0.923			
EXTEXP1									0.885		
EXTEXP2									0.937		
EXTEXP3									0.923		
EXTEXP4									0.843		
EXTEXP5									0.848		
DDD1										0.787	
DDD2										0.878	
DDD3										0.812	
LCKINTEXP3											0.811
LCKINTEXP4											0.650
LCKINTEXP5											0.855

APPENDIX F

CROSS LOADINGS

	RELADV	CMPLX	CMPTBL	CSTSAV	TMBLF	TMPRT	ORGSCLK	MIMPRS	EXTEXP	DDD	LCKINTEXP
RELADV1	0.907	0.309	0.517	0.505	0.382	0.292	0.044	0.362	0.296	0.272	0.138
RELADV2	0.915	0.295	0.457	0.479	0.336	0.227	0.025	0.370	0.205	0.199	0.155
RELADV3	0.914	0.249	0.405	0.506	0.288	0.259	-0.022	0.446	0.251	0.298	0.111
RELADV4	0.841	0.340	0.450	0.404	0.384	0.290	-0.037	0.442	0.262	0.327	0.138
RELADV5	0.927	0.310	0.476	0.547	0.351	0.248	0.023	0.442	0.243	0.325	0.176
CMPLX1	0.405	0.915	0.325	0.195	0.117	0.114	-0.083	0.181	0.183	0.055	0.291
CMPLX2	0.182	0.818	0.124	0.080	-0.020	0.001	-0.033	0.107	-0.038	-0.033	0.325
CMPLX3	0.014	0.563	-0.008	0.029	-0.044	-0.031	-0.035	0.079	-0.220	-0.042	0.324
CMPLX4	0.058	0.642	-0.044	0.040	-0.063	-0.084	-0.067	0.066	-0.140	-0.006	0.339
CMPTBL1	0.453	0.288	0.844	0.228	0.446	0.371	0.041	0.305	0.295	0.404	0.068
CMPTBL2	0.429	0.192	0.897	0.312	0.381	0.433	-0.003	0.296	0.291	0.366	-0.003
CMPTBL3	0.466	0.229	0.848	0.352	0.397	0.356	0.039	0.283	0.285	0.398	0.075
CMPTBL4	0.368	0.109	0.771	0.338	0.325	0.289	0.057	0.239	0.226	0.370	0.060
CSTSAV1	0.372	0.252	0.101	0.643	0.088	-0.036	0.162	0.130	0.003	0.017	0.154
CSTSAV2	0.554	0.116	0.382	0.872	0.296	0.246	0.128	0.385	0.234	0.281	0.236
CSTSAV3	0.217	-0.007	0.294	0.711	0.146	0.094	0.178	0.154	0.029	0.180	0.078
TMBLF1	0.364	0.024	0.463	0.294	0.967	0.582	0.112	0.201	0.362	0.359	-0.051

