

Digital Twin Framework for Solar Power Plants in Kazakhstan

**Master of Engineering Management
2022 – 2024**

Submitted in partial fulfillment of the requirements for the degree of Master of
Engineering Management



School of Engineering and Digital Sciences

In partnership with

Graduate School of Business

**Shyndaulet Kassymkhanov
Gani Kazbek
Yerassyl Alipov
Nurbek Meirbekov**

April, 2024

Digital Twin Framework for Solar Power Plants in Kazakhstan

by
Shyndaulet Kassymkhanov
Gani Kazbek
Yerassyl Alipov
Nurbek Meiirbekov

Submitted in partial fulfillment of the requirements for the degree of Master of
Engineering Management

School of Engineering and Digital Sciences
Graduate School of Business
Nazarbayev University

Supervised by
Prof. Essam Shehab
Altay Zhakatayev

April, 2024

Declaration

We, Gani Kazbek, Nurbek Meirbekov, Shyndaulet Kassymkhanov and Yerassyl Alipov, declare that the research contained in this thesis, unless otherwise formally indicated within the text, is the author's original work. The thesis has not been previously submitted to this or any other university for a degree and does not incorporate any material already submitted for a degree.

Signature(s):

Gani Kazbek



Nurbek Meirbekov



Shyndaulet Kassymkhanov



Yerassyl Alipov



Date: 28.04.2024

Abstract

The renewable energy sector (RES) is a fast-growing industry that provides sustainable and clean energy. The adoption and integration of such technologies positively affect the environment and boost the country's economy. The global trend for sustainability opens new opportunities to integrate sophisticated technologies and improve existing energy infrastructure. This project aims to propose a detailed Digital Twin (DT) framework indicating important implementation steps and providing insights into DT technology that improves operational efficiency, optimizes performance, and creates a user-friendly platform for real-time monitoring.

This study has produced several deliverables through collaboration with the domestic company TechnoGroupService (TGS), which specializes in constructing and operating solar power plants in Kazakhstan. Deliverables include a detailed investigation of the 50MW solar power plant's operational process, infrastructure, and data transfer models. That detailed plant information helped to build a strong foundation for a comprehensive DT implementation framework that describes all the steps and procedures required to obtain a working DT platform.

The framework provides a DT architecture consisting of six main phases, from plant infrastructure exploration to DT application description. The entire framework was then used to create a detailed roadmap for the implementation of DT in the context of Kazakhstan, taking into account the peculiarities of the region and adapting international standards to the realities of the country. As one part of the development phase, an online web platform has been developed. These include real-time monitoring indicating the current status of the solar plant which is further used in reactive-maintenance capabilities. The results offered have been reviewed by industry experts and capstone supervisors to ensure applicability, reliability, and actuality.

Keywords: Digital Twin, framework, reactive maintenance, solar power plant.

Acknowledgments

Every individual whose assistance has been invaluable in ensuring the accomplishment of our capstone project deserves our deepest appreciation.

Our supervisor, Professor Essam Shehab, has provided us with unflinching direction, support, and encouragement throughout the entirety of the project. For this, we are thankful. His insightful counsel and professional guidance have been crucial in helping us to accomplish our objectives and shape our research.

Additionally, we wish to express our gratitude to Zhomart Mominbayev and LLP “Techno Group Service” for their considerable support and assistance. We appreciate the chance to work with such a prestigious company and the knowledge and resources they were ready to supply us so that we could complete our research.

We also acknowledge Altay Zhakatayev for his help and advice on the project.

We are grateful for their aid and backing in furnishing us with the essential facts and facilitating our advancement in our undertaking. We greatly appreciate their assistance, since it has greatly enhanced the quality of our findings.

Finally, we would like to sincerely thank the School of Engineering for their contribution to this collaborative project. We appreciate the opportunity to work on a real-world industry project, and the experience acquired has been vital to our learning and development.

We want to thank everyone who helped make our capstone project a success once more from the bottom of our hearts. We sincerely appreciate your help, inspiration, and insightful advice.

Contents

Declaration	iii
Abstract	iv
Acknowledgments	v
Contents	vi
List of Tables	viii
List of Figures	x
List of Abbreviations	xi
CHAPTER 1: INTRODUCTION	1
1.1 Background	1
1.2 Research Motivation	1
1.3 Aim and Objectives	3
1.4 Scope of the Project	3
1.5 Structure of the Report	4
CHAPTER 2: LITERATURE REVIEW	6
2.1 Introduction	6
2.2 Digital Twin Technology	7
2.3 Types of Digital Twin	9
2.4 Application of Digital Twin in Solar Power Plants	11
2.5 Implementation of DT Technology	13
2.6 Research Gap Analysis	17
CHAPTER 3: RESEARCH METHODOLOGY	19
3.1 Introduction	19
3.2 Project Development Phases	20
CHAPTER 4: ANALYSIS OF THE CURRENT PRACTICES AT TGS COMPANY	22
4.1 Introduction	22
4.2 Company Overview	22
4.3 50 MW Solar Power Plant Analysis	23
4.3.1 50 MW Power Plant General Characteristics	23
4.3.2 Infrastructural Layer Analysis	24
4.3.3 Communication Layer Analysis	27
4.3.4 Database Structure	31
CHAPTER 5: FRAMEWORK DEVELOPMENT	35
5.1 Introduction	35
5.2 Digital Twin Framework	35
5.2.1 Interactions between the Physical Plant and the Digital Twin	37

5.2.2 Model Layer	38
5.2.3 Application Layer	39
5.4 Justification for the Design Choices in the Framework	42
CHAPTER 6: IMPLEMENTATION OF DT AND PROTOTYPE DEVELOPMENT	43
6.1 Introduction	43
6.2 Digital Twin Implementation Process Flow	43
6.3 UML Diagram	46
6.4 Web platform - Minimum Viable Product (Prototype)	48
6.5 Prototype Validation and Stakeholders Feedback	51
CHAPTER 7: CONCLUSION AND FUTURE WORK	54
7.1 Introduction	54
7.2 Conclusions	54
7.3 Limitations of the project	55
7.4 Future Work	55
References	57
Appendix List	61

List of Tables

Table 2.1 Attributes description with standards.....	14
Table 2.2 ISO -23247 reference architecture domains.....	15
Table 4.1 Characteristics of power plant.....	23
Table 6.1 Experts validation.....	51
Table A1 Communication Equipment Description.....	61

List of Figures

Figure 1.1 Scope of the project.....	4
Figure 1.2 The structure of the report.....	5
Figure 2.1 Concept map for literature review.....	6
Figure 2.2 Literature review methodology.....	7
Figure 2.3 Types of Digital Twin.....	9
Figure 2.4 Models of Digital Twin.....	10
Figure 2.5 Aerial imagery of solar plants that indicate shading.....	12
Figure 2.6 A functional reference model for the digital twin.....	16
Figure 2.7 D-Arc technology stack.....	17
Figure 3.1 Methodology of the Project.....	19
Figure 4.1 TGS Company’s Goals Board.....	23
Figure 4.2 Predicted Energy Output of a Plant by Month.....	24
Figure 4.3 Operation Diagram of a 50 MW Solar Plant.....	26
Figure 4.4 Executive scheme of laying HDPE pipe.....	27
Figure 4.5 Executive scheme No. 3 laying fiber optic cable-8-G.652D-2.7KN.....	28
Figure 4.6 Communication Structure.....	29
Figure 4.7 Top view of Balkhash Solar Power Plant.....	33
Figure 4.8 Data Structure Model.....	34
Figure 5.1 Framework for Digital Twin.....	36

Figure 6.1 Digital Twin Implementation Process Flow	45
Figure 6.2 D-Arc description for the case and UML sequence flow process using a transformer.....	47
Figure 6.3 The Prototype Demo for a 50 MW Solar Power Plant (Average Reading).....	49
Figure 6.4 The Prototype Demo for a 50 MW Solar Power Plant (Meteo Data).....	50
Figure 6.5 The Prototype Demo for a 50 MW Solar Power Plant (Block 4 Combiner Box 18 Data).....	50
Figure B1 “Trends” table.....	62
Figure B2 “Variables” table.....	63
Figure B3 “Messages” table.....	63

List of Abbreviations

DT	Digital Twin
SPP	Solar Power Plant
TGS	Techno Group Service
ML	Machine Learning
AI	Artificial Intelligence
CB	Combiner Box
SCADA	Supervisory Control and Data Acquisition
ASCEM	Automated System of Commercial Electricity Measurement
PV	Photovoltaics
ANN	Artificial Neural Network
NN	Neural Network
ISO	International Standard Organization
OME	Observable Manufacturing Element
NDA	Non-Disclosure Agreement
ID	Identification
RES	Renewable Energy Sources
AC	Alternating Current
DC	Direct Current
IoT	Internet of Things
UML	Unified Modelling Language
MCS	monitoring and Control System
ESMS	Engineering System Monitoring System
PROFIBUS	Process Field Bus
BMBF	German Department of Education and Research
IEC	International Electrotechnical Commission
FE	Functional Element
MR	Mixed Reality
MVP	Minimal Viable Product
API	Application Programming Interface

CHAPTER 1: INTRODUCTION

1.1 Background

Kazakhstan has become a significant player in the worldwide market for renewable energy thanks to its enormous landmasses and plentiful solar and wind resources. With ambitious aims to diversify its energy mix and reduce dependency on fossil fuels, the Kazakh government has developed several regulations and incentives to encourage the development of renewable energy. The country's convenient legal landscape, along with its strategic positioning between major energy routes, has been the reason for considerable investments in renewable energy projects, especially solar and wind (Lim, 2021).

The renewable energy sector of Kazakhstan could be described as an emerging market with an increasingly growing number of large-scale projects characterized by advanced new digital technologies. TechnoGroupService (TGS) is among the leading companies that provide a wide spectrum of services for renewable energy projects, including design, construction, maintenance and consulting functions. Since TGS has built a solid expertise and knowledge in developing green energy projects the company is in a position to provide consulting and management services in terms of engineering design and construction phases.

TGS has established a reputation as a company that thrives for excellence and innovativeness in the renewable energy sector by accomplishing high-quality projects on schedule and under budget.

1.2 Research Motivation

The idea of developing a digital twin solution for solar power plants in Kazakhstan is rooted in the digital transformation happening in the energy sector of the country. The main purpose of the digital twin is to provide real-time information and sophisticated analytics of a power plant's current state and performance. Being one of the strategic sectors of the economy, the introduction of cutting-edge technologies into the renewable energy sector (RES) could play a vital role in attracting foreign investments and growing interest in Kazakhstan's renewable energy initiatives. The topicality of implementing innovative solutions is reinforced by the fact that Kazakhstan possesses huge solar and wind resources (Akhmetkaliyeva, 2020). There are

many challenges in relation to the operational efficiency and effective analytics of RES projects that could be addressed by digital twin solutions. It becomes especially prominent in terms of real-time monitoring, reactive maintenance and deep analytics of solar and wind power plants. The motivation for writing this capstone project is to provide a novel solution to solve the above-mentioned issues by introducing a comprehensive digital twin framework adopted to the operational realities of Kazakhstan's solar power plants.

Traditional monitoring and maintenance procedures and plant management often demonstrate a lack of efficiency, which could lead to insufficient performance and growing costs of maintenance. By introducing a digital twin solution this project will provide a smooth transition from the theoretical domain to the real power plant conditions by equipping the stakeholders and plant management with effective tools to monitor, predict and prevent the state of the plant and perform advanced analytics.

Moreover, the development of a digital twin architecture will substantially assist in increasing the integration of RES projects into Kazakhstan's energy sector. The provided framework will help decision-makers in resource allocation and optimization to maximize solar power plant's operational efficiency via accurate monitoring and analytics. Furthermore, digital twin could provide a basis for data-driven decision-making in long-term perspectives and successful integration of RES projects into the national electrical grid. Energy forecasting applications enabled by digital twin would contribute to achieving a more balanced generation and consumption of all objects in the energy sector.

Additionally, in the long-term perspective, this research initiative will become one of the pioneering digital innovation solutions in the renewable energy sector of Kazakhstan, creating an impulse to implement new digital technologies into the energy sector. This project promotes interdisciplinary cooperation and knowledge transfer (exchange), while also making the renewable energy industry more competitive by becoming the first to establish a digital twin framework specifically designed for Kazakhstan's solar power plants. Strategic partnerships with academic institutions, industry representatives, and policymakers, this research initiative will produce knowledge that can assist in implementing subsequent policy initiatives and technological developments that aim to initiate the transition toward a safer and sustainable energy ecosystem.

In conclusion, the development of a digital twin framework for Kazakhstan's solar power plants is an important step in achieving the region's highest renewable energy potential. This project aims to provide stakeholders with the expertise and skills required for effective management of the complex nature of solar energy production by addressing crucial operational challenges and promoting innovation. In the process, it aspires to open the way for a more prosperous and sustainable energy future.

1.3 Aim and Objectives

The primary aim of this Capstone project is to develop a Digital Twin (DT) framework for Kazakhstani solar power plants. The proposed framework will propose a comprehensive toolset for deploying DT technology for solar facilities in Kazakhstan and provide insights for applications that can enhance efficiency, maintenance, and operational procedures. Real-time data from existing solar power plants has been obtained in collaboration with TGS company which specializes in the construction and operation of renewable energy sources (RES).

The objectives of the project are to:

1. Capture comprehensive requirements for deploying digital twin for solar power plants;
2. Analyze the operational processes and procedures specific to a 50MW solar power plant located in the Balkhash region;
3. Develop a DT architecture and a framework for implementing the DT in solar power plants;
4. Develop a road map for implementing digital twin technology in the 50 MW solar plant;
5. Develop and validate the prototype via experts' judgment.

1.4 Scope of the Project

The scope of the project, as shown in Figure 1.1, mainly focuses on understanding a 50 MW solar power plant's operational process with an evaluation of the plant's infrastructure to identify areas that can be improved through integration with digital twins. Key areas to focus on include SCADA integration with the DT platform, corrective maintenance, performance analysis, real-time data transfer and DT implementation framework. However, this project will

not delve into building platforms from scratch. Instead, it will explore open-source options. Mixed reality is out of scope due to complicated programming and hardware demands. Additionally, predictive maintenance, which requires substantial data collection, analysis, and failure mechanism expertise, will not be the project’s focus. While cybersecurity will not be a primary concern during the prototype phase, it will be addressed in the framework. Financial analysis of the digital twin implementation will be omitted from the project scope.

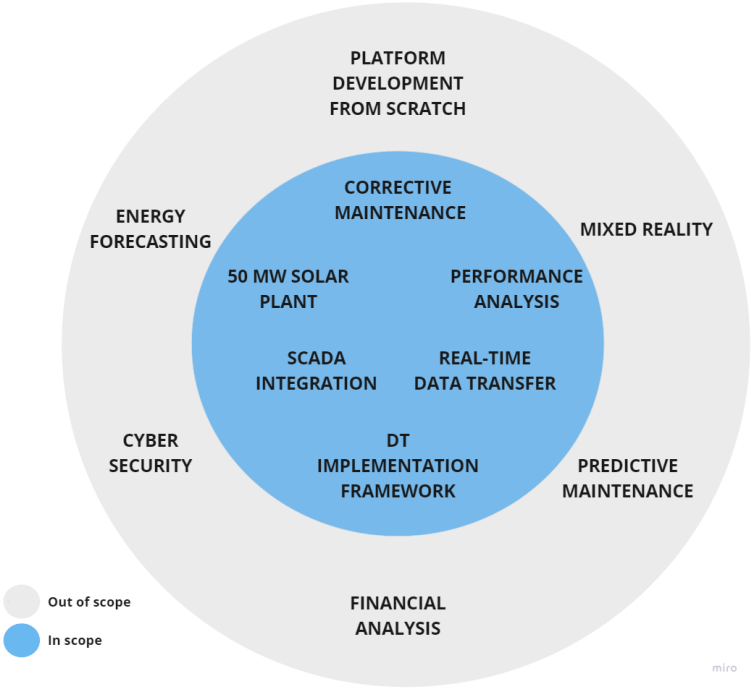


Figure 1.1 Scope of the project

1.5 Structure of the Report

This report consists of seven chapters and its structure is shown in Figure 1.2. The report commences with an Introduction chapter that mainly provides background information on the renewable energy sector and a preliminary introduction of the chosen company. In addition, in this chapter, the project’s aim, objectives, motivation and scope are presented. Next, in Chapter 2 a comprehensive literature review with digital twin detailed descriptions, its application and implementation technologies are provided. The next chapter is the methodology of the project, where the research approach and sequences are presented. Furthermore, Chapter 4 describes the current situation in TGS company with a comprehensive plant analysis. Chapter 5 presents the

synthesis of Chapters 2 and 4 and one of the final deliverables of this project a digital twin development framework. Following the framework development, Chapter 6 presents the prototype developed for a 50 MW solar power plant. Chapter 6 also provides information about the validation of the prepared framework and prototype with expert judgment and approval from stakeholders, including industrial and capstone supervisors. The final chapter concludes the whole project and paves the way for further study.

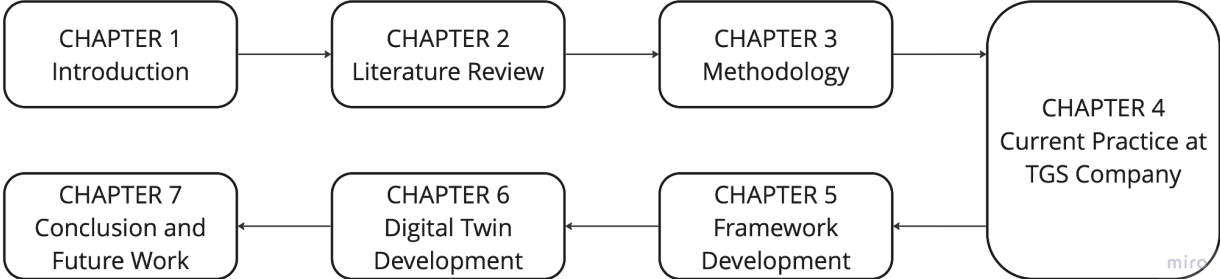


Figure 1.2 The structure of the report

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This chapter provides a literature review on solar power plants and digital twin technology. The concept map of the literature review indicates key research areas that have been studied through academic databases. The research field includes the 50 MW solar power plant located in the Balkhash region, digital twin, digital twin applications, key components for DT implementation, data transfer, and existing standards for DT.

