2023

Conceptual Architecture For Digital Twin Implementation In A Smart City Framework

Wesley Conwell
University of Alabama at Birmingham

CONCEPTUAL ARCHITECTURE
FOR DIGITAL TWIN IMPLEMENTATION IN A SMART CITY FRAMEWORK

by

WESLEY L. CONWELL JR.

LEON JOLOLIAN, COMMITTEE CHAIR
JUSTIN HILL
KARTHIKEYAN LINGASUBRAMANIAN
MURAT TANIK
NUO XU

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA

2023

CONCEPTUAL ARCHITECTURE FOR DIGITAL TWIN IMPLEMENTATION IN A SMART CITY FRAMEWORK

WESLEY L. CONWELL JR

INTERDISCIPLINARY ENGINEERING

ABSTRACT

More than 9 billion people are expected to live on the planet in 2050, about two-thirds in cities. This substantial population growth in urban areas increases demands on cities' resources, services, and infrastructure. Hence, many cities leverage intelligent technologies to ensure urban communities' efficiency in integrating massive infrastructure components and services. Cities define this reconstruction of their operational framework as a smart city. Effective resource consumption management in a smart city integrative structure presents a problem with solutions concerning the distribution of city services. This dissertation aims to establish a conceptual framework to effectively manage resource consumption in a smart city using digital twin architecture for implementation. The smart city space's energy component serves as this research's focus. Modeling a smart city digital twin establishes the framework to mitigate risks and manage real-world problems.

This research uses the case study method to test the smart energy subset of the smart city model development in conjunction with 1) actual data from residential customers in Austin, TX, 2) the Microsoft Azure platform to replicate IoT, digital twin, and machine learning aspects, and 3) review of pertinent literature related to the research question. Insights gained from this applied research effort can broaden understanding for application to smart city domains beyond energy, such as health, sustainability, and education.

DEDICATION

To my dear parents, Wesley Conwell Sr. and Cecelia Conwell

&

My lovely wife, Natalie, and wonderful daughter (Laura) and sons (Caleb and Ellis)

&

My brother (Timothy) and sister (Kimberli)

ACKNOWLEDGEMENTS

I am grateful to numerous individuals who have assisted me throughout my dissertation journey. I sincerely thank Dr. Leon Jololian, my committee chair and co-advisor, for his invaluable guidance and unwavering support throughout this endeavor. I would only have successfully concluded this journey with his invaluable support and motivation. I express my sincere appreciation and indebtedness to my esteemed advisor, Dr. Murat M. Tanik, for his exceptional guidance and mentorship throughout my academic journey. It is with immense gratitude that I acknowledge his instrumental role in facilitating my connection with Dr. Jololian, a collaboration for which I am eternally thankful. I also want to thank my other committee members: Dr. Karthikeyan Lingasubramanian, for his valuable advice, guidance, and support. Our initial discussions on contributing to the field of science have greatly influenced my intellectual growth and understanding. Dr. Nuo Xu was my initial mentor within the Interdisciplinary Engineering program, representing the business school perspective. I have acquired substantial insight through our numerous lunchtime, office, and classroom discussions. Dr. Justin Hill provided invaluable assistance in grounding my research ideas related to applied research.

Lastly, the guidance and support provided by Dr. Bharat Soni and Dr. Dale Callahan are also highly valued. Dr. Soni played a pivotal role in facilitating my enrollment in the doctoral program, while Dr. Callahan was a source of inspiration that motivated me to embark on the journey of pursuing a PhD.

TABLE OF CONTENTS

ABSTRACT iii
DEDICATION iv
ACKNOWLEDGEMENTS v
LIST OF TABLES ix
LIST OF FIGURES x
LIST OF ABBREVIATIONS xii

CHAPTER

1. INTRODUCTION 1
   1.1 Background 1
   1.2 Motivation for the Study 1
   1.3 Research Question 2
   1.4 Significance of the Study 3
   1.5 Contributions of the Study 4
   1.6 Delimitations and Limitations 4
   1.7 Organization of the Dissertation 5

2. LITERATURE REVIEW 6
   2.1 Introduction 6
   2.2 Smart Cities 10
      2.2.1 Foundational Research 10

2.2.2 Smart Cities Definitions 13
2.2.3 Smart Cities Standards 15
2.2.4 Smart Cities Use Cases 18
2.3 Digital Twin 23
2.4 Internet of Things..34
2.5 Big Data 36
2.6 Machine Learning 38
2.7 Summary of Literature Review and Implication 41
2.8 Summary of Literature Review and Implication 42
2.9 Conceptual Framework of the Study 44

3. METHODOLOGY 48
3.1 Introduction..48
3.2 Research Philosophy..50
3.2.1 Research Philosophy Assumptions 50
3.2.2 Pragmatism 51
3.2.3 Applied Research 53
3.3 Research Approach 54
3.4 Research Methodology 56
3.5 Research Strategy...56
3.6 Research Time Horizon..58
3.7 Data Management and Analysis...58
3.8 Role of the Researcher 60
3.9 Summary of Methodology and Research Design 60

4. FINDINGS...61
4.1 Introduction..61
4.2 Case Study 1 – Descriptive Twin...61

4.2.1 Case Study 1 Approach 62
4.2.2 Case Study 1 Detailed Activities...63
4.2.3 Case Study 1 Results..65

4.3 Case Study 2 – Predictive Twin 73
4.3.1 Case Study 2 Approach 73
4.3.2 Case Study 2 Detailed Activities...74
4.3.3 Case Study 2 Results..80
4.4 Case Study Summary 87

5. CONCLUSION 88
5.1 Introduction 88
5.2 Addressing the Research Questions 89
5.3 Contributions of the Study 91
5.4 Limitations of the Study..93
5.5 Future Direction 95

REFERENCES 96
APPENDIX A Definition of Key Terminology...102
APPENDIX B Applications and Tools..105

LIST OF TABLES

Table *Page*

Table 1:Philosophical assumptions of pragmatism, Saunders (2019) 52

Table 2: Research Approach Selection 55

Table 3: Bldg 661 Summary Statistics 76

LIST OF FIGURES

Figure *Page*

Figure 1: Citavi Library Documents....7

Figure 2: Citavi Literature Focused List....7

Figure 3: IEEE Smart City Standards....17

Figure 4 Sensing Energy Consumption in Vienna....19

Figure 5: Toyota Prototype Smart City....21

Figure 6: Digital Twin Lifecycle....24

Figure 7: Information as time, energy, material trade-off....25

Figure 8: The evolution from Industry 1.0 to Industry 4.0.....27

Figure 9: Self-management and healing in a DT system....28

Figure 10: Proposed X73-based Digital Twin (DT) System Architecture....29

Figure 11: ViLo: The Virtual London Platform by CASA....30

Figure 12: IoT and smart city integration....35

Figure 13: Smart City Data Analytics Connection....38

Figure 14: Machine Learning Smart City Representation....40

Figure 15: Notable Taxonomy of Machine Learning Algorithms....41

Figure 16: Conceptual Framework 44

Figure 17: Conceptual Levels of Development 46

Figure 18: Research Onion , Adapted from Saunders (2023) 49

Figure 19: Example Smart Home Connected Community 62

Figure 20 - RealEstateCore Ontology 65

Figure 21: Digital Twin Graph 73

Figure 22: Microsoft Azure Auto ML 74

Figure 23: Bldg 661 Visual Analysis 75

Figure 24: Figure 24: House ID Markers for Pecan Street Dataset 78

Figure 25: Case Study 2 Data Variables 79

Figure 26: Microsoft Azure List of Experiments 80

Figure 27: Case Study AutoML Jobs 81

Figure 28: Case Study 2 Properties 82

Figure 29: July Case2 Study Inputs, Outputs, Best Model Summary 83

Figure 30: Case Study 2 Data Guardrails 84

Figure 31: Case Study 2 Model Performance 84

Figure 32: Case Study 2 Aggregate Feature Performance 85

Figure 33: Model Deployment Simulation 86

Figure 34: Mock Dashboard 86

LIST OF ABBREVIATIONS

AMI	ADVANCED METERING INFRASTRUCTURE
CASA	CENTRE FOR ADVANCED SPATIAL ANALYSIS
COVID-19	CORONAVIRUS DISEASE 2019
CO2	CARBON DIOXIDE
CPS	CYBER-PHYSICAL SYSTEMS
DT	DIGITAL TWIN
DTDL	DIGITAL TWIN DEFINITION LANGUAGE
EDA	EXPLORATORY DATA ANALYSIS
ESA	EUROPEAN SPACE AGENCY
ETSI	EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE
EU	EUROPEAN UNION
GDP	GROSS DOMESTIC PRODUCT

IBML	INTERNET OF THINGS, BIG DATA, MACHINE LEARNING
ICT	INFORMATION AND COMMUNICATIONS TECHNOLOGY
IEEE	INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS
IoT	INTERNET OF THINGS
ISO	INTERNATIONAL ORGANIZATION FOR STANDARDIZATION
ISOCARP	INTERNATIONAL SOCIETY OF CITY AND REGIONAL PLANNERS
M2M	MACHINE TO MACHINE
ML	MACHINE LEARNING
NIST	NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
NOAA	NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
OSI	OPEN SYSTEMS INTERCONNECTION
PAAS	PLATFORM AS A SERVICE
UAB	UNIVERSITY OF ALABAMA AT BIRMINGHAM
UN	UNITED NATIONS
VILO	VIRTUAL LONDON

CHAPTER 1

Introduction

1.1 Background

Cities are a human invention, born from the human need for security, the convenience of living together, easier management of resources, better quality of life, and smaller mobility distances; they are engines of growth in the global economy. According to the World Economic Forum's ranking, the top five most important cities – New York, Tokyo, Los Angeles, London, and Shanghai – are expected to contribute $8.5 trillion in GDP by 2035. Notably, the 2020 global COVID-19 pandemic laid bare the significance of technology and its vital importance for a city to reinvent itself and stay relevant. With the advent and implementation of smart cities, the lessons learned from the COVID-19 pandemic resolve to build resilient and future-proof cities to withstand the toughest challenges. To better understand the relevant issues and problems, more profound insight into defining smart cities is required to ensure the appropriate application of critical technologies. Despite considerable research in the field, debates remain concerning the definition of a smart city. Prof Vito Albino relates a smart city as a fuzzy concept used in ways that are not always consistent (Albino et al., 2015).

There is neither a single template for framing a smart city nor a one-size-fits-all definition. First used in the 1990s, the term smart city focused on the significance of new

Information and Communications Technology (ICT) regarding modern infrastructures within cities. Smart city researcher Michael Batty considered that “the concept of the smart city emerged during the last decade as a fusion of ideas about how information and communications technologies might improve the functioning of cities, enhancing their efficiency, improving their competitiveness, and providing new ways in which problems of poverty, social deprivation, and poor environment might be addressed” (Batty et al., 2012b). More than ever, increasing numbers of people elect to live in cities, which poses unprecedented challenges for city stakeholders in addressing city inhabitants' quality of life. Frameworks for sustainable city development are needed to address the many challenges cities worldwide face today. Due to the nature of smart city objectives being highly local and even regional, different cities require different “smart” solutions. Nevertheless, research to date has not convincingly demonstrated optimized solutions to run a smart city.

1.2 Motivation for the Study

This interest in this research topic stemmed from a desire to learn about Big Data impacts in the electric utility industry. After further exploration of the literature, my curiosity was piqued as consistent themes emerged in the areas of smart cities, machine learning, the Internet of Things, and Big Data. Notably, the COVID-19 pandemic of 2020 provided municipalities around the world additional motivation to invest in new technologies for several areas, such as public transportation, contact tracing, and enforcing social distancing. This seminal moment in history addressing the challenges of the pandemic provided an ideal time for a new understanding of smart-city infrastructure.

Additionally, climate change impacts, macroeconomic displacement, and immigration flows drive the need for a paradigm shift in operational and managerial structure in which cities provide critical services to citizens and visitors.

Leveraging the capabilities of a digital twin framework presents an opportunity to more effectively manage resources/consumption in a smart city integrative structure. This research study focuses on developing the architecture for the digital twin implementation of a smart city framework to solve complex city operation/management problems. The uniqueness of this research moves forward in seeking to incorporate the concepts of the digital twin. The digital twin concept is becoming increasingly popular with researchers and professionals in numerous industries to visualize, model, and work with complex urban systems. Implementation of this concept is achieved by coupling physical systems with comprehensive digital representations that automatically update to match their physical counterparts' state.

1.3 Research Questions

The research questions intended to be answered by this thesis are: 1) How can current state-of-the-art technologies be leveraged for monitoring and managing resources within a smart city framework? 2) What methods can allow virtually any city to implement smart city concepts in their city operations and management using simulation? To answer these questions, this research will make the following contributions: 1) Establishment of design and specification of basic core functionalities for a smart city digital twin framework to enhance resource consumption in a smart city or any other resource-dependent construct, 2) Development of a maturity model structure for smart city application, 3) Implementation of a case study in the smart energy sub-domain of smart city using a digital twin construct.

1.4 Significance of the Study

Going forward, cities must effectively deal with the next pandemic or another major crisis – transforming smart cities can help. By 2050, 68% of the global population will live in cities (World Economic Forum, 2021). Given the complex, interconnected, and dynamic nature of urban systems, the actual value of this transformation to smart cities can be achieved only when cities grow as an intelligent "system of systems" that maximize the value of the data (Amit Midha, 2021) As more cities define themselves as smart within their own domains and begin to communicate with other cities and governments both domestically and globally, there will be a significant hurdle to translate their efforts to businesses and constituents who need to interact with their smart services. A 2017 paper by Hatem Sta of Tunisia asserts the reality of information in cities, "Information is almost always tainted with various kinds of imperfection: imprecision, uncertainty, ambiguity" (Sta, 2017). The complexity of smart city implementation drives the need for an overall architecture to manage the imperfection of information. As urban centers grow in population and complexity, cities rely increasingly on ICTs and smart solutions to reach sustainable growth (Monzon, 2015). Consequently, cities' services and infrastructures are being stretched to their scalability, environment, and security limits as they adapt to support this population growth. Along with cities' growth, innovative solutions are crucial for improving productivity (increasing operational efficiencies) and reducing management costs. The development of a broad-based, solution-focused architecture helps manage the complexity. The interdisciplinary model developed through this conceptual architecture will establish the framework to manage smart city complexity and be general enough to be used for crafting solutions to the representation and combination of all kinds of imperfect information.

1.5 Contributions of the Study

The approach researched in this effort focused on establishing a framework for smart city resource consumption across varying contexts. With that as the impetus, this study identifies 1) the design and specification of core functionalities for a smart city digital twin framework to enhance resource consumption and 2) case study development in the smart energy sub-domain of a smart city using a digital twin construct.

1.6 Delimitations and Limitations

This study is centered upon a particular contextual framework – the facilitation of smart city implementation through utilizing digital twin development methodology. Smart city (Appendix A) has many definitions. Albino's paper *Smart Cities: Definitions, Dimensions, Performance, and Initiatives* (Albino et al., 2015) is the world's most cited paper referencing the definition of a smart city, and it concludes with a lack of a universal definition. However, the article considers the multi-faceted nature of smart cities, including dimensions, measures of performance, and citizen experiences. For this study, the focus on smart cities in the problem domain centers on measures of performance aspect concerning the smart environment connection to energy – smart energy for this research. The data utilized for this study came from residential energy sources and is not intended to represent commercial or industrial energy sources.

Several factors may constrain the applicability of this study. One such factor pertains to the challenge of comparing projects within disparate problem domains, which can impede the generalizability of findings. Additionally, variations in the assessments of networked systems (such as smart cities) across different contexts may further limit the extent to which the study's results can be extrapolated.

1.7 Organization of the Dissertation

This dissertation is presented in five chapters following a standard format for dissertations ((Corbin & Strauss, 2008)). This first chapter describes the research problem, the motivation for the study, the significance of the study, and the organization of this dissertation. The second chapter describes the literature review for the study. Chapter three explains the use of research methodology and classical Grounded Theory and how the data were collected. The fourth chapter presents the findings from the study and the analysis of the data collected in the study. The fifth chapter summarizes the study's significant findings and suggestions for future research.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

The research questions for this study emphasize the applied research focus of digital twin implementation in a smart city, energy problem domain. Over the years of working through the research process, a notable amount of time was exhausted to review and analyze literature research sources. UAB Library's extensive resources through Compendex and Engineering Village proved invaluable; Google Scholar also provided tremendous assistance. Appendix B includes a listing of resources included in the literature review process. In congruence with the abstract nature of the architecture's development across several technologies and ideas, the Citavi literature database contains 778 collected resources of journal articles, book references, research lab presentations, patent references, and other literature scoured during the years of the study's effort. That exhaustive search was honed into a relevant literature exposition in establishing the applied research design and case study development. The literature tool Litmaps supplied a visual illustration of the Citavi library of documents. Figure 1 illustrates the expanse of the evaluated sources and connections.

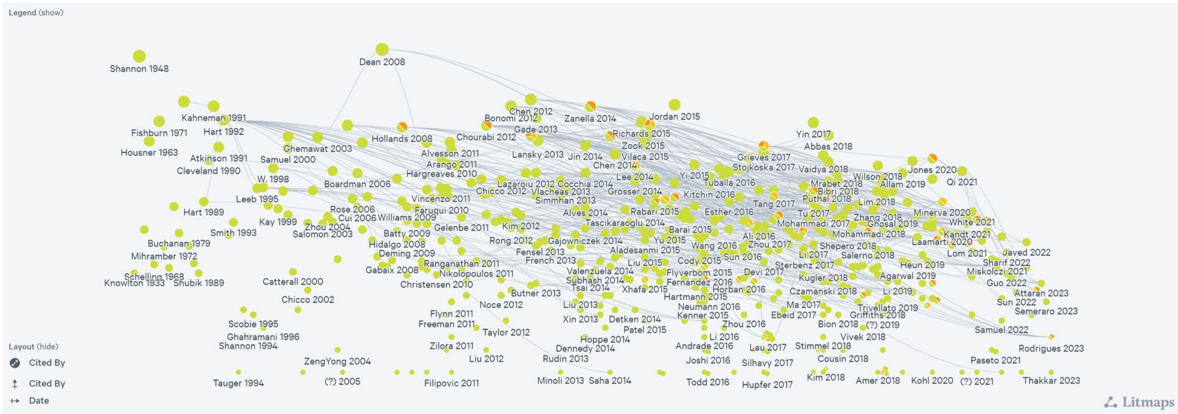

Figure 1: Citavi Library Documents

That expanse of documents from Figure 1 was culled through evaluation and review to reveal the documents noted in Figure 2

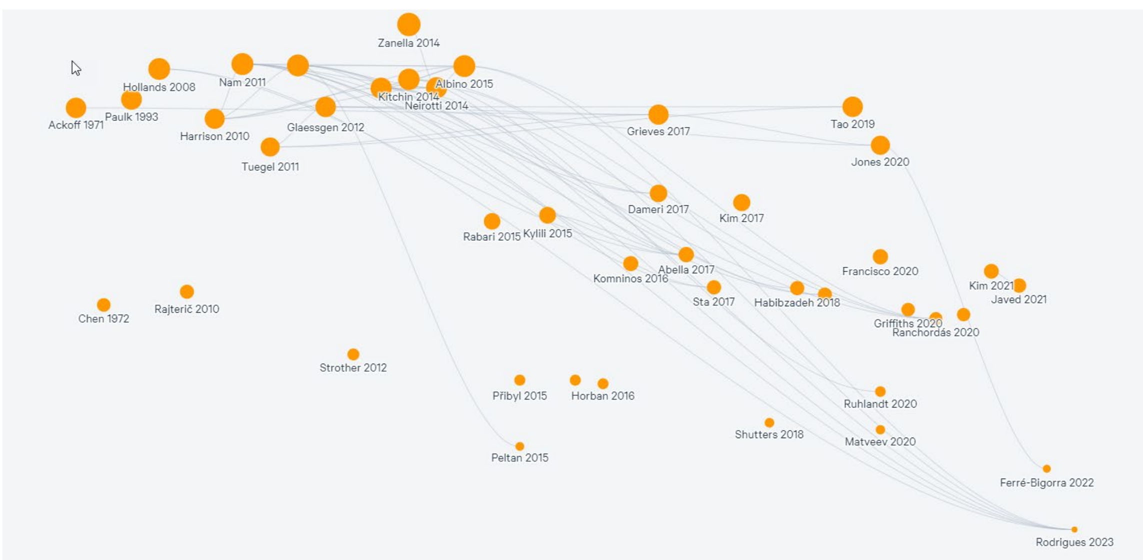

Figure 2: Citavi Literature Focused List

The literature firmly supports the notion that technology drives the concept of a smart city. In turn, a critical question for consideration is what technologies can be used for managing resources in a smart city? Machine learning, the Internet of Things (IoT), and Big Data are three critical technologies at the cusp of a long run of global impact for years to come. The opportunities generated from these technologies are borne out of a confluence of events dispensed through the global deployment of robust networks, supercomputing processing power, and virtually unlimited storage capacity (Friedman, 2005). In evaluating the literature relative to the nature of the conceptual framework, the literature review chapter of this study is outlined as follows: the first review of the applicable technology research areas – smart cities, digital twins, IoT, Big Data, and machine learning. Reviewing these technologies will help lay out the conceptual basis for the research. In addition to the technologies noted, the impact of behavioral economics is included in the model for literature discussion in considering the interdisciplinary interaction of business components. Secondly, those findings are summarized, and the requisite implications are discussed. Finally, the conceptual model will be illustrated, and its components and parts will be explained.