	RELADV	CMPLX	CMPTBL	CSTSAV	TMBLF	TMPRT	ORGLCK	MIMPRS	EXTEXP	DDD	LCKINTEXP
TMBLF2	0.383	0.085	0.432	0.210	0.969	0.617	0.066	0.267	0.376	0.293	-0.063
TMPRT1	0.287	0.002	0.413	0.193	0.631	0.921	0.005	0.170	0.325	0.430	0.027
TMPRT2	0.279	0.091	0.378	0.141	0.609	0.953	0.086	0.224	0.295	0.392	0.102
TMPRT3	0.249	0.078	0.423	0.126	0.489	0.927	0.013	0.215	0.272	0.385	0.080
ORGLCK1	0.043	-0.038	0.070	0.203	0.090	0.050	0.980	-0.187	0.120	0.084	-0.021
ORGLCK2	-0.132	-0.195	-0.103	0.082	0.056	-0.026	0.625	-0.238	-0.009	-0.047	-0.055
MIMPRS1	0.433	0.129	0.397	0.346	0.218	0.181	-0.116	0.853	0.242	0.200	0.122
MIMPRS2	0.329	0.209	0.232	0.232	0.158	0.152	-0.190	0.907	0.157	0.108	0.275
MIMPRS3	0.450	0.125	0.261	0.300	0.265	0.243	-0.274	0.923	0.232	0.249	0.298
EXTEXP1	0.303	0.077	0.312	0.138	0.338	0.262	0.084	0.247	0.885	0.278	-0.112
EXTEXP2	0.255	0.074	0.346	0.132	0.357	0.295	0.117	0.215	0.937	0.298	-0.049
EXTEXP3	0.154	0.066	0.282	0.100	0.300	0.279	0.064	0.164	0.923	0.269	-0.063
EXTEXP4	0.194	0.049	0.223	0.138	0.303	0.223	0.069	0.179	0.843	0.132	-0.041
EXTEXP5	0.334	0.038	0.286	0.163	0.396	0.357	0.127	0.254	0.848	0.345	0.000
DDD1	0.256	-0.019	0.357	0.201	0.283	0.378	0.002	0.148	0.218	0.787	0.074
DDD2	0.253	0.013	0.405	0.169	0.350	0.425	0.096	0.200	0.260	0.878	-0.022
DDD3	0.270	0.042	0.370	0.212	0.213	0.284	0.048	0.173	0.256	0.812	0.130
LCKINTEXP3	0.111	0.250	0.082	0.213	0.018	0.052	-0.005	0.187	-0.075	0.106	0.811
LCKINTEXP4	0.107	0.296	0.101	0.065	0.002	0.138	-0.142	0.257	0.115	0.064	0.650
LCKINTEXP5	0.151	0.324	-0.022	0.217	-0.138	0.012	0.035	0.180	-0.129	0.014	0.855

REFERENCES

- Abbasi, A., Sarker, S., & Chiang, R. H. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, 17(2), 3.
- Agarwal, R., & Dhar, V. (2014). Big data, data science, and analytics: The opportunity and challenge for IS research. *Information Systems Research*, 25(3), 443-448.
- Agrawal, K. (2015). Investigating the determinants of big data Analytics (BDA) adoption in Asian emerging economies. In *AMCIS 2015* (pp. 4425-4435).
- Ajah, I. A., & Nweke, H. F. (2019). Big data and business analytics: Trends, platforms, success factors and applications. *Big Data and Cognitive Computing*, 3(2), 32.
- Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In *Action Control* (pp. 11-39). Springer, Berlin, Heidelberg.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179-211.
- Akter, S., D'ambra, J., & Ray, P. (2011). An evaluation of PLS based complex models: the roles of power analysis, predictive relevance and GoF index. *AMCIS 2011 Proceedings - All Submissions*. 151.
- Alreck, P. L., Settle, R. B., & Robert, S. (1995). *The survey research handbook*. McGraw-Hill/Irwin.
- Al-Sai, Z. A., Abdullah, R., & Husin, M. H. (2020). Critical success factors for big data: a systematic literature review. *IEEE Access*, 8, 118940-118956.
- Baesens, B., Bapna, R., Marsden, J. R., Vanthienen, J., & Zhao, J. L. (2016). Transformational issues of big data and analytics in networked business. *MIS Quarterly*, 40(4), 807-818.
- Baker, J. (2012). The technology–organization–environment framework. *Information Systems Theory*, 231-245.