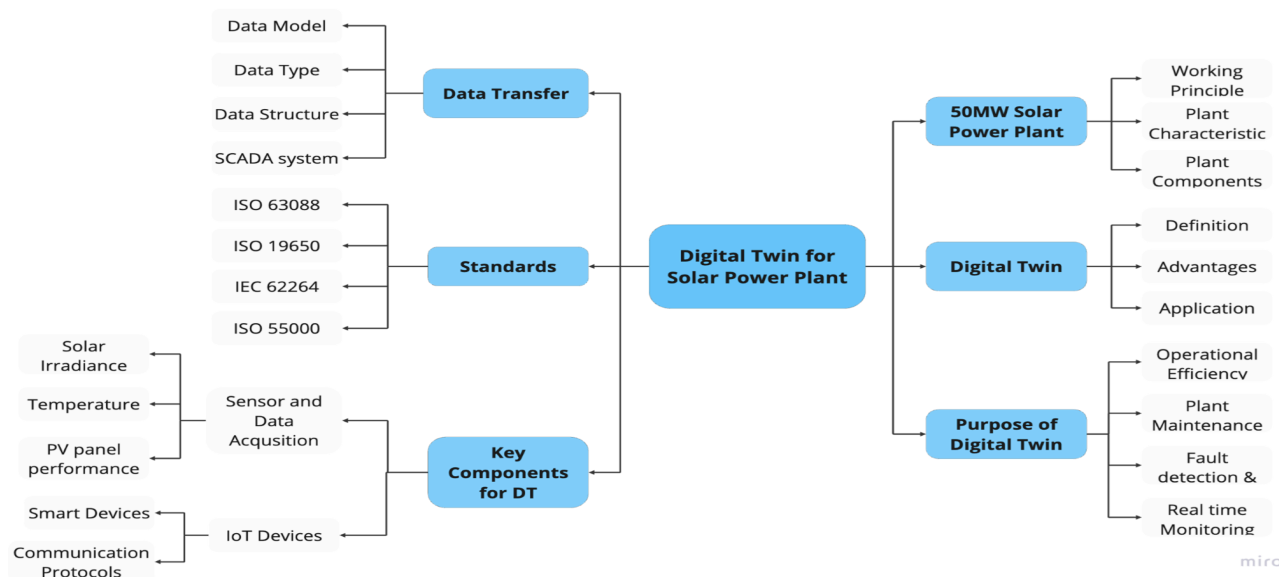


Figure 2.1 Concept map for literature review

Using a thorough analysis of existing data, scholarly articles and industry reports the knowledge gap related to digital twin implementation in Kazakhstan's renewable energy sector was identified. To obtain the most relevant and up-to-date information a comprehensive search of academic databases such as Scopus and Google Scholar was conducted. Figure 2.2 shows a literature review methodology for the project. To stick close to the research scope various keywords such as “digital twin”, “solar plant”, “application”, and “data transfer” were used. Such an approach helped to obtain the most relatable papers and limit the number of research articles to 449 from both databases at the initial search. Additionally, several filters were applied to obtain the most recent research papers written in English. Finally, the title and abstract

scanning, following the introduction and conclusion skimming provided 31 research articles to read.

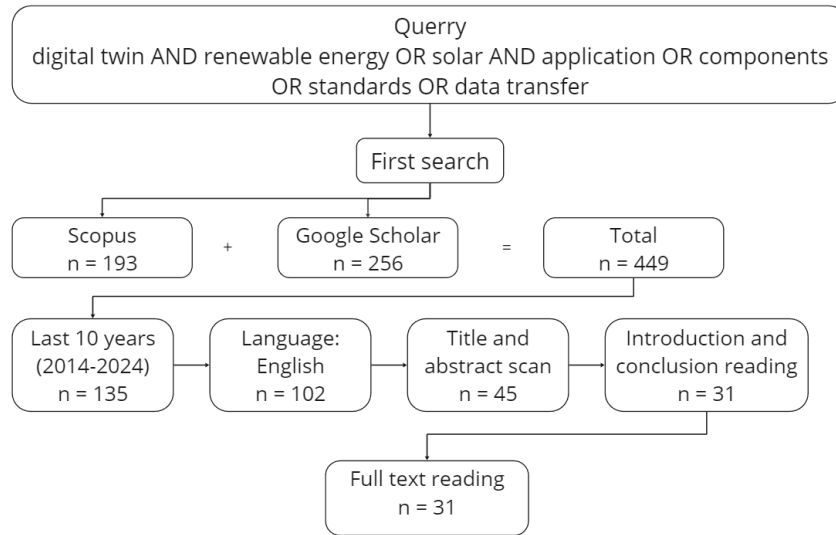


Figure 2.2 Literature review methodology

2.2 Digital Twin Technology

Several countries followed the trend of digital transformation and launched their strategies such as “Industry 4.0” in Germany, “Society 5.0” in Japan, “Made in China 2025” in China, and more generally “The Factory of Future” in Europe (Semeraro et al., 2021). Digital twin technology has developed as an innovation in the energy business, serving as a virtual representation of physical assets and processes. Digital twins, which were originally developed in the manufacturing sector, then spread to other industries, including the energy sector. This section explains the basic idea behind digital twin technology, which is a dynamic digital copy that imitates the behavior and actions of real-world objects. Fundamental insight into the development of digital twins from conception to broad adoption can be gained by following their historical history. Since their inception in the 2000s, digital twins have progressed from simple simulations to complex, data-driven models that enable various businesses to maximize efficiency, reduce hazards, and simplify the decision-making process (Kritzinger et al., 2018).

The term "digital twin" refers to a real-time, virtual representation of a physical object or activity. As a result, a digital twin grid is a virtual replica of actual systems, processes, and

equipment that can carry out computer modeling, simulation, intelligent data analysis, and machine learning to help users improve conventional electric grid planning and decision-making processes (Jiang et al., 2022).

To deliver a thoughtful response to the complex power grid system, the digital twin has developed modeling, simulation, and optimization capabilities. DT serves both visual and predictive purposes for the electric grid. It may also be used to forecast its expected behavior. The DT modeling approach can also be applied to simulation-based systems that support decision-making using real-time data. Real-time data collection, transmission, and processing from real space is accomplished by the virtual model (Grieves, 2014).

The goal of DT is to offer accurate information on the physical twin without relying on modeling or simulation of system behavior (Cimino et al., 2019). Noteworthy are the important developments in cyber-physical systems, the Internet of Things (IoT), and artificial intelligence that are presented as novel revolutions in smart power plants. In addition to these technologies, real-time data is constantly gathered, handled, and evaluated to produce a digital representation of a physical system and a precise understanding of its present and potential future states. As a result, the precise, dynamic, and contemporary virtual representation of the intended system is prepared for secure real-time analysis and control. The idea of DT modeling is included in this description and is taken into consideration by numerous industrial sectors, particularly the power and energy sectors. Real-time and historical data can be handled safely and effectively with the help of a DT power system. This system also helps with maintenance, design, and operation management. Since DT-based power systems are novel and intricate in concept, research on them is still in its early phases.

Many definitions of DTs, which reflect the skills and knowledge of analysts across a range of areas, are given in articles. Digital twins, in a nutshell, are representations of actual physical systems that are linked to the real system via a communication link. Information is continuously exchanged over this communication channel, allowing control, monitoring, and optimization of the physical object. This idea was first presented in the aerospace industry, and as improved technology developed, it was expanded to other industries (Li et al., 2021). The manufacturing sector uses the DT to monitor and control the conditions of its products and production lines (Jafari et. al, 2020).

2.3 Types of Digital Twin

The digital twins are widely used across different production phases by manufacturing specialists (Kritzinger et al., 2018). Digital twins are intended to help with the optimization of working processes, namely the processes involving predictive analytics, adoption of smart automation, etc.

As mentioned above, digital twins may be categorized into a variety of types based on the coverage scale (Andrade, 2023). While some reproduce the whole factory, others illustrate a production line's smallest unit. The four primary types of digital twins are the following:

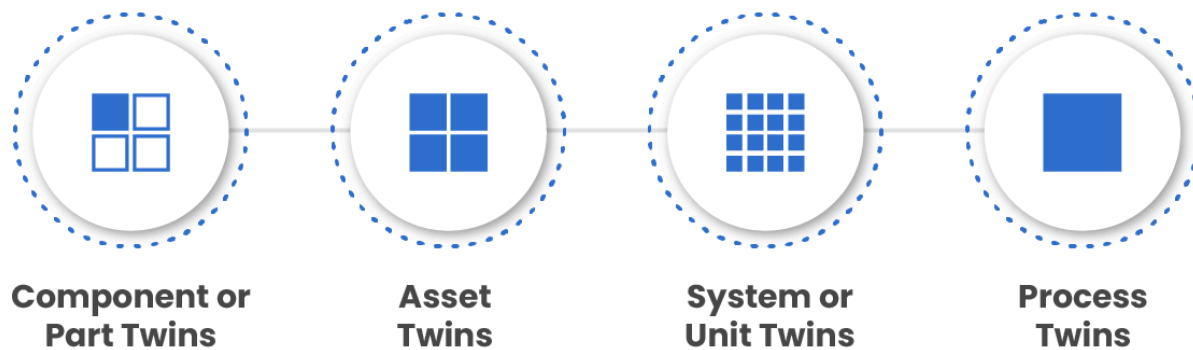


Figure 2.3 Types of Digital Twin

1. **Component Twin:** Likewise referred to as parts twin, it is considered the foundational level of digital twin technology. It complies with the system's smallest elements, e.g. valves, switches, or sensors. The digital illustration of the components allows users to observe components' characteristics to assess their performance, endurance, stability, and efficiency rates in real time.

2. **Product Digital Twin or Asset Digital Twin:** The given kind of DT consists of different component twins or it extracts information details from part twins to model complex assets, namely buildings, pumps, and engines. The asset twin examines how individual components work together to form a coherent whole, providing information about the operation of the machinery and possible areas for improvement.

3. **System Twin:** System twins show how various assets combine to generate functional units by mimicking the assets at the system level. System Twin provides a comprehensive factory viewpoint to increase productivity or find new business prospects. System twins are

mostly used for visualizing the assets associated with certain operations, allowing for real-time simulation and monitoring that is important for process visibility and strategic decision-making.

4. **Process Digital Twin:** Process twins are the most advanced type of digital twin. Process twins combine multiple system twins into a single entity to investigate intersystem coordination and synchronization. This method allows for a flexible examination of outputs by providing an in-depth understanding of the workflows and processes inside plants and factories. Process twinning enables data-driven decision-making by allowing changes to production temperatures or raw material feeding rates without interfering with manufacturing operations. Top management teams may test different business hypotheses and efficiently monitor key performance indicators, encouraging well-informed decision-making rather than relying too much on intuition or past experience.

Digital twins can also be classified into three types based on the data exchange principle (Andrade, 2023). All of them have in common the digital object and the existing physical asset. These are the 3 main models:

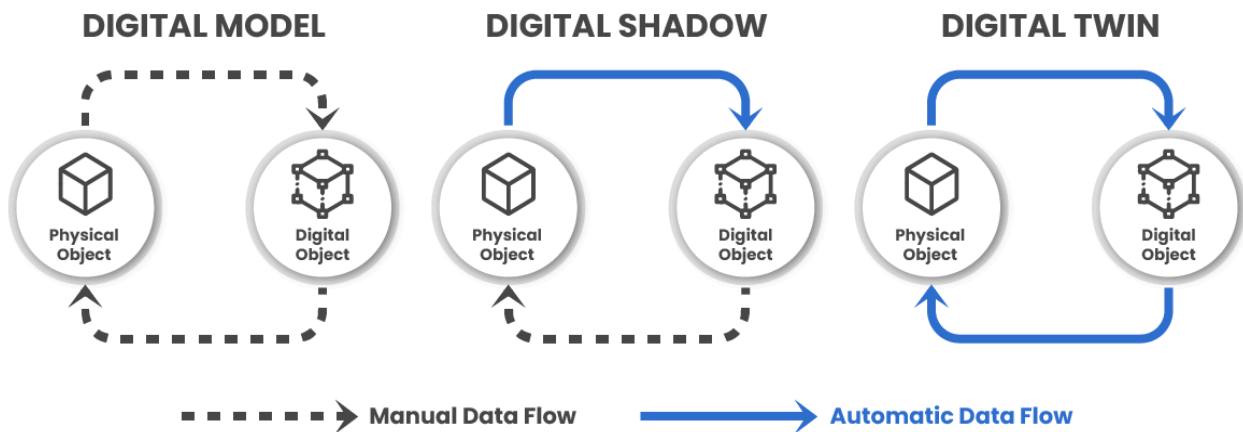


Figure 2.4 Models of Digital Twin

1. **Digital Model:** The Digital Model has a manual data exchange from the physical to the digital model. This approach is used as a foundation to build a Digital Twin system. As the data flow is done manually, the changes in the state of the physical or digital model have no synchronization in any of them.

2. Digital Shadow: In contrast to the Digital Model, the Digital Shadow automatically transfers data from the physical to the digital object. While changes in the physical object can impact the digital representation, the reverse doesn't work.
3. Digital Twin: The main advantage of Digital Twin is the automatic flow of data from the physical to the digital asset and vice versa. As the Digital Shadow, the digital twin can mimic the physical object, but also the changes in the digital object cause the changes in the physical.

2.4 Application of Digital Twin in Solar Power Plants

The application of digital twins in the renewable energy sector creates multiple opportunities from real-time monitoring to fault detection and power generation forecast. Such a powerful technology can be used to deeply analyze the plant performance and identify any faults in advance. Digital twin helps to create proactive actions that will decrease maintenance costs and increase the plant's performance through its life cycle. Specifically, for solar power plants digital twins can be used in several ways: real-time monitoring; predictive and corrective maintenance; fault detection; power generation forecasting and increase in operational efficiency.

A combination of physical systems and smart sensors for measuring different parameters can reveal real-time performance. The use of digital twins can provide an opportunity to monitor the plant's real-time performance, power generation, and health-state monitoring. Current technologies provide an opportunity for cloud-based platforms to obtain real-time information and can be used to evaluate energy resources (Livera et al, 2022). Such data can be used for further data analytics to expand plant performance and identify certain parameters that can be optimized or used for fault detection. The digital version of the solar power plant can decrease the plant's maintenance cost as all information can be obtained directly from the plant itself and used to identify systems that require repair or service works in advance, leading to a decrease in the system's downtime (J.V.S. do Amaral a & Falekas, 2023).

Photovoltaic (PV) panels are a widespread and reliable technology, although PV panels may encounter certain problems during operation such as power fluctuation due to faults in inverters and PV cells; contamination of solar panels with dust, snow, and other types of dirt; PV panel overheating and hotspots (Osmani et al., 2020). Therefore, fault diagnosis is one of the crucial procedures that ensure an uninterrupted plant operation. There are various methods of

fault diagnosis aimed at the identification of anomalies in the different system components. Faults can be categorized based on the system components whether it is PV cells, wiring, or inverters. The PV panels work in an open environment and are exposed to environmental influences such as high-speed wind, high temperatures, and solid precipitation. Given the size of solar power plants, identifying specific failure modes is time-consuming and expensive, so a digital twin platform with integrated aerial imagery or computer vision with high-resolution cameras can detect specific anomalies and alert operators when faults are detected with a specific cell positioning (Vlaminck et al., 2022).

Aerial imagery can be done with drone photography of the plant site and used to identify anomalies on the PV panel's surface. In addition, one of the problems with PV panels is hotspots that reduce panel performance, hence leading to a decrease in overall plant productivity. The excessive proportional heating of a particular cell element compared to the surrounding cells reveals the anomaly within the cell or whole panel that is usually caused by the shading of a certain point on the solar panel. Such shading can cause reverse bias and concentration of current on specific cells, leading to a temperature rise in that region of the panel (Pruthviraj et al., 2023). Therefore, the use of drone thermal photography or static infrared cameras can identify such hotspots and warn operators of excessive temperature rise, which will prevent whole panel degradation. Moreover, a visual inspection of the plant site and the solar panel's condition can spot areas covered with sand, snow, and other types of dirt, increasing maintenance efficiency and decreasing operational costs.

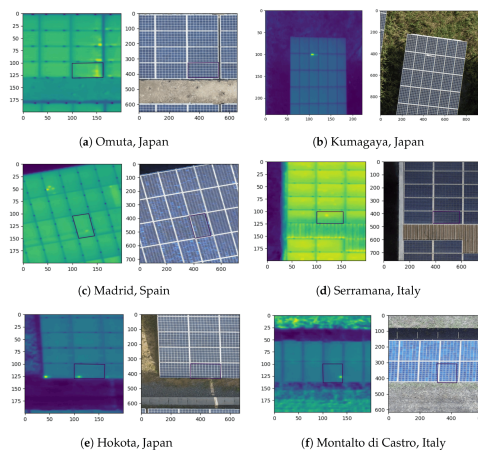


Figure 2.5. Aerial imagery of solar plants that indicate shading (Vlaminck et al., 2022).

Sophisticated development of machine learning algorithms and artificial neural networks made great improvements in fault diagnosis procedures. The acquired data through multiple

sensors located on the plant not only show the plant's current state but also can be used for analysis of historical data. It is important to note there is a difference between fault detection and fault identification as different procedures and approaches should be used for each of them. Fault detection is a process with a binary outcome that indicates the presence or absence of failure, meanwhile, fault identification is a precise classification of fault type. (Jain et al., 2020)

Machine learning (ML) algorithms perform best for fault detection as they can identify certain patterns in big amounts of data and recognize abnormality in the system's behavior (Jain et al., 2020). The digital twin development implies the use of various sensors and IoT devices that can transmit and store data, ML algorithms can quickly show faults and warn users. In addition, machine learning applications can be expanded to analyze power generation and used to forecast a plant's output based on meteorological information. Artificial neural networks (ANN) are a subset of machine learning algorithms that work best in fault identification due to complex pattern recognition and constant training on real-time data. In addition, ANN can be used for predictive maintenance as it can automatically extract certain features such as temperature, vibration, or current fluctuations without manual measurements (Ye et al., 2020). Such ability to identify the specific correlations between different parameters enhances the prediction and proactively spots anomalies that can not be apparent to plant operators. Deep neural network algorithms can predict the remaining useful life of equipment on the station based on historical maintenance data. Such proactive planning can assist the maintenance crew in identifying the weak spots and ensuring target component replacement, which positively affects the plant's uptime.

2.5 Implementation of DT Technology

Overall, the crucial part of exploring the digital twin and its use case in real life is searching for guidelines and step-by-step descriptions on how to do it. The cornerstone of the digital twin is the set of standards that guide its implementation and development. Standards provide the framework, system architecture, and guidelines necessary for harnessing the power of digital twins through operations.

ISO 23247 series defines the development framework for various implementation cases that consist of 4 parts (Aheleroff et al., 2021). The concept of digital twins is at a very early stage of development and is not complete. Therefore, existing standards made for specific areas or

general use such as data transfer, system control, etc. are providing a good foundation for upcoming improvement in Industry 4.0 (Shao, 2021). ISO 23247 series covers general principles, architecture, digital representation, and information exchange. According to Guadong Shao ISO 23247 series has introduced observable manufacturing elements (OME) that represent physical objects and processes in digital dimensions (Shao et al., 2023). Standard defines 8 types of OMEs by its 7 attributes that construct relations with other attributes in the system domains providing us a framework for the digital twin. A detailed description of attributes for the OME can be found in Table 2.1 below.