This research study's genesis stemmed from a desire to learn about Big Data impacts in the electric utility industry. After further exploration of the literature, consistent themes emerged in the areas of smart cities, machine learning, the Internet of Things, and Big Data. Notably, the COVID-19 pandemic of 2020 provided municipalities around the world additional motivation to invest in new technologies for several areas, such as public transportation, contact tracing, and enforcing social distancing. This seminal moment in history addressing the challenges of the pandemic provided an ideal time for a new understanding of smart-city infrastructure.

Since I have worked for 32 years in the energy space, my background compelled me to focus on energy in my research. Major cities occupy only 5% of the earth's land and 75% of the world's fossil fuel usage. In turn, rapid urbanization and explosive increases in energy demand in cities are of concern. Because buildings account for most of city energy consumption, they have become a key focus for innovative city initiatives. The opportunity for real-time intelligent planning and urban energy management lies at the intersection of smart cities and building energy efficiency. This case study considers energy efficiency in residential housing stock as the building component for evaluation. Out of this smart energy case study, we can envision new services enabling higher quality of life and reducing energy consumption and CO2 emissions. Although this research focused on energy, the architecture is abstract and can conceptually be employed for health, education, traffic, and other domains.

2.2 Smart Cities

2.2.1 Foundational research

A large growing body of literature has investigated how best to define a smart city. There is no universal standard for defining a smart city, so that we will evaluate the literature across several context measures. We begin with framing smart cities in the historical context of technology. The last 250 years relative to technology can be grouped into four revolutions. The first revolution was associated with mechanical steam power until about 1830; the second with electrical power until about 1920. The third technological revolution is associated with the invention of computers and information technology and ended in the late 20th century. Arguably, we are in the fourth revolution, essentially the world of machine intelligence, the smart city, digital health care, and the massive proliferation of information devices (Batty, 2018). A smart city ultimately represents a complex system of systems that leverages cutting-edge ICT for the benefit of a community, integrated with complex subsystems to fulfill its operational objectives. In the smart cities space, several researchers developed unique solutions concerning the management of smart cities. First, Strohbach established that increasing amounts of valuable data sources, advances in the Internet of Things, and Big Data technologies offer new potential to deliver analytical services to citizens and urban decision-makers. However, a gap exists in combining the current state of the art in an integrated framework that would help reduce development costs and enable new services. Strohbach's work demonstrated how such an integrated Big Data analytical framework for the Internet of Things and Smart City applications would work. The contributions were noted as threefold: (1) an overview of Big Data and Internet of Things technologies, including a summary of their relationships; (2) a case study in the smart grid domain illustrating the high-level requirements towards such an analytical Big Data framework, and (3) an initial version of the framework mainly addressing the volume and velocity challenge (Martin Strohbach, Holger Ziekow, Vangelis Gazis, 2015).

Secondly, globally, China is advancing towards the forefront of smart city solutions and ideas because of its sheer size. Chen's 2016 work noted that to build sustainable, livable, and safe cities, the Chinese government began experimenting with its first Smart City project in 90 cities in late 2012. The project aimed to transform digital cities into smart cities, and as of 2016, more than 290 cities were involved. China emphasizes smart cities' data from the data life cycle perspective (data collection, data analysis, and data-driven smart services). Chen's research highlighted three unique challenges in China's smart city efforts and presented three promising future directions to characterize a potential road map for smart city development. The three challenges were noted as 1) Overwhelming amounts of data not being open for broader analysis, 2) structural deficiency in operating through centralized, "top-down" approaches, and 3) reactive responses to significant events. Chen's research proposed solutions to include: 1) a directional approach to structurally move from data gathering to data opening and sharing for better use; 2) advocacy for bottom-up ideas arising from residents and following local culture. Concepts include leveraging opportunities from the crowd power of citizen sensing from mobile phones, urban activity, and environment monitoring without deploying unique infrastructures, and 3) moving from reactive responses to proactive responses using advanced big data prediction algorithms. Compared with reactive responses, the proactive approach attempts to prevent concerning events from taking place. Advanced data predictive tools allow city governors to foresee future events; alarms can then be triggered to warn the public to take necessary measures in advance, such as re-planning travel routes to avoid roads with high congestion (Chen et al., 2016).

Thirdly, Mohammadi's 2018 research provided crucial insight into the development of smart cities and their fast-paced deployment, resulting in the generation of large quantities of data at unprecedented rates. Unfortunately, most of the generated data is wasted without extracting potentially helpful information and knowledge because of the lack of established mechanisms and standards that benefit from such data's availability. Smart cities' highly dynamic nature calls for new operation strategies that are flexible and adaptable to cope with the dynamicity of data to perform analytics and learn from real-time data. Mohammadi and his team shed light on the challenge of underutilizing the big data generated by smart cities from a machine-learning perspective. They proposed a three-level learning framework for smart cities that matches the hierarchical nature of big data generated by smart cities to provide different levels of knowledge abstraction. Fundamentally, the framework benefits from semi-supervised deep reinforcement learning, where a small amount of data that has users' feedback serves as labeled data. In contrast, without such users' feedback, an immense amount serves as unlabeled data. The framework utilizes a mix of labeled and unlabeled data to converge toward better control policies instead of wasting the unlabeled data. (M. Mohammadi & Al-Fuqaha, 2018). Lastly, Mohammadi and Taylor's 2018 study references rapid urbanization challenges, in which cities are determined to implement advanced socio-technological changes and transform into smarter cities. The success of such transformation dramatically relies on a thorough understanding of the city's states of spatiotemporal flux. Understanding such fluctuations in context and interdependencies among various entities across time and space is crucial for cities to maintain their smart growth. Here, Mohammadi introduced a Smart City Digital Twin paradigm that can enable increased visibility into cities' human-infrastructure technology interactions, in which spatiotemporal fluctuations of the city are integrated into an analytics platform at the real-time intersection of reality-virtuality. (N. Mohammadi & Taylor, 2017).

Overall, the smart city literature relays many challenges and opportunities for managing smart cities. The advantages of urban living cause continuous migration from rural to urban environments, leading to larger cities. Growing populations in urban areas demand more resources and better services. The research suggests a gap in the literature on optimizing municipal resources and increasing the quality of services while city populations are growing (Jalali, 2017). Therein lies the opportunity for smart cities to provide solutions. This research study claims that the digital twin architecture presents a preeminent implementation platform for a smart city. The next section of the study discusses the digital twin background with insight from critical scholars in the field.

2.2.2 Smart Cities Definitions

Scholars have sought to define a smart city for quite some time. In the paper "*Towards an effective framework for building smart cities*," Professors Lee and Hancock establish that the smart city concept originated from various definitions, including those of the 'intelligent city,' and 'digital city' (Lee et al., 2014) The work of Prof Hatem Sta suggests that “a Smart City is a modern city that uses smart information infrastructure to ensure the sustainability and the competitiveness of the different urban functions by integrating different dimensions of urban development and investments in order to reduce the environmental impact and to improve the quality of citizens’ lives” (Sta, 2017) Entire theses have been written regarding smart cities such as that of Ali (Ali, 2019) and (Williams, 2020) Overall, the smart city concept is a tool structured to tackle urban challenges facing citizens and stakeholders (Fernandez-Anez et al., 2018).

Regardless, with no shared definition to date, cities are self-defining as smart without a consensus on what it means to be smart (Kummitha & Crutzen, 2017). Although motivations may vary, studies suggest cities take on innovative city initiatives as part of a strategic vision for sustainability (Kummitha & Crutzen, 2017). Further, these self-defined smart cities embark on smart city initiatives with no shared construct or model to describe or measure their smart city programs (Lombardi, 2011). As a result, smart city initiatives are often isolated to “sectoral improvements” without a vision across sectors (Gori, Parcu, & Stasi, 2015). The lack of a shared model that defines what it means to be a smart city poses further challenges to a smart city’s interaction with the public, private industry, and other governments when attempting to provide ICT-driven and smart services. Although each city may define what “smart” means to them individually, its innovative services will remain isolated to their city limits and vertical programs. Establishing a universal smart city definition enables city services to be more portable and interoperable outside the city limits by establishing a standard definition and vernacular around which public and industry innovation can be seeded. As outlined by Gori et al. (2015), the problem then becomes how to define the smart city as a matter of "horizontally cumulative elements" (Khatoun & Zeadally, 2016).

Defining a smart city framework spanning academic, business, and government domains leveraging existing technologies to solve city problems is crucial. Smart city implementation draws an increase in ICT penetration within its complex systems. This consideration leads to the challenge of interoperability chasms between data silos in terms of visibility, syntax, protocols, semantics, security, licensing, and trust. Even beyond this, the nature of leveraging these resources for business and societal value is not clearly understood, resulting in the idiom "drowning in data"(Howell, 2017).

For purposes of this study, a smart city is defined as a complex system of systems that leverages cutting-edge ICT for the benefit of a community. This complex system is integrated with complex subsystems to fulfill operational objectives. This research will leverage the technology and standards already existing in the smart city domain and integrate the technologies to yield the wasted data in a smart city framework. The next section of the study delves into more significant insight into smart city standards in exploring varying definitions of operation.

2.2.3 Smart Cities - Standards

No discussion of the literature on smart cities is complete without consideration of standard-setting bodies weighing in on the debate in defining a smart city.

International Organization for Standardization (ISO) Technical Committee (TC) ISO/TC 268/SC focuses on sustainable cities and communities. ISO standard ISO 37101 is the core standard in this arena and is supported by a series of standard indicators such as those referenced by ISO 37122 (indicators for smart cities). ISO 37122 emphasizes sustainable cities and communities to make cities inclusive, safe, resilient, and sustainable. IEEE seeks to establish standards tied to the technologies that drive our understanding of a smart city. Figure 3 illustrates ten critical technologies IEEE standards reference in enabling smart city technologies for communities, including 1) smart grid, 2) learning technologies, 3) smart home, 4) eGovernance, 5) Cyber security, 6) 5G, 7) Internet of Things, 8) Energy Efficiency, 9) eHealth, and 10) Intelligent transportation.

Figure 3: IEEE Smart City Standards

Lastly, a standard considered during the literature review involved Microsoft Corporation partnering with the European Telecommunications Standards Institute (ETSI). In this regard, ETSI put forward ETSI TS 103 410-4 V1.1.1 (2019-05)) to establish technical specifications for a smart cities domain ontology and semantics (ETSI, 2019).

Next, following clarification of the definition of a smart city and evaluation of relevant standards, an assessment of various use cases was researched in the literature. Although Batty, Sta, Lee, Khatoun, and other scholars established a firm footing in the domain, research opportunity lies in the interdisciplinary applied research contraction of the resource consumption aspects tied to cities of varying resources and capabilities. The following chapters relay foundational research for the associated conceptual framework technologies.

2.2.4 Smart Cities – Use Cases

Many use cases abound in implementing smart city concepts. To establish sustainable, livable, and safe cities, the smart cities framework is a desirable option for solving managerial and operational problems in cities. Some entities establish smart cities constructed from scratch, and others implement retrofit projects in existing municipalities; there is also some regard for futuristic views of smart cities. Regarding retrofit projects, cities in European and Asian countries are vanguards in implementing the smart city concept for their betterment. In 1965, the UN and the Council of Europe established the International Society of City and Regional Planners (ISOCARP) to bring together a global association of experienced urban planners to vet ideas around the concept we now know as smart cities.

First, consider use cases in existing infrastructure arrangements. Geneva (Switzerland) developed a smart-infrastructure project incorporating fiber-optic and smart-grid networks. The project instituted a citizen data pilot covering a small urban area with a 100 MBps network. Participation was free of charge, and the initial pilot offered voice and TV services (Anthopoulos, 2017). Vienna (Austria) established a smart city project portfolio mainly focused on energy efficiency and innovation enhancement. As an illustration, Figure 4 lists a public display of energy consumption in one of the local Vienna commercial buildings. In providing the public visualization, Vienna citizens garner a broader sense of their energy consumption in real-time. Vienna's smart city interlinks energy, mobility, buildings, and infrastructure.

Figure 4 Sensing Energy Consumption in Vienna

With respect to Asia, just after the transfer of sovereignty over Hong Kong from the United Kingdom to China, the smart city configuration of Hong Kong was conceptualized in 1998. It was followed by four strategic updates, morphing the city into a large-scale project. Several initiatives were undertaken, from open data projects and smart government service deployment to IoT and new district development. Hong Kong has succeeded in smart infrastructure and smart service deployment, mainly based on cyber-physical integration as connected to a digital twin.

Next, consider smart cities designed as such from inception; Masdar City is considered a model by many (Griffiths & Sovacool, 2020). The Abu Dhabi government announced 2006 its intent to spend $22 billion to build Masdar City. This carbon-neutral zero-waste city would demonstrate the state-of-the-art in sustainable city design. After further consideration, the plan changed to an investment of $10 billion with targeted completion over the course of 20 to 25 years. For Masdar City, its critical objectives focused on sustainability, such as a 40% reduction in the energy consumption of its buildings (relative to comparable buildings in Abu Dhabi).

New Songdo (South Korea) is another from-scratch, model smart city. New Songdo (or Songdo International Business District) was developed on reclaimed wetland from the Yellow Sea, covering an area of 48.26 km^2 and is located close to the city of Seoul. It is known as the South Korean smart city flagship and part of a broader national smart city program (NIA, 2007). The plan was completed in 2015 with a $38 U.S. billion budget. Utilizing key AI technologies, Songdo focuses on sustainable planning and embedding ICT in the city's physical systems.

Lastly, there is a consideration for futuristic-focused smart cities. Toyota revealed plans to build a prototype "city" of the future on a 175-acre site at the base of Mt. Fuji in Japan (Figure 5). Sometimes known as the Woven City, it will be a fully connected ecosystem powered by hydrogen fuel cells.

Figure 5 – Toyota Prototype Smart City

Envisioned as a "living laboratory," the Woven City will serve as a home to full-time residents and researchers who can test and develop technologies such as autonomy, robotics, personal mobility, smart homes, and artificial intelligence in a real-world environment. Building a complete city from the ground up presents a unique opportunity to develop future technologies, including a digital operating system for the city's infrastructure. Toyota will extend an open invitation to collaborate with other commercial and academic partners and invite interested scientists and researchers from around the world to work on their own projects in this one-of-a-kind, real-world incubator. The city is planned to be fully sustainable, with buildings made primarily of wood to minimize carbon footprint, using traditional Japanese wood joinery and robotic production methods. The rooftops will be covered in photo-voltaic panels to generate solar power and power generated by hydrogen fuel cells. Toyota plans to weave in the outdoors throughout the city with native vegetation and hydroponics. Residences will be equipped with the latest human support technologies, such as in-home robotics, to assist with daily living. The homes will use sensor-based AI to check occupants' health, care for basic needs, and enhance daily life. Only fully autonomous, zero-emission vehicles will be allowed on the thoroughfares to move residents through the city.

The definitions, standards, and use cases represent an integration of known components regarding smart cities. The gaps connect to the abstraction of a framework noted in the research questions denoted in Chapter 1.3.

2.3 Digital Twin

Digital twin technology has been a concept practiced since the 1960s. NASA used basic twinning ideas during this period for the space program. They created physically duplicated systems at ground level to match the systems in space. NASA developed a digital twin to assess and simulate conditions on board Apollo 13. Most notably, the explosion of oxygen tanks in the April 1970 mission of Apollo 13 became a famous rescue mission, technical issues needing to be resolved from up to 200,000 miles away. However, a key to the rescue mission was that NASA had a digital twin model of Apollo 13 on Earth, which allowed engineers to test possible solutions from the ground level (Jones, 2021). The digital twin concept gained modern recognition in 2002 after the Challenge Advisory group hosted a presentation for Michael Grieves at the University of Michigan on technology. The presentation involved the development of a product lifecycle management center.

While the terminology has changed over time, the basic concept of the Digital Twin model has remained relatively stable from its inception in 2002. It is based on the idea that a digital informational construct about a physical system could be created as an entity on its own. This digital information would be a "twin" of the information embedded within the physical system itself and be linked with that physical system through the entire lifecycle (Figure 6) of the system. Reference to lifecycle implies that this is not a static representation; the systems are linked throughout the entire system lifecycle. The virtual and natural systems would be connected as the system went through the four phases of creation, production (manufacture), operation (sustainment/support), and disposal.

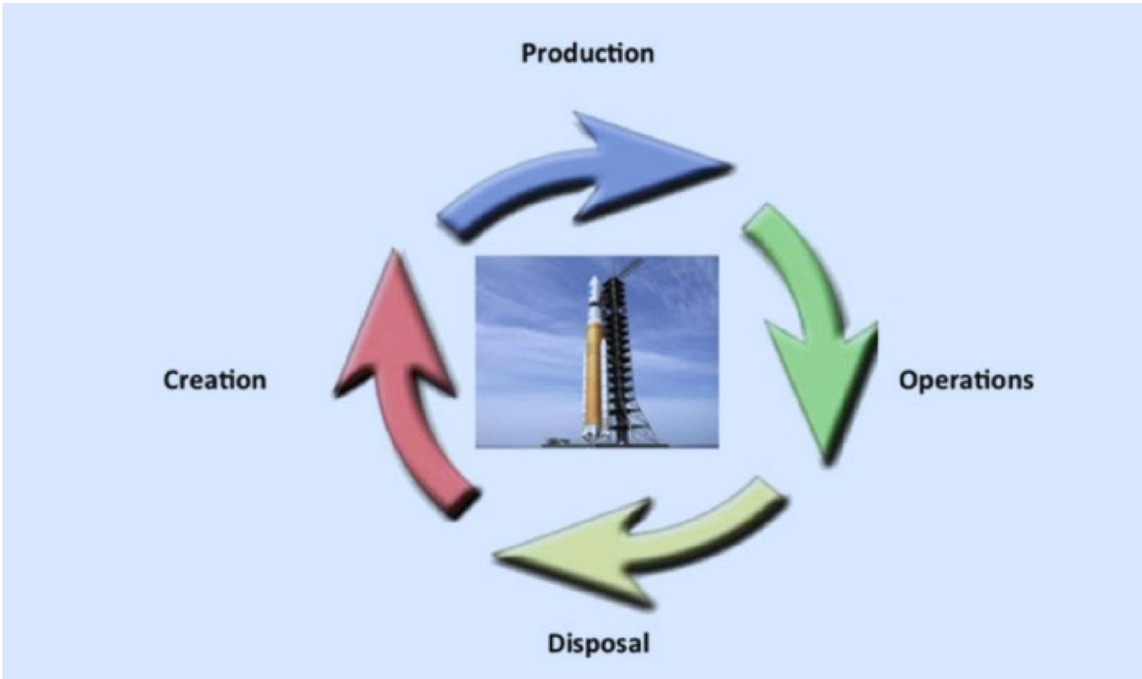

Figure 6: Digital Twin Lifecycle

Grieves' research establishes the value of a digital twin as the core premise that information replaces wasted physical resources, i.e., time, energy, and material. For example, consider any task, such as designing an airplane, manufacturing a fuel tank, or managing a smart city. That task can be represented quantitatively as the sum of the cost of all resources required to complete the task (illustrated in Figure 7).

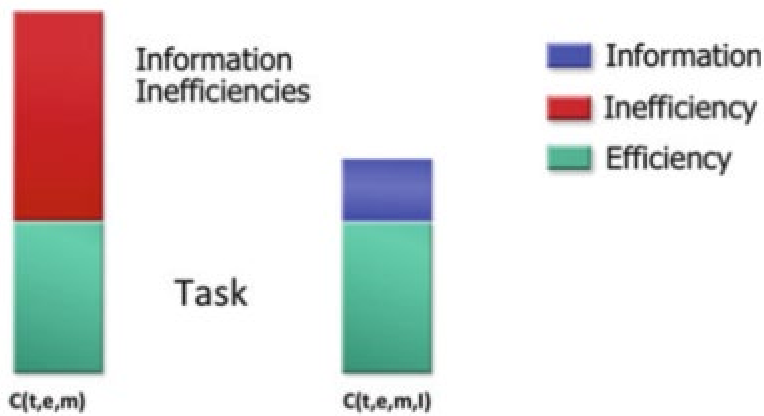

Figure 7 - Information as time, energy, material trade-off

In the illustration, the costs are monetized as labor time, material, and energy costs over the task's life. The tasks are divided into two parts on the left side of the figure. The bottom part is the minimum physical resources required to perform the task. If perfect knowledge were assumed in the situation, the minimum amount of resources needed to perform the task would be known, and then precisely perform the task, minimizing the necessary resources. The upper part of the bar represents the waste of resources we incur in performing the task. Because we do not have perfect knowledge, resources are wasted in various ways. Next, we have the same task on the right side of Figure 7. The resources needed to perform the task without waste are precisely the same. However, on the upper part of the bar, we have information that allows us to replace some, but not usually all, of the wasted resources necessary to perform this task. Because information is never free to acquire or use, there is always some use of resources in developing and using information. Hence, our fundamental assumption is that the cost of information will be less than the cost of the wasted physical resources. This is usually a reasonable assumption for repetitive tasks and tasks of complexity. The main idea is that using information can significantly reduce the waste of physical resources (Grieves & Vickers, 2016).