- Banker, R. D., & Kauffman, R. J. (1988). Strategic contributions of information technology: an empirical study of ATM networks. *NYU Working Paper No. IS-88-25*
- Barcikowski, R. S. (1981). Statistical power with group mean as the unit of analysis. *Journal of Educational Statistics*, 6(3), 267-285.
- Barclay, D., Higgins, C., & Thompson, R. (1995). The partial least squares (PLS) approach to casual modeling: personal computer adoption and use as an Illustration. *Technology Studies*, 2: 285–309
- Baroudi, J. J., & Orlikowski, W. J. (1989). The problem of statistical power in MIS research. *MIS Quarterly*, 13(1), 87-106.
- Barton, D., & Court, D. (2012). Making advanced analytics work for you. *Harvard Business Review*, 90(10), 78-83.
- Basole, R. C., Seuss, C. D., & Rouse, W. B. (2013). IT innovation adoption by enterprises: Knowledge discovery through text analytics. *Decision Support Systems*, 54(2), 1044-1054.
- Becker, J. M., Klein, K., & Wetzels, M. (2012). Hierarchical latent variable models in PLS-SEM: guidelines for using reflective-formative type models. *Long Range Planning*, 45(5-6), 359-394.
- Booz Allen Hamilton. (2015). *The field guide to data science*. Retrieved from <https://www.boozallen.com/s/insight/publication/field-guide-to-data-science.html>
- Booz Allen Hamilton. (2017). *Data science playbook*. Retrieved from <https://www.boozallen.com/s/insight/publication/data-science-playbook.html>
- Boudreau, M. C., Gefen, D., & Straub, D. W. (2001). Validation in information systems research: A state-of-the-art assessment. *MIS Quarterly*, 25(1), 1-16.
- Bradford, M., & Florin, J. (2003). Examining the role of innovation diffusion factors on the implementation success of enterprise resource planning systems. *International Journal of Accounting Information Systems*, 4(3), 205-225.

- Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). Strength in numbers: How does data-driven decision making affect firm performance?. *Available at SSRN 1819486*.
- Buchanan, E. A., & Hvizdak, E. E. (2009). Online survey tools: Ethical and methodological concerns of human research ethics committees. *Journal Of Empirical Research on Human Research Ethics*, 4(2), 37-48.
- Burt, R. S. (1987). Social contagion and innovation: Cohesion versus structural equivalence. *American Journal of Sociology*, 92(6), 1287-1335.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81.
- Cao, L. (2017). Data science: a comprehensive overview. *ACM Computing Surveys (CSUR)*, 50(3), 1-42.
- Cao, L. (2017). Data science: challenges and directions. *Communications of the ACM*, 60(8), 59-68.
- Capgemini, (2019). *The data-powered enterprise*. Capgemini.
- Chau, P. Y., & Tam, K. Y. (1997). Factors affecting the adoption of open systems: an exploratory study. *MIS Quarterly*, 21(1), 1-24.
- Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4), 1165-1188
- Chen, W., & Hirschheim, R. (2004). A paradigmatic and methodological examination of information systems research from 1991 to 2001. *Information Systems Journal*, 14(3), 197-235.
- Chin, W. W. (1998). The partial least squares approach to structural equation modeling. *Modern Methods for Business Research*, 295(2), 295-336.
- Chin, W., Cheah, J. H., Liu, Y., Ting, H., Lim, X. J., & Cham, T. H. (2020). Demystifying the role of causal-predictive modeling using partial least squares structural equation modeling in information systems research. *Industrial Management & Data Systems*. 120(12), 2161-2209

- Choudrie, J., & Dwivedi, Y. K. (2005). Investigating the research approaches for examining technology adoption issues. *Journal of Research Practice*, 1(1), D1-D1.
- Churchill, G.A. (1991), *Marketing research: Methodological foundations*, 5th ed., Dryden Press, Orlando, FL.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155.
- Corea, F. (2016). *Big data analytics: A management perspective* (Vol. 21). Springer.
- Crocetta, C., Antonucci, L., Cataldo, R., Galasso, R., Grassia, M. G., Lauro, C. N., & Marino, M. (2021). Higher-order PLS-PM approach for different types of constructs. *Social Indicators Research*, 154(2), 725-754.
- Cuadrado-Gallego, J. J., & Demchenko, Y. (2020). *The data science framework. A view from the EDISON project*.
- Davenport, T. H. (2013). Analytics 3.0. *Harvard Business Review*, 91(12), 64-72.
- Davenport, T. H. (2018). From analytics to artificial intelligence. *Journal of Business Analytics*, 1(2), 73-80.
- Davenport, T. H., & Patil, D. J. (2012). Data scientist: the sexiest job of the 21st century. *Harvard Business Review*, 90(10), 70–76.
- Davis, F. D. (1986). *A technology acceptance model for empirically testing new end-user information systems: Theory and results* (Doctoral dissertation, Massachusetts Institute of Technology).
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.
- 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.
- De Mauro, A., Greco, M., & Grimaldi, M. (2015, February). What is big data? A consensual definition and a review of key research topics. In *AIP Conference Proceedings* (Vol. 1644, No. 1, pp. 97-104). American Institute of Physics.