Table 2.1. Attributes description with standards

# of attribute	Attribute	Description	Standard
1	Identifier	Unique identity of OME	ISO 8000-115
2	Characteristic	General/Model feature of OME	IEC 62264-2, ISO 13584-42, or IEC 61360
3	Schedule	Time domain description used for manufacturing process/ Working hours	ISO 8601 series
4	Status	Current state of OME For instance: in process, complete, tested	VDMA 24582
5	Location	Location of OME	ISO 6709, GPS
6	Report	Task/processes completed by OME	Quality information framework, MTConnect
7	Relationship	Link/Connection to other OME	IEC 62264-2

One thing to mention is that solar power plants may not be considered traditional manufacturing since no goods are produced, but all the operational elements, processes, and constructions mainly come from manufacturing. In general view, ISO 23247 is a standard for

creating digital twins and as its name (Automation Systems and Integration) states it is used for the automation of systems.

After describing general principles, the “Reference Architecture” ISO 23247 series introduces four main domains needed for building an architecture that can be seen in Table 2.2. Each domain has different functional entities that provide functionality within the system and come as a user entity, core entity, data collection and control entity, cross-system entity, and specific functional OME entities (Shao, 2021). ISO 23247 gives a reference model for the functionality of digital twins which can be seen in Figure 2.6 below. The last third and fourth parts of ISO 23247 rely on existing IEC 62264 -2 (Enterprise system integration) to get Digital representation and information exchange and provide technical requirements (Shao, 2021).

Table 2.2. ISO 23247 Reference architecture domains

Domain	Description
Observable manufacturing domain	Describes operating environment and works with system control and data gathering
Data collection and system control domain	Connects OME to digital twin by adjustment
Core domain	Manages digital twin
User domain	Interacts with any DT application or service

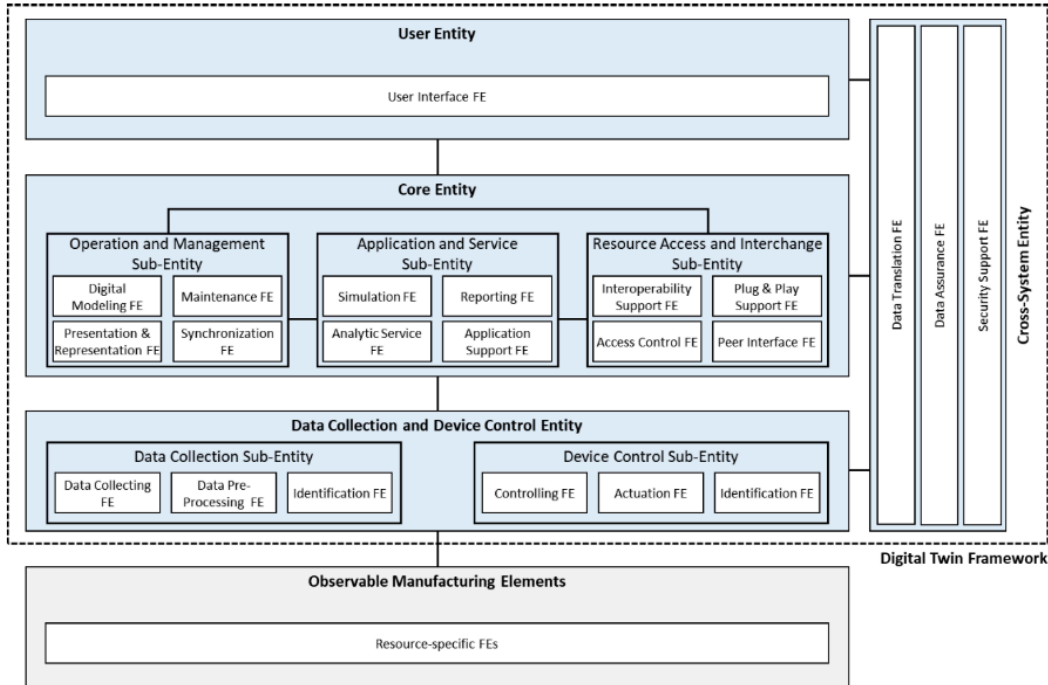


Figure 2.6. A functional reference model for the digital twin

The next step after reviewing the base for the development of the digital twin is exploring architectures for implementation in solar power plants. All existing literature does not specify what equipment to use and directs users to refer to international standards. However, Gourisetti et al. (2023) in their paper imply its framework called the digital twin architecture stack (D-Arch) that gives insight into the real implementation of the digital twin. The authors aim to provide a structural approach for designing digital twins in the energy sector by providing additional depth and specificity to the existing foundation of international standards.

D-Arch technology stack is categorized into physical and digital dimensions that consist of six horizontal layers and interlayer vertical connections. D-Arch covers all the domains previously discussed with the same functional entities in more detailed categorization as can be seen in Figure 2.7. In the physical layer, we have infrastructure and peripheral layers. They correspond to infrastructure and physical systems and devices. This also includes environmental and operational entities such as interactions with the physical system, and built-in/external data acquisition systems with sensors and actuators. For the acquisition systems as part of the peripheral layer, different systems such as Volttron, Cymbiote and SerialTap can be used(Gourisetti, 2023). However, the adaptation of such technologies is difficult due to the low level of compatibility between systems and sensors/equipment at the infrastructure level.

The next four layers correspond to the digital dimension or data orchestration level that has consecutive operations and data-driven processes. The communication layer is responsible for the connection of physical systems with digital systems wired or wireless, ensuring that connection protocols are provided for secure and continuous exchange of information. The model layer consequently takes data from the physical layer and transforms it into digital. The end product from the layer is the transformation of data into the observable digital model by the use of software with precise digital representation. Modeling provides real-time monitoring and reactive maintenance by visualized data and digital twin system analysis. The output layer provides testing and validation of composed representations to avoid erroneous data and conclusions(Gourisetti, 2023). The application layer is mainly composed of data-driven actions by AI/ML for further applications. Layers combined lead to a digital twin that can interact with the user and user applications.

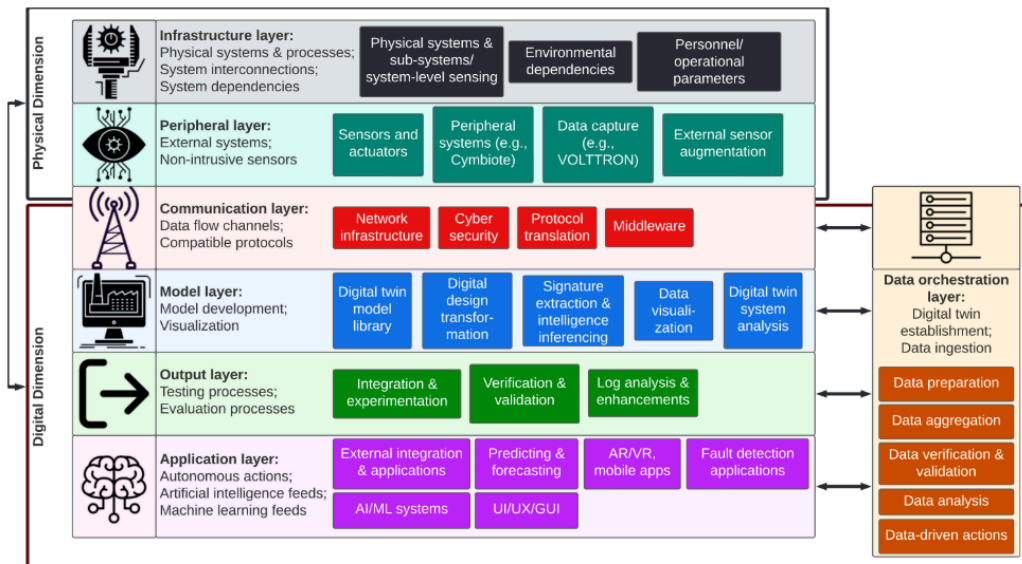


Figure 2.7. D-Arc technology stack

2.6 Research Gap Analysis

The research gap analysis was conducted to evaluate the current level of digital twin development in the renewable energy industry in Kazakhstan. The analysis revealed that there is little information regarding the DT technology and its application in Kazakhstan’s solar power

plants. In addition, research gap analysis helped to point out specific areas to focus on for DT development.

- Existing literature shows a lack of research in the field of digital twin solutions in solar power plants in Kazakhstan. Limited research on the development of DT for solar energy has been carried out in different parts of the world, but it is focused on small-scale solar power plants. This situation creates a significant gap as the equipment, data transmission and operating procedures at large solar power plants vary significantly. In addition, the environmental conditions and competencies of the company are also different from existing solutions and cannot be directly applied in Kazakhstan.
- Due to the difference in solar power plant scale and operating environment, the existing literature does not provide information on the operational processes of Kazakhstan's solar power plants. Such information is crucial since it provides a foundation for future digital twin platform applications. The plant's facilities, scale, equipment and hardware are essential components for digital twin. Without a comprehensive understanding of various plant processes, it is challenging to develop effective digital twin models tailored to the local context.
- One of the most critical gaps identified is the lack of a digital twin implementation framework and roadmap in Kazakhstan. The absence of clear guidance and procedures for developing digital twin solutions complicates the planning and implementation of such solutions in Kazakhstan's renewable energy industry. This lack of structured direction hampers efforts to enhance operational efficiency, optimize overall plant performance, and utilize advanced technologies for improving maintenance practices.

In conclusion, research gap analysis suggests a comprehensive analysis of Kazakhstani solar plants. Explore the facilities and operations of big-scale solar plants for effective integration with digital twin solutions. In addition, a detailed framework with essential components of DT should be prepared to fill the existing gaps.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Introduction

This project uses a real operating facility to create a Digital Twin framework for a 50MW solar power plant. The goal is to build a fundamental framework that can be beneficial as an innovative tool for the renewable energy sector of Kazakhstan. To achieve this goal, concurrent mixed-method research has been conducted with deep analytical methods. The project consists of four distinct phases commencing with Phase One: Project Initiation and Planning, followed by Phase Two: Research and Data Acquisition, then progressing to Phase Three: Framework and Prototype Development and concluding with Phase Four: Expert Validation and Project Closure. The adopted methodology is illustrated in Figure 3.1

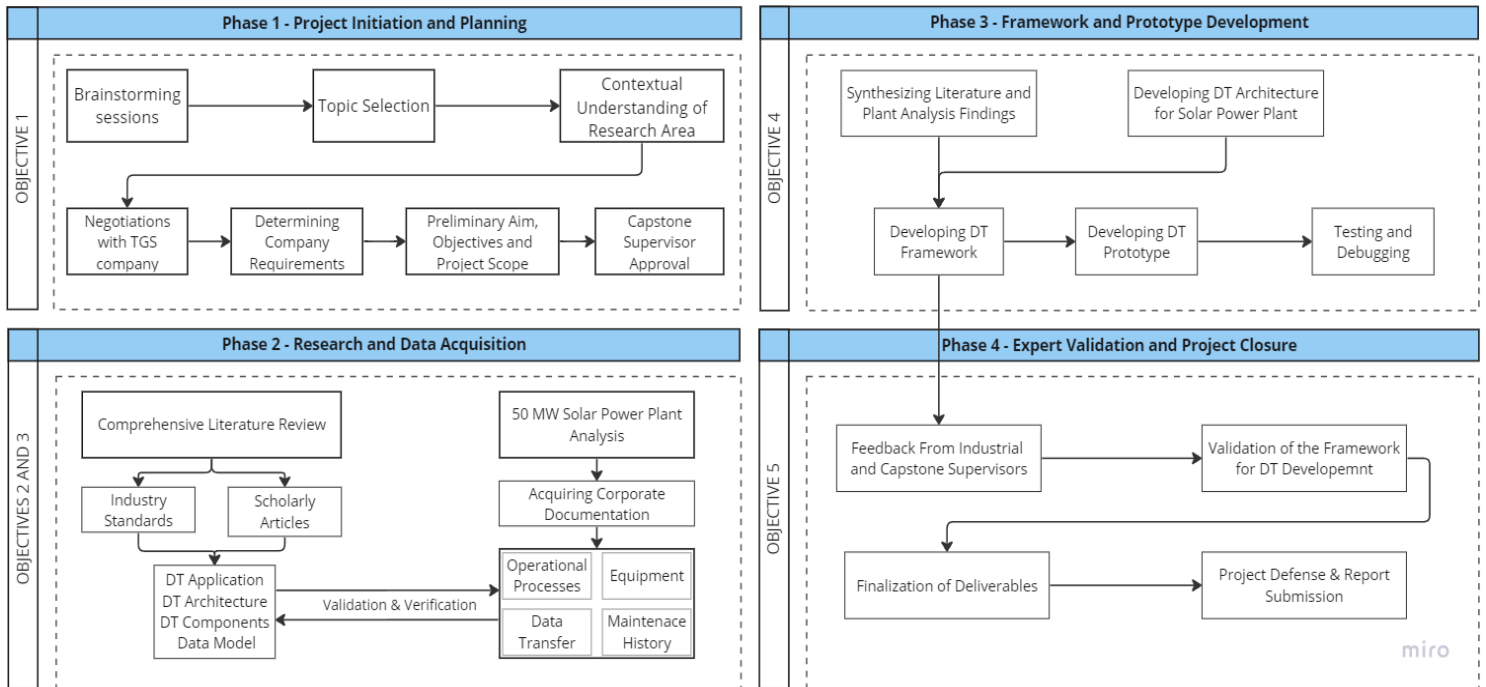


Figure 3.1 Methodology of the Project

3.2 Project Development Phases

Phase 1. Initiation and Planning

Project initiation and planning are the initial stages where the preliminary aim, objectives, and project scope are identified. In addition, after the topic selection and identification of the research area, negotiations with the TGS company were initiated to gather company requirements. Through meetings with company representatives, the direction of the project was identified to benefit both parties, and a non-disclosure agreement (NDA) was signed. After that consultation with a capstone supervisor was held to identify deliverables and establish the project's final aim and objectives. Moreover, after document analysis, a preliminary literature review and several meetings were held to identify team members' capabilities and narrow the scope to achieve the best possible results. Based on these findings a Client Brief was prepared to get the approval from the company and the capstone supervisor. At this point of phase one initial objective of the project was successfully achieved.

Phase 2. Research and Data Acquisition

The second phase of research is focused on conducting a comprehensive literature review and obtaining the company's internal documentation. The literature review provides data regarding the industry trends and common standards for DT implementation in the renewable energy sector. Meanwhile, to understand the TGS operation and solar power plant characteristics, the company delivered several packages of internal documents required for plant analysis. Plant analysis is done to investigate the operational processes, data transfer, and plant infrastructure. This gives an overview of interactions and correlations between the physical grid system and its counterparts. Instantaneously, data flow from plant analysis and literature review is set to reach validation and verification. This helps to assess the fitting level of the theoretical approach and construct a fact base for the further development of the digital twin framework.

Phase 3. Framework and Prototype Development

Phase Three is a crucial stage of the project where the main output is generated. The synthesis of previous stages provided detailed information regarding the components and sequences of DT for solar power plants. Alongside analysis and adoption of acquired documents, the DT architecture was prepared to become a foundation for the framework. The framework

development was based on the detailed analysis of data models, data transfer from the plant's equipment, available infrastructure and company competencies. As a result of this work, the framework for digital twin development for solar power plants in Kazakhstan has been meticulously prepared. This framework integrates all stages of DT development with detailed information on the technologies and data involved.

Consequently, based on the proposed framework the prototype development was initiated. The TGS company already had a platform that was integrated with the SCADA system. Such a foundation helped to use the existing web platform to develop the DT prototype. The prototype includes the 2D plant representation and contains information regarding the crucial equipment such as inverters, combiner boxes, PV panels, and sensors and provides the exact location. Such an approach helped to mimic the real solar power plant on the screen and integration with the company's server provides real-time information from each component. After that, testing and debugging were conducted to ensure stable work of the prototype.

Phase 4. Expert Validation and Project Closure

The final stage of the project was to validate the framework and prototype through expert judgment and stakeholder assessment. During the whole project, constant meetings and presentations were done to reveal progress and obtain valuable feedback from both industrial and capstone supervisors. The last stage of the project was the preparation of all deliverables and the submission of the capstone report.

CHAPTER 4: ANALYSIS OF THE CURRENT PRACTICES AT TGS COMPANY

4.1 Introduction

In this chapter, a comprehensive analysis of the operational processes, infrastructure of the plant, communication structure, and data organization is presented. To develop a properly functioning Digital Twin the analysis of the infrastructure of the plant should be performed. It is more convenient to apply the individual approach for each of the plants in terms of infrastructure and communication. The reason is that each solar or wind power plant project could have different design solutions and decisions for the choice of the equipment, the layout of the equipment, the communication between the components, the data structure and the software used. After getting a holistic and detailed understanding of the power plant operations and software the preparation of a framework for this power plant is to be performed.

4.2 Company Overview

TechnoGroupService LLP is a domestic engineering group operating under the TGS brand, dedicated to implementing projects in the field of renewable energy sources (RES) using adaptive innovative solutions and maximizing domestic content.

The group of companies includes:

- TechnoGroupService LLP - the construction and installation division.
- Center for Green Technologies LLP - guaranteed energy supply and energy forecasting.

The company specializes in designing facilities powered by renewable energy sources of various capacities and types, including hybrid solutions. During the project stage, the company offers its clients solutions for optimizing both capital and operational expenditures. Services include project exploration, conceptual design, project documentation, working documentation, cost estimates, and author's supervision. The company has engineered proprietary load-bearing metal structures for photovoltaic modules and established their production.

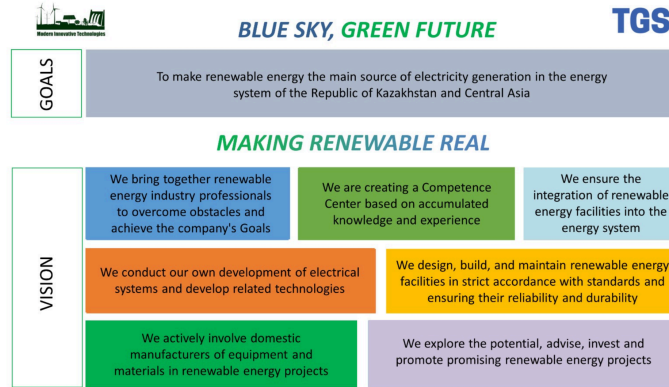


Figure 4.1 TGS Company’s Goals Board

According to the description of goals and vision in Figure 4.1, the company has clear ambitions to develop the Renewable energy sector of Kazakhstan by innovating, integrating and building strong competence. The company provides a wide range of services and strives for innovation leadership. TGS-energy puts strong emphasis and effort on IT direction to develop IT solutions for the renewable energy sector of Kazakhstan.