Although digital twins have been highly familiar since 2002, only as recently as 2017 has it become one of the top strategic technology trends. The Internet of Things enabled digital twins to become cost-effective so they could become relevant across multiple business domains. First, in medicine, some have advocated for a medical digital twin. In this instance, a digital twin is a virtual or digital replica of a living organism or a part of that organism, like a heart or lungs. A digital twin would be created from data points such as imaging records, in-person measurements, lab results, and genetics. All of that data is then used to map out an exact digital replica of an individual. The medical digital twin could prevent over 250,000 deaths yearly caused by medical mistakes, according to Johns Hopkins (Minevich, 2021). Secondly, in geoscience, the European Space Agency (ESA) is developing a digital twin of Earth. ESA scientists are configuring an Earth digital twin over the next decade, constantly feeding real-world data into the model. Next, EU scientists will use neural networks (computer algorithms) to identify patterns in Earth's weather systems and make more accurate predictions (Hart, 2021).

In this study, my research will yield applicable insight concerning a smart city digital twin. A smart city digital twin is a virtual platform that utilizes data and IoT technology to replicate and emulate changes happening in an actual city's infrastructure systems to provide insight that could help improve sustainability, resilience, and livability (Tech, n.d.). Unlike the lifecycle product consideration in the typical manufacturing digital twin, a city does not experience the standard "cradle to grave" lifecycle constructs. Operationally, cities evolve in response to economic opportunity and external threats (Austin et al., 2020). Georgia Tech researchers Mohammadi and Taylor are credited with pioneering work in this space. Their paper "Smart City Digital Twins" lays out ideas on the determination of cities to implement advanced socio-technological changes and transform them into smarter cities (N. Mohammadi & Taylor, 2017). A key focus of this research's contribution is to develop an architecture that can monitor a smart city with a digital twin as the backbone of the implementation. As referenced in Figure 8Figure 8, the digital twin is part of the continuum advanced by the 4th industrial revolution, bringing forth the age of cyber-physical systems (CPS). CPS is the core foundation of Industry 4.0 in which physical and software components are deeply intertwined, each operating on different spatial and temporal scales and interacting in myriad ways that change with context (Xu et al., 2016).

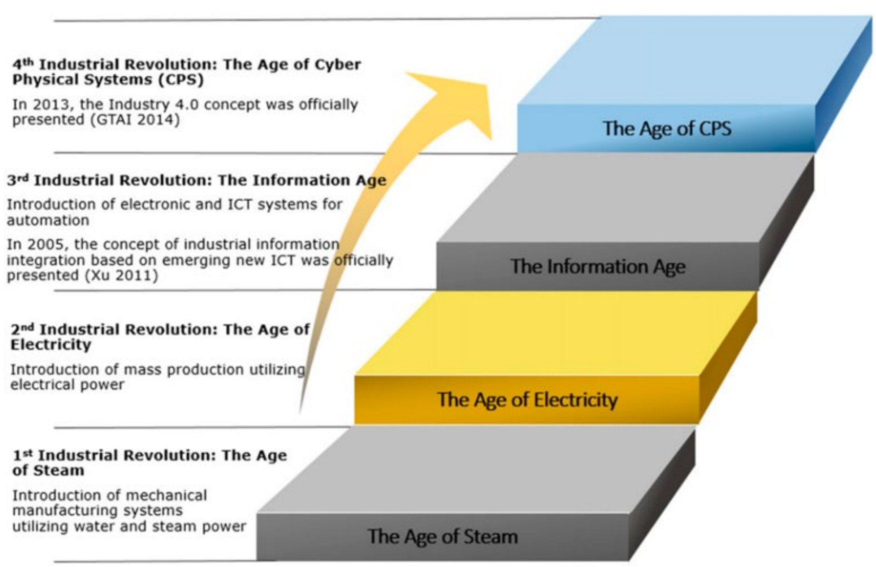

Figure 8: The evolution from Industry 1.0 to Industry 4.0.

From another conceptual perspective, in the paper titled "Digital Twin in the IoT Context: A Survey on Technical Features, Scenarios, and Architectural Models," a digital twin smart city model was proposed that incorporated functionalities for self-management to account for the possibility of smart city applications to scale up hundreds of thousands or millions of objects to be controlled, governed and orchestrated (Minerva et al., 2020) As a visual representation to illustrate the application, Figure 9depicts the model and the situation in which one object detects some malfunctioning. It can warn the system of the issue.

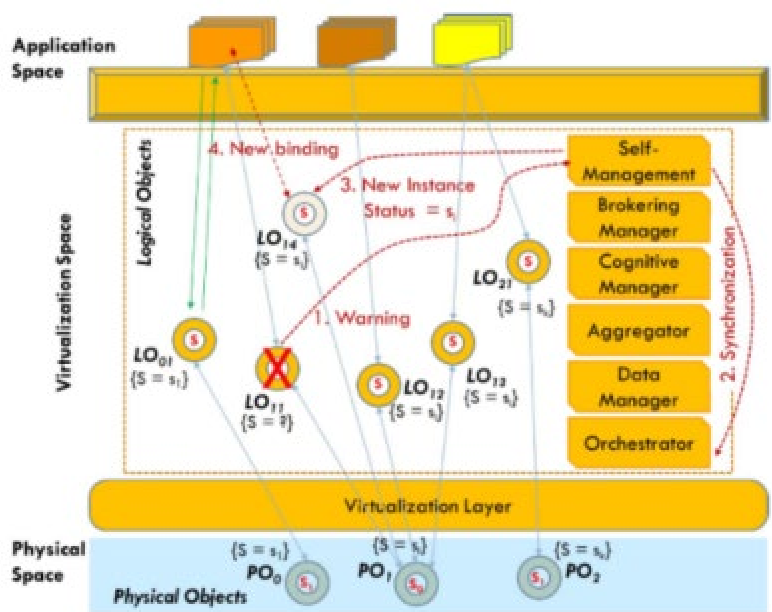

Figure 9: Self-management and healing in a DT system

This model relates to the methodology in which smart city sensors or other applications can be implemented by leveraging the digital twin framework. In the realm of applied research, autonomous vehicles are representative of the Figure 9 representation of a digital twin.

Laamarti proposed a standardized digital twin framework for smart health and well-being in a smart city in another notable application. This study utilizes IEEE/ISO standard 11073 to produce a framework for interoperability and reduce technological complexity. The proposed IEEE/ISO standard 11073-based architecture constitutes a foothold in a smart city's smart health domain, yielding a system that caregivers can use in the health and well-being system. To better illustrate the architecture, Figure 10references the proposed smart health architecture for the smart city domain (Laamarti et al., 2020).

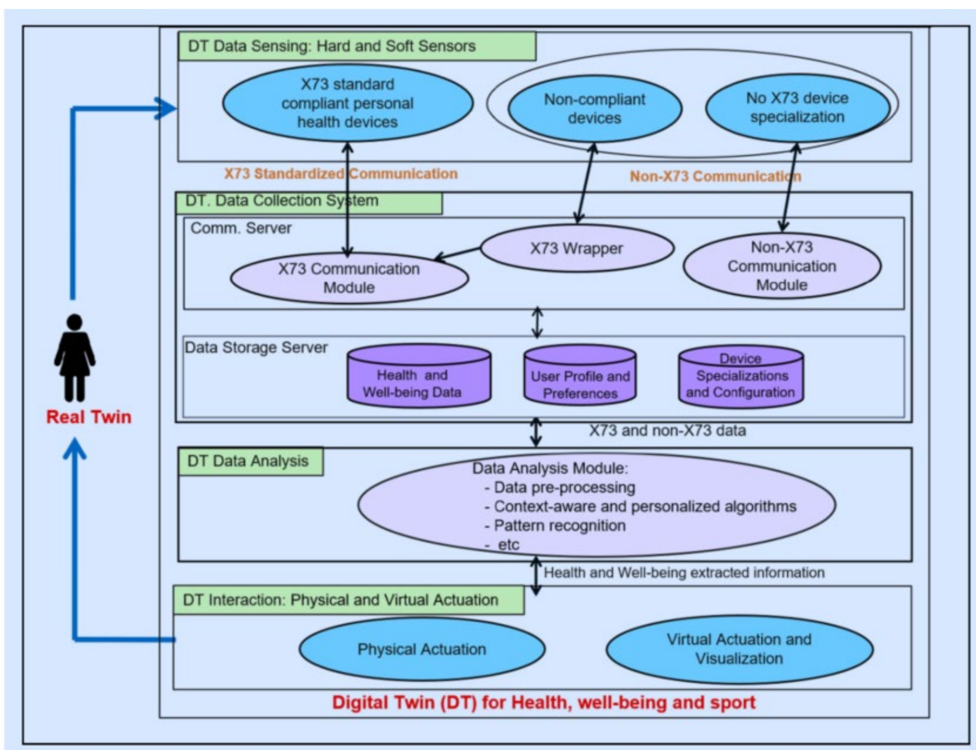

Figure 10: Proposed X73-based Digital Twin (DT) System Architecture

Most notably, in the digital twin literature review, the research moved from the broadly assessed perspective closer to the desired area of focus in the smart energy space. A digital twin application (Figure 11)in London "twins" the Queen Elizabeth Park for the Virtual London (VilLo) project. This park area was heavily developed during the London summer Olympics in 2012. Researchers developed a 3D digital twin of a neighborhood in London to test the digital twin concept, as noted in Figure 11 (Dawkins, 2018).

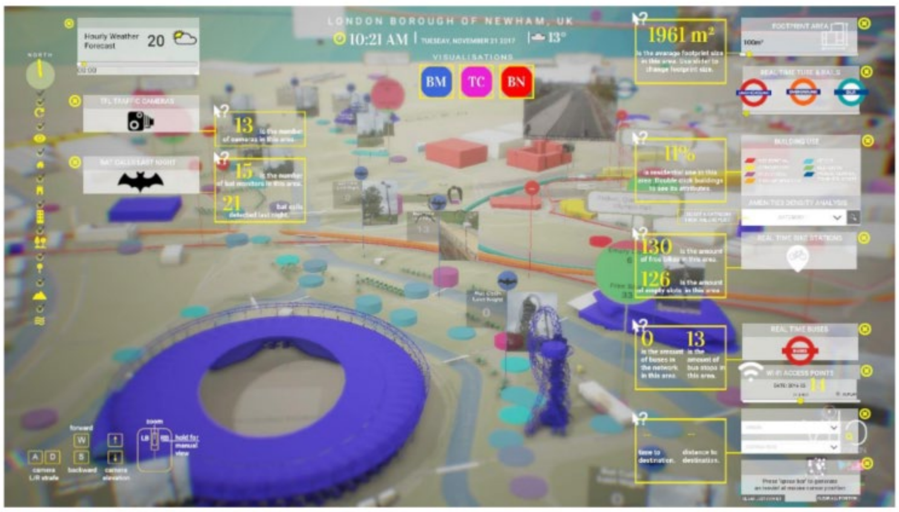

Figure 11: ViLo: The Virtual London Platform by CASA

Critics question the construct of the digital twin in smart cities in that there is no widely accepted standard for development. Starting in the manufacturing space with CAD representations of physical objects, the digital twin concept moved to broader systems across other domains. One area from which to draw lessons for establishing digital twin standards connects us to applications in the smart health space. The ISO/IEEE 11073 Personal Health Device standards emerged in 2008 to facilitate communication between personal health devices [*such as computer systems and smartphones*] and health care managers. The standards aim to promote health data exchange while providing plug-and-play real-time interoperability. The minimum requirement for personal health devices and managers to be X73 compliant is to adhere to the X73 communication model. As such, the X73 standards ensure that a health device and a compliant manager can complete the data transfer successfully.

As one would expect, expanding the digital twin push for some measure of standards, industry titans joined the conversation. Microsoft developed Azure Digital Twins leveraging its Azure cloud computing service. Azure Digital Twins is created as a platform as a service (PaaS) offering that enables the creation of digital models of entire environments (George, 2020). These environments could be buildings, factories, farms, energy networks, railways, stadiums, and more—even entire cities. Also, IBM developed a digital twin exchange for easy access to digital twin data for equipment, facilities, and IoT; it uses real-time data to enable understanding, learning, and reasoning; applications range from the Duke Energy renewables wind turbine fleet to Europe's largest shipping port – Port of Rotterdam (IBM, 2022).

Nevertheless, accounting for the digital twin applications researched in the literature, this dissertation explores an architecture applicable across multiple domains within the smart city framework. This dissertation research seeks to establish that framework. The following three sections of the literature review reflect insight into the three critical technologies required for the conceptual architecture.

Strohbach (2015) relates that an increasing amount of valuable data sources, advances in the Internet of Things and Big Data technologies and the availability of a wide range of machine learning algorithms offer new potential to deliver analytical services to citizens and urban decision-makers. However, a gap exists in combining the current state of the art in an integrated framework that would help reduce development costs and enable new services. Mohammadi (2018) suggests that developing smart cities and their fast-paced deployment generates large quantities of data at unprecedented rates. Unfortunately, most of the generated data is wasted without extracting potentially helpful information and knowledge because of the lack of established mechanisms and standards that benefit from the availability of such data. Moreover, the highly dynamic nature of smart cities calls for a new generation of machine learning approaches that are flexible and adaptable to cope with the dynamicity of data to perform analytics and learn from real-time data. As a means of smart city implementation, I explored research in the area of the digital twin. The seminal work of Grieves (2017) lays out the framing of stripping out the information from a physical object to determine a commensurate object known as the digital twin. Saddik's (2018) research built on Grieves'ideas and moved the digital twin concept away from the primary focus in manufacturing to facilitate the means to monitor, understand, and optimize the functions of all physical entities and for humans to provide continuous feedback in improving quality of life and well-being (El Saddik, 2018) Next, the conceptual idea moves forward by connecting the work to future smart cities dependent on developing systems that can address the computational demands of expanded digitized data and related advanced software in fields such as health and wellness, security and safety, transport and energy, and mobility and communications. In turn, the convergence of technologies and scientific knowledge promises to boost citizens' well-being and quality of life.

2.4 Internet of Things

A literature review duly noted that machine-to-machine (M2M) communication is the precursor to what we now reference as the Internet of Things. The modern implementation of the idea started in the 1970s and was known as pervasive computing. The Internet of Things (IoT) enabled digital twins to become cost-effective so they could become as imperative to business as they are today (Miskinis, n.d.). IoT attempts to solve the challenge of interoperability. It has made significant advances, but much work remains to achieve the level of integration required to unlock the value of advanced software in these domains. With respect to the smart city framework, Jin(Jin et al., 2014) best defined, “Smartness of a city driven and enabled technologically by the emergent Internet of Things (IoT) — a radical evolution of the current Internet into a ubiquitous network of interconnected objects that not only harvests information from the environments (sensing) and interacts with the physical world (actuation/command/control), but also uses existing Internet standards to provide services for information transfer, analytics, and applications.” Viewed through the lens of the Open Systems Interconnection (OSI) model as connected to Ali’s smart city definition, Jin illustrates a view of IoT and smart city integration, as noted in Figure 12:

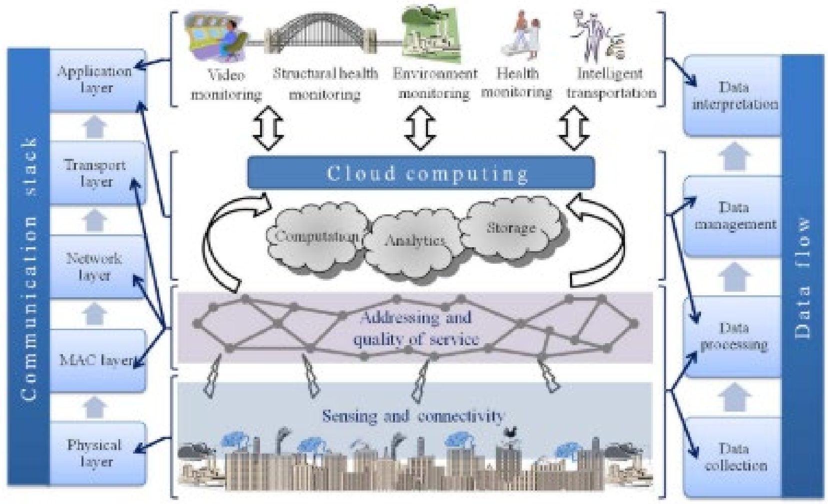

Figure 12: IoT and smart city integration

At a base level, IoT is representative of the physical layer in a smart city's infrastructure; rising to the application layer provides overall framing for the architecture.

Within this applied design research effort, IoT parallels the smart meters of the energy case study (4.3 Case Study 2 – Predictive Twin). The work of Ahuja and Khosla (Ahuja & Khosla, 2019) asserts that smart meter data analytics can optimize, manage, and address peak demand issues for smart grids. Utilities better understand consumers' consumption patterns for providing personalized services to consumers and their own consumption ratings. The many benefits included: 1) Consumption analysis and pattern discovery facilitated for consumers and utilities to take advantage of energy-efficient decision-making, 2) Forecasting of individual customer's energy consumption, and 3) Anomaly detection possible.

IoT is an extension of the Internet into expansively many more devices. In turn, this infrastructure ultimately connects billions of devices to the Internet for processing and analysis, tying in our connection of the smart city to Big Data. In capturing the Big Data aspects of IoT, the quality of data and usage are considered the next concept in the critical technical components of the conceptual model.

2.5 Big Data

A major disrupter in this current digital revolution is data. During the era of prominence for John D. Rockefeller (mid to late 19th century), his company, Standard Oil, disrupted the oil industry and made crude oil the most traded commodity in the world. Data has been touted in many circles as the "new oil." As the fuel of the future, data is a prime mover of global economic growth and change (The Economist, 2017). Nevertheless, Big data is a term that has been a buzzword in the zeitgeist for well over a decade. Big Data is generated from the IoT sensors at the physical layer of the smart city concept. Data is an elemental unit that serves as the engine for this effort. Data is a commodity that typically is not an asset listed on a firm's balance sheet, but most firms will attest to the fact that data is one of the most valuable assets at a company.

Zhang (Zhang et al., 2018) established that Big Data is the "…*amount of data beyond technology's capability to store, manage, and process efficiently as the data size increases along with the evolvement of ICT technologies*. " Constructing a knowledge base of Big Data requires the discernment of Big Data's key characteristics. The literature succinctly relays those key characteristics through the 5V Big Data model – volume, variety, velocity, veracity, and value. Volume refers to the massive datasets generated by the various interactive components of Big Data. Within the context of the smart grid, AMI meters are the prime generators of the volume. Variety references the types of data we can use in the Big Data construct, distinctly categorized as structured (e.g., spreadsheets, formatted dataset) and unstructured (e.g., messages, social media conversations, digital images, sensor data, video or voice recordings). Velocity "…refers to the speed at which new data is generated and the speed at which data moves around (Zhang et al., 2018). If desired, an electric utility with one million customers can use its AMI meters to poll sample data from those customers every 15 minutes; this process would generate 35 billion records in 15 minutes. Veracity "…refers to the messiness or trustworthiness of the data." Accuracy issues can arise in handling such massive datasets, which elicits challenges in the validity of business intelligence results. Lastly, value is the fifth 'v' of the Big Data characteristic model. Within the value proposition of Big Data, the most critical aspect is whether meaningful benefits can be accrued from the large datasets and whether crucial stakeholders understand the value of representing their interests.

In terms of smart city impact, Hashem's paper, "The Role of big data in smart city" ((Hashem et al., 2016)) reflects how Big data is the fuel for the engine in data analytics for solving smart city challenges; it elucidates the landscape of the smart technologies with big data and cloud computing, in which various smart applications exchange information using embedded sensor devices and other devices integrated with the cloud-computing infrastructure (Hashem et al., 2016). Figure 13 ties together the disparate components of Big Data and a smart city concept.

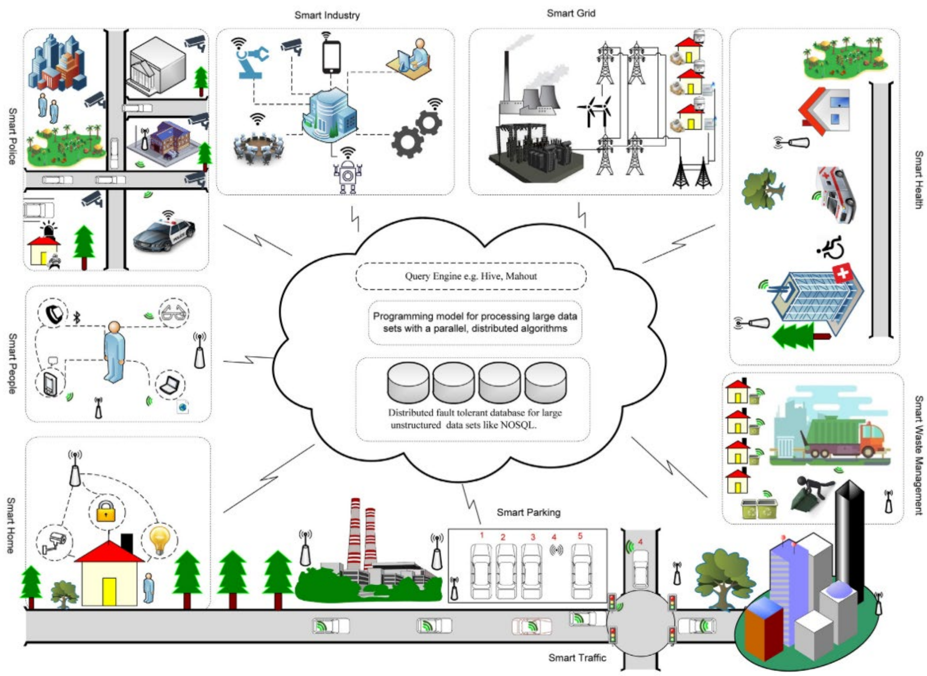

Figure 13: Smart City Data Analytics Connection

Figure 13 additionally ties the relevance of machine learning and data analytics into a clear view. As investigated by Rouse (Rouse, n.d.), the literature conveys data analytics as the process of examining large and varied data sets—that is, Big Data—to uncover hidden patterns, unknown correlations, market trends, customer preferences and other helpful information that can help organizations make more-informed business decisions". In turn, Big Data provides the fuel needed for machine learning to drive the engine of insight in a smart city.