- Dhar, V. (2013). Data science and prediction. *Communications of the ACM*, 56(12), 64-73.
- Diamantopoulos, A. & Siguaw, J. A. (2000). *Introducing LISREL: A guide for the uninitiated*. Sage.
- Dijkstra, T. K., & Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Quarterly*, 39(2), 297-316.
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147-160.
- Donoho, D. (2017). 50 years of data science. *Journal of Computational and Graphical Statistics*, 26(4), 745-766.
- Economist (2010). Data, data everywhere. *The Economist Newspaper Limited*. Retrieved from <https://www.economist.com/special-report/2010/02/27/data-data-everywhere>
- Economist (2017). The world's most valuable resource is no longer oil, but data. *The Economist: New York, NY, USA*. Retrieved from <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>
- Evans, J. R., & Mathur, A. (2018). The value of online surveys: A look back and a look ahead. *Internet Research*. 28(4), 854-887
- Ewusi-Mensah, K., & Przasnyski, Z. H. (1991). On information systems project abandonment: an exploratory study of organizational practices. *MIS Quarterly*, 15(1), 67-86.
- Falk, R. F., & Miller, N. B. (1992). *A primer for soft modeling*. University of Akron Press.
- Fortune. (2015). *McKinsey buys formula one data analytics firm QuantumBlack*. Retrieved from <https://fortune.com/2015/12/14/mckinsey-buys-quantumblack/>
- Galliers, R. D., & Land, F. F. (1987). Choosing appropriate information systems research methodologies. *Communications of the ACM*, 30(11), 901-902.

- Gallivan, M. J. (2001). Organizational adoption and assimilation of complex technological innovations: development and application of a new framework. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, 32(3), 51-85.
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144.
- Gartner (2012). *The nexus of forces: Social, mobile, cloud and information*. Retrieved from <https://www.gartner.com/en/documents/2049315>.
- Goodhue, D. L., Lewis, W., & Thompson, R. (2012). Does PLS have advantages for small sample size or non-normal data?. *MIS Quarterly*, 36(3), 981-1001.
- Grover, V. (1993). An empirically derived model for the adoption of customer-based interorganizational systems. *Decision Sciences*, 24(3), 603-640.
- Grover, V., Chiang, R. H., 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.
- Grover, V., Lee, C. C., & Durand, D. (1993). Analyzing methodological rigor of MIS survey research from 1980–1989. *Information & Management*, 24(6), 305-317.
- Günther, W. A., Mehrizi, M. H. R., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems*, 26(3), 191-209.
- Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, 53(8), 1049-1064.
- 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), 107-123.
- Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, 19(2), 139-152.
- 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