4.3 50 MW Solar Power Plant Analysis

4.3.1 50 MW Power Plant General Characteristics

The framework and implementation were performed on a 50 MW Solar Power Plant located in the Balkhash region. The key characteristics of the plant are represented in Table 4.1

Table 4.1 Characteristics of Power Plant

Characteristics	Description
Total power of the plant	50 MW
Latitude	46.941222
Longitude	75.001417
PV module	Si-mono model JKM525M-72HL4-TV
PV module power	525Wp
Number of PV modules	94 150 pcs.

Configuration of the PV modules	in series (25 modules)
Total number of Blocks	8
Inverter model	SG3125HV-20 a power of 3.125 MVA
The rated power of the transformer in the block	6.250 MVA
Annual electricity production	83 million kWh
Number of employees	28

In Figure 4.2 we can see the predicted energy output of the plant for the whole year with estimated collection losses and system losses. It is observed that during the winter season, there are minimal losses both from the system and collection sides. However, during the summer and spring seasons, the losses are at the maximum. A better understanding of these losses and their root causes could be achieved by using a digital twin technology for gathering and processing real-time data from the plant and analyzing historical data conveniently.

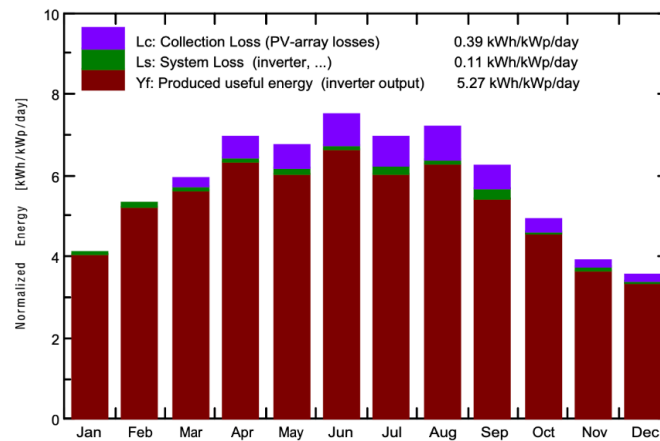


Figure 4.2 Predicted Energy Output of a Plant by Month

4.3.2 Infrastructural Layer Analysis

After a thorough analysis of the operations of this plant, we developed the following conceptual diagram that represents the key processes of the plant. The organization of the plant is a complex task due to the necessity to properly arrange 100,000 panels and achieve an efficient flow of electricity across the components. Figure 4.3 displays the key physical components of the plant, including PV modules, Combiner Boxes, Blocks, Inverters, Modules, Transformers, a

Substation, and a Block-Module Building. Overall, this diagram is a top-level representation of the plant operations that includes only the key components and electricity flow without any details about the electrical and electronic components and data flow. Two main entities flow across the components: electricity and information (data). During the analysis, it was found that the organization of the communication between components in terms of data flow significantly differs from the structure of the plant in terms of electricity flow. Taking this into consideration the data model and communication scheme of the plant are developed separately and will be described later. In Figure 4.3 the focus is established on the electricity flow across the components.

The electricity generation starts on the surface of PV modules, where the sun energy or irradiation is converted into electrical direct current (DC). The current flows from the modules into combiner boxes. These boxes serve as structural nodes that accumulate and centralize the electricity coming from PV panels. Each box can accumulate around 1200 PV panels. Consequently, the electricity is transferred to the Blocks. Block is a key component of the plant since the inverters and other equipment are located inside it. The inverters convert the DC voltage into AC voltage. Each inverter has 2 modules that help to accumulate the electricity coming from a large number of combiner boxes. There are 8 main blocks on the plant and each one demonstrates different behavior since these structures could be considered as individual solar power plants. From the blocks, the electricity is delivered to the transformers where the voltage is increased to be transmitted to the grid and reduce the losses. Increasing the voltage is a common practice and method to transmit electricity for long distances without substantial losses in the cables. Finally, the current is delivered to the substation where it is transmitted to the national grid. This conceptual representation gives a preliminary and holistic understanding and vision of how the 50 MW plant operates and what are the key processes.

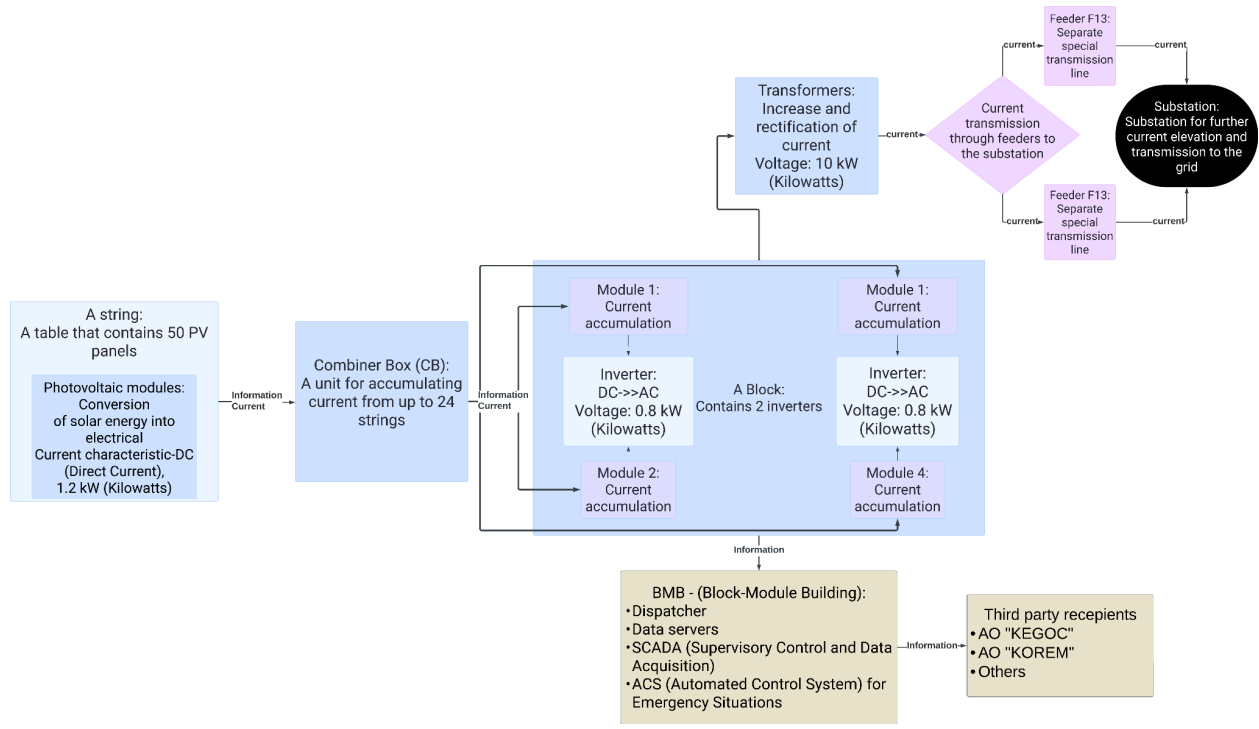


Figure 4.3 Operation Diagram of a 50 MW Solar Plant

Apart from the conceptual representation, the topology of the organization of key components (the layout) is represented in Figure 4.4 below which represents exactly how the panels, combiner boxes and blocks are connected. It was found that each set of 50 panels is organized on the table or string. Each combiner box is connected to 12 tables or 600 PV panels. However, that is not always the case and some combiner boxes have less number of connected tables. As it can be seen from the scheme Block 1 accumulates 21 combiner boxes but some blocks have up to 24 combiner boxes. High-density polyethylene with a 40 mm diameter is used for covering the electrical cables inside since the cables are put under the ground.

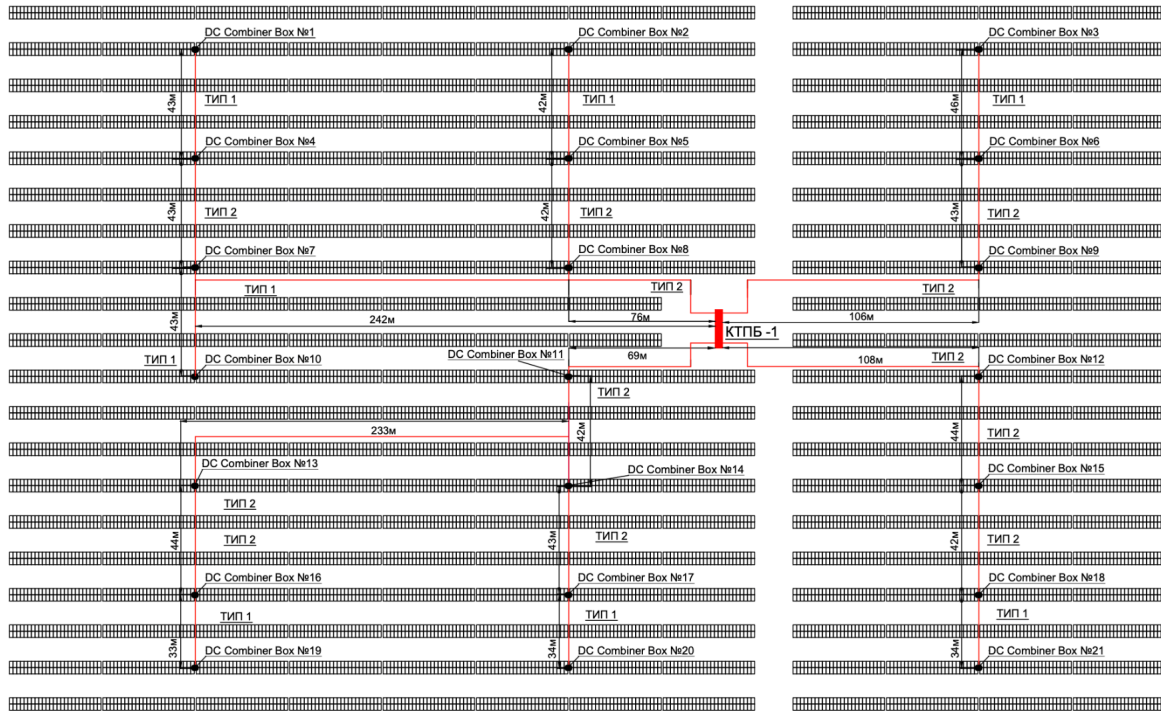


Figure 4.4 Executive scheme of laying HDPE pipe

4.3.3 Communication Layer Analysis

Organization of communication on the plant

In this section, the analysis of the data flow on the plant is presented to provide a full picture and understanding of how the data is transmitted across the elements of the plant. There are 2 main dimensions to be discussed in this section: hardware (equipment) and software. The communication consists of the electronic components, cables, and physical components along with the software that reads, processes, and interprets the signals and data from the equipment. The analysis starts with the top-level description of the communication structure followed by a detailed and narrow description of the communication.

Figure 4.5 reveals the layout of fiber optic cable connections between the blocks. It is observed that the connection of the blocks is linear, starting from Block 1 and ending with Block 2. There is a separate connection between the module building and a substation. An important detail of the scheme is a description of the IP addresses of different components. Each block has its own IP address and inside the blocks, each of 2 inverters has its own IP address. This means that the data is gathered from each inverter separately. There is a unique IP address, named

“meteo” that represents the data from the meteorological station, which will be described in the following section. Most probably the meteorological station is located on block 1 and there is only one such meteorological station on the whole station, which means the meteorological data will be generated only from 1 point. This could also be a serious limitation of the digital twin since the area of the station that should be covered is substantial and one meteorological station will not cover the whole area of the station.

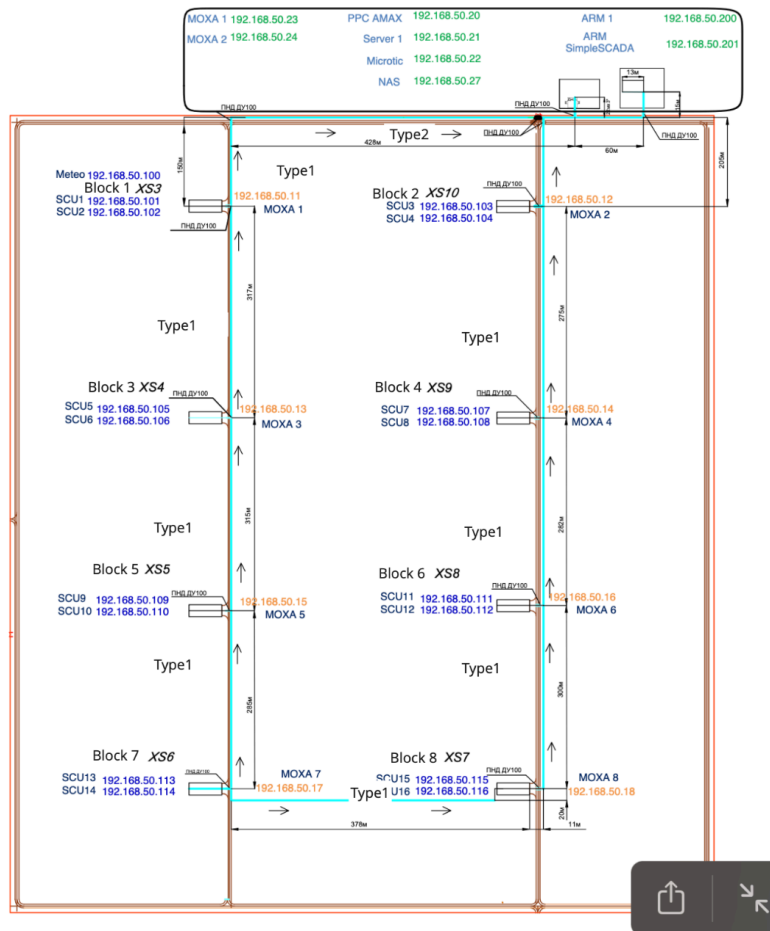


Figure 4.5 Executive scheme No. 3 laying fiber optic cable-8-G.652D-2.7KN

Apart from the physical configuration of the optical cable connection between the blocks, the important element of the station is the structure of the 2 main systems: Engineering Systems Monitoring System and Monitoring & Control System (MCS). These are specialized systems for monitoring and controlling the state of the plant and its key components. There are specialized standards that should be met to provide all the necessary functionality and safety on the plant and

these systems are not exceptions. In Figure 4.6 below there is a detailed communication structure of a 50MW solar power plant.

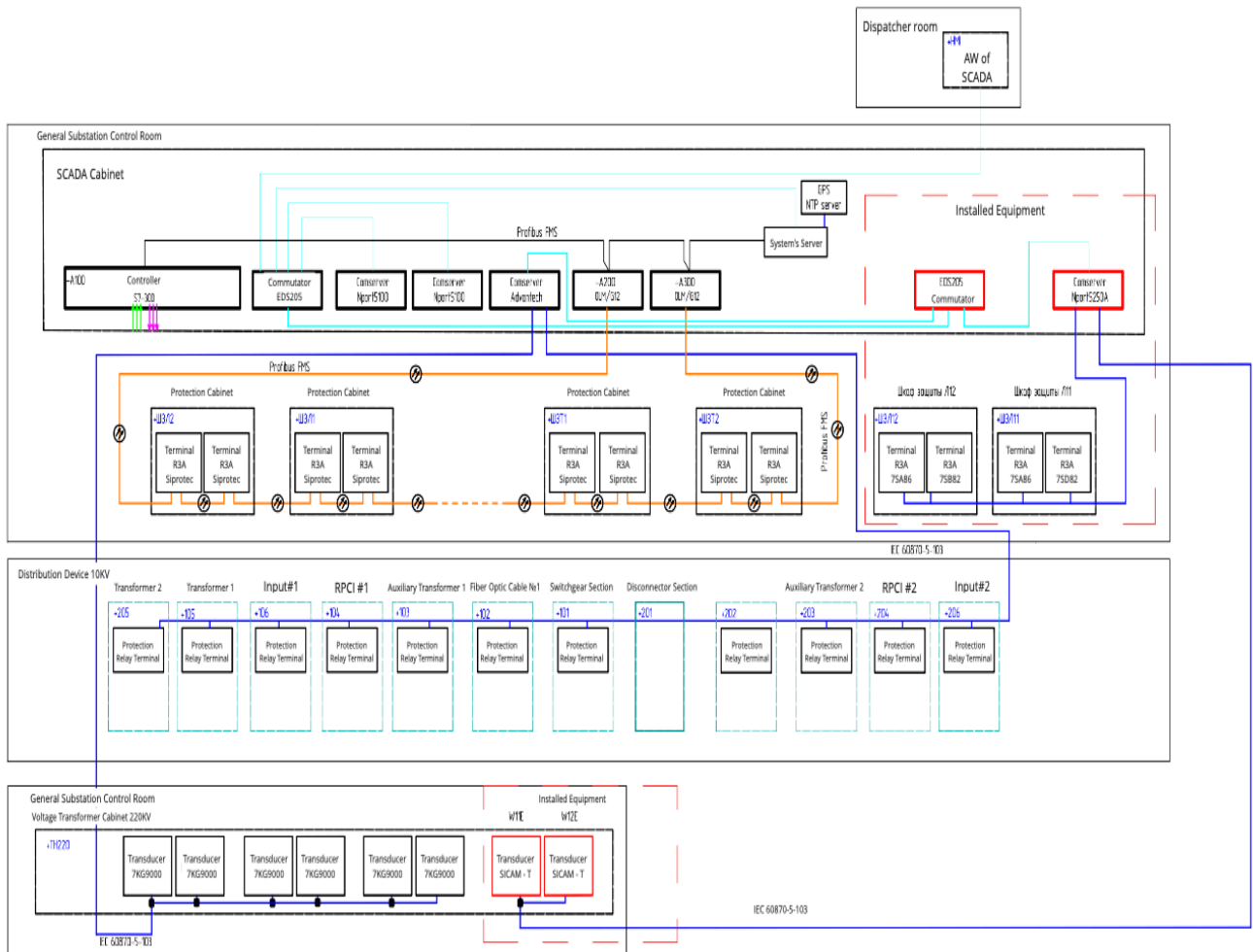


Figure 4.6 Communication Structure

The communication structure consists of 4 principal components: Dispatcher Room, General Substation Control Room, Distribution Device, and Voltage transformer Cabinet. The interactions between these components and elements inside the components are quite complicated since the data flow is not linear. Physically, several components are located in the same physical space but have completely different functions.

The flow of data could be described as bottom-up starting from the bottom layer. This component is the Voltage Transformer Cabinet. Inside the cabinet, a series of Transducers 7KG9000 are located. These digital metering transducers convert the physical measurements

from the measuring devices into digital signals that are transmitted and processed further. The signals from transducers are transmitted to the COM server to be further processed.