2.6 Machine Learning

The field of machine learning is sufficiently young that it is still rapidly expanding, often by inventing new formalizations of machine-learning problems driven by practical applications. Despite its practical and commercial successes, machine learning remains a young field with many underexplored research opportunities. Two vanguards in the discipline defined machine learning via two unique perspectives. In his 1959 paper “Some Studies in Machine Learning Using the Game of Checkers,” former Stanford University professor A.L. Samuel defined machine learning as a "field of study that gives computers the ability to learn without being explicitly programmed." In his 1997 classic book, “Machine Learning,” former Department Head of Carnegie Mellon University Machine Learning Department, Tom Mitchell, provides an understanding of machine learning from the perspective that “A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.” Insights, from the Big Data attributed to a smart city, can be revealed leveraging the tools of machine learning. Notably, there are many challenges concerning smart cities for machine learning. In the seminal paper, “Enabling Cognitive Smart Cities Using Big Data and Machine Learning: Approaches and Challenges,” Mohammadi probed the convergence of the Internet of Things (IoT), smart city technologies, Big Data, and artificial intelligence techniques as adeptly illustrated in Figure 14.

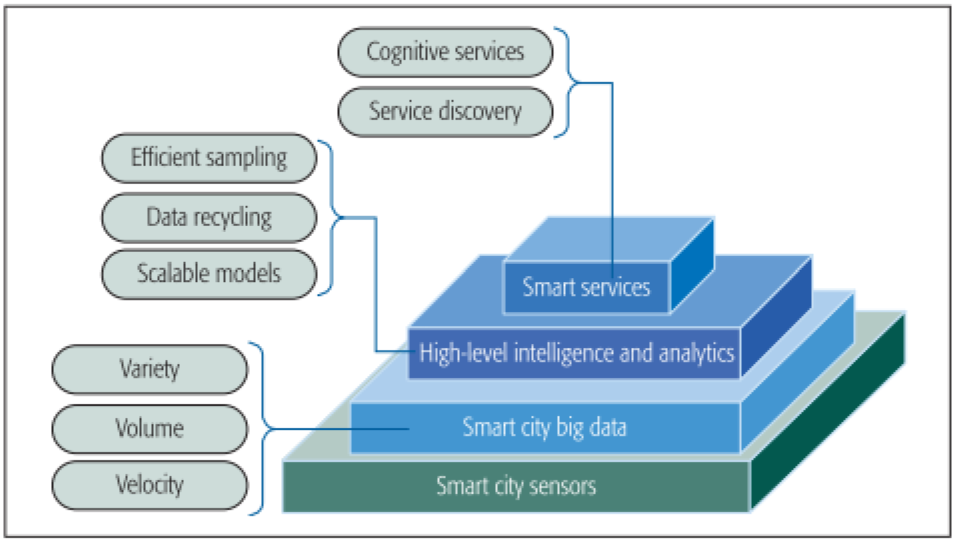

Figure 14: Machine Learning Smart City Representation

A further contribution of Mohammadi's study is the presentation of a taxonomy (Figure 15) of machine learning algorithms explaining how different techniques are applied to the data to extract higher-level information. This taxonomy is vital because it presents a base of algorithms for experimentation for consideration in this dissertation's architecture development process.

Overview of applying machine learning algorithms to Internet of Things use cases.

Machine learning Algorithm	IoT, Smart City use cases	Metric to Optimize
Classification	Smart Traffic	Traffic Prediction, Increase Data Abbreviation
Clustering	Smart Traffic, Smart Health	Traffic Prediction, Increase Data Abbreviation
Anomaly Detection	Smart Traffic, Smart Environment	Traffic Prediction, Increase Data Abbreviation, Finding Anomalies in Power Dataset
Support Vector Regression	Smart Weather Prediction	Forecasting
Linear Regression	Economics, Market analysis, Energy usage	Real Time Prediction, Reducing Amount of Data
Classification and Regression Trees	Smart Citizens	Real Time Prediction, Passengers Travel Pattern
Support Vector Machine	All Use Cases	Classify Data, Real Time Prediction
K-Nearest Neighbors	Smart Citizen	Passengers' Travel Pattern, Efficiency of the Learned Metric
Naive Bayes	Smart Agriculture, Smart Citizen	Food Safety, Passengers Travel Pattern, Estimate the Numbers of Nodes
K-Means	Smart City, Smart Home, Smart Citizen, Controlling Air and Traffic	Outlier Detection, fraud detection, Analyze Small Data set, Forecasting Energy Consumption, Passengers Travel Pattern, Stream Data Analyze
Density-Based Clustering	Smart Citizen	Labeling Data, Fraud Detection, Passengers Travel Pattern
Feed Forward Neural Network	Smart Health	Reducing Energy Consumption, Forecast the States of Elements, Overcome the Redundant Data and Information
Principal Component Analysis	Monitoring Public Places	Fault Detection
Canonical Correlation Analysis	Monitoring Public Places	Fault Detection
One-class Support Vector Machines	Smart Human Activity Control	Fraud Detection, Emerging Anomalies in the data

Figure 15: Notable Taxonomy of Machine Learning Algorithms

Emerging from the exploration of IoT, Big Data, and machine learning, the development of a conceptual model establishes a platform for researching the digital twin implementation of a smart city.

2.7 Behavioral Economic Factors

In modeling, considerations for the impact of error, uncertainty, and unexpected values are part of the implementation process. In establishing a framework for this proposed research, I have sought to model some measure of uncertainty for the human element in the process by considering behavioral economic factors. Behavioral economics began as an academic attempt at modeling irrational consumer choices, challenging the notion of the rational consumer of traditional economics. It is a relatively new field combining insights from psychology, judgment, decision-making, and economics to understand human behavior better (Gino, 2017). The term “behavioral economics” was in use as early as 1958. These days, as it is typically employed, “behavioral economics” refers to the attempt to increase the explanatory and predictive power of economic theory by providing it with more psychologically plausible foundations (Angner, Erik; Loewenstein, 2006). Several smart city projects worldwide have employed the Internet of Things, big data, and machine learning to nudge citizens to save water and energy, live healthily, use public transportation, and participate more actively in local affairs. Combining technology with behavioral insights may allow smart cities to nudge citizens more systematically to achieve their sustainability goals and promote civic engagement. In this study, I plan to incorporate this modeling aspect as a unique development point for research contribution.

2.8 Summary of Literature Review and Implication

The United Nations projects that by 2050, 68% of the world’s population will live in urban centers (Nations, 2018). There exists no debate that every city has opportunities to improve existing operations and services provided to citizens. Batty’s research espouses the value of an optimally functioning smart city, outlining that “…information and communications technologies might improve the functioning of cities, enhancing their efficiency, improving their competitiveness, and providing new ways in which problems of poverty, social deprivation, and poor environment might be addressed”(Batty et al., 2012a) One critical problem for smart city applications is lack of a clear strategic vision for technology integration that fulfills its more significant purpose in the development of thriving and sustainable growth of a city. To combat this issue, studies by scholars such as Howell and Alsamani demonstrate the efficacy of Big Data, IoT, and machine learning as tools for improved system utilization in smart city applications

(Howell, 2017), (Alsamani, 2019). A second vital issue to address in implementing many smart city efforts is the wasted value of generated data. Most of the generated data is wasted without extracting potentially helpful information and knowledge because of the lack of established mechanisms and standards that benefit from the availability of such data (Mohammadi, 2018). Implementation of a smart city in a digital twin framework provides the architecture platform needed to leverage the valuable data generated by a smart city more effectively

As referenced in the literature, the very definition of a smart city is not firmly settled upon by researchers. Nevertheless, for purposes of this study, it is defined as the integrated framework of service, operational, and human technology interactions that allow for the efficient and beneficial needs of citizenry in a populationally condensed urban setting. The literature has more consensus in defining the digital twin as the computer-generated representation of an actual process or entity; it is quite suitable for this dissertation research application for implementation in a smart city framework. This research will also benefit by leveraging the digital twin efforts of Mohammadi (2017) for modeling Atlanta, Georgia, and Dawkins (2018) for modeling London, England.

2.9 Conceptual Framework of the Study

Regarding an overview of the framework, the starting point is a representative smart city digitized via the Microsoft Azure platform (Microsoft, 2021), then optimized as encompassed in the architecture noted in Figure 16.

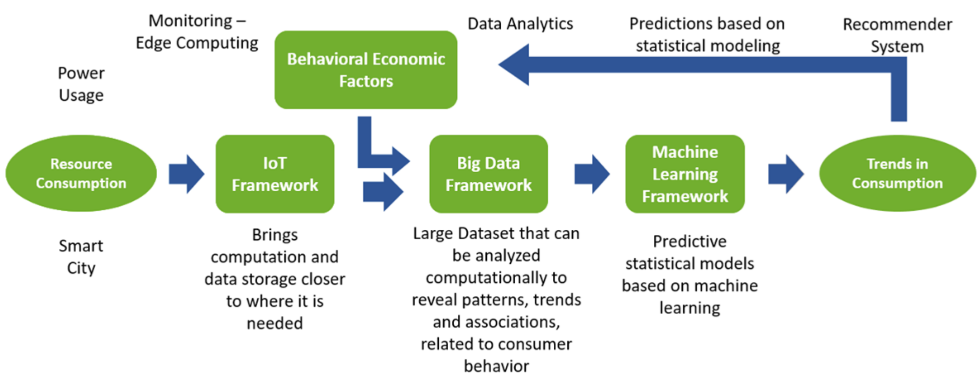

Figure 16: Conceptual Framework

First, the monitoring tier senses components associated with resource consumption. The framework's approach starts with the Internet of Things element connecting the sensors necessary to feed the data required to inform the development of the digital twin. Internet of Things sensors are the digital foundation of a smart city framework. This section of the framework includes the application of edge computing, where the sensors in a subsystem of the smart city tie to the larger, more extensive system with distributed sensors in the field, where computation may occur at a distant location or near the device. Second, a middle layer analyzes data to detect real-time anomalies or patterns. Storage encompasses this aspect as a component of the Big Data framework. Also, this is where data can reveal patterns and trends unseen in previously constructed operational frameworks. The third layer of the conceptual framework involves machine learning to discover the classification of the type of usage for relevant resources, predictive models, and other aspects of discovery feeding into the recommender subnetwork to assess trends and consumption valuations. The architecture connects to the recommender system's interdisciplinary component associated with behavioral economic factors. Those factors additionally serve as input to the Big Data framework. The behavioral economic factors are important because they inject the conceptual measure of uncertainty and irrationality into the framework related to human decision-making.

Generally, developers of digital twin models seek varying implementation approaches and not a specific prescribed and formulaic method. This research proposes establishing a digital twin model leveraging a universal framework based on the three tiers of bedrock technologies utilizing existing standards. First-tier accounts for monitoring and the sensing of smart city resources. The second (middle) tier captures all relevant smart city data in a real-time environment to detect anomalies or patterns of significance relayed in the Big Data framework. ---And then on the back end, the third tier, where machine learning is leveraged for knowledge discovery, as well as classification of the type of usage for notable parameters being measured.

In expanding the accessibility of smart city component implementation, this research proposes a structure within which maturity models from varying contexts yield the framing as points of entry ((Paulk, 2009; Rajteric, 2010)). Figure 17presents the engagement of smart city implementation across varying levels of resources and capabilities.

Digital Twin Conceptual Levels of Development

❑Level 1 Descriptive Twin:	editable version of design – a visual replica of a built asset
❑Level 2 Informative Twin:	twin captures and aggregates defined data and verifies data to make sure that systems work together
❑Level 3 Predictive Twin:	leverage operation data to gain insights
❑Level 4 Comprehensive Twin:	simulates future scenarios and considers “what-if "questions
❑Level 5 Autonomous Twin:	ability to learn and act on behalf of users

Figure 17: Conceptual Levels of Development

Level 1 is the Descriptive Twin level. It is a live, editable design version — a visual replica of a built asset. Users specify what information they want to be included and what data they want to extract. It conveys the most basic level, converting analog data to digital format, and acts as the base for all the other upper levels.

Level 2 denotes the Informative Twin level, with an added operational and sensory data layer. The twin captures and aggregates specific data, then verifies it to ensure that systems work together.

Level 3 is the Predictive Twin level. This twin can use operational data to acquire insights. At this level, insights from the data answer questions about why certain things or situations occur, and data analysis becomes a standard component at this stage.

Level 4 is the Comprehensive Twin level, and this twin simulates future scenarios and considers “what-if” questions. At this level, systems can respond to future scenarios and decide accordingly; artificial intelligence algorithms are also prevalent.

Level 5 is the Autonomous Twin level. This twin can learn and act on behalf of users; the machine recommends the ideal solution and takes action. This level of implementation represents the ultimate sophistication of the conceptual architecture.

These five levels of development allow for implementation to be as straightforward or complex as resources are available for a city.

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter describes the research methodology used in this study. Methodology refers to the overarching research strategy that outlines the approach and procedures to be employed in conducting a study. As mentioned above, the framework encompasses a set of beliefs and philosophical assumptions that influence the comprehension of research inquiries and serve as the foundation for selecting research methodologies. The research methodology is a crucial component as it ensures the alignment between the selected tools, techniques, and underlying philosophical principles. One approach to constructing a research methodology is founded on the theoretical framework known as the "research onion" (Figure 18), which was introduced by University of Birmingham professor Mark Saunders (Saunders et al., 2019).

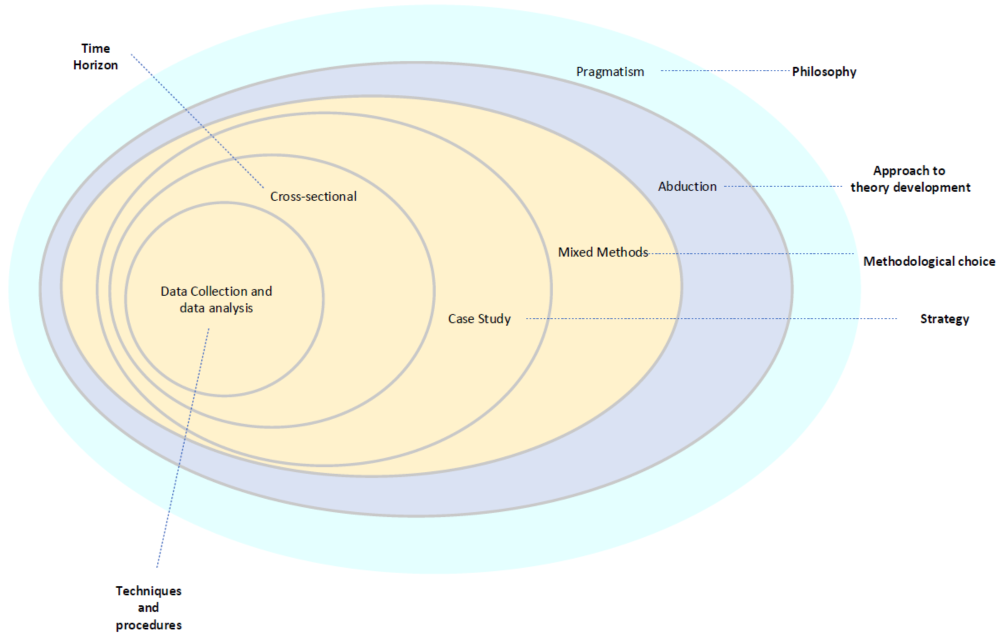

Figure 18: Research Onion , Adapted from Saunders (2023)

A framework for research is the interconnection of philosophical worldview, designs, and research methods. The research onion comprehensively depicts the primary research layers or phases for developing a robust research methodology (Raithatha, 2017). Although the research onion model was traditionally designed for business studies, it applies similarly to this interdisciplinary engineering research effort.

The research methodology commences by first outlining the fundamental philosophy, selecting appropriate approaches, methods, and strategies, and establishing time horizons. These elements collectively guide the research logic toward developing the research design, which encompasses the primary techniques and procedures for data collection and analysis (Figure 18).

3.2 Research Philosophy

In planning a study, researchers need to think through the philosophical worldview assumptions that they bring to the study, the research design that is related to this worldview, and the specific methods or procedures of research that translate the approach into practice (Creswell, J. W., & Creswell, 2017). As a part of developing the research, a researcher must make explicit the more significant philosophical ideas espoused. Research philosophy refers to a system of beliefs and assumptions about knowledge development. In this research, it is referred to as a philosophical worldview. We see worldviews as a general philosophical orientation about the world and the nature of research that a researcher brings to a study. Individuals develop worldviews based on their discipline orientations, research communities, advisors and mentors, and past research experiences. In the knowledge development process, selecting a research philosophy sets out the worldview for conducting research.

3.2.1 Research Philosophy Assumptions

Every stage of research imposes several types of assumptions. The assumptions are categorized as ontological, epistemological, and axiological. The research philosophy serves as a foundational element of the research process, as it defines ontology, epistemology, and axiology. Ontology refers to understanding reality, epistemology pertains to the nature and sources of knowledge or facts, and axiology encompasses the values, beliefs, and ethics that guide the research. These assumptions inevitably shape the understanding of research questions, methods used, and interpretation of findings (Crotty 1998). A well-thought-out and consistent set of assumptions constitutes a credible research philosophy and shapes the choice of research question. In turn, this underpins the methodological choice, research strategy, data collection procedures, analysis techniques, reporting of findings, discussion, and conclusion. This process allows for the design of a coherent research project in which all elements of research fit together.

First, ontology refers to assumptions about the nature of reality; these assumptions shape how research objects are seen and studied. For this research, the reality studied was resource consumption concerning electric utility energy. Second. epistemology refers to assumptions about what constitutes acceptable, valid, and legitimate knowledge and how we can communicate knowledge to others (Burrell and Morgan 2016). Whereas ontology may initially seem rather abstract, the relevance of epistemology is more prominent. The interdisciplinary context of business and engineering means that different types of knowledge – ranging from numerical data to textual and visual data, can all be considered legitimate. This set of assumptions answers the questions: 1) What constitutes good-quality data? and 2) What kinds of contributions to knowledge can be made? Lastly, axiology refers to the role of values and ethics in the research process. One of the vital axiological choices faced as a researcher is the extent to which one views the impact of values and beliefs on his research as a positive contribution.

3.2.2 Pragmatism

In alignment with the noted philosophy assumptions, this research fits closely with the pragmatism worldview. Pragmatism asserts that concepts are only relevant where they support action (Kelemen and Rumens 2008). Pragmatism originated in the late-nineteenth–early-twentieth-century USA in the work of philosophers Charles Pierce, William James, and John Dewey. It considers theories, concepts, ideas, hypotheses, and research findings not in an abstract form but in terms of their roles as instruments of thought and action and their practical consequences in specific contexts. Reality matters to pragmatists as practical effects of ideas and knowledge are valued for enabling actions to be carried out successfully.

Ontology (nature of reality or being)	Epistemology (what constitutes acceptable knowledge)	Axiology (role of values)	Typical methods
	Pragmatism		
Complex, rich, external 'Reality' is the practical consequences of ideas Flux of processes, experiences and practices	Practical meaning of knowledge in specific contexts 'True' theories and knowledge are those that enable successful action Focus on problems, practices and relevance Problem solving and informed future practice as contribution	Value-driven research Research initiated and sustained by researcher's doubts and beliefs Researcher reflexive	Following research problem and research question Range of methods: mixed, multiple, qualitative, quantitative, action research Emphasis on practical solutions and outcomes

Table 1:Philosophical assumptions of pragmatism, Saunders (2019)

For a pragmatist, research starts with a problem and aims to contribute practical solutions that inform future practice. Researcher values drive the reflexive process of inquiry, initiated by doubt and a sense that something is wrong or out of place, and recreates belief when the problem has been resolved (Elkjaer and Simpson 2011). In pragmatist research, the most crucial determinant for this research design and strategy was the research problem addressed and the research question. The research question, in turn, would likely incorporate the pragmatist emphasis on practical outcomes.