- Hair, J., Hollingsworth, C. L., Randolph, A. B., & Chong, A. Y. L. (2017). An updated and expanded assessment of PLS-SEM in information systems research. *Industrial Management & Data Systems*, 117(3), 442-458
- Hair, Joe, et al. *A primer on partial least squares structural equation modeling (PLS-SEM)*. 2nd ed., SAGE Publications, 2017.
- Halper, F., & Krishnan, K. (2013). TDWI big data maturity model guide. *TDWI Research*, 2014, 1-20.
- Hameed, M. A., Counsell, S., & Swift, S. (2012). A conceptual model for the process of IT innovation adoption in organizations. *Journal of Engineering and Technology Management*, 29(3), 358-390.
- Haveman, H. A. (1993). Follow the leader: Mimetic isomorphism and entry into new markets. *Administrative Science Quarterly*, 593-627.
- Henseler, J. (2007). *A new and simple approach to multi-group analysis in partial least squares path modelling*. Retrieved from <https://repository.ubn.ru.nl/bitstream/handle/2066/160875/160875.pdf>
- Henseler, J., & Chin, W. W. (2010). A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling. *Structural Equation Modeling*, 17(1), 82-109.
- Henseler, J., & Fassott, G. (2010). Testing moderating effects in PLS path models: An illustration of available procedures. In *Handbook of partial least squares* (pp. 713-735). Springer, Berlin, Heidelberg.
- Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: updated guidelines. *Industrial Management & Data Systems*, 116(1), 2-2
- Iansiti, M., & Nadella, S., (2022). Democratizing Transformation. *Harvard Business Review*.
- Inamdar, Z., Raut, R., Narwane, V. S., Gardas, B., Narkhede, B., & Sagnak, M. (2020). A systematic literature review with bibliometric analysis of big data analytics adoption from period 2014 to 2018. *Journal of Enterprise Information Management*, 34(1), 101-139

- Jenkins, A. M. (1985). Research methodologies and MIS research. *Research Methods in Information Systems*, 2(1), 103-117.
- Ju, T. L., Chen, Y. Y., Sun, S. Y., & Wu, C. Y. (2006). Rigor in MIS survey research: In search of ideal survey methodological attributes. *Journal of Computer Information Systems*, 47(2), 112-123.
- Karahanna, E., Straub, D. W., & Chervany, N. L. (1999). Information technology adoption across time: a cross-sectional comparison of pre-adoption and post-adoption beliefs. *MIS Quarterly*, 23(2), 183-213.
- Karanja, E., & Zaveri, J. (2013). A comprehensive review of survey- based research in MIS. *Journal of Systems and Information Technology*. 15(2), 159-188
- Kock, N., & Hadaya, P. (2018). Minimum sample size estimation in PLS- SEM: The inverse square root and gamma- exponential methods. *Information Systems Journal*, 28(1), 227-261.
- 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. *The International Journal of Logistics Management*, 29(2), 676-703.
- Lewis, B. R., Templeton, G. F., & Byrd, T. A. (2005). A methodology for construct development in MIS research. *European Journal of Information Systems*, 14(4), 388-400.
- 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.
- Lohmöller, J. B. (1989). Predictive vs. structural modeling: Pls vs. ml. In *Latent variable path modeling with partial least squares* (pp. 199-226). Physica, Heidelberg.
- Lucas Jr, H. C. (1989). *Methodological issues in information systems survey research*. NYU Working Paper No. IS-89-046, Available at SSRN: <https://ssrn.com/abstract=1289691>
- MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, 35(2), 293-334.

- Malhotra, M. K., & Grover, V. (1998). An assessment of survey research in POM: from constructs to theory. *Journal of Operations Management*, 16(4), 407-425.
- Marcoulides, G. A., & Saunders, C. (2006). Editor's comments: PLS: a silver bullet?. *MIS Quarterly*, 30(2), iii-ix.
- Maroufkhani, P., Ismail, W. K. W., & Ghobakhloo, M. (2020). Big data analytics adoption model for small and medium enterprises. *Journal of Science and Technology Policy Management*. 11(4), 483-513
- McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: the management revolution. *Harvard Business Review*, 90(10), 60-68.
- McKinsey (2011) *Big Data: The Next Frontier for Innovation, Competition, and Productivity*, Retrieved from <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation>
- McKinsey (2016). *The age of analytics: competing in a data-driven world*. Retrieved from <https://www.mckinsey.com/business-functions/quantumblack/our-insights/the-age-of-analytics-competing-in-a-data-driven-world>
- McKinsey (2022). *The data-driven enterprise of 2025*. Retrieved from <https://www.mckinsey.com/business-functions/quantumblack/our-insights/the-data-driven-enterprise-of-2025>
- Memon, M. A., Cheah, J. H., Ramayah, T., Ting, H., Chuah, F., & Cham, T. H. (2019). Moderation analysis: issues and guidelines. *Journal of Applied Structural Equation Modeling*, 3(1), 1-11.
- Mikalef, P., Pappas, I., Krogstie, J., & Pavlou, P. A. (Eds.). (2020). *Big data and business analytics: A research agenda for realizing business value*. Norway: Elsevier.
- Mitchell, T. R. (1985). An evaluation of the validity of correlational research conducted in organizations. *Academy of Management Review*, 10(2), 192-205.
- Molinillo, S., & Japutra, A. (2017). Organizational adoption of digital information and technology: a theoretical review. *The Bottom Line*. 30(1), 33-46.

- Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research*, 2(3), 192-222.
- Newsted, P. R., Huff, S. L., & Munro, M. C. (1998). Survey instruments in information systems. *MIS Quarterly*, 22(4), 553.
- Nohria, N., & Gulati, R. (1996). Is slack good or bad for innovation?. *Academy of Management Journal*, 39(5), 1245-1264.
- Oliveira, T., & Martins, M. F. (2010). Information technology adoption models at firm level: review of literature. In *The European Conference on Information Systems Management* (p. 312). Academic Conferences International Limited.
- Oliveira, T., Thomas, M., & Espadanal, M. (2014). Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors. *Information & Management*, 51(5), 497-510.
- Orlikowski, W. J., & Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research*, 2(1), 1-28.
- Palvia, P., Daneshvar Kakhki, M., Ghoshal, T., Uppala, V., & Wang, W. (2015). Methodological and topic trends in information systems research: A meta-analysis of IS journals. *Communications of the Association for Information Systems*, 37(1), 30.
- Palvia, P., Leary, D., Mao, E., Midha, V., Pinjani, P., & Salam, A. F. (2004). Research methodologies in MIS: an update. *The Communications of the Association for Information Systems*, 14(1), 58.
- Palvia, P., Mao, E., Salam, A. F., & Soliman, K. S. (2003). Management information systems research: what's there in a methodology?. *Communications of the Association for Information Systems*, 11(1), 16.
- Petter, S. (2018). " Haters gonna hate": PLS and information systems research. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, 49(2), 10-13.
- Petter, S., & Hadavi, Y. (2021). With Great Power Comes Great Responsibility: The Use of Partial Least Squares in Information Systems Research. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, 52(SI), 10-23.

- Phillips-Wren, G., Iyer, L. S., Kulkarni, U., & Ariyachandra, T. (2015). Business analytics in the context of big data: A roadmap for research. *Communications of the Association for Information Systems*, 37(1), 23.
- Pinsonneault, A., & Kraemer, K. (1993). Survey research methodology in management information systems: an assessment. *Journal of Management Information Systems*, 10(2), 75-105.
- Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, 1(1), 51-59.
- 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.
- Riedl, R., & Rueckel, D. (2011). Historical development of research methods in the information systems discipline. *AMCIS 2011 Proceedings - All Submissions*. 28
- Ringle, C. M., Sarstedt, M., & Straub, D. W. (2012). Editor's comments: a critical look at the use of PLS-SEM in MIS Quarterly. *MIS Quarterly*, 36(1), iii-xiv.
- Rogers, E. M. (1995). *Diffusion of Innovations*. New York: Free Press.
- Roldán, J. L., & Sánchez-Franco, M. J. (2012). Variance-based structural equation modeling: Guidelines for using partial least squares in information systems research. In *Research methodologies, innovations and philosophies in software systems engineering and information systems* (pp. 193-221). IGI global.
- Rovit, E. (2020). *Acquisitions by consulting companies*. Retrieved from <https://www.rocketblocks.me/blog/consulting-company-acquisitions.php>
- Sangle, S. (2011). Adoption of cleaner technology for climate proactivity: a technology–firm–stakeholder framework. *Business Strategy and the Environment*, 20(6), 365-378.
- Sarstedt, M., & Mooi, E. (2014). *A concise guide to market research*. Springer
- Sarstedt, M., Hair Jr, J. F., Cheah, J. H., Becker, J. M., & Ringle, C. M. (2019). How to specify, estimate, and validate higher-order constructs in PLS-SEM. *Australasian Marketing Journal (AMJ)*, 27(3), 197-211.