The second component of the scheme is a distribution device. Inside the distribution device, several relay protection terminals are located. The function of these terminals is to facilitate the connection of sensor inputs, such as current and voltage transformers, to the protection relay. The terminal collects these signals, allowing the relay to monitor electrical parameters and detect faults. Protection relay terminals are an indispensable component to ensure the safe operation of the system.

The third component is the General Substation Control Room. This Room has several sections: SCADA Cabinet, Installed Equipment, and Profibus RMS. PROFIBUS (Process Field Bus) is a standard for fieldbus communication in automation technology, allowing the integration of a wide range of devices such as sensors, actuators, and controllers across different manufacturers and systems. It was developed and promoted by BMBF (German Department of Education and Research) and adopted by Siemens. Profibus is one of the advanced standards in data communication due to such advantages as high interoperability, real-time communication, flexibility and adoption for maintenance and diagnosis. As it is observed, the communication on a 50 MW plant is done using the Profibus, which indicates the advanced level of communication and highly standardized approach. According to Profibus, there are terminals from Siprotec Siemens used on the plant for further control and processing of the signals from the metering devices and other equipment. Finally, the SCADA cabinet hosts all the receiving equipment, such as controllers and COM servers. The COM servers receive signals from the Voltage Transformer Cabinet and Distribution Device. The PROFIBUS Optical Link Module (OLM) G12 V4.0 is a module from the PROFIBUS family that receives signals from the Siprotec Terminals. As it is observed the Profibus standards are well sustained by using the Profibus equipment and protocols to achieve high quality of communication.

The final and most straightforward component is a dispatcher room. Here the Automated Workplace of SCADA is located. The graphical interface of SCADA and control equipment that operators use to monitor and control the operation of the plant is located in the dispatcher's room. This is an output point of the whole system where the data from the controllers and COM servers are received to be displayed and applied by human operators.

Overall, the architecture of the communication system and all the equipment used follow advanced Profibus and IEC 60870-5-103 standards that imply high speed, reliable, real-time communication with smart options and advanced analytics capabilities.

4.3.4 Database Structure

In this section, the description of the data, including all the transferred parameters and their characteristics is presented. In the previous sections, the analysis of “how” the data are organized and transferred was conducted. In this section, the analysis of “what” to be transferred is performed.

There were several key overarching requirements for a database to successfully develop a digital twin. The first requirement for a database is the operability of a large number of parameters. It was decided that for the digital twin project and possible future projects, it is more visionary to get the most possible number of parameters. From the power plant, it is possible to retrieve 6125 parameters. Such several parameters are quite hard to handle since each parameter has its archivation period (from several seconds to 10-15 minutes). Another constraint of building a database is the relationships between parameters: some parameters are child parameters and others are parents, which means some parameters will be parts of other ones. For instance, each inverter has the parameters that represent modules and each module consists of combiner boxes that also have their unique parameters.

Taking all the above into consideration, there were 2 possible structures of the database for the project.

Structure #1

The difference in both structures is the way how to define an index (a leading column). For the first database, the ID (unique identifier) of each parameter is set as the index. In this case, all 6125 parameters are represented by their unique ID. However, each ID will have a large number of timestamps related to this particular ID. The issue with this structure is that the timestamps for the parameters will not be uniform and the same for the parameters due to differences in archivation periods and the volume (size) of the database will be substantial due to frequent database updates according to the timestamps.

Structure #2

The second proposal for a database is setting the timestamp column as an index. In this scenario, all the parameters will be processed according to the timestamp. This means that the timestamps will become unique and the same for all the parameters. The difficulty of this method of database setting is finding the uniform archivation period for all the parameters, which will inevitably lead to empty (NULL) spaces in the database since some parameters will have larger archivation periods than others. Such a structure will occupy less space and volume and be easy to use in Machine Learning Applications and production analytics. Having a timestamp as an index makes it more convenient to analyze the dynamic performance of different parameters and compare them across time.

However, after analyzing both structures and discovering additional options a new solution for a database structure was developed. It was decided that for the digital twin purposes, a more winnable approach would be a synergy of both Structures 1 and 2. The structure of the database will resemble Structure 1, where the ID is set as an index (leading column) but the archivation period for all variables will be 1 minute. This means that all the IDs will be updated with a frequency of 1 minute. If some of the parameters should be updated less frequently, the database will use the previous value until the new value is updated. Such an approach will add uniformity and a high frequency of archivation for all parameters while sustaining unique IDs for the parameters. To execute the chosen structure the database will have 3 main tables: trends, variables, and messages. A “Trends” table represents the actual values of the parameters (see Appendix B). The columns of the table are ID, Timestamp, Value, and Quality of Code. The last column shows how good the received value of the parameter is and whether there is any issue with this value.

The second table “Variables” acts as a definer or acronym table, which shows what is the meaning of each ID and which parameter this ID represents (see Appendix B). The columns are the ID, the corresponding name of the parameter, the description of the parameter, and 2 additional columns latitude and longitude. The purpose of these columns will be explained later. The last table “Messages” demonstrates the messages to be sent using the database. The columns in this table are the text of the message, the timestamp of other operational columns, which will be described later (see Appendix B). The general purpose of this table is to send different messages to the operator or engineer in case of emergency for reactive maintenance practices.

As it was mentioned above there are 2 additional columns in the variables table: latitude and longitude. The primary application of the digital twin is real-time monitoring of the solar power plant. It was decided to use a 2D visualization of the solar power plant (top view), where all the components are located in the exact positions as in the real solar power plant. To mimic the positions of all the components on the digital 2D visualization it was decided to use the exact coordinates of all components on the real power plant. This approach will make it possible to achieve the highest accuracy in locating the components on the 2D scheme of the plant. Google Maps presented in Figure 4.7 was used to find the coordinates (latitude and longitude) as it has sufficient resolution and updated images of the plant.

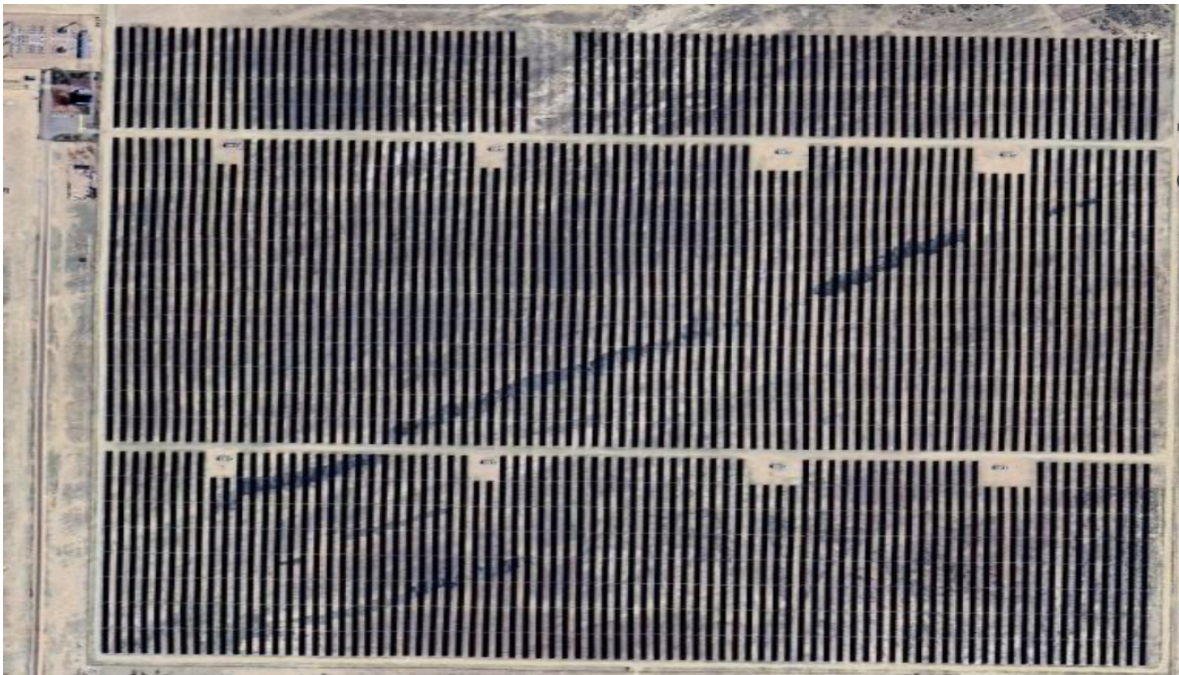


Figure 4.7 Top view of Balkhash Solar Power Plant

Summarizing 50 MW solar power plant analysis we reviewed overall characteristics, infrastructural layer, communication layer, and database structure. Using operational diagrams and connection structure the data structure model has been depicted in Figure 4.8. It simplifies the overall understanding and represents critical elements/hardware and their connections. Mainly, it connects operational elements such as panels, solar control units, combiner boxes, and transformers with communicational elements such as dispatcher rooms, general substation control rooms, distribution devices, and voltage transformer cabinets. Also, it explains how data

occurs on cloud-based servers. Consequently, Figure 4.8 is a middle-out approach combining both bottom-up and top-down reviews of a 50 MW solar power plant. Accordingly, the data structure model has been used in the next steps of the development of the digital twin framework.

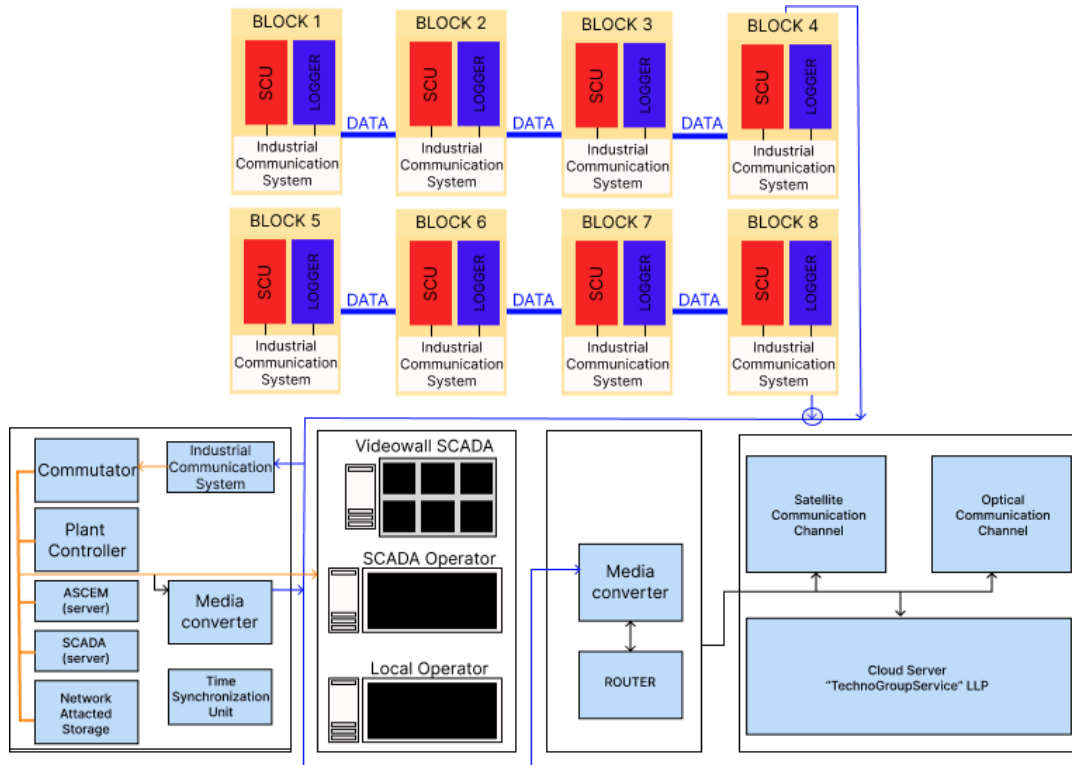


Figure 4.8 Data Structure Model

CHAPTER 5: FRAMEWORK DEVELOPMENT

5.1 Introduction

This chapter presents the developed detailed comprehensive framework for implementing Digital Twin in Solar Power Plants in Kazakhstan. The proposed framework would be utilized for the implementation of DT in a 50 MW Solar Power Plant near the Balkhash region. The framework was developed following the findings from the literature review and the analysis of the operational practices of a 50 MW Solar Power Plant. The synergy of these 2 domains (practical and theoretical) will help to achieve a balance between integrating experience from other research works and adapting to the realities of the Renewable Energy Sector in Kazakhstan. The framework covers all the principal layers of the proposed architecture with a detailed description of each of the layers, including the modeling and application layers and presents the wide range of DT utilization ways for divergent purposes.

5.2 Digital Twin Framework

The components of a Digital Twin for a Solar Power Plant are as follows:

A solar power plant's digital twin is a complicated system made up of multiple essential parts, each of which has a digital version that precisely replicates its behavior and physical attributes. Together, these linked parts allow for precise forecasts and simulations.

Solar panels are the first component that will be discussed. In terms of size, direction, and efficiency, a solar panel's digital twin is an exact copy of its physical form. Additionally, it incorporates solar intensity and angle of incidence into its simulations of behavior under diverse meteorological conditions. This makes it possible to forecast the panel's power output with accuracy at any given time.

Next come the **inverters**, which convert the direct current (DC) generated by the solar panels into alternating current (AC) that can be fed into the grid or stored in batteries. An inverter's digital twin models how well it works and how reliable it is, taking into account things like temperature and load. This makes it possible to make predictions about how well the inverter will work and find potential problems before they get worse.

The digital twin also incorporates the representation of the **grid connection**. This comprises the electrical lines, transformers, and other infrastructure used to link the solar power

plant to the grid. This infrastructure's electrical flow is simulated by the digital twin, which accounts for load, losses, and grid stability. As a result, the power output may be efficiently managed, guaranteeing a steady and dependable supply of electricity. The digital twin will not cover the grid connection in this paper.

Weather data plays a crucial role in the digital twin. This contains information on temperature, wind speed, sunshine intensity, and other relevant variables. These data are used by the digital twin to model how the solar power plant operates in different weather scenarios. This makes it possible to anticipate the plant's power output with accuracy and facilitates planning for times when there is less sunlight.

In summary, every element of the digital twin is essential to replicating the features and actions of its real-world power plant. Together, they make it possible for the digital twin to provide precise simulations and forecasts, which facilitates the solar power plant's effective management and operation. This demonstrates the enormous potential that digital twins have in the field of renewable energy.

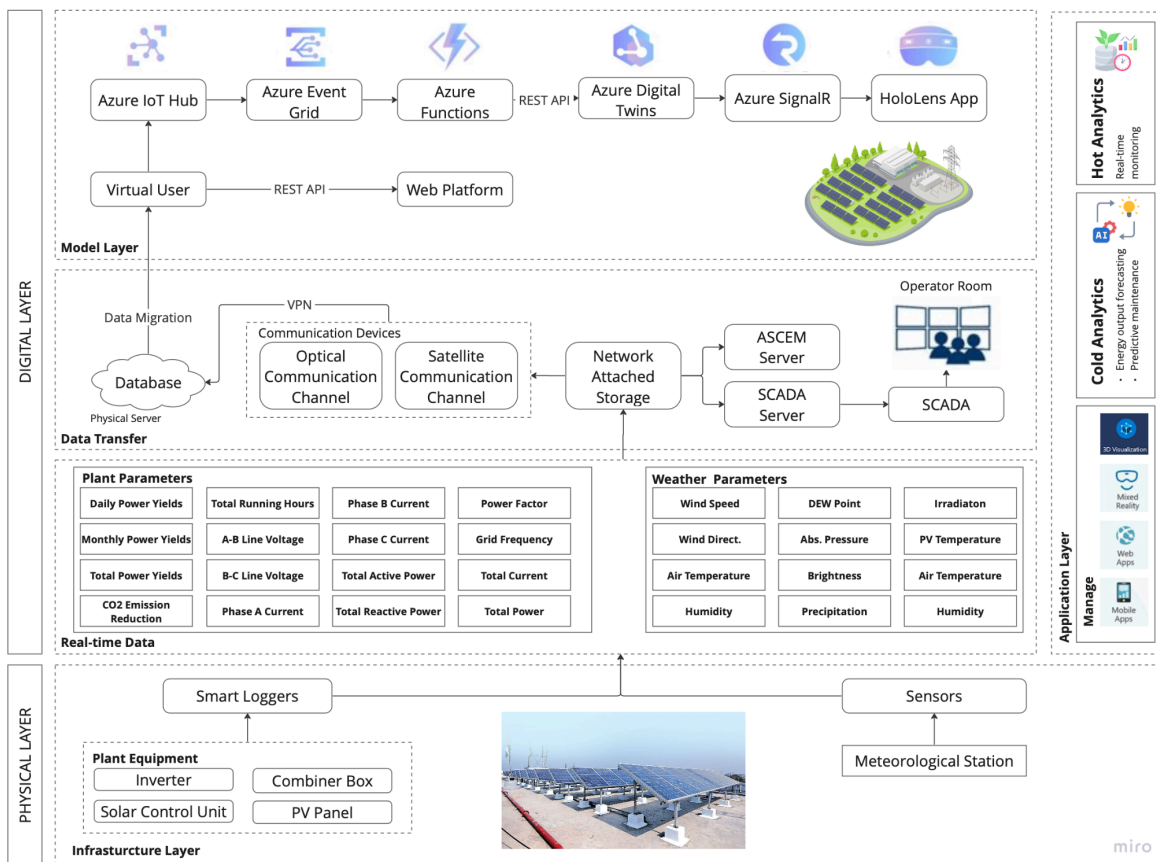


Figure 5.1. Framework for Digital Twin

5.2.1 Interactions between the Physical Plant and the Digital Twin

The fundamental element of the digital twin technology is the dynamic, bidirectional interaction between the actual solar power plant and its digital twin. A network of sensors integrated into the physical plant, in addition to sophisticated algorithms operating on the digital counterpart, facilitates this interaction.

This contact begins with data being transferred from the physical plant to the digital twin. The solar power plant's numerous components are equipped with sensors that are used to continuously monitor several factors mentioned in the previous part. The digital twin receives this data in real-time, generating a dynamic representation of the real plant in the virtual environment. The smart loggers collect information from the equipment and send it to the local network attached storage. Then data goes to the SCADA and ASCEM servers to be visualized in the operator room. In parallel, the data obtained from the solar power plant is transferred to the cloud database, called “LLP TechnoGroupService”, for further processing. To guarantee that it appropriately depicts the actual state of the physical plant, the digital twin updates its internal models and simulations using this data regularly.

The second phase of the interaction is the transfer of control commands from the digital twin to the actual plant. The digital twin can create control commands for many components of the physical plant using its simulations and forecasts. For instance, it might modify the solar panels' orientation to maximize solar absorption or control the battery's charge and discharge cycles to extend its life. In the physical plant, these control orders are carried out by actuators after being transmitted back. As a result, the performance and efficiency of the solar power plant can be optimized by the digital twin through proactive management of its operations. In this paper, this approach will not be covered as there are no actuators in the field.