Suppose a research problem does not suggest unambiguously adopting one particular knowledge or method. In that case, this only confirms the pragmatist's view that working with different types of knowledge and methods is perfectly possible. Pragmatists recognize that there are many different ways of interpreting the world and undertaking research, that no single point of view can ever give the entire picture, and that there may be multiple realities. This does not mean that pragmatists always use multiple methods. Instead, they use the method that enables credible, well-founded, reliable, and relevant data to be collected to advance the research (Kelemen and Rumens 2008). Following consideration of philosophy, an approach to theory development was considered.

3.2.3 Applied Research

In connection with the overall worldview of pragmatism, this research effort was aligned with applied research as compared to that of basic research. Basic research is grounded firmly in the experimental method and aims to create new knowledge about how fundamental processes work. Applied research is also rooted in the experimental method, but it uses scientific methodology to develop information to clarify or confront an immediate societal problem. Demonstrating the conceptual model using the case study strategy employed the scientific methodology to establish smart city architecture.

3.3 Research Approach

The research approach to theory development follows the research philosophy selection of the research onion (Figure 18) and commonly encompasses one of three methods: deduction, induction, or abduction. Deduction involves commencing the research process with an established theory, posing a question or hypothesis, and collecting data to validate or refute the hypothesis. On the other hand, induction entails initiating the research with observation and data collection, progressing towards description and analysis to formulate a theory. Abduction refers to observing empirical phenomena, followed by research that generates a best guess or conclusion based on the available evidence. The work of Saunders additionally yields credence in this consideration of five areas of focus for selecting a research approach: 1) logic, generalizability, use of data, theory, and philosophical underpinning. In this regard, the researcher selected induction as the research approach, as discerned in Table 2: Research Approach Selection

.

Category	**Induction Rationale**	**Connection to Research**
Logic	In inductive inference, known premises are used to generate untested conclusions.	Known arguments digital twin implementation allow for inductive

		logic application to the research case studies.
Generalizability	Generalizing from the specific to the general	The structure of the research development case allows for generalization to other domains.
Use of data	Data collection explores a phenomenon, identifies themes and patterns, and creates a conceptual framework.	Research utilization of AMI data applies in this area.
Theory	Theory generation and building	Builds on a broader concept of systems engineering theory
Philosophical underpinning	Pragmatism	Connects to applied research design effort

Table 2: Research Approach Selection

Induction best fits the research effort's aims in evaluating the five categories of the research approach. Qualitative researchers typically work inductively, building patterns, categories, and themes from the bottom by organizing the data into increasingly more abstract information units. This inductive process illustrates working back and forth between the themes until the researchers have established a comprehensive set of themes.

(Creswell, John W.; Creswell, J. David. Research Design (p. 181). SAGE Publications. Kindle Edition.

3.4 Research Methodology

The methodology employed encompassed many aspects of qualitative research. Qualitative research explores the meaning individuals or groups ascribe to a social or human problem; smart city research squarely falls within this categorization. The process of qualitative research involves emerging questions and procedures, data typically collected in the participant's setting, data analysis inductively building from particulars to general themes, and the researcher making interpretations of the meaning of the data.(Creswell, J. W., & Creswell, 2017)

However, some quantitative methods were employed during the research process as well. As such, the research methodology utilized in the study was deemed mixed methods.

3.5 Research Strategy

The next step in determining the methodology was to assess a strategy to conduct the applied research design; the case study strategy was selected. A case study is an empirical inquiry investigating a contemporary phenomenon in depth and within its real-life context, especially when the boundaries between phenomenon and context are not evident (Yin, 2017). Case study research is also concerned with studying the phenomenon in context, so the findings generate insight into how the phenomenon occurs within a given situation (Farquhar, Jillian Dawes. Case Study Research for Business (p. 6). SAGE Publications. Kindle Edition

Case studies are critical elements of the research and validation process since they allow a theoretical process or proposed model to be executed, measured, and refuted against reality. Case study research designs or approaches can be chosen based on their type, characteristics, or disciplinary orientation. One’s selection of a particular research design is determined by how well it allows full investigation of a specific research question.

Types of case study research designs include exploratory, explanatory, and descriptive (Yin, 2003, 2013). Exploratory designs seek to define the research questions of a subsequent study or determine the research procedures' feasibility. These designs are often a prelude to additional research efforts and involve fieldwork and information collection before defining a research question. Explanatory designs seek to establish cause-and-effect relationships. Their primary purpose is to determine how events occur and which ones may influence particular outcomes. Descriptive designs illustrate or explain a phenomenon's critical features within its context. This thesis utilized the exploratory design type.

3.6 Research Time Horizon

Considering time horizons, an important question to be asked in designing research is, 'Do I want my research to be a "snapshot" taken at a particular time, or do I want it to be more akin to a diary or a series of snapshots and be a representation of events over a given period?' This consideration will, of course, depend on selected research questions. The 'snapshot' refers to the cross-sectional time horizon, while the longitudinal time horizon encompasses the 'diary' perspective. Cross-sectional studies involve the study of a particular phenomenon (or phenomena) at a particular time. This is attributable to the fact that most research projects undertaken for academic study are necessarily time-constrained. In that regard, the case studies employed for this effort leveraged a cross-sectional time horizon.

3.7 Data Management and Analysis

The techniques and procedures employed in this study encompass various aspects of data collection and analysis. These include the utilization of secondary data sources and the selection of appropriate sample groups. When considering how to obtain data to answer research question(s) or meet objectives, researchers are often expected to consider further analyses of data collected initially for some other purpose. Such data are secondary data and include raw data and published summaries. Once obtained and evaluated as suitable to help answer the research question, these data can be further analyzed to provide additional or different knowledge, interpretations, or conclusions (Bishop and Kuula-Luumi 2017; Bulmer et al. 2009). Considering the constraints involved in this study, secondary data provided distinct advantages. First, fewer resource requirements were a clear advantage.

For many research questions and objectives, a key advantage of using secondary data is the enormous saving in resources compared to collecting your own (primary) data. In general, using secondary data is much less expensive and time-consuming than collecting the data on an individual basis, especially when data can be downloaded in a format compatible with a researcher's analysis software. Secondly, secondary data has the advantage of being unobtrusive because it has already been collected and removes the expense of collecting firsthand. Lastly, a more accessible ethical review can be achieved. Secondary data are often in the public domain, and many data sets have already been anonymized.

The conceptual model of this research focuses on the residential customer class of retail electricity customers. The necessary data comes from Pecan Street Research (Austin, TX). Launched in 2009, Pecan Street's research network serves as the planet's only real-world electricity-gas-water test bed. It includes 1,115 active homes and businesses, 250 solar homes, and 65 electric vehicle owners. Each home's energy generation and use is measured at intervals ranging from one second to one minute and can be analyzed down to the circuit level. This high-resolution data provides insight into how energy is used, generated, and stored. Data is organized into schemas by type of data (electricity, water, gas, static data) rather than by geographic location. Data is available for each home and correlated to a unique Data ID.

3.8 Role of the Researcher

One of the characteristics of a qualitative approach to research is reflexivity, where the researcher reflects on their role in the study and how their background, experiences, and beliefs may shape the direction of the study (Creswell, John W.; Creswell, 2018). The researcher has been involved in the energy industry for 33 years and smart city research for six years. For the last ten years, the researcher worked as a Project Manager and Senior analyst, modeling and analyzing the cost of retail energy service and avoided cost analysis in the renewable energy space.

The researcher observed that this background and experience provide insight into connecting the theory surrounding smart cities with the application in the energy case study.

3.9 Summary of Methodology and Research Design

The methodological approach chosen for this study was the Qualitative approach utilizing the Case Study strategy for implementation. In this respect, therein lies an appropriate approach concerning the philosophy of pragmatism. Ultimately, the intended aim is the development of resource consumption digital twin model applications across multiple domains in addition to the electric utility industry.

CHAPTER 4

FINDINGS

4.1 Introduction

This chapter analyzes two case studies to demonstrate the implementation of the conceptual framework. Through the model explanations, data analyses, and documentation, the study allows one to construct a system for managing city-wide systems of systems constrained by resource consumption. Nevertheless, as referenced in Chapter 2, this study focuses on smart cities and smart energy as a sub-system for specific consideration. The case studies Findings chapter provides the practical implications of the innovation strategies implemented by digital twin configurations.

4.2 Case Study 1 – Descriptive Twin

Case Study 1 research involves the development of a Level 1 implementation for the smart city digital twin prototype of the conceptual model; smart energy is the subset of the smart cities research space. The Level 1 structure establishes the Descriptive Digital Twin, allowing users to understand related components and the relationship between the components that make up the system.

4.2.1 Case Study 1 Approach

This research captured the representation of the 25 homes of the Pecan Street AMI meter dataset from Austin, TX. Notably, this residential home representation reflects a digital twin of a connected community such as those specified in Figure 19 's Alabama Power Reynold's Landing project (Energy, 2023a) or Wisconsin's Connecting Communities for Sustainable Solutions project (Energy, 2023b)

Figure 19: Example Smart Home Connected Community

To distill the Level 1 representation, Azure Digital Twin Definition language served as the resource to translate the virtual to the physical representation of the digital twin. Microsoft's Digital Twins Definition Language (DTDL) provides the platform for modeling the noted sensors. DTDL uses a variation of the JavaScript Object Notation (JSON) format, namely JavaScript Object Notation for Linked Data (JSON-LD). JSON-LD is a lightweight Linked Data format and provides a way to help JSON data interoperate at Web-scale. Microsoft developed DTDL to describe models that include

IoT devices (such as AMI meters), Digital Twins, and asset Digital (Nath, S. V., Van Schalkwyk, P., & Isaacs, 2021).

4.2.2 Case Study 1 Detailed Activities

Case Study 1 disburses a Level 1 implementation of the framework. The expectation was to develop a structure that could be implemented for a city of any size in the smart city construction. The objective ties back to the framing of an overall architecture for this case study, not the complete construction of a Descriptive Twin. To assess the residential neighborhood components in the Azure platform, the research leveraged the work of existing smart city / smart building ontologies. Smart Cities ontologies were explored to leverage a range of benefits, including flexibility and seamless integration with other components of the Azure platform. Developers have the ability to describe twins within the context of DTDL by specifying the telemetry emitted by the twins, the properties they report or synchronize, and the commands they are capable of responding to (Russom, n.d.). However, given that the focus of Case Study 1 is a residential customer base, the Digital Twins Definition Language-based RealEstateCore ontology for smart buildings was best suited for appropriate representation (Hammar & Wallin, n.d.; Szcodronski, n.d.). RealEstateCore is a widely adopted linguistic framework utilized for modeling and managing buildings, thereby streamlining the process of creating novel services. The ontology is comprehensive and extensive, offering ease of understanding and practical relevance through established collaborations and solutions within the industry. This technology has been implemented in substantial real estate portfolios for the past few years. It has undergone multiple iterations in response to feedback and experiential knowledge gained from practical use. The primary objective of RealEstateCore is not to establish a novel standard but rather to serve as a unifying factor and facilitate interoperability with existing building industry standards like Brick Schema, Project Haystack, W3C Building Topology Ontology (W3C BOT), and others.

The RealEstateCore structure illustrated in Figure 20 consists of four primary interfaces: 1) ***Space*** refers to a connected portion of the hysical world that possesses a three-dimensional spatial extent and can encompass or include sub-spaces. Case Study 1 demonstrates ***Space*** in the context of a city. The "Building" structure of the ontology distinguishes the residential house in the smart energy case. 2) ***An asset*** is formally defined as an object within a building but does not constitute an essential component of the building's structure, such as architectural elements, furniture, equipment, systems, etc. Case Study 1 extends the interpretation to include limited outside components such as solar panels and batteries. 3) ***Capability*** refers to the inherent ability of an entity, such as a Space, an Asset, or a Logical Device, to generate or receive data. There are distinct subclasses that exhibit specialized behavior in this context. Sensor entities are responsible for collecting data from the physical environment, Actuator entities are designed to receive commands from a digital twin platform, and Parameter entities are responsible for configuring specific capabilities or systems. Concerning Case Study 1, ***Capability*** relates to our research's IoT device of the experiment, the AMI meter. 4) ***LogicalDevice*** A physical or logical entity characterized as an electronic device or software that engages in communication and interaction with a digital twin platform. A logical device encompasses various forms, such as an integrated circuit embedded within a smart HVAC unit. Regarding Case Study 1, the Logical Device relates to utilizing the Microsoft software programs leveraged for this research portion – Azure DTDL, Excel, and Visual Studio Code.

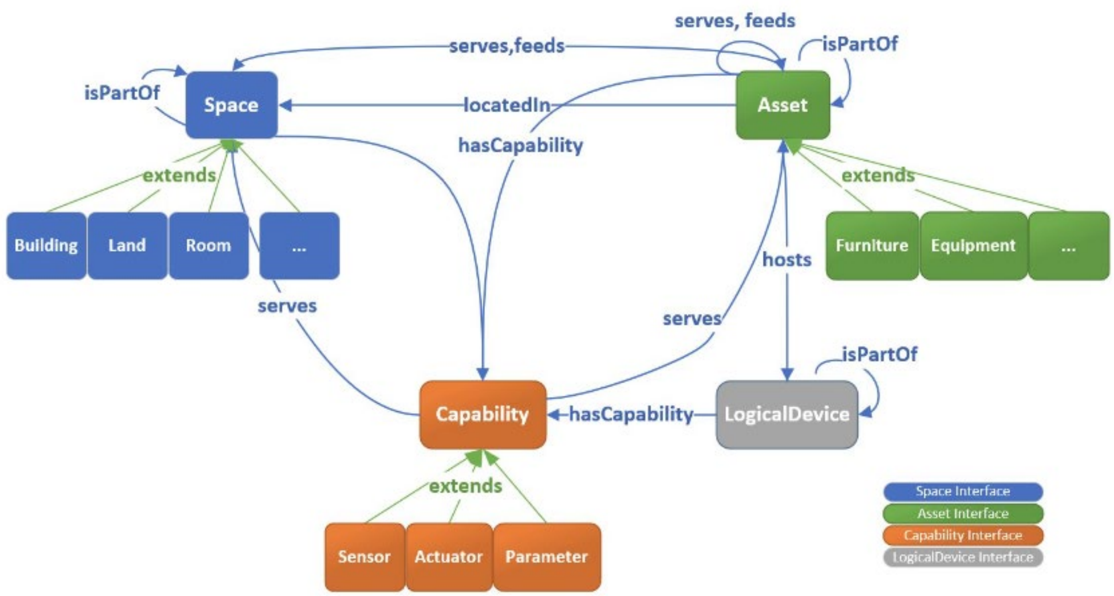

Figure 20 - RealEstateCore Ontology

4.2.3 Case Study 1 Results

In modeling with DTDL, a domain vocabulary was established as the foundation for this smart energy case, and from the RealEstateCore ontology emerged the components allowing for descriptions of 1) important entities and concepts, 2) entity relationships and connections, and 3) data and behavior. The JSON-LD models were developed for city, building, AMI meter, and solar panel concepts within the Azure model structure. Using JSON models reduced system complexity and provided input for future visual analysis. The following pages list the JSON code used to establish the model for the smart energy environment.

JSON Code - Descriptive Twin 1 – City of Austin

{
  "@id": "dtmi:digitaltwins:rec_3_3:core:City;1",
  "@type": "Interface",
  "contents": [
    {
      "@type": "Relationship",
      "displayName": {
        "en": "operated by"
      },
      "name": "operatedBy",
      "target": "dtmi:digitaltwins:rec_3_3:core:Agent;1"
    },
    {
      "@type": "Relationship",
      "description": {
        "en": "Member entities of a collection; e.g., a Campus includes some Space, an Apartment includes some Room, etc. Inverse of: includedIn"
      },
      "displayName": {
        "en": "includes"
      },
      "minMultiplicity": 0,
      "name": "includes",
      "target": "dtmi:digitaltwins:rec_3_3:core:Space;1"
    }
  ],
  "description": {
    "en": "A city represents a collection of location entities. The constituent locations may have differing legal ownership and utilization purposes, but they are generally perceived

as a coherent unit or sub-region within a city or other region. E.g., a university campus, a hospital campus, a corporate campus, etc."

  },

  "displayName": {

    "en": "Austin_City"

  },

  "extends": "dtmi:digitaltwins:rec_3_3:core:SpaceCollection;1",

  "@context": "dtmi:dtdl:context;2"

}

JSON Code - Descriptive Twin 2 – Austin House 661


```
{
 "@id": "dtmi:digitaltwins:rec_3_3:core:House;1",
 "@type": "Interface",
 "contents": [
  {
   "@type": "Component",
   "displayName": {
    "en": "address_661"
   },
   "name": "address",
   "schema": "dtmi:digitaltwins:rec_3_3:addressing:Address;1"
  },
  {
   "@type": "Relationship",
   "description": {
    "en": "Parthood traversal property linking Houses to the Housing Components that they are made up of. Inverse of: isPartOfHouse, componentOfHouse"
   },
   "displayName": {
    "en": "has house component"
   },
   "name": "hasHouseComponent",
   "target": "dtmi:digitaltwins:rec_3_3:core:HousingComponent;1"
  }
 ],
 "description": {
  "en": "A confined housing structure."
 },
```

```
  "displayName": {
    "en": "Housing"
  },
  "extends": "dtmi:digitaltwins:rec_3_3:core:Space;1",
  "@context": "dtmi:dtdl:context;2"
}
```

JSON Code - Descriptive Twin 3 – AMI Meter 661


```
{
  "@id": "dtmi:digitaltwins:rec_3_3:asset:MeterEquipmentGroup;1",
  "@type": "Interface",
  "contents": [
    {
      "@type": "Relationship",
      "description": {
        "en": "Member entities of a collection; e.g., a City includes some Space, a House
includes some Room, etc. Inverse of: includedIn"
      },
      "displayName": {
        "en": "includes"
      },
      "minMultiplicity": 0,
      "name": "includes",
      "target": "dtmi:digitaltwins:rec_3_3:asset:Meter;1"
    }
  ],
  "displayName": "Meter",
"description": "Physical asset that performs the metering role of the usage point. Used for
measuring consumption and detection of events.",
"comment": "Adapted from CIM and https://github.com/smart-data-
models/dataModel.EnergyCIM/ data models",
"contents": [
    {
        "@type": "Property",
        "name": "connectionCategory",
        "comment": "A code used to specify the connection category, e.g. low voltage,
where the meter operates.",
```

"schema": "string",

"writable": true

},

{

"@type": "Property",

"name": "formNumber",

"comment": "Meter form designation per ANSI C12.10 or other applicable standard. An alphanumeric designation denoting the circuit arrangement for which the meter is applicable and its specific terminal arrangement.",

"schema": "string",

"writable": true

},

{

"@type": "Relationship",

"name": "meterReading",

"target": "dtmi:digitaltwins:ngsi_ld:cim:energy:MeterReading;1",

"displayName": "meterReading"

}

],

"@context": "dtmi:dtdl:context;2"

}

JSON Code - Descriptive Twin 4 – Austin Solar Panel 661

{

"@id": "dtmi:digitaltwins:rec_3_3:asset:ElectricalGenerationStorageEquipment;1",

"@type": "Interface",

"displayName": {

```
    "en": "Solar_Panel"
  },
  "extends": "dtmi:digitaltwins:rec_3_3:asset:ElectricalEquipment;1",
  "@context": "dtmi:dtdl:context;2"
}
```

Evaluation of relationships within the digital twin virtual-to-physical (Attaran & Celik, 2023)construct presents pertinent benefits of the configuration. A vital aspect of this representation is the digital twin graph (Figure 21).

Model_ID	Twin_ID	RelationshipName	TargetTwin_ID	TwinInitData
dtmi:CaseStudy1:City;1	CityAustin			{"Date": 2018}
dtmi:CaseStudy1:House;1	House_ID661	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID1642	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID2335	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID2361	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID2818	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID3039	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID3456	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID3538	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID4031	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID4373	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID4767	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID5746	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID6139	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID7536	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID7719	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID7800	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID7901	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID7951	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID8156	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID8386	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID8565	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID9019	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID9160	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID9278	CityAustin	part of	
dtmi:CaseStudy1:House;1	House_ID9922	CityAustin	part of	
dtmi:CaseStudy1:AMI;1	AMI661	House_ID661	IsLinkedTo	{"Demand": 0}
dtmi:CaseStudy1:AMI;1	AMI1642	House_ID1642	IsLinkedTo	{"Demand": 0}
dtmi:CaseStudy1:AMI;1	AMI2335	House_ID2335	IsLinkedTo	{"Demand": 0}
dtmi:CaseStudy1:AMI;1	AMI2361	House_ID2361	IsLinkedTo	{"Demand": 0}
dtmi:CaseStudy1:AMI;1	AMI2818	House_ID2818	IsLinkedTo	{"Demand": 0}
dtmi:CaseStudy1:AMI;1	AMI3039	House_ID3039	IsLinkedTo	{"Demand": 0}
dtmi:CaseStudy1:AMI;1	AMI3456	House_ID3456	IsLinkedTo	{"Demand": 0}
dtmi:CaseStudy1:AMI;1	AMI3538	House_ID3538	IsLinkedTo	{"Demand": 0}

Figure 21: Digital Twin Graph

4.3 Case Study 2 – Predictive Twin

4.3.1 Case Study 2 Approach

Case Study number two captures the research focus on conceptualizing a predictive twin using the conceptual framework. Given the high failure rate of AI projects, the lack of understanding that businesses have regarding how machine learning works, and the long length of time each project takes, Microsoft and other companies have worked to develop solutions that allow faster development and a higher success rate. One such solution is automated machine learning (AutoML). By automating a lot of the work that data scientists do and harnessing the power of cloud computing, AutoML on Azure allows data scientists to work faster and even non-specialists to build AI solutions. AutoML on Azure transforms data, builds models, and tunes hyperparameters.