- Sarstedt, M., Hair, J. F., Pick, M., Liengard, B. D., Radomir, L., & Ringle, C. M. (2022). Progress in partial least squares structural equation modeling use in marketing research in the last decade. *Psychology & Marketing*, 39(5), 1035-1064.
- Schelling, T.C. (1978) *Micromotives and Macrobehavior*. New York: Norton.
- Shmueli, G., Ray, S., Estrada, J. M. V., & Chatla, S. B. (2016). The elephant in the room: Predictive performance of PLS models. *Journal of Business Research*, 69(10), 4552-4564.
- Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J. H., Ting, H., Vaithilingam, S., & Ringle, C. M. (2019). Predictive model assessment in PLS-SEM: guidelines for using PLSpredict. *European Journal of Marketing*. 53(11), 2322-2347
- 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.
- Straub, D. W. (1989). Validating instruments in MIS research. *MIS Quarterly*, 13(2), 147-169.
- 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.
- Surbakti, F. P. S., Wang, W., Indulska, M., & Sadiq, S. (2020). Factors influencing effective use of big data: A research framework. *Information & Management*, 57(1), 103146.
- Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. 5, pp. 481-498). Boston, MA: Pearson.
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics*. Northridge. Cal.: Harper Collins.
- Taherdoost, H. (2016). Sampling methods in research methodology; how to choose a sampling technique for research. *How to Choose a Sampling Technique for Research* (April 10, 2016).

- Tallon, P. P., & Pinsonneault, A. (2011). Competing perspectives on the link between strategic information technology alignment and organizational agility: insights from a mediation model. *MIS Quarterly*, 35(2), 463-486.
- Teo, H. H., Wei, K. K., & Benbasat, I. (2003). Predicting intention to adopt interorganizational linkages: An institutional perspective. *MIS Quarterly*, 27(1), 19-49.
- Thong, J. Y., Yap, C. S., & Raman, K. S. (1996). Top management support, external expertise and information systems implementation in small businesses. *Information Systems Research*, 7(2), 248-267.
- Tojib, D. R., & Sugianto, L. F. (2006). Content validity of instruments in IS research. *Journal of Information Technology Theory and Application (JITTA)*, 8(3), 5.
- Tornatzky, L. and Fleischer, M. (1990) *The Process of Technology Innovation*, Lexington, MA, Lexington Books
- Westland, J. C. (2010). Lower bounds on sample size in structural equation modeling. *Electronic Commerce Research and Applications*, 9(6), 476-487.
- Williams, M. D., Dwivedi, Y. K., Lal, B., & Schwarz, A. (2009). Contemporary trends and issues in IT adoption and diffusion research. *Journal of Information Technology*, 24(1), 1-10.
- Wimmer, H., & Aasheim, C. (2019). Examining Factors that Influence Intent to Adopt Data Science. *Journal of Computer Information Systems*, 59(1), 43-51.
- Wold, H. (1974). Causal flows with latent variables: partings of the ways in the light of NIPALS modelling. *European Economic Review*, 5(1), 67-86.
- Zhu, K., & Kraemer, K. L. (2005). Post-adoption variations in usage and value of e-business by organizations: cross-country evidence from the retail industry. *Information Systems Research*, 16(1), 61-84.
- Zimmer, J. C., Arsal, R., Al-Marzouq, M., Moore, D., & Grover, V. (2010). Knowing your customers: Using a reciprocal relationship to enhance voluntary information disclosure. *Decision Support Systems*, 48(2), 395-406.