The contact between the real plant and its digital twin is continuous, rather than one-time. The digital twin receives new data generated by the physical plant as it functions and changes. In response, the digital twin creates new control commands and updates its simulations. With time, the digital twin's models and forecasts will improve as a result of this ongoing feedback loop, which lets it learn from how the physical plant operates.

5.2.2 Model Layer

According to the proposed framework the modeling layer is used to visualize the data and model the application scenarios in a user-friendly way with high user interaction. The modeling layer is an environment for real-time monitoring and reactive maintenance. The modeling layer is also used for analytics to display the graphs, figures and tables for hot analytics. The non-direct application of a modeling layer is energy forecasting and predictive maintenance since these applications do not require visualization options and frontend functionality. To realize the modeling layer that will meet all the requirements for the digital twin there are several solutions to be employed.

The first solution is developing a Web Platform from scratch. The backend choice could be pretty straightforward since the proposed structure of the database provides space for flexibility and operability. However, the frontend development could be challenging since the modeling requires 2D visualization of the plant or geospatial data of the plant for convenient representation. To develop such microservices there could be different alternatives in terms of libraries to be used or general approaches to the visualization of the plant. Integration of geospatial data could be an easier way compared to the 2D visualization (scheme of the plant). The front end would also include analytics, which means the development of the graphs, tables, and key metrics for the plant's performance.

Another solution would be using open-source software, such as Azure Digital Twin as a SaaS (Software as a Service). This platform provides a ready-to-use fundament for applying digital twin solutions. However, the functionality and flexibility of such a solution could be limited to the capabilities of the platform. Additionally, the expertise to use these platforms could also be an obstacle.

The Mixed Reality (MR) application is developed using the Unity game engine, a powerful tool visualizing the solar power plant model in 3D. Additionally, the different tools of Microsoft Azure are used for data transfer. The data pipeline for the HoloLens application starts with cloud storage, which is subsequently sent to Azure IoT Hub for safe cloud connection. Azure Event Grid receives the incoming data and coordinates its flow, causing Azure Functions to begin processing. These functions are capable of carrying out a wide range of tasks, including data analysis and action triggers based on incoming information. Next, Azure Digital Twins uses the processing data to build a digital copy of the real world, which makes immersive experiences

possible. Azure SignalR enables real-time communication between the HoloLens and other clients or backend services for seamless updates and notifications. This integrated flow makes sure that the HoloLens application is always responsive, flexible, and able to provide users with rich, contextual experiences in real-time.

5.2.3 Application Layer

Hot Analytics

Real-time monitoring

The application layer of a DT includes the software, various algorithms, and analytics that use data collected from the physical plant. Combining and analyzing the data collected provides a comprehensive understanding of the plant's operation and may be applied to more complex analyses like predictive maintenance and power generation forecasting.. One of the main applications of DT is to evaluate performance in real-time due to the continuous flow of information from sensors and IoT devices located throughout the plant. Real-time monitoring, taking into account the landscape and plant scale, helps control the key parameters of each equipment and obtain information about the operating status of energy production. Moreover, DT empowers high data granularity that helps to obtain information from each panel, hence comprehensive monitoring and analysis of each element of the system became possible. Any deviations and anomalies in certain PV panels can be easily detected, hence facilitating reactive and proactive maintenance. Such an immediate approach positively affects the overall plant's performance as a significant decrease in downtime can be achieved.

Reactive maintenance

Reactive maintenance in the context of the DT framework involves addressing the plant's component failure or any issue after its occurrence. The biggest challenge at the current time is the huge plant scale that complicates failure identification and requires a thorough inspection of each component separately. Such an approach increases maintenance time and leads to long downtime. The DT approach facilitates reactive maintenance by providing instant data and equipment status indicating specific points of failure within the system. For instance, if a certain string experiences some voltage and current fluctuations the DT can alert operators to the problem, and an urgent maintenance crew can be dispatched to the specific failure point and diagnose the issue. In addition, DT offers access to extensive historical data that can be used to

analyze past failures and develop preventive strategies based on past maintenance activities. Reactive maintenance can be used as a primary tool for the analysis and introduction of DT for plant personnel.

Cold Analytics

Predictive maintenance

Predictive maintenance is an important part of the digital twin system for solar power plants. It saves money and makes operations more efficient. In contrast to reactive maintenance, predictive maintenance uses machine learning algorithms and data analytics to forecast and prevent equipment failures before their occurrence. Solar power plants use predictive maintenance to monitor equipment state and parameters like temperature, voltage, and current to identify any issues and predict possible failures.

Predictive maintenance reduces energy production downtime. By the use of historical data and equipment behavior, predictive maintenance algorithms can detect possible problems, and provide operators insights regarding the maintenance. Such a proactive method prevents unwilling downtime and keeps the solar power plant running, increasing energy generation, operational efficiency, and profitability. Maintenance activities can be prioritized based on asset conditions rather than timetables to maximize resource allocation and eliminate superfluous chores. Maintaining equipment only when needed minimizes labor and material expenses and increases its lifespan.

Predictive maintenance in the digital twin framework requires advanced data analytics and machine learning capabilities. Operators must collect and analyze vast amounts of data from inverters, transformers, and weather stations. Finally, machine learning algorithms analyze data patterns and trends to design and construct prediction models.

Artificial Intelligence

AI systems can optimize many plant activities by analyzing massive volumes of data gathered from sensors, weather forecasts, previous performance records, and other sources. ML algorithms can predict energy-generating patterns based on weather forecasts and historical data, allowing operators to optimize plant operations for energy output and grid integration.

There are a few different ways that can be used for AI integration into the digital twin application layer. One typical method is to create ML models that continuously learn and adjust based on real-time data gathered from the solar plant. These models can identify irregularities,

forecast malfunctions in machinery, and enhance efficiency by adapting to fluctuating environmental circumstances. Another option is to integrate AI algorithms into monitoring and control systems to automate decision-making such as panel layout adjustments and energy storage system management based on demand.

Integration of AI into the digital twin framework promises several benefits, but there are obstacles to overcome. Reliable and accurate sensors and monitoring data are essential for AI-driven applications. When it comes to training and validating AI models, one of the most significant obstacles is the availability and quality of the data that is necessary. Artificial intelligence in plant management may raise regulatory and privacy problems, especially when collecting and processing sensitive data, as the energy sector is a strategically important infrastructure.

Additional Functionality

Aerial Imagery

Digital Twin allows the integration of different tools and the integration of drones and cameras of different spectrums is no exception. Aerial imagery can be a powerful tool for reactive and predictive maintenance as it provides a wide range of data from the identification of dirt or vegetation on the panels to the locating hotspots that severely affect PV panel lifespan. The incorporation of aerial imagery especially with drones flying around can assist the maintenance crew by capturing any changes in plant equipment, identifying PV panel contamination, and inspecting for potential defects or damages on solar panels. Apart from maintenance, high-resolution cameras can be used to evaluate meteorological factors, cloud cover and atmospheric conditions to predict power generation. For instance, based on cloud coverage the tilt angle can be adjusted in advance and utilize its full potential to generate power.

Infrared camera inspection

The installation of cameras with thermal vision can greatly assist in detecting thermal anomalies that cannot be visible to the naked eye. The solar power plant is exposed to external and internal heating, such severe operating conditions can lead to the degradation of key components. The utilization of infrared cameras provides an opportunity to detect thermal anomalies, overheating, loose connection, or faulty cells and identify any areas of concern that seriously affect the lifespan of the equipment. Integration of thermal cameras to a digital twin system enhances monitoring and maintenance capabilities due to real-time insights. Captured

infrared images then can be processed by AI and compared against historical pictures and performance benchmarks to identify certain correlations and plan future maintenance activities.

5.4 Justification for the Design Choices in the Framework

The suggested framework for a digital twin of a solar power plant is designed to maximize three critical attributes: accuracy, adaptability, and scalability. Each of these features is very important to how well and how often the digital twin works.

The accuracy of a digital twin ensures that the virtual model matches reality. To achieve great accuracy, an advanced real-time data transmission mechanism is used to mimic every solar plant component. This allows the digital twin to correctly depict the actual plant's current state, and the data can be used for making predictions.

Flexibility is another essential quality that lets the digital twin adapt to changes in the physical plant or its operating environment. The suggested framework can be easily changed because it is made up of separate modules. The digital twin architecture allows its layers to function either separately or in tandem. This system's built-in flexibility makes it easy to make changes and updates. For example, only the digital twin of the solar panel needs to be updated if a new kind of solar panel is installed in the physical plant; the rest of the framework stays the same.

Scalability is the third important attribute. It lets the digital twin handle more complex and large solar power plants. With its distributed architecture and efficient data management mechanisms, the proposed framework can easily handle large datasets. The suggested architecture stores field data in the cloud database and distributes the computational load across several processors or servers to handle large data volumes and complex simulations. The digital twin's efficient data management system allows it to store and interpret large volumes of sensor data from the physical plant without seeing a drop in performance.

In summary, the proposed framework for the digital twin has the precision, adaptability, scalability, real-time responsiveness, and prediction accuracy of the digital twin. These characteristics make the proposed framework an efficient tool for monitoring and optimizing the performance of solar power plants.

CHAPTER 6: IMPLEMENTATION OF DT AND PROTOTYPE DEVELOPMENT

6.1 Introduction

After developing the framework for implementing Digital Twins at solar power plants, the next logical step is to translate all the insights at an operating 50 MW solar power plant in the Balkhash region. To test the quality of the proposed framework and apply all the developed architecture in real-life conditions the prototype of DT with real-time data transfer was developed for a 50 MW power plant.

6.2 Digital Twin Implementation Process Flow

Integration of digital twins into existing solar power plants consists of several major stages and should follow a certain sequence to successfully iterate all the aspects into one working system. Therefore, during this project, the digital twin implementation process flow has been created to outline important stages with specific tasks to accomplish. In addition, the estimated time for each stage was identified based on the company's experience and competencies. Figure 6.1 shows the detailed digital twin creation process consisting of 11 major stages from initiation to evaluation and future scalability stages.

After initiation, the project preparation phase begins with negotiations with TGS Company to set the objective and define the scope of the desired project. The duration of this phase is estimated to be 3 weeks as some time is needed to conduct a preliminary literature review and assess the company's capability. This phase is followed by the Infrastructure Audit and Data Collection and Analysis phases which are performed simultaneously. The Infrastructure Audit phase evaluates key plant components, analyzes equipment and communication structures to obtain detailed information for further integration with digital twin platforms. Since all documentation is held within the company this stage's duration is estimated at 2 weeks. Meanwhile, the Data Collection and Analysis phase focuses on the assessment of current data collection processes and investigation requirements for digital twin integration. In addition, due to rapid technological advancement in this sphere, it is essential to update information regarding the sensors and IoT devices. Therefore, after analyzing all the requirements for data communication and its components, an improvement plan is provided as necessary.

The next stages Risk Assessment & Mitigation Planning mainly focus on risk identification and developing mitigation strategies, meanwhile, Regulatory & Standards Compliance ensures the alignment with standards and regulations of the energy sector of Kazakhstan. The subsequent phase is Technical Requirements and Design which consists of sub-stages such as System Architecture Design, Sensors and IoT Integration Plan and Software and Platform Selection. This phase builds a foundation for future digital twin models as the DT architecture design, critical parameters identification, sensor deployment and DT platform evaluation are performed.

After this, the Development and Testing phase continues the progress of the project. Sub-stages of this phase include digital twin development consisting of backend development for the web platform, database integration and frontend development that provides DATA visualization. The developed platform then goes through prototype testing and improvement after receiving feedback. One of the final phases is Implementation and Integration where the DT platform is deployed and integrated with existing plant facilities and control platforms. After a successful deployment of the platform, the DT can be used for real-time monitoring of the plant's operations, and provide valuable data for reactive maintenance by observing any anomalies in the power generation. To improve the performance of DT the plant's staff and operating engineers should be familiarized and trained to use DT and evaluate its work for further development.

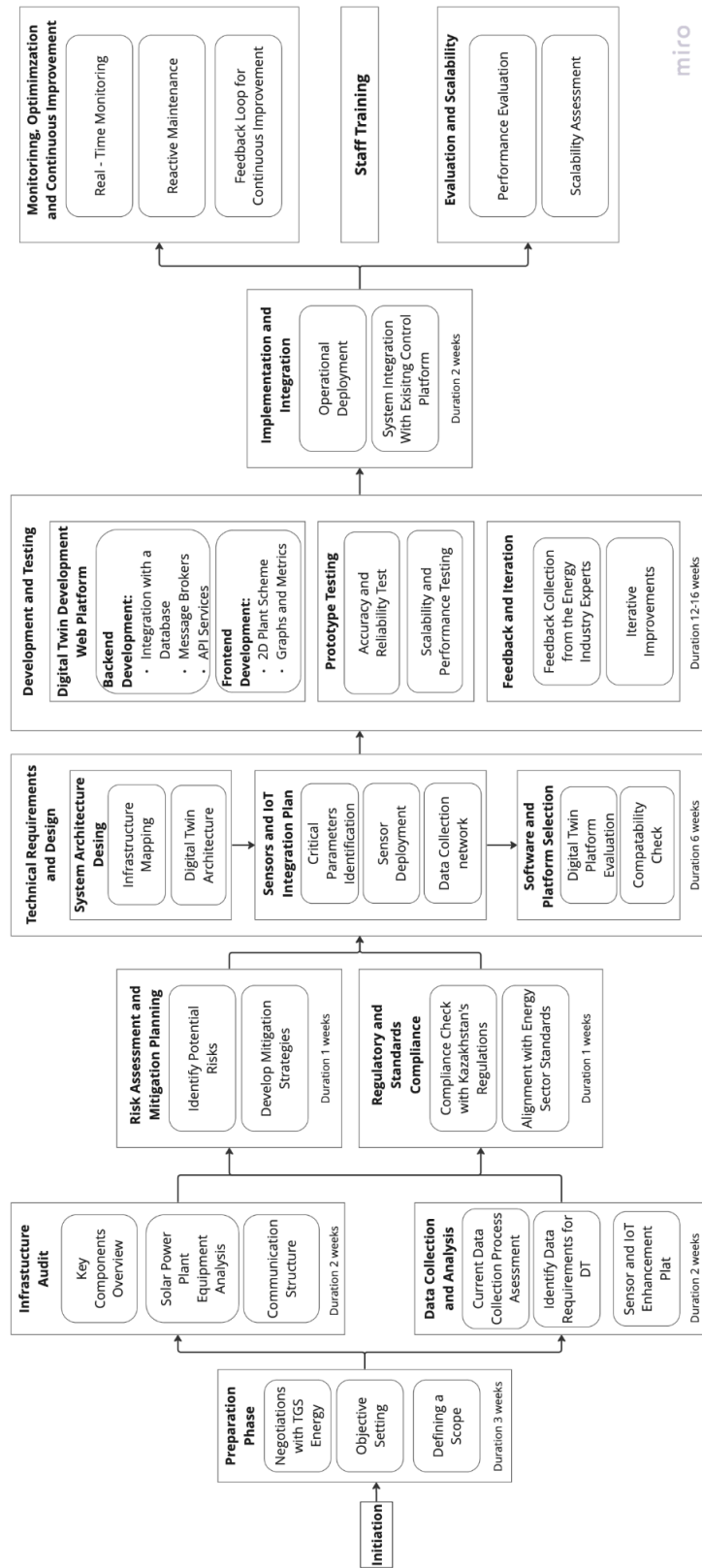


Figure 6.1 Digital Twin Implementation Process Flow

6.3 UML Diagram

Following the D-Arch technology stack, the existing solar power plant system has been taken apart and explained by the use of Unified Modelling Language (UML) (Gourisetti, 2023) in simple ways. Section embarks an explanatory note on informational data flow and Digital Twin Framework that lays as an essential requirement to start the development of the user application layer.

The Unified Modeling Language (UML) Class diagram standard served as the model for information flow and object model diagrams used in this section. Basically, in software development processes, UML class diagrams are frequently utilized to display systems, processes, and architectures. Since the digital twin is a composite technology with software and equipment development, the UML class diagram can be useful for further explanatory processes. Class diagrams can be used for the depiction of physical objects, buildings, and systems or software-related objects (Gourisetti, 2023). Usually, rectangular figures in UML indicate developing/maintaining objects in programming, but in this case, rectangles are used for hardware parts, systems, and subsystems. Regularly class objects consist of 3 parts:

1. Building block/object name
2. Potential parameters and data types
3. Operations executed by block/object

Figure 6.2 has been produced to illustrate the (1) main infrastructure elements, (2) peripheral elements as sensors and actuators, (3) communication layer objects, (4) systems for visualization, (5) interactions within the system, and (6) systems and subsystems. UML is a language for software developers to understand and explain how applications work. It implies what the system needs to do before building the digital system. Therefore, UML is useful for the deployment of system architecture and framework in further sections.

Following Figure 4.8 the physical dimension of the system has been constructed and supplemented by a digital dimension that can be seen in Figure 6.1 *Note: only critical physical objects have been depicted.* The system itself starts from solar panels that are connected to the solar control unit on each existing 8 blocks. Also, each block has its own weather station to gather the data.

The solar control unit collects electrical parameters, environmental factors, and possibly other relevant data. Meteo stations, on the other hand, collect meteorological and environmental

parameters. Parameters collected from the Solar Control unit and Meteorological station are inherited by the combiner box whose primary function is to connect strings, measure parameters, merge output current, and ensure overcurrent protection. Those collected data move further to Logger 3000, or a data logger that monitors and manages the data coming from the combiner box/photovoltaic system. All the data respectively spread to industrial communication systems or MOXA that in sequence transform data to SCADA and other communication protocols. When data can be found on cloud servers, the physical dimension is replaced by digital. There Azure related applications are made and in parallel web applications could be developed.