Following the AutoML approach allows for a faster fail rate and moves forward to a state for decision-making between adding more data or dropping the project. Instead of wasting time tuning models without a chance of working, AutoML presents a definitive answer after only a single AutoML run ().

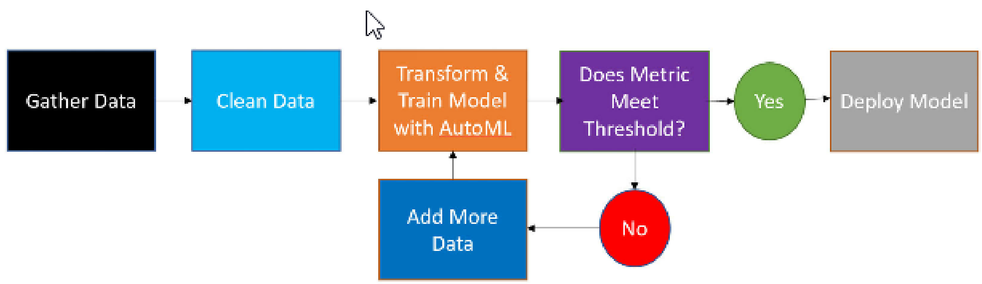

Figure 22: Microsoft Azure Auto ML

The Microsoft Azure AutoML module provided the most suitable mechanism to validate and test the conceptual model.

4.*3.2* Case Study 2 *Detailed Activities*

The dataset's exploratory data analysis (EDA) was conducted using Microsoft Excel. Data analysis is ready to begin once data has been entered, checked, and errors corrected. Tukey's exploratory data analysis (EDA) is a valuable approach in these initial stages. This approach emphasizes using graphs to explore and understand data. Tukey (2020) also emphasizes the importance of using data to guide choices of analysis techniques. As expected, it is essential to remember the research question(s) and objectives when exploring data. Fortunately, exploratory data analysis allows the flexibility to introduce previously unplanned analyses to respond to new findings. It, therefore, formalizes the common practice of looking for other relationships in data for which research was not initially designed to test. This should not be discounted, as it may suggest other fruitful avenues for analysis.

As an illustration, results for Pecan Street building 661 are noted below:

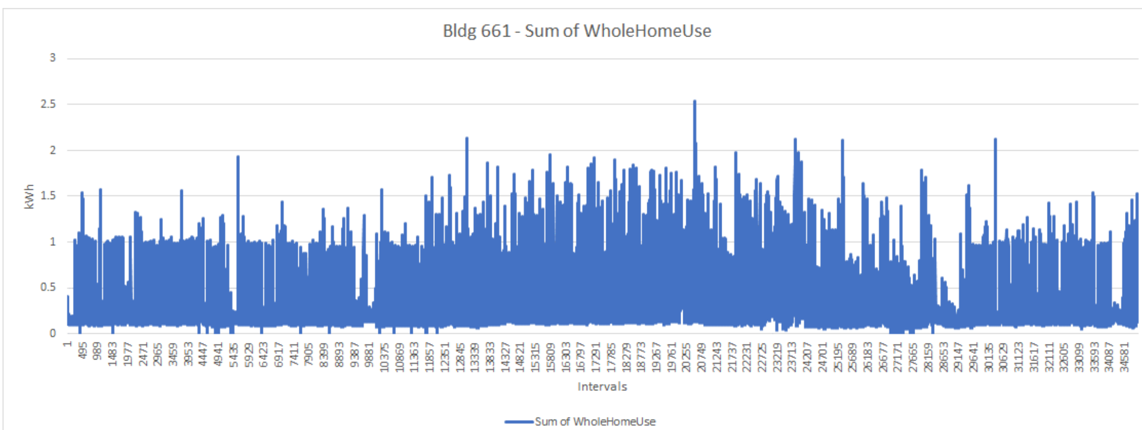

Figure 23: Bldg 661 Visual Analysis

Figure 23 provides a visual analysis of one of the Austin, TX homes referenced in the study, yielding detailed and compact 15-minute intervals for 12 months of data.

Table 3 continues the exploratory data analysis effort and establishes a layout of summary statistics in determining the reasonableness of the dataset.

Sum of WholeHomeUse		*Sum of grid_kWh*		*Sum of solar2_kWh*		*Sum of solar_kWh*	
Mean	0.4	Mean	0.2	Mean	0	Mean	0.2
Standard Error	0.0	Standard Error	0.0	Standard Erro	0	Standard Err	0.0
Median	0.2	Median	0.1	Median	0	Median	0.0
Mode	0.0	Mode	0.0	Mode	0	Mode	0.0
Standard Deviation	0.3	Standard Deviation	0.5	Standard Dev	0	Standard De	0.4
Sample Variance	0.1	Sample Variance	0.2	Sample Variai	0	Sample Vari	0.1
Kurtosis	1.6	Kurtosis	0.6	Kurtosis	#DIV/0!	Kurtosis	0.8
Skewness	1.5	Skewness	-0.2	Skewness	#DIV/0!	Skewness	1.5
Range	2.5	Range	3.3	Range	0	Range	1.4
Minimum	0.0	Minimum	-1.2	Minimum	0	Minimum	0.0
Maximum	2.5	Maximum	2.1	Maximum	0	Maximum	1.4
Sum	12945.3	Sum	5286.4	Sum	0	Sum	7658.8
Count	35032.0	Count	35032.0	Count	35032	Count	35032.0

Table 3: Bldg 661 Summary Statistics

Following the exploratory data analysis stage, evaluation of the Azure model implementation ensued.

Although the exact implementation details can vary, the general structure of a machine-learning project stays relatively constant:

1. Data cleaning and formatting
2. Exploratory data analysis
3. Feature engineering and selection
4. Establish a baseline and compare several machine learning models on a performance metric
5. Perform hyperparameter tuning on the best model to optimize it for the problem.
6. Evaluate the best model on the testing set
7. Interpret the model results to the extent possible
8. Draw conclusions and write a well-documented report

In exploring the conceptual framework for Case Study 2, prototype models were analyzed utilizing the Microsoft Azure Digital Twins Platform and Microsoft Excel. Azure Digital Twins is a platform as a service (PaaS) offering that enables the creation of twin graphs based on digital models of entire environments. Models are defined in a JSON-like language called Digital Twins Definition Language (DTDL), and they describe twins in terms of their state properties, telemetry events, commands, components, and relationships. Microsoft Excel is a spreadsheet analysis and visualization tool I have used for over 20 years for business analysis; it served as a proxy for the Big Data source and EDA tool for the research effort.

Several datasets were considered to develop the conceptual model's proof of concept. Datasets from Southern Company (NYSE: SO) and Exelon Corporation (NasdaqGS: EXC) were initially evaluated as preferred dataset choices. After careful consideration, the dataset made available by Pecan Street proved the optimum choice. Pecan Street Inc. is a 501(c)(3) research and development organization. Its mission aims, and goals fit succinctly with that of my research. The Dataport product from Pecan Street is the world's largest residential energy use data resource. With datasets available from Texas, California, New York, and Colorado, the dataset from Austin, Texas, was selected for this research effort. Twenty-five homes (Figure 24) from the Austin dataset were selected for the conceptual design.

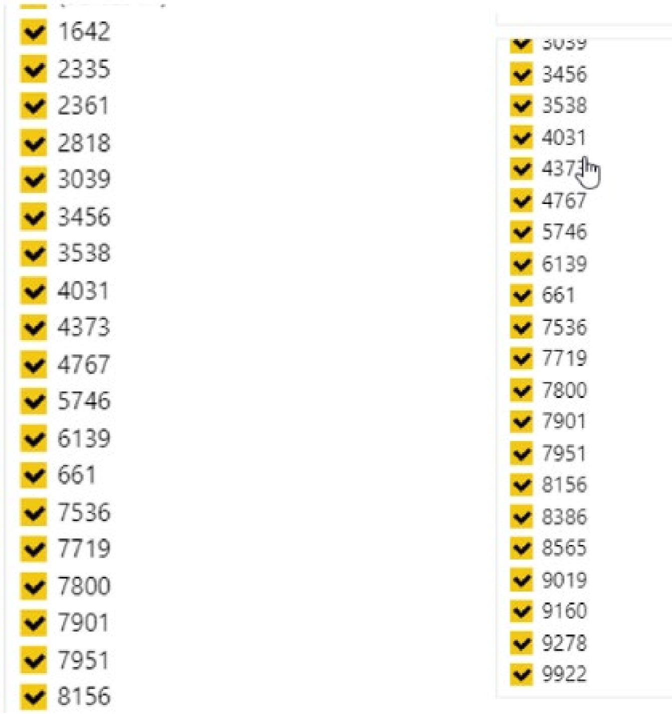

Figure 24: Figure 24: House ID Markers for Pecan Street Dataset

Pecan Street's metadata for each home provides up to 120 fields of variables to be assessed in the analysis of the model. In Case Study2, six variables (Figure 25) encompassed inclusion for the Azure Auto ML predictive model: 1) electric vehicle energy consumption (kWh), 2) household energy consumption (kWh), 3) solar panel energy production (kWh), 4) net energy consumption (kWh), 5) hourly temperature, and 6) simulated variable (*with a range of 0 to 2*) representative of human decisions in energy consumption. Variables one through four were derived from the Pecan Street dataset. Variable number five (hourly temperature) data was obtained via NOAA's database for Past Weather by Zip Code. The variable number six data listed as "SimBehavData" was simulated using a random number generator to account for some randomness noted in human behavior.

Int_MnthCnt	Date	Sum of car1_kWh	Sum of grid_kWh	Sum of solar_kWh	Sum of NetConsump	Sum of HLY-TEMP-NORMAL	SimBehavData
1	7/1/18 12:00 AM	0.00075	0.1595	-0.0005	0.159	77.4	1.03938858
2	7/1/18 12:15 AM	0.00075	0.70225	-0.0005	0.70175	77.4	1.098815374
3	7/1/18 12:30 AM	0.00075	0.14925	-0.0005	0.14875	77.4	1.191308268
4	7/1/18 12:45 AM	0.00075	0.635	-0.0005	0.6345	77.4	1.973943409
5	7/1/18 1:00 AM	0.00075	0.1585	-0.0005	0.158	76.1	0.735467126
6	7/1/18 1:15 AM	0.00075	0.13125	-0.0005	0.13075	76.1	1.523744749
7	7/1/18 1:30 AM	0.00075	0.6545	-0.0005	0.654	76.1	0.916548529
8	7/1/18 1:45 AM	0.00075	0.1445	-0.0005	0.144	76.1	1.757895193
9	7/1/18 2:00 AM	0.00075	0.1725	-0.0005	0.172	75.3	0.160127893
10	7/1/18 2:15 AM	0.00075	0.619	-0.0005	0.6185	75.3	0.729407564
11	7/1/18 2:30 AM	0.00075	0.149	-0.0005	0.1485	75.3	1.136328746
12	7/1/18 2:45 AM	0.00075	0.1485	-0.0005	0.148	75.3	0.407740319
13	7/1/18 3:00 AM	0.00075	0.64625	-0.0005	0.64575	74.7	0.243547383
14	7/1/18 3:15 AM	0.00075	0.143	-0.0005	0.1425	74.7	0.352399173
15	7/1/18 3:30 AM	0.00075	0.13125	-0.0005	0.13075	74.7	1.785292069
16	7/1/18 3:45 AM	0.00075	0.13175	-0.0005	0.13125	74.7	0.697321294
17	7/1/18 4:00 AM	0.00075	0.6085	-0.0005	0.608	74.3	0.452357895
18	7/1/18 4:15 AM	0.00075	0.1795	-0.0005	0.179	74.3	0.018526154
19	7/1/18 4:30 AM	0.00075	0.1515	-0.0005	0.151	74.3	1.666967389
20	7/1/18 4:45 AM	0.00075	0.15	-0.0005	0.1495	74.3	1.449073049
21	7/1/18 5:00 AM	0.00075	0.14975	-0.0005	0.14925	73.9	0.226635731
22	7/1/18 5:15 AM	0.00075	0.60525	-0.0005	0.60475	73.9	0.383210854
23	7/1/18 5:30 AM	0.00075	0.149	-0.0005	0.1485	73.9	1.426093161
24	7/1/18 5:45 AM	0.00075	0.13975	-0.0005	0.13925	73.9	0.214693266
25	7/1/18 6:00 AM	0.00075	0.1315	-0.0005	0.131	74.1	0.945018413
26	7/1/18 6:15 AM	0.00075	0.1315	-0.0005	0.131	74.1	0.263132393
27	7/1/18 6:30 AM	0.00075	0.14275	-0.00175	0.141	74.1	0.196659512
28	7/1/18 6:45 AM	0.00075	0.5995	0.00675	0.60625	74.1	1.228140037
29	7/1/18 7:00 AM	0.00075	0.12775	0.02325	0.151	75.9	1.425939872
30	7/1/18 7:15 AM	0.00075	0.1045	0.04575	0.15025	75.9	1.879063078
31	7/1/18 7:30 AM	0.00075	0.08675	0.0625	0.14925	75.9	0.06369556
32	7/1/18 7:45 AM	0.00075	0.05675	0.09275	0.1495	75.9	1.159306408
33	7/1/18 8:00 AM	0.00075	0.05025	0.09975	0.15	78.8	1.814818463
34	7/1/18 8:15 AM	0.00075	0.4665	0.1375	0.604	78.8	1.148619651
35	7/1/18 8:30 AM	0.00075	-0.02875	0.162	0.13325	78.8	1.854096229

Figure 25: Case Study 2 Data Variables

After the exploratory data analysis, the dataset was uploaded to the Azure Services platform for the subsequent implementation step. Microsoft generously makes available Azure for Students subscriptions for research and experimentation.

Ten experiments were conducted over several months to evaluate varying models for conceptual proof of concept. The "WesDissertation_Model" (Figure 26) from the Azure Machine Learning workspace was selected to represent the conceptual model intentions.

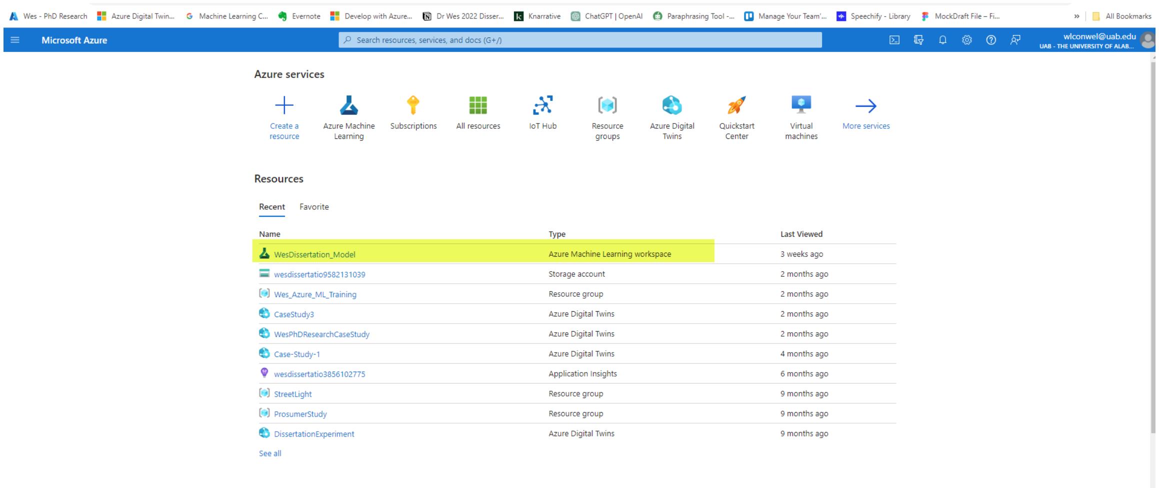

Figure 26: Microsoft Azure List of Experiments

4.3.3 Case Study 2 Results

Within the experimentation of the *WesDissertation_Model* (Figure 26) from the Azure Machine Learning workspace, five Automated ML jobs were conducted. The Automated ML jobs required varying levels of compute time ranging from 41 minutes to 24 hours. For purposes of developing the Predictive twin prototype within the time and resource constraints of the research, the dataset parameters were limited to one month of 15-minute interval data for each of the six variables. In turn, the *Wes_CaseStudy2_AutoML* model (Figure 27) using the *4767_JulyData_Behav* experiment was selected to demonstrate the concept. This predictive model was structured to forecast energy consumption for a subsequent period.

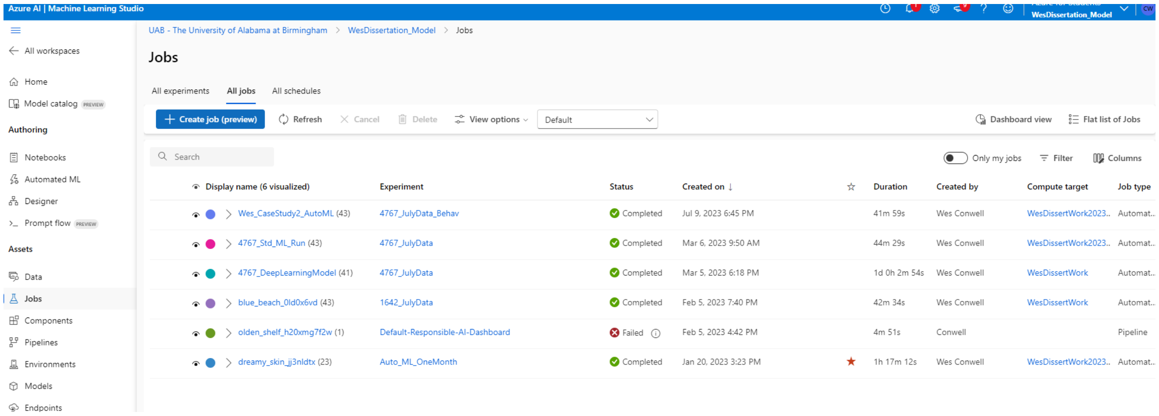

Figure 27: Case Study AutoML Jobs

Following the successful completion of the *4767_JulyData_Behav* experiment of the Azure AutoML job, insightful information was provided to the researcher for a detailed evaluation of the model capabilities. Starting with the *Overview* (Figure 28) results, envisioning research as a true smart cities modeler, one could discern key properties such as Compute target, Compute duration, and Arguments.

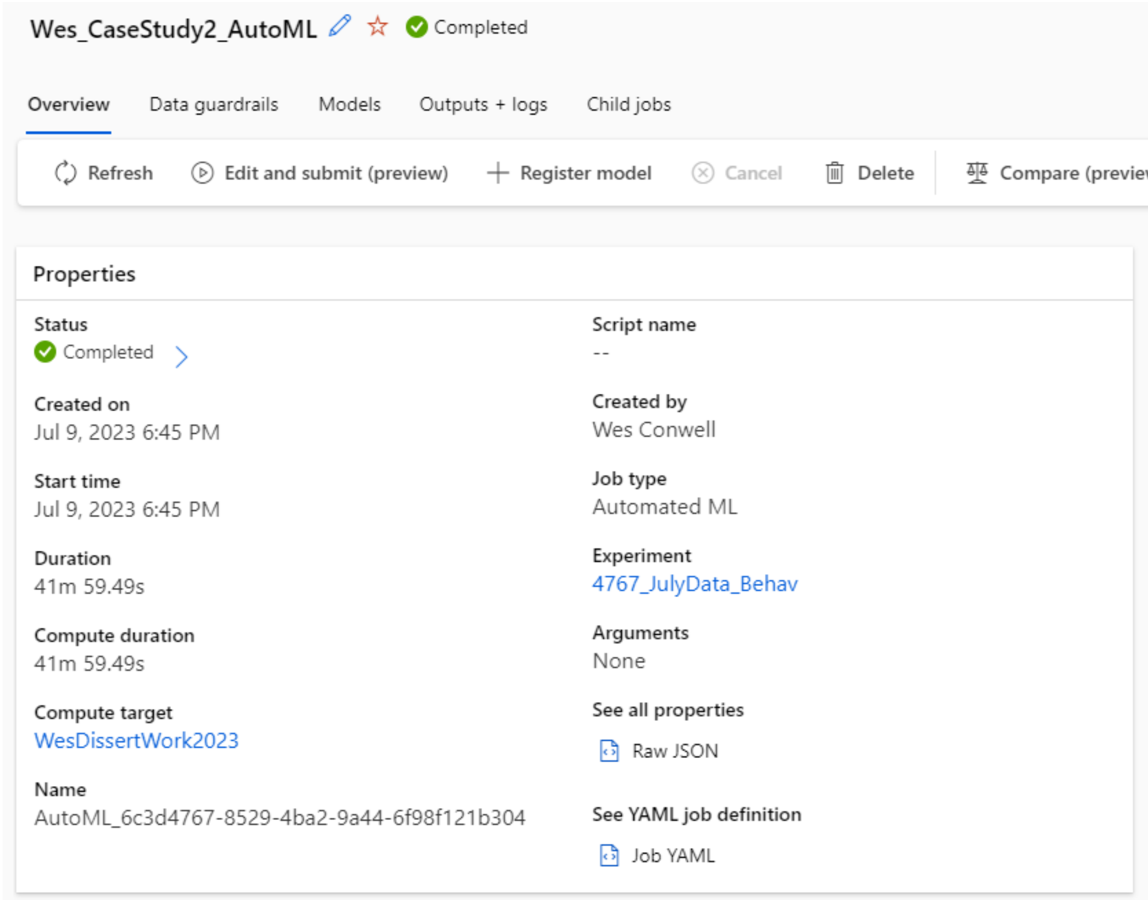

Figure 28: Case Study 2 Properties

The *Overview* (Figure 29) also includes for the smart cities researcher an easily accessible layout of Inputs [*training data*], Outputs [*best model, full training dataset*], and Best Model summary [the optimized and recommended ML algorithm and hyperparameters]. During the training process, Azure Machine Learning generates multiple pipelines simultaneously, which experiment with various algorithms and parameters on behalf of the user. The service employs iterating through machine learning algorithms combined with feature selections. Each iteration generates a model that is evaluated based on its training score. The higher the score for the optimized metric, the greater the degree to which the model accurately represents the data. The cessation of the experiment will occur upon reaching the predefined exit criteria.