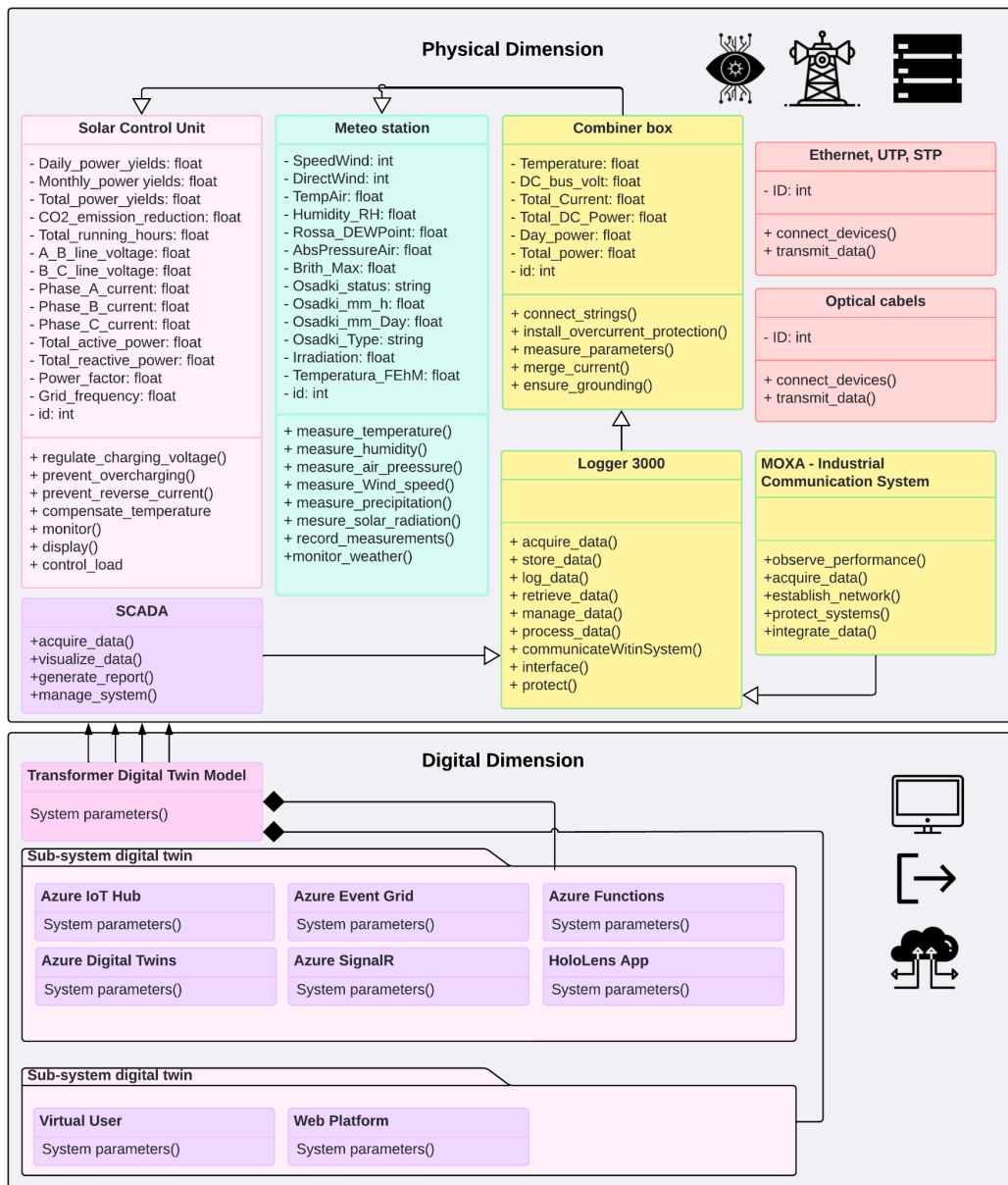


Figure 6.2 D-Arc description for the case and UML sequence flow process using a transformer

6.4 Web platform - Minimum Viable Product (Prototype)

One of the central components of the digital twin is a modeling layer. As it is highlighted in the proposed architecture, on the modeling layer the visualization of the solar power plant for real-time monitoring should be performed. Since TGS Energy has already developed its web platform for the client to demonstrate the required analytics (production graphs, weather parameters, key metrics) it was decided to add a module for real-time monitoring of the plant.

The current platform was written on Python Flask as a backend choice and React as a frontend choice. The architecture of the platform is designed in a way that the microservices can be easily integrated into the platform. However, a significant challenge is the database used in the current platform and a database developed for a digital twin. These 2 databases are completely different in structure, since the database for a Digital Twin employs the combined Structures 1 and 2, while the database of the current platform is closer to the Structure 2 Type.

From the perspective of the front-end development adding a module of real-time monitoring is a complex task compared to the backend development. On the front, the 2D visualization of the plant along with positioning all the components on this visualization and displaying the parameters of each component in a user-friendly way should be performed. For the backend the integration of the services between the database and the frontend, the external services (weather forecasts using API) and all the analytics demonstrated on the front should be performed.

The architecture of the developed prototype is the following:

- FastAPI framework for backend development (microservices) written in Python
- PostgreSQL database for data storage and querying
- React with Google Maps API for frontend development

For the demo prototype development (MVP) the limited functionality and preliminary design were performed. On the frontend side, the integration of Google Maps API for displaying the top view of the power plant was carried out. On top of the map of the power plant around 185-200 unique data points were located. For the demo version of the prototype, it was decided to include the data from the blocks and combiner boxes along with the meteorological data. The data coming from each of the strings should be included in the fully functional platform during the later stages of the development. At that point, the expected number of data points is around 7000. The figures below are screenshots of the demo version of the prototype. The prototype is a

fully functional module with real-time (1-minute sampling period) data transfer from all the combiner boxes and blocks of the power plant. In figure 6.3 the average readings of the whole station (general statistics) are depicted. Some of the values are total sum values of the components (In Figure 6.3 the updated meteo parameters are demonstrated. In Figure 6.4 the information about Combiner Box #18 from the Block 4 Inverter 1 is displayed. All the presented data is real generated data with a 1-minute update frequency.

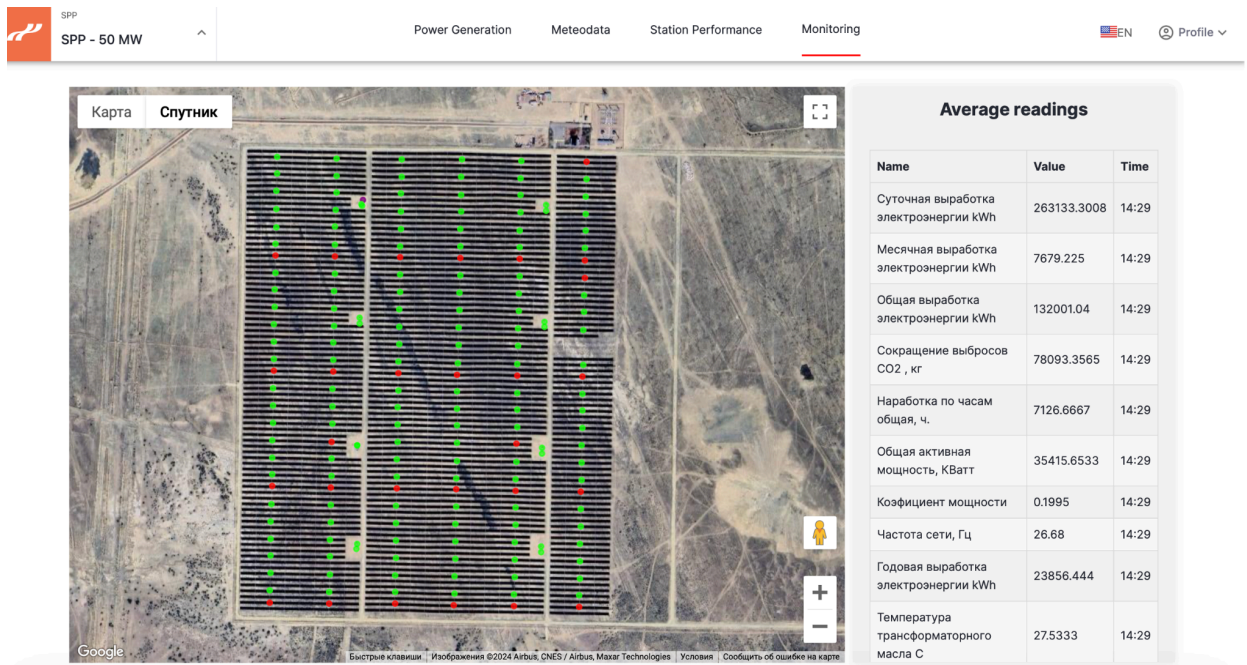


Figure 6.3 The Prototype Demo for a 50 MW Solar Power Plant (Average Readings)

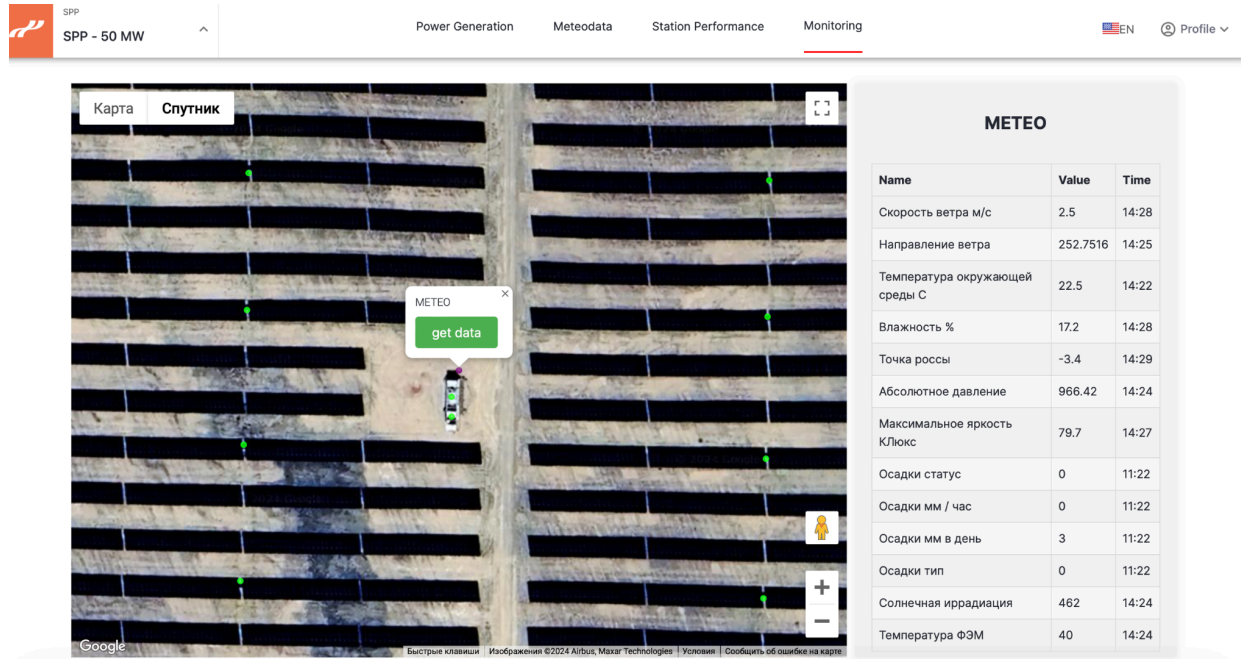


Figure 6.4 The Prototype Demo for a 50 MW Solar Power Plant (Meteo Data)

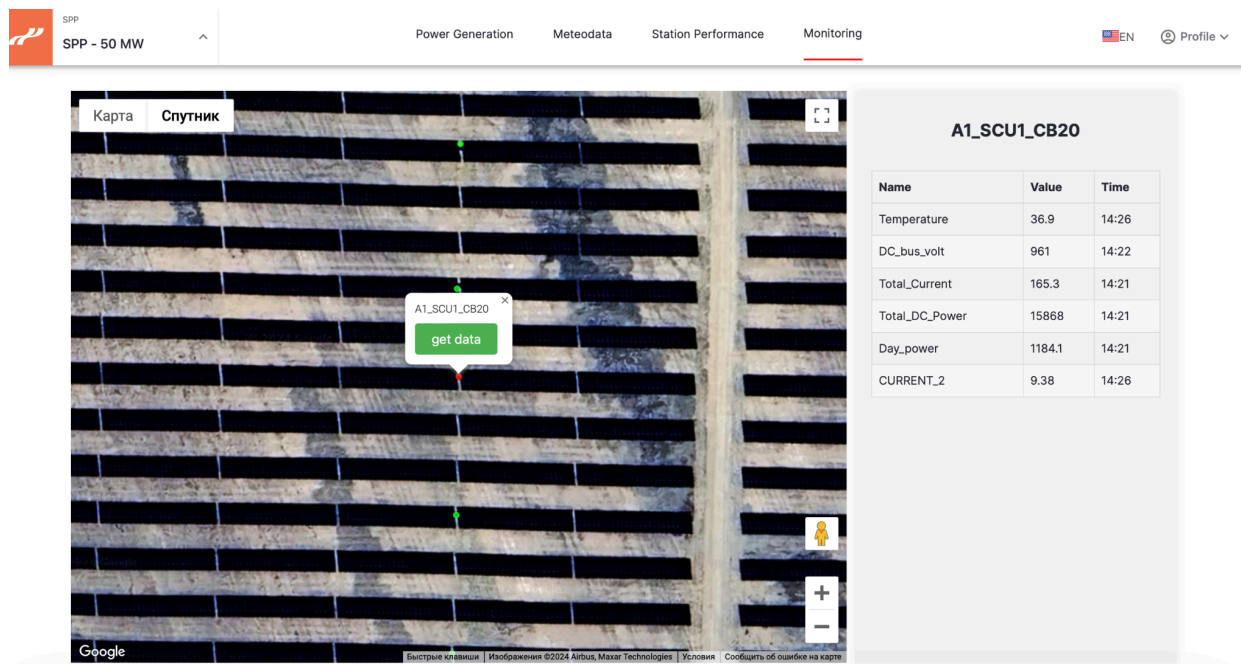


Figure 6.5 The Prototype Demo for a 50 MW Solar Power Plant (Block 4 Combiner Box 20 Data)

For the further development of the prototype, the following steps are planned:

1. Integration of the monitoring module into the existing web platform
2. Increase the number of data points to be displayed on the platform
3. Development of the microservices for an analytics module (production graphs, weather parameters graphs, key metrics)

6.5 Prototype Validation and Stakeholders Feedback

Following the development of the prototype, the presentation of the demo version to the stakeholders of the project was organized by the team. The stakeholders included the representatives of TGS Energy and Nazarbayev University committee. The presentation consisted of an overview of the prototype, the architecture choice (backend and frontend) description and a product presentation. The gathered feedback will be used for incremental improvements to the product in later development stages.

The expert's feedback and proposed improvements were developed and depicted on Table 6.1.

Table 6.1. Experts' validation and proposed improvements

Object under validation	Expert	Expert's position	Feedback	Improvement for the project
Prototype	Expert 1	Top Manager of TGS Energy	The interface of the real-time monitoring module could be improved to make it more attractive	Add general information about the power plant's performance (total power) on the dashboard without pressing any buttons by default
Prototype	Expert 2	Engineer on the plant	It is necessary to add messages in different scenarios and emergency situations. The functionality of the	Implement the notifications based on several key parameters (total DC Power, temperature of the oil) to

			platform should be more interactive with the end user and provide various assistance (alerts, notifications, visual representations) to help the user to find out any issues as soon as possible	timely notify the end user about the potential issue.
Architecture of the Platform	Expert 3	IT specialist	The architecture of the platform should be flexible enough to withstand the adoption of more than one power plant (scalability of the product) since the digital twin technology could be implemented in other power plants using the same platform	The architecture of the platform will consist of many small microservices each performing a different function and using different data.
Prototype	Academic Supervisor 1	Professor, School of Engineering and Digital Sciences	Not all the parameters are translated to English and Kazakh by design but it would be more convenient for international partners and institutions to translate all the parameters to English and Kazakh	All parameters will be translated into English and Kazakh.
Prototype	Academic Supervisor 2	Professor, School of Engineering	Time variation parameters should be added. The current status of power	The status of the combiner box (its color) will be determined by

		and Digital Sciences	generation can be depicted by different colors.	comparing the value of the total power of the combiner box with the average power output of all combiner boxes
Prototype	Academic Supervisor 3	Professor, School of Engineering and Digital Sciences	All parameters should have the corresponding units in the description	All parameters will be shown with corresponding units. Also, all parameters will have 4 decimal places after the decimal point.

CHAPTER 7: CONCLUSION AND FUTURE WORK

7.1 Introduction

In the continuing pursuit of green energy, renewable energy sources have become a light of hope. Kazakhstan, with its huge potential for solar energy resources, is a leader in Central Asia. However, the generation of solar energy requires constant monitoring and highly efficient operation management. This is where the concept of digital twins steps in and transforms the way power plants work. This project examined the current state of the art in the field of the development of digital twins and their applications. During this study, we have thoroughly explored this technology and created a prototype of a digital twin for a 50 MW solar power plant in the Balkhash region.

7.2 Conclusions

In conclusion, the framework proposed for the implementation of DT technology adapted to the unique characteristics of Renewable Energy Sources in Kazakhstan demonstrates significant potential. With its adaptability to the renewable energy landscape in solar plants, this framework serves as an effective tool for enhancing operational efficiency and performance optimization. Moreover, the development of a Web Platform emerges as a crucial component, offering a user-friendly interface for real-time monitoring and reactive maintenance strategies. This visualization tool not only facilitates immediate decision-making but also enables stakeholders to flexibly respond to emerging issues, thereby minimizing downtime and maximizing productivity. Furthermore, the incorporation of both cold and hot analytics within the proposed DT framework presents a robust approach towards predictive maintenance and energy forecasting. By leveraging historical data and real-time insights, organizations can anticipate potential faults, mitigate risks, and optimize resource utilization, thereby enhancing overall system reliability and performance.

The project successfully met all the established objectives and delivered main outputs in the form of a DT implementation framework, detailed implementation roadmap and prototype for a 50 MW solar power plant. In addition, this work thoroughly described the operational procedures of a 50MW solar power plant and indicated the main aspects in terms of equipment,

sensors and IoT devices to consider for DT implementation.

7.3 Limitations of the project

This project lacks a description of the cybersecurity aspect of DT. However, cybersecurity is an essential part of DT in terms of data privacy. Integration of cybersecurity measures ensures the safety of sensitive data and prevents unauthorized access and data manipulation. By overlooking cybersecurity, this project does not focus on potential risks associated with unauthorized access and data leakage.

During this project, the main focus was only on one solar plant that was chosen as a benchmark among other solar facilities in Kazakhstan. Such a narrow focus cannot guarantee full compatibility and integration with different existing plants due to a lack of standardization in design. RES is a wide sector to implement digital twin modeling. This case scenario covers only one existing solar power plant. Consequently, there is the possibility that the digital twin framework could not be implemented with other solar power plants due to hardware and software differences. Also, the digital twin framework made for this scenario does not cover wind, water, and geothermal renewable energy sources indicating no transformability.

7.4 Future Work

For future work, the following set of actions is scheduled in a time horizon of one year. In terms of the prototype development, the further development of the web platform by integrating the monitoring module into the existing platform and adding analytics tools should be performed in 2 months. One of the most prominent development stages of the web platform is the Integration of AI tools for cold analytics. It is planned to implement Machine Learning Algorithms for energy output predictions and predictive maintenance. Additionally, in a longer time horizon, the interesting integration example will be Drone Imagery. Applying Drone Imagery to generate more real-time data for plant inspection is scheduled to be performed in a 6-7 months period.