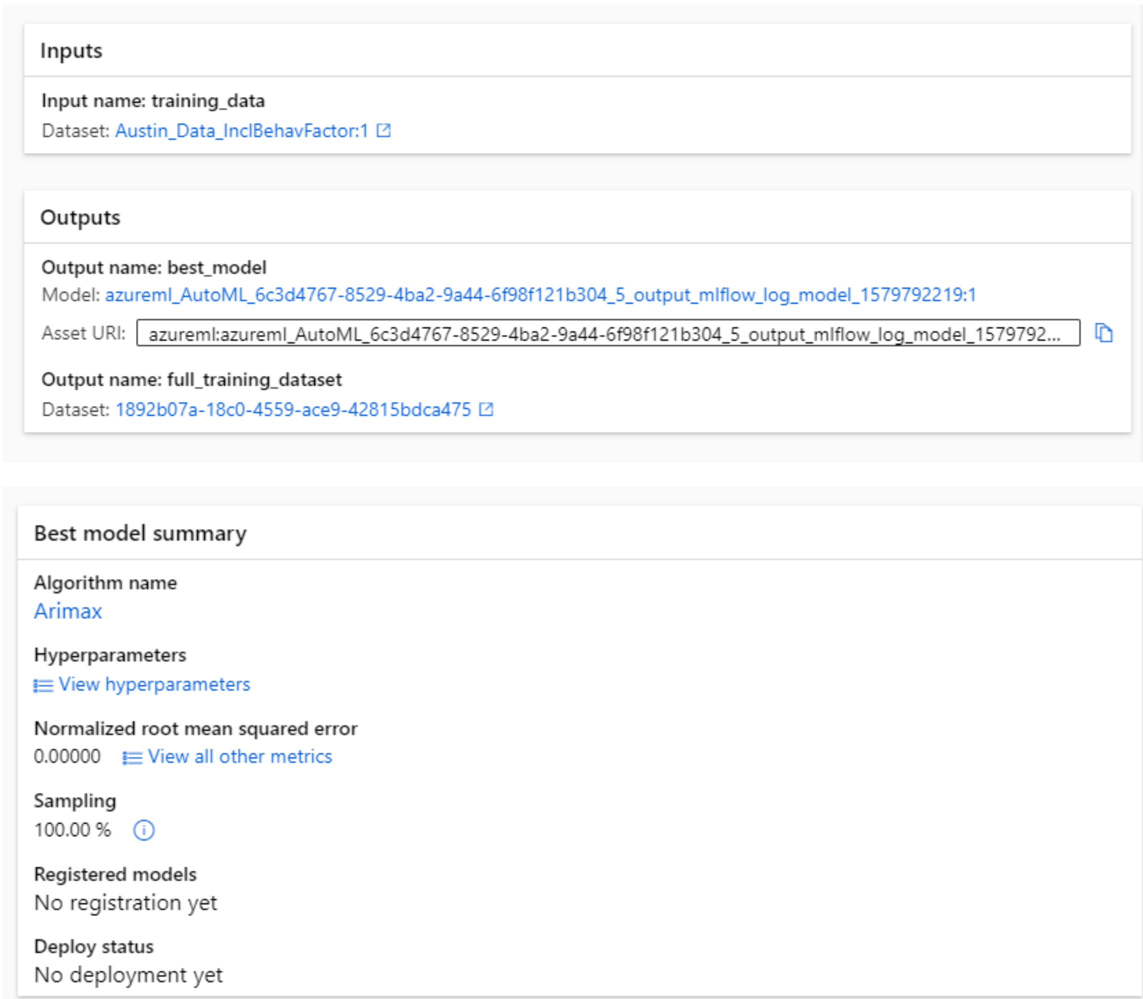

Figure 29: July Case2 Study Inputs, Outputs, Best Model Summary

Utilization of Data guardrails (Figure 30) yields yet another advantageous check to ensure high-quality data is used to train the model.

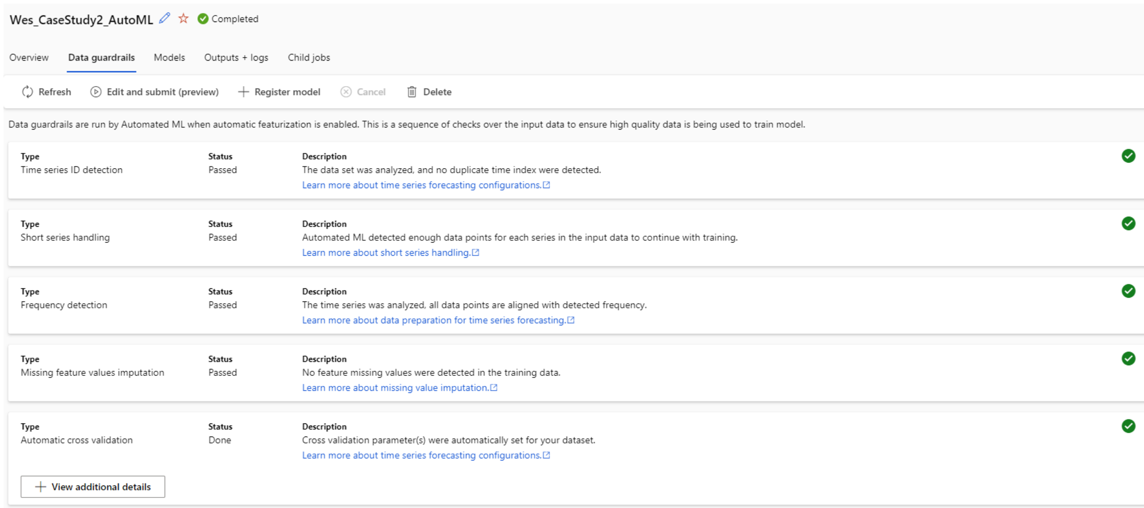

Figure 30: Case Study 2 Data Guardrails

The Data Guardrails passed the noted tests (Figure 30), so the next model performance was evaluated. The model exhibited reasonable results illustrated in Azure AutoML generated box and whisker plots (Figure 31).

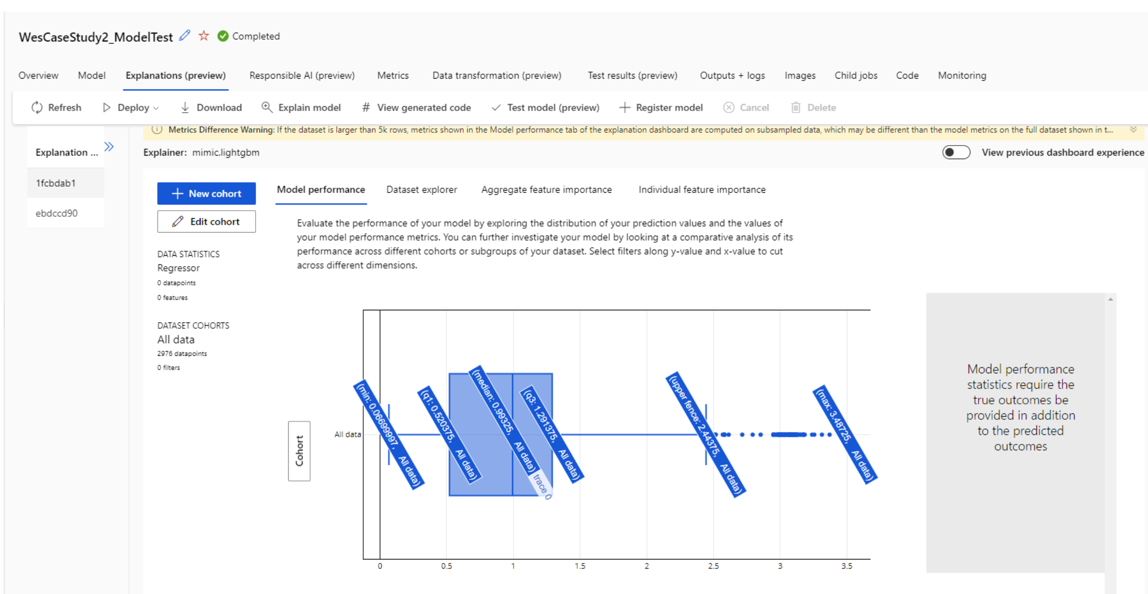

Figure 31: Case Study 2 Model Performance

The Azure Auto ML results also allow exploring the top-k most essential features that impact overall model predictions (Figure 32).

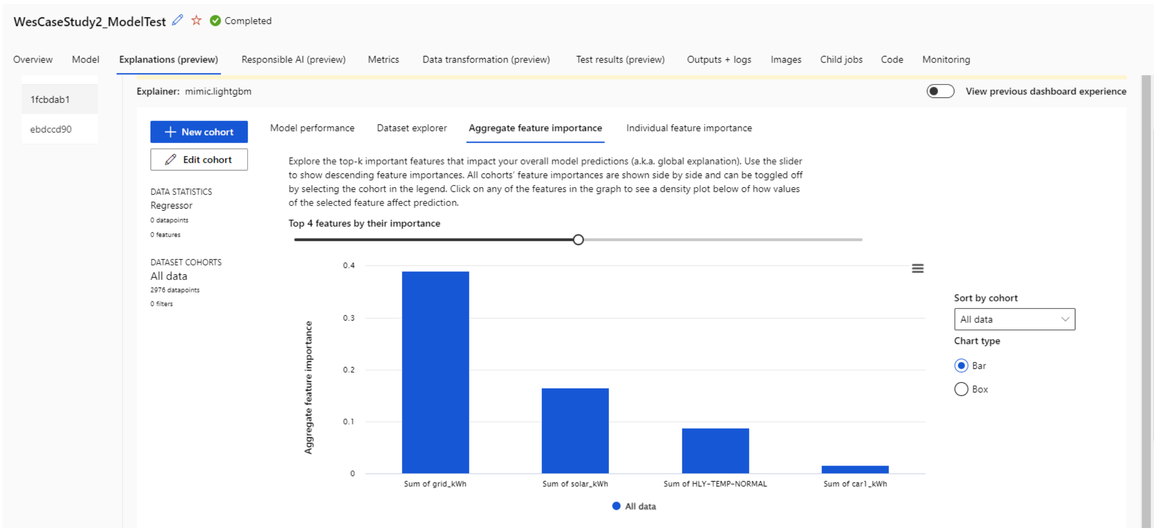

Figure 32: Case Study 2 Aggregate Feature Performance

The last section of the conceptual model demonstration, from Azure Auto ML, provides information concerning the capability for model deployment in the field. The experiment conducted for this case replicated a simulation for field deployment. Figure 33 displays a link to the Arimax algorithm to be deployed in field implementation.

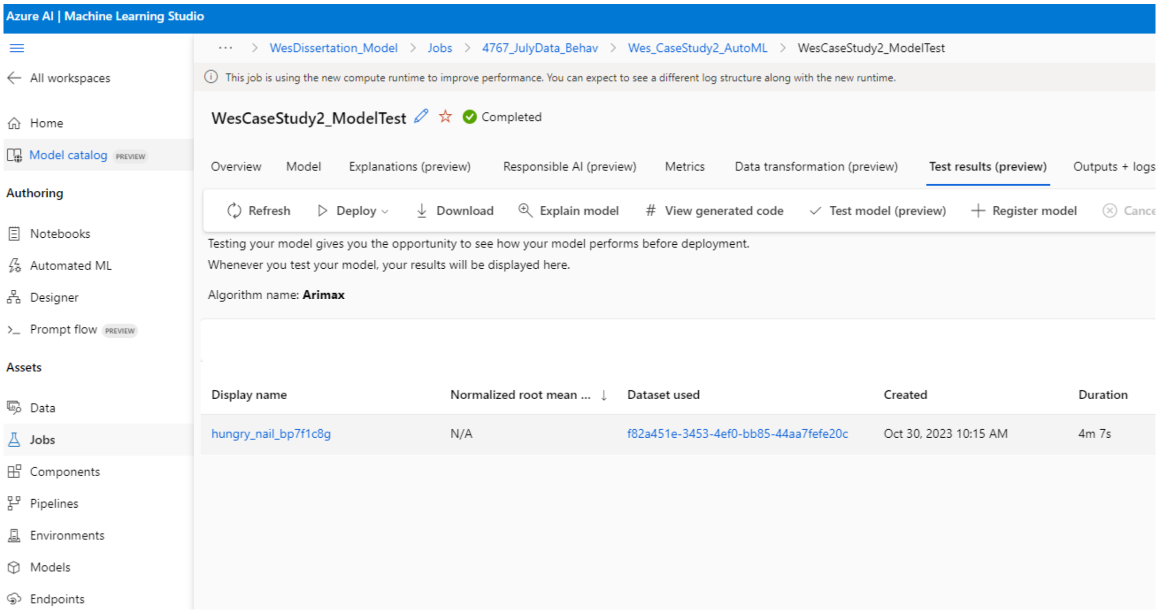

Figure 33: Model Deployment Simulation

To close out the case study proof of concept, Figure 34 demonstrates a mock dashboard developed for a smart cities end user for this Case Study 2 smart energy domain case.

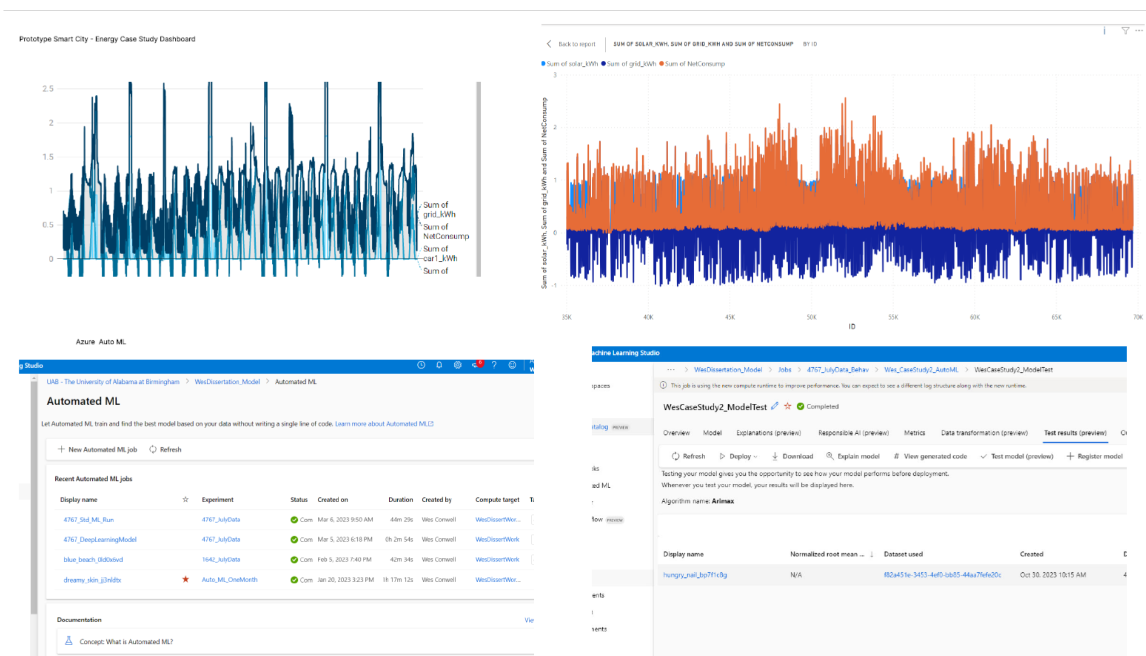

Figure 34: Mock Dashboard

4.4 Case Study Summary

This chapter presented the findings of this study developed in accordance with tenets of the pragmatic philosophy and case study strategy. The Case Study strategy employs the most suitable methodology in demonstrating the abstraction of the conceptual framework. Case Study 1 contributes to the smart energy domain for a city with limited resources and capabilities to implement a Level 1 smart city implementation. It can be implemented at a minimum cost level such that it can be implemented with the processing power of a smartphone. Case Study 2 demonstrates that with more resources and capabilities, a Level 3 implementation can be achieved with access to a platform such as Microsoft Azure. Ultimately, each case study is technology agnostic and provides access to an architecture with allowance for implementation across multiple domains.

CHAPTER 5

CONCLUSION

5.1 Introduction

The purpose of this chapter is to provide a synthesis of the critical points and contributions of the research. The research was developed from four broad but disparate subject areas: smart cities, the Internet of Things, Big Data, and Machine Learning. The literature review culminated in the development of an initial conceptual framework. The research adopts the manufacturing and construction applications of the digital twin and applies them to smart cities. In turn, the central argument of the thesis stressed the utilization of a digital twin as applied to varying resource consumption domains of a smart city. As a result of this study, the smart city digital twin (SCDT) architecture can enable cities of varying sizes to apply the conceptual model to resource consumption problems.

The rest of the chapter is structured as follows: Section 5.2 summarizes the research questions and objectives of the research. The contributions to the field are presented in section 5.3, while the study's limitations are presented in section 5.4. A view of future directions for research in Section 5.5 completes this chapter.

5.2 Addressing the Research Questions

This research investigated the impact of leveraging existing state-of-the-art technologies for problem-solving in smart cities. This effort culminated in developing an empirically informed conceptual framework (Figure 16). At an organizational level, the smart city digital twin framework (SCDTF) captures the impact of using Internet of Things sensors, big data tools, and machine learning algorithms in developing a smart city to address urban problems. Smart city initiatives are an attempt to improve urban performance by using data along with information communication technologies (ICT) to provide more efficient services to citizens, monitor and optimize existing infrastructure, increase collaboration among different economic actors, and encourage innovative business models in both the private and public sectors (Marsal-Llacuna et al., 2015)

The purpose of this section is to re-evaluate the research questions, which were central to achieving the research aim. Also, the section content itemizes each research question and explains how the questions were considered within the research context.

Research Question 1:

How can current state-of-the-art technologies be leveraged for monitoring and managing resources within a smart city framework?

Addressing this research question required critical evaluation of the interactions between the core technological concepts - smart cities, big data, the Internet of Things, and Machine Learning. Exploration around the boundaries of the core concepts and reading for gaps in the literature culminated in the development of a conceptual framework. The conceptual framework demonstrated the interactions between the core concepts and was pivotal to determining the boundaries of the concepts as applied in the case studies. Therefore, by incorporating disparate aspects of the literature, the conceptual framework provides a unique perspective to problem-solving for smart cities across varying domains.

Research Question 2:

What methods can allow virtually any typical US city to implement smart city concepts in their city operations and management using simulation?

Addressing this research question in terms of the method followed with an exploration of methods in varying research labs and industry applications. Digital twin was deemed the optimal approach for this applied research effort. While simulations and digital twins share the use of digital models to replicate the processes of a system, it is essential to note that a digital twin is distinct in that it provides a virtual environment that offers a significantly enhanced capacity for analysis and investigation. The distinction between digital twins and simulation primarily pertains to the extent or magnitude of representation. While a simulation generally focuses on analyzing a specific process, a digital twin has the capability to conduct numerous valuable simulations to examine multiple processes simultaneously.

There are additional distinctions beyond what has been mentioned. For instance, simulations typically do not derive advantages from the utilization of real-time data. Digital twins are conceptualized with a bidirectional exchange of information, wherein object sensors initially transmit pertinent data to the system processor, followed by the subsequent dissemination of insights generated by the processor back to the source object.

Digital twins possess the capability to investigate a broader array of subjects from numerous perspectives, surpassing the capabilities of conventional simulations. This capability is made possible by digital twin access to enhanced and regularly updated data across various domains and the augmented computational power offered by virtual environments. Consequently, digital twins exhibit more potential to enhance products and processes.

5.3 Contributions of the Study

The contributions to knowledge are twofold: to theory and practice. Through the empirically informed conceptual framework for smart cities, problem solving with the use of the Internet of Things, big data, and machine learning (IBML) technologies, the research provides the academy and practitioners with an understanding of efficient mechanisms to coordinate their problem-solving activities for the operations of smart cities of varying sizes. Primarily, the theoretical contributions capture 1) the Establishment of design and specification of basic core functionalities for a smart city digital twin framework to enhance resource consumption in a smart city or any other resource-dependent construct and 2) the Development of a maturity model structure for smart city application. Additionally, the theoretical contribution of this thesis centers on developing an empirically informed conceptual framework utilizing modern technologies in the realm of organizational problem-solving for smart cities. The research sets the groundwork for future research to assess organizational problem-solving in other domains such as healthcare, sports, education, and transportation. Most importantly, having adopted a smart city context, the research offers insight into the complexities of working with the Internet of Things, big data, and machine learning components in creating smart cities.

Developing an empirically informed conceptual framework (see Section 2.9) was the original contribution to theory. This contribution was central in demonstrating the impact of leveraging advanced technologies for organizational problem-solving for the applications of smart cities. Past studies have not addressed how city authorities identify the most efficient mechanisms to coordinate their problem-solving activities for smart city applications leveraging IOT, big data, and machine learning. The two-stage framework presented in Figure 16 and Figure 17 serves as the final output of the empirical analysis contributing to theory with: i) a generic framework underpinned by Gray's criteria on successful frameworks (Gray, 2021), ii) a context-driven conceptualization on application of the framework for smart cities of varying sizes The development of the empirically informed conceptual framework was predicated by an initial conceptual framework, which integrated the three disparate areas of research - smart cities, big data and organizational problem solving - to demonstrate their critical inter-relationships (see Chapter 2, 3 and 4). Previous studies have not integrated these three disparate research areas into this conceptual framework. Therefore, the SMDT framework was valuable in exploring and integrating the concepts used in the case study phase of the research.

The main contribution to practice emanates from the applied research demonstrated by the case studies in Sections 4.2 and 4.3. This contribution communicates the implementation of the case study in the smart energy sub-domain of the smart city using a digital twin construct. Through the conceptual framework, practitioners are equipped with a context-driven perspective on using IBML in problem-solving for smart cities and beyond. IBML and its constituent elements can help solve problems organizations face. This structure allows organizations to expand the boundaries of known problems and potentially unearth undetected ones. In other words, with rich levels of characteristic and interactive data from various sources, organizations can make sense of problems that have bedeviled them. The research equips practitioners with an understanding of the impact of IBML in organizational problem-solving for developing smart cities.

5.4 Limitations of the study

This section aims to highlight the limitations of the research, which can serve as avenues for further research. The research study provided an architectural foundation for understanding the use of IBML, especially regarding the development of smart cities. However, there are several limitations in this study, which, in turn, serve as areas for possible future research. Cybersecurity/privacy impacts garner the utmost attention in today's hyperconnected age. Data from Pecan Street was anonymized for the study, but privacy was not researched explicitly. In the energy industry, the total retail perspective can generate broader insight. This research study limited scope to a residential customer focus due to limited time and system capabilities in graduate research available resources. COVID-19 impacts on smart cities substantiate long-term effects with hybrid work, increased residential working space, and modified business operations. However, COVID-19 impacts outside the research's original scope were perceived as more compelling future work opportunities. Lastly, a simulation for the system case was included for the interdisciplinary research effort and not full test implementation. Full implementation presents a tremendous opportunity but requires much more robust resources, expense, and capability to deploy in the field.

5.5 Future Research

A natural progression of this work is to expand the research case study towards Level 5 autonomous development. As referenced in Chapter 2.9, this study delineated the general architecture for a smart city platform as a beachhead across multiple domains. The Level 5 development presents the pinnacle development of the framework in optimizing its capabilities. Aside from the energy case study, further research will allow for consideration in the domains of finance, health, education, and others.

Another area of future research would be the application of the framework in comparable domains in countries of the African continent. Within smart cities this presents a tremendous opportunity to impact generations for centuries. The United Nations predicts that by 2050, Africa will be the most populous continent, with growth projection as the world's fastest-urbanizing region (Kuyoro et al., 2023). The framework established in this research lays out a structure that presents a smart city platform to assist with this tremendous growth.

.

REFERENCES

Ahuja, K., & Khosla, A. (2019). Data analytics criteria of IoT enabled smart energy meters (SEMs) in smart cities. *International Journal of Energy Sector Management*, *13*(2), 402–423. https://doi.org/10.1108/IJESM-11-2017-0006

Albino, V., Berardi, U., & Dangelico, R. M. (2015). Smart cities: Definitions, dimensions, performance, and initiatives. *Journal of Urban Technology*, *22*(1), 3–21. https://doi.org/10.1080/10630732.2014.942092

Ali, F. (2019). *Characterising Smart Cities*. University of Delaware.

Alsamani, B. (2019). *Design a Framework for Ideating and Developing Smart Services and Smart Kiosk Lounge Platform Using IOT Artificial Intelligence*. Claremont Graduate University.

Amit Midha, P. (2021). *How 'work from anywhere' can build the city of the future*. World Economic Forum. https://www.weforum.org/agenda/2021/04/future-of-smart-cities-covid-19-digital/

Angner, Erik; Loewenstein, G. (2006). Behavioural economics. In *Handbook of the Philosophy of Science,* (Vol. 5). https://doi.org/10.1111/2057-1615.12303

Anthopoulos, L. (2017). Smart utopia VS smart reality: Learning by experience from 10 smart city cases. *Cities*, *63*, 128–148. https://doi.org/10.1016/j.cities.2016.10.005

Attaran, M., & Celik, B. G. (2023). Digital Twin: Benefits, use cases, challenges, and opportunities. *Decision Analytics Journal*, *6*(January), 100165. https://doi.org/10.1016/j.dajour.2023.100165

Batty, M. (2018). *Inventing Future Cities* (Kindle Edi). MIT Press.

Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Ouzounis, G., & Portugali, Y. (2012a). Smart cities of the future. *European Physical Journal: Special Topics*, *214*(1), 481–518. https://doi.org/10.1140/epjst/e2012-01703-3

Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Ouzounis, G., & Portugali, Y. (2012b). THE EUROPEAN PHYSICAL JOURNAL Smart cities of the future. *Eur. Phys. J. Special Topics The European Physical Journal Special Topics*, *214*, 481–518. https://doi.org/10.1140/epjst/e2012-01703-3

Chen, C., Wang, S., & Guo, B. (2016). The Road to the Chinese Smart City. *IT Pro - IEEE Computer Society*, *January/February*, 14–17.

Corbin, J. M., & Strauss, A. L. (2008). *Basics of Qualitative Research 3e.* SAGE Publications.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches.* (Fifth). SAGE Publications.

Creswell, John W.; Creswell, J. David. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (Kindle Edi). SAGE Publications.

Dawkins, O. (2018). Living with a Digital Twin: Operational management and engagement using IoT and Mixed Realities at UCL's Here East Campus on the Queen Elizabeth Olympic Park. *Giscience and Remote Sensing*.

El Saddik, A. (2018). Digital Twins: The Convergence of Multimedia Technologies. *IEEE Multimedia*, *25*(2), 87–92. https://doi.org/10.1109/MMUL.2018.023121167

Energy, D. of. (2023a). *Alabama Power Reynolds Landing Project*. https://connectedcommunities.lbl.gov/projects/project-name-alabama-power-smart-neighborhoodr-reynolds-landing

Energy, D. of. (2023b). *Wisconsin's Connecting Communities for Sustainable Solutions project*. https://connectedcommunities.lbl.gov/connecting-communities-sustainable-solutions

ETSI. (2019). *Extension to SAREF ; Part 4 : Smart Cities Domain* (Vol. 1). https://www.etsi.org/deliver/etsi_ts/103400_103499/10341004/01.01.01_60/ts_10341004v010101p.pdf

Fernandez-Anez, V., Fernández-Güell, J. M., & Giffinger, R. (2018). Smart City implementation and discourses: An integrated conceptual model. The case of Vienna. *Cities*, *78*(December 2017), 4–16. https://doi.org/10.1016/j.cities.2017.12.004

Friedman, T. (2005). *The world is flat: A brief history of the twenty-first century*. Macmillan.

George, S. (2020). *Azure Digital Twins now generally available: Create IoT solutions that model the real world*. Microsoft Azure - IoT Blog. https://azure.microsoft.com/en-us/blog/azure-digital-twins-now-generally-available-create-iot-solutions-that-model-the-real-world/

Gray, D. (2021). What Makes Successful Frameworks Rise Above the Rest. *MIT Sloan Management Review*, *62*(4), 1–6. https://sloanreview.mit.edu/article/what-makes-successful-frameworks-rise-above-the-rest/#:~:text=Clarity%2C Memorability%2C and Integration&text=These important factors contribute to,easier to understand and use.

Grieves, M., & Vickers, J. (2016). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches* (pp. 85–113). Springer International Publishing. https://doi.org/10.1007/978-3-319-38756-7_4

Griffiths, S., & Sovacool, B. K. (2020). Rethinking the future low-carbon city: Carbon neutrality, green design, and sustainability tensions in the making of Masdar City. *Energy Research and Social Science*, *62*(August 2019), 101368. https://doi.org/10.1016/j.erss.2019.101368

Hammar, K., & Wallin, E. (n.d.). *RealEstateCore*. https://service.tib.eu/webvowl/#iri=https://w3id.org/rec/full/3.3/

Hart, M. (2021). *Scientists Are Building a 'Digital Twin' of Earth*. European Space Agency. https://www.yahoo.com/now/scientists-building-digital-twin-earth-190140384.html

Hashem, I. A. T., Chang, V., Anuar, N. B., Adewole, K., Yaqoob, I., Gani, A., Ahmed, E., & Chiroma, H. (2016). The role of big data in smart city. *International Journal of Information Management*, *36*(5), 748–758. https://doi.org/10.1016/j.ijinfomgt.2016.05.002

Howell, S. K. (2017). *Towards a semantic web of things for smart cities* [Cardiff University]. http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.716058%0Ahttp://orca.cf.ac.uk/100995/1/2017HowellSPhD.pdf%0Ahttp://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.716058%0Ahttp://orca.cf.ac.uk/100995/1/2017HowellSPhD.pdf

IBM. (2022). *What is a digital twin?* https://www.ibm.com/topics/what-is-a-digital-twin

Jalali, R. (2017). *Community-Oriented Architecture for Smart Cities*. University of ontario Institue of Technology.

Jin, J., Gubbi, J., Marusic, S., & Palaniswami, M. (2014). An information framework for creating a smart city through internet of things. *IEEE Internet of Things Journal*, *1*(2), 112–121. https://doi.org/10.1109/JIOT.2013.2296516

Jones, S. (2021). *What is a Digital Twin? How Intelligent Data Models Can Shape the Built World*. Autodesk. https://redshift.autodesk.com/articles/what-is-a-digital-twin

Khatoun, B. Y. R., & Zeadally, S. (2016). Smart cities: Concepts, Architectures, Research Opportunities. *ACM*, 46–57.

Kuyoro, M., Leke, A., White, O., Woetzel, J., Jayaram, K., & Hicks, K. (2023). *Reimagining economic growth in Africa: Turning diversity into opportunity*. https://www.mckinsey.com/mgi/our-research/reimagining-economic-growth-in-africa-turning-diversity-into-opportunity

Laamarti, F., Badawi, H. F., Ding, Y., Arafsha, F., Hafidh, B., & Saddik, A. El. (2020). An ISO/IEEE 11073 Standardized Digital Twin Framework for Health and Well-Being in Smart Cities. *IEEE Access*, *8*, 105950–105961. https://doi.org/10.1109/ACCESS.2020.2999871

Lee, J. H., Hancock, M. G., & Hu, M. C. (2014). Towards an effective framework for building smart cities: Lessons from Seoul and San Francisco. *Technological Forecasting and Social Change*, *89*, 80–99. https://doi.org/10.1016/j.techfore.2013.08.033

Marsal-Llacuna, M. L., Colomer-Llinàs, J., & Meléndez-Frigola, J. (2015). Lessons in urban monitoring taken from sustainable and livable cities to better address the Smart Cities initiative. *Technological Forecasting and Social Change*, *90*(PB), 611–622. https://doi.org/10.1016/j.techfore.2014.01.012

Martin Strohbach, Holger Ziekow, Vangelis Gazis, and N. A. (2015). Towards a Big Data Analytics Framework for IoT and Smart City Applications. *Modeling and Processing for Next-Generation Big-Data Technologies*, *4*(JANUARY 2014), 283–317. https://doi.org/10.1007/978-3-319-09177-8

Microsoft. (2021). *Azure digital twins documentation. (n.d.). Developer tools, technical documentation, and coding examples | Microsoft Docs.* https://docs.microsoft.com/en-us/azure/digital-twins/

Minerva, R., Lee, G. M., & Crespi, N. (2020). Digital Twin in the IoT Context: A Survey on Technical Features, Scenarios, and Architectural Models. *Proceedings of the IEEE*, *108*(10), 1785–1824. https://doi.org/10.1109/JPROC.2020.2998530

Minevich, M. (2021). *Digital Twin technology may save lives and solve the biggest challenges facing humanity*. https://thehill.com/opinion/technology/554881-digital-twin-technology-may-save-lives-and-solve-the-biggest-challenges/

Miskinis, C. (n.d.). *The mysterious history of digital twin technology and who created it*. Retrieved March 17, 2019, from https://www.challenge.org/insights/digital-twin-history/

Mohammadi, M., & Al-Fuqaha, A. (2018). Enabling Cognitive Smart Cities Using Big Data and Machine Learning: Approaches and Challenges. *IEEE Communications Magazine*, *56*(2), 94–101. https://doi.org/10.1109/MCOM.2018.1700298

Mohammadi, N., & Taylor, J. E. (2017). Smart city digital twins. *2017 IEEE Symposium Series on Computational Intelligence, SSCI 2017 - Proceedings*, *2018-Janua*, 1–5. https://doi.org/10.1109/SSCI.2017.8285439

Monzon, A. (2015). Smart Cities Concept and Challenges: Bases for the Assessment of Smart City Projects. In *Communications in Computer and Information Science* (Vol. 579, pp. 17–31). SCITEPRESS.

Nath, S. V., Van Schalkwyk, P., & Isaacs, D. (2021). *Building industrial digital twins: Design, develop, and deploy digital twin solutions for real-world industries using Azure digital twins.* (S. Editing, Ed.; Kindle). Packt Publishing Ltd.

Nations, U. (2018). *United Nations World Urbanization Prospects - 2018 Revision*.

Paulk, M. C. (2009). A History of the Capability Maturity Model for Software. *The Software Quality Profile*, *1*(1), 5–19.

Raithatha, Y. (2017). *Understanding the economic impact terrorism has on the destination decision making: Northern Irish tourists.* Dublin Business School.

Rajteric, I. (2010). Overview of Business Intelligence Maturity Models. *Management*, *15*(1), 47–67.

Rouse, M. (n.d.). *Big data analytics*. Retrieved March 1, 2017, from https://searchbusinessanalytics.techtarget.com/definition/big-data-analytics

Russom, M. B. (n.d.). *Digital Twins Definition Language (DTDL) ontology for Smart Cities*. https://github.com/Azure/opendigitaltwins-smartcities/tree/main

Saunders, M. N. K., Lewis, P., & Thornhill, A. (2019). Research Methods for Business Students. In *Research Methods for Business Studies*. https://www.researchgate.net/publication/330760964_Research_Methods_for_Busine ss_Students_Chapter_4_Understanding_research_philosophy_and_approaches_to_theor y_development

Sta, H. Ben. (2017). Quality and the efficiency of data in "Smart-Cities." *Future Generation Computer Systems*, *74*, 409–416. https://doi.org/10.1016/j.future.2016.12.021

Szcodronski, R. (n.d.). *Digital Twins Definition Language-based RealEstateCore ontology for smart buildings*. https://github.com/Azure/opendigitaltwins-building

Tech, G. (n.d.). *Georgia Tech Smart Cities and Inclusive Innovation*. Retrieved November 2, 2021, from https://smartcities.gatech.edu/taxonomy/term/543

The Economist. (2017). The data economy: Fuel of the future. *The Economist*, *423*, 13–16.

Williams, S. (2020). *Defining Smart Cities : A Qualitative Study to Define the Smart City Paradox*. Robert Morris University.

Xu, H., Huang, H., Khalid, R. S., & Yu, H. (2016). *Distributed Machine Learning based Smart-grid Energy Management with Occupant Cognition*.

Yin, R. K. (2017). *Case study research and applications* (Kindle). Sage Publishing.

Zhang, Y., Huang, T., & Bompard, E. F. (2018). Big data analytics in smart grids: a review. *Energy Informatics*, *1*(1), 8. https://doi.org/10.1186/s42162-018-0007-5

A

Definition of Key Terminology

APPENDIX A

AMI

AMI is not a single technology implementation but a fully configured infrastructure that must be integrated into existing and new utility processes and applications. This infrastructure includes home network systems, including communicating thermostats and other in-home controls, smart meters, communication networks from the meters to local data concentrators, back-haul communications networks to corporate data centers, meter data management systems (MDMS), and, finally, data integration into existing and new software application platforms.

Framework

Framework denotes a basic structure, methodology, or set of principles that allow research projects, industry projects, and other work to be organized and executed methodically and more efficiently.

Ontology

Defines a common vocabulary for researchers who need to share information in a domain. It includes machine-interpretable definitions of basic concepts in the domain and relations among them.

Smart City

A complex system of systems that leverages innovative ICT to benefit a community, integrated with complex subsystems to fulfill its operational objectives.

Smart Energy

An approach in which smart electricity, thermal, and gas grids are combined with storage technologies and coordinated to identify synergies between them to achieve an optimal solution for each sector and the overall energy system.

B

Applications and Tools

APPENDIX B

A list of tools used for this study is established in this appendix as a reference for others. These are the tools I found helpful in this research and met my needs.

Adobe Acrobat Pro

Adobe Acrobat Pro was critical in reviewing and managing PDF research journal articles and e-books.

Citavi

Citavi is a program for reference management and knowledge organization for Microsoft Windows published by Swiss Academic Software in Wädenswil, Switzerland. I cataloged hundreds of articles, books, and reports in Citavi for summaries and quick reviews.

Creately

Creately is a software as a service (SaaS) visual collaboration tool with diagramming and design capabilities designed by Cinergix. I used the cloud edition of Creately to create flowcharts, organization charts, and project charts.

Evernote

Evernote is a note-taking and task-management application developed by the Evernote Corporation. It is intended for archiving and creating notes with embedded photos, audio, and saved web content. Over the years of my dissertation process, Evernote was to capture years of notes and observations.

Figma

Figma is a collaborative web application for interface design, with additional offline features enabled by desktop applications for macOS and Windows. The feature set of Figma focuses on user interface and user experience design, emphasizing real-time collaboration utilizing a variety of vector graphics editors and prototyping tools. Figma was used to up a prototype dashboard for the smart energy case study.

Goodnotes

Goodnotes is an iPad iOS app that replaces paper notebooks with a digital note-taking representation. The Goodnotes tool was used to organize conference notes, research articles, and to-do lists.

Google Earth Pro

Google Earth is a computer program that renders a 3D representation of Earth based primarily on satellite imagery. The Pro version includes add-on software for movie making, advanced printing, and precise measurements and is currently available for Windows, macOS, and Linux. It was used for developing a movie-making segment of a presentation.

Kindle e-Book Reader

Kindle e-book reader is the Amazon Corporation software that allows for reading Kindle books and other electronic documents across various digital platforms. For dissertation research purposes, this was a valuable resource to quickly access textbooks, scientific references, and numerous writing resources.

Litmaps

Litmaps is a cloud-based resource that creates interactive literature maps. A user creates maps for research by searching the Litmaps literature database, linking a researcher's reference manager, or through automatic generation from seed articles. The Litmaps tool was used to develop a visual representation of the literature review and assist with evaluating other relevant literature.

Mendeley

Mendeley is reference manager software founded in 2007 by Ph.D. students Paul Foeckler, Victor Henning, and Jan Reichelt and acquired by Elsevier in 2013. It manages and shares research papers and generates bibliographies for scholarly articles. It was used as the primary reference manager in the research process and integrated seamlessly to include the reference section of the final dissertation, Microsoft Word document.

Microsoft Azure

Azure is a cloud computing platform run by Microsoft. Its machine learning and digital twin modules were essential for the Case Study development process.

Microsoft OneDrive

Microsoft OneDrive is a file-hosting service operated by Microsoft. OneDrive was a critical component in the dissertation workflow process. All journal articles, presentations, e-books, and digitally storable materials were stored with OneDrive

Notion

Notion is a productivity and note-taking web application developed by Notion Labs Inc. Notion was used to store all notes, references, and to-do lists during the dissertation process.

Otter.ai

Otter.ai is a speech-to-text transcription application that uses artificial intelligence and machine learning. This service was used to transcribe the audio of virtual meetings.

Scrivener

Scrivener is a word processing program and outliner for writers. It provides a management system for attached documents, notes, and metadata. It allows the organization of notes, concepts, research, and whole documents. Scrivener was used to capture my Research Journal, Reading Journal, Years of meeting notes, Literature review summaries, and planning items.

SnagIt

Snagit is screen capture and screen recording software for Windows. It was used to capture figures and tables used in the research process.