For the further development of the proposed framework, a comparative analysis with other existing DT projects for solar power plants should be performed. However, such projects could be in the early development stages as the DT concept is considered an emerging

technology. Performing the benchmarking analysis would help to assess the performance of the development DT and its applicability, reusability and replicability.

In terms of broadening the application of DT, the visionary step will be developing DT for Wind Power Plants. Moving beyond the scope of the solar power plant will substantially widen the applicability of DT in the Renewable Energy Sector.

References

- A. Livera *et al.* (2022). Intelligent Cloud-Based Monitoring and Control Digital Twin for Photovoltaic Power Plants. *2022 IEEE 49th Photovoltaics Specialists Conference (PVSC)*. Philadelphia, PA, USA, 2022, pp. 0267-0274, doi: 10.1109/PVSC48317.2022.9938505.
- Aheleroff, S., Xu, X., Zhong, R. Y., & Lu, Y. (2021). Digital Twin as a service (dtaas) in industry 4.0: An architecture reference model. *Advanced Engineering Informatics*, 47, 101225. <https://doi.org/10.1016/j.aei.2020.101225>
- Akhmetkaliyeva, S. (2020). A promising green energy resource in Kazakhstan: Solar Power. ERI. <https://www.eurasian-research.org/publication/a-promising-green-energy-resource-in-kazakhstan-solar-power/>
- Andrade, A. (2023, January 26). The 3 levels of the Digital Twin Technology. *Vidya*. <https://vidyatec.com/blog/the-3-levels-of-the-digital-twin-technology-2/>
- Cimino, C., Negri, E., & Fumagalli, L. (2019). Review of digital twin applications in manufacturing. *Computers in industry*, 113, 103130.
- Gourisetti, S. N., Bhadra, S., Sebastian-Cardenas, D. J., Touhiduzzaman, M., & Ahmed, O. (2023). A theoretical open architecture framework and Technology Stack for Digital Twins in Energy Sector Applications. *Energies*, 16(13), 4853. <https://doi.org/10.3390/en16134853>
- Grieves, M. (2014). Digital twin: manufacturing excellence through virtual factory replication. *White paper*, 1-7.
- Heluany, J., & Vasileios, G., (2023). A review on digital twins for power generation and distribution. *International Journal of Information Security*. 23. 10.1007/s10207-023-00784-x.
- He, X., Ai, Q., Qiu, R. C., & Zhang, D. (2019, September 16). Preliminary exploration on digital twin for Power Systems: Challenges, framework, and applications. *arXiv.org*. <https://arxiv.org/abs/1909.06977>

- Huang, J., Zhao, L., Wei, F., & Cao, B. (2021). The application of digital twin in the power industry. *IOP Conference Series: Earth and Environmental Science*, 647(1), 012015. <https://doi.org/10.1088/1755-1315/647/1/012015>
- Ismail, F. B., Al-Faiz, H., Hasini, H., Al-Bazi, A., & Kazem, H. A. (2024). A comprehensive review of the dynamic applications of the digital twin technology across diverse energy sectors. *Energy Strategy Reviews*, 52, 101334. <https://doi.org/10.1016/j.esr.2024.101334>
- J.V.S. do Amaral a, & Falekas, G. (2023, October 16). Energy Digital Twin Applications: A Review. *Renewable and Sustainable Energy Reviews*. <https://doi.org/10.1016/j.rser.2023.113891>
- Jain, P., Poon, J., & Spanos, C. (2020, January). A Digital Twin Approach for Fault Diagnosis in Distributed Photovoltaic Systems. <https://ieeexplore.ieee.org/document/8692656/>
- Jiang, Z., Lv, H., Li, Y. et al. (2022). A novel application architecture of digital twins in smart grid. *J Ambient Intell Human Comput* 13, 3819–3835. <https://doi.org/10.1007/s12652-021-03329-z>
- Khaled Osmani, Ahmad, N. I., Bahaidarah, H. M. S., & Hammad, B. (2020, August 22). A review on maintenance strategies for PV Systems. *Science of The Total Environment*. <https://doi.org/10.1016/j.scitotenv.2020.141753>
- Kritzinger, W., Karner, M., Traar, G., Henjes, J., & Sihn, W. (2018). Digital Twin in manufacturing: A categorical literature review and classification. *Ifac-PapersOnline*, 51(11), 1016-1022.
- Kumari, N., Sharma, A., Tran, B., Chilamkurti, N., & Alahakoon, D. (2023). A comprehensive review of Digital Twin Technology for grid-connected microgrid systems: State of the art, potential and challenges faced. *Energies*, 16(14), 5525. <https://doi.org/10.3390/en16145525>
- Li, L., Aslam, S., Wileman, A., & Perinpanayagam, S. (2021). Digital twin in the aerospace industry: A gentle introduction. *IEEE Access*, 10, 9543-9562.
- Lim, N. (2021). Renewable Energy Market in Kazakhstan: Potential, Challenges, and Prospects. <https://www.pwc.com/kz/en/publications/esg/may-2021-eng.pdf>

- Liu, M., Fang, S., Dong, H., & Xu, C. (2021). Review of digital twin about concepts, technologies, and Industrial Applications. *Journal of Manufacturing Systems*, 58, 346–361. <https://doi.org/10.1016/j.jmsy.2020.06.017>
- M. Jafari, A. Kavousi-Fard, T. Chen and M. Karimi. (2023). A Review on Digital Twin Technology in Smart Grid, Transportation System and Smart City: Challenges and Future. *IEEE Access*. vol. 11, pp. 17471-17484. <https://doi.org/10.1109/ACCESS.2023.3241588>.
- Pruthviraj, U., Kashyap, Y., Baxevanaki, E., & Kosmopoulos, P. (2023, April 2). Solar photovoltaic hotspot inspection using unmanned aerial vehicle thermal images at a solar field in South India. *MDPI*. <https://www.mdpi.com/2072-4292/15/7/1914>
- Semeraro, C., Lezoche, M., Panetto, H., & Dassisti, M. (2021). Digital twin paradigm: A systematic literature review. *Computers in Industry*, 130, 103469.
- Shao, G. (2021). Use Case Scenarios for Digital Twin Implementation Based on ISO 23247. <https://doi.org/10.6028/nist.ams.400-2>
- Shao, G., Frechette, S., & Srinivasan, V. (2023). An analysis of the new ISO 23247 series of standards on Digital Twin Framework for manufacturing. Volume 2: Manufacturing Equipment and Automation; Manufacturing Processes; Manufacturing Systems; Nano/Micro/Meso Manufacturing; *Quality and Reliability*. <https://doi.org/10.1115/msec2023-101127>
- Sleiti, A., Kapat, Jayanta, & Vesely, L. (2022, March 16). Digital Twin in the energy industry: Proposed robust digital twin for power plant and other complex capital-intensive large engineering systems. *Energy Reports*. <https://www.sciencedirect.com/science/article/pii/S2352484722005522>
- Steindl, G., Stagl, M., Kasper, L., Kastner, W., & Hofmann, R. (2020, December 13). Generic Digital Twin Architecture for Industrial Energy Systems. *MDPI*. <https://www.mdpi.com/2076-3417/10/24/8903>
- Vlaminck, M., Heidbuchel, R., Philips, W., & Luong, H. (2022, February 7). Region-based CNN for anomaly detection in PV power plants using aerial imagery. *MDPI*. <https://www.mdpi.com/1424-8220/22/3/1244>

Yalçın, T., Paradell Solà, P., Stefanidou-Voziki, P., Domínguez-García, J. L., & Demirdelen, T. (2023, June 29). Exploiting digitalization of solar PV plants using machine learning: Digital Twin concept for Operation. *MDPI*. <https://www.mdpi.com/1996-1073/16/13/5044>

Ye, Q., Liu, S., & Liu, C. (2020, August 1). A deep learning model for fault diagnosis with a deep neural network and feature fusion on multi-channel sensory signals. *MDPI*. <https://www.mdpi.com/1424-8220/20/15/4300>

Yei, W., & Panos, P. (2022, April 2). Energy Digital Twin Technology for Industrial Energy Management: Classification, challenges and future. *Renewable and Sustainable Energy Reviews*. <https://www.sciencedirect.com/science/article/pii/S136403212200315>

Zhou, M., Yan J., & Feng D. (2019). Digital Twin and its application to power grid online analysis. *CSEE Journal of Power and Energy Systems*. <https://doi.org/10.17775/cseejpes.2018.01460>

Appendix List

Appendix A

Table A1. Communication Equipment Description

Type of trench	Trace			Cable	
	Start	End	Section of cable laying	Brand	Length, m
Type1-Type 2	Cabinet BMZ optical cross XS cabinet SHAS-1	Block-1. optical cross XS3	in the pipe in the trench	OKKL-8(6-652D -2.7kN)	600
Type 1	Block-1. optical cross XS3	Block-3. optical cross XS4	in the pipe in the trench	OKKL-8(6-652D -2.7kN)	317
Type 1	Block-3. optical cross XS4	Block-5. optical cross XS5	in the pipe in the trench	OKKL-8(6-652D -2.7kN)	315
Type 1	Block-5. optical cross XS5	Block-7. optical cross XS6	in the pipe in the trench	OKKL-8(6-652D -2.7kN)	285
Type 1	Block-7. optical cross XS6	Block-8. optical cross XS7	in the pipe in the trench	OKKL-8(6-652D -2.7kN)	398
Type 1	Block-8. optical cross XS7	Block-6. optical cross XS8	in the pipe in the trench	OKKL-8(6-652D -2.7kN)	381
Type 1	Block-6. optical cross XS8	Block-4. optical cross XS9	in the pipe in the trench	OKKL-8(6-652D -2.7kN)	212
Type 1	Block-4. optical cross XS9	Block-2. optical cross XS10	in the pipe in the trench	OKKL-8(6-652D -2.7kN)	275
Type 1	Block-2. optical cross XS10	Cabinet BMZ optical cross XS	in the pipe in the trench	OKKL-8(6-652D -2.7kN)	227

		cabinet SHAS-1			
Type 2	Cabinet BMZ optical cross XS cabinet SHAS-1	Substation 220kV. OPU building optical cross XS11. Communication cabinet	in the pipe in the trench	OKKL-8(6-652D -2.7kN)	110

Appendix B

id	t	v	q
1	2024-03-04 18:13:45.827	599.599975585938	0
2	2024-03-04 18:13:44.367	600.5	0
3	2024-03-04 18:14:03.089	599.599975585938	0
4	2024-03-04 18:14:32.351	601.5	0
5	2024-03-04 18:13:36.291	599.599975585938	16
6	2024-03-04 18:13:36.291	599.599975585938	16
7	2024-03-04 18:13:36.291	599.599975585938	16
8	2024-03-04 18:13:36.291	599.599975585938	16
9	2024-03-04 18:13:42.575	600.299987792969	0
10	2024-03-04 18:13:44.367	600.5	0
11	2024-03-04 18:13:45.827	599.599975585938	0
12	2024-03-04 18:13:47.510	600.099975585938	0
13	2024-03-04 18:13:48.990	600	0
14	2024-03-04 18:13:50.536	600.299987792969	0
15	2024-03-04 18:13:52.077	600	0
16	2024-03-04 18:13:53.453	600.099975585938	0
17	2024-03-04 18:13:54.931	600.200012207031	0
18	2024-03-04 18:13:56.411	600.099975585938	0
19	2024-03-04 18:13:58.269	600.299987792969	0
20	2024-03-04 18:13:59.660	600.200012207031	0
21	2024-03-04 18:14:01.575	599.799987792969	0
22	2024-03-04 18:14:03.089	599.599975585938	0
23	2024-03-04 18:14:04.702	600.099975585938	0
24	2024-03-04 18:14:06.210	600.299987792969	0
25	2024-03-04 18:14:07.848	599.799987792969	0
26	2024-03-04 18:14:09.497	600.700012207031	0
27	2024-03-04 18:14:11.218	601.200012207031	0
28	2024-03-04 18:14:12.603	601	0

Figure B1. "Trends" table

id	name	description	latitude	longitude
1	30297 SpeedWind	Скорость ветра м/с	<null>	<null>
2	30298 DirectWind	Направление ветра	<null>	<null>
3	30299 TempAir	Температура окружающей среды С	<null>	<null>
4	30300 Humidity_RH_	Влажность %	<null>	<null>
5	30301 Rossa_DEWPoint_	Точка россы	<null>	<null>
6	30302 AbsPressureAir	Абсолютное давление	<null>	<null>
7	30303 Brith_Max	Максимальное яркость КЛюкс	<null>	<null>
8	30304 Osadki_status	Осадки статус	<null>	<null>
9	30305 Osadki_mm_h	Осадки мм / час	<null>	<null>
10	30306 Osadki_mm_Day	Осадки мм в день	<null>	<null>
11	30307 Osadki_Type	Осадки тип ()	<null>	<null>
12	30308 Irradiation	Солнечная иррадиация	<null>	<null>
13	24094 METEO_SpeedWind	Скорость ветра м/с	<null>	<null>
14	24095 METEO_DirectWind	Направление ветра	<null>	<null>
15	24096 METEO_TempAir	Температура окружающей среды С	<null>	<null>
16	24097 METEO_Humidity_RH_	Влажность %	<null>	<null>
17	24098 METEO_Rossa_DEWPoint_	Точка россы	<null>	<null>
18	24099 METEO_AbsPressureAir	Абсолютное давление	<null>	<null>
19	24100 METEO_Brith_Max	Максимальное яркость КЛюкс	<null>	<null>
20	24101 METEO_Osadki_status	Осадки статус	<null>	<null>
21	24102 METEO_Osadki_mm_h	Осадки мм / час	<null>	<null>
22	24103 METEO_Osadki_mm_Day	Осадки мм в день	<null>	<null>
23	24104 METEO_Osadki_Type	Осадки тип	<null>	<null>
24	24105 METEO_Irradiation	Солнечная иррадиация	<null>	<null>
25	24106 METEO_Temperatuna_FeHm	Температура ФЭМ	<null>	<null>
26	24107 A1_SCU1_Daily_power_yields	Суточная выработка электроэнергии kWh	<null>	<null>
27	24108 A1_SCU1_Monthly_power_yields	Месачная выработка электроэнергии kWh	<null>	<null>
28	24109 A1_SCU1_Total_power_yields	Общая выработка электроэнергии kWh	<null>	<null>

Figure B2. “Variables” table

t	gid	mid	k	n	v
2024-03-04 18:12:53.563	-6	0	2	Second_test	Проект запущен
2024-03-04 18:13:04.286	-5	1	2	127.0.0.1 (307733914C6C4531802F306EEF28EE27)	Клиент подключен
2024-03-04 19:13:47.858	-5	2	2	127.0.0.1 (307733914C6C4531802F306EEF28EE27)	Клиент отключен, Длительность сессии: 01:00:00
2024-03-04 19:13:47.878	-6	3	2	Second_test	Проект остановлен
2024-03-04 22:54:40.591	-6	0	2	Second_test	Проект запущен
2024-03-04 23:55:34.111	-6	1	2	Second_test	Проект остановлен
2024-03-05 12:14:53.283	-6	0	2	Second_test	Проект запущен
2024-03-05 12:15:09.440	-5	1	2	127.0.0.1 (8B420AAF703B4706B1856D2A6A89BF49)	Клиент подключен
2024-03-05 13:15:47.049	-5	2	2	127.0.0.1 (8B420AAF703B4706B1856D2A6A89BF49)	Клиент отключен, Длительность сессии: 01:00:00
2024-03-05 13:15:47.049	-6	3	2	Second_test	Проект остановлен
2024-03-06 16:48:28.103	-6	0	2	Second_test	Проект запущен
2024-03-06 16:48:55.115	-5	1	2	127.0.0.1 (1C8F9A77A61C48CE942681D0A0E06478E)	Клиент подключен
2024-03-06 17:49:19.917	-5	2	2	127.0.0.1 (1C8F9A77A61C48CE942681D0A0E06478E)	Клиент отключен, Длительность сессии: 01:00:00
2024-03-06 17:49:19.937	-6	3	2	Second_test	Проект остановлен
2024-03-14 08:51:13.095	-6	0	2	Second_test	Проект запущен
2024-03-14 08:51:13.100	-5	1	2	127.0.0.1 (1D4E912A5CE349D180A28A75C40F23A3)	Клиент подключен
2024-03-14 09:23:37.734	-5	2	2	127.0.0.1 (1D4E912A5CE349D180A28A75C40F23A3)	Клиент отключен, Длительность сессии: 00:33:00
2024-03-14 09:23:37.751	-6	3	2	Second_test	Проект остановлен
2024-03-14 11:07:42.647	-6	0	2	Second_test	Проект запущен
2024-03-14 11:07:42.659	-5	1	2	127.0.0.1 (9F3E1C95B4A84914B3830BE66297D498)	Клиент подключен
2024-03-14 11:38:35.198	-5	2	2	127.0.0.1 (9F3E1C95B4A84914B3830BE66297D498)	Клиент отключен, Длительность сессии: 00:31:00
2024-03-14 11:38:35.220	-6	3	2	Second_test	Проект остановлен
2024-03-14 11:39:53.676	-6	0	2	Second_test	Проект запущен
2024-03-14 11:39:53.681	-5	1	2	127.0.0.1 (4EECB8447EB5415A9E590E374188B658)	Клиент подключен
2024-03-14 11:50:52.161	-5	2	2	127.0.0.1 (4EECB8447EB5415A9E590E374188B658)	Клиент отключен, Длительность сессии: 00:11:00
2024-03-14 11:50:52.472	-6	3	2	Second_test	Проект остановлен
2024-03-15 20:45:11.310	-6	0	2	Second_test	Проект запущен
2024-03-15 21:46:04.115	-6	1	2	Second_test	Проект остановлен
2024-03-16 03:42:30.106	-6	0	2	Second_test	Проект запущен
2024-03-16 03:42:30.111	-5	1	2	127.0.0.1 (135B981594D74892830650B989C8F6C7)	Клиент подключен
2024-03-16 03:43:42.573	-5	2	2	127.0.0.1 (135B981594D74892830650B989C8F6C7)	Клиент отключен, Длительность сессии: 00:01:00
2024-03-16 03:43:42.812	-6	3	2	Second_test	Проект остановлен
2024-03-16 17:38:37.685	-6	0	2	Second_test	Проект запущен
2024-03-16 17:45:39.654	-5	1	2	127.0.0.1 (440AB67C3DCC48558145BF8B792738B5)	Клиент подключен
2024-03-16 18:39:30.536	-5	2	2	127.0.0.1 (440AB67C3DCC48558145BF8B792738B5)	Клиент отключен, Длительность сессии: 00:53:00
2024-03-16 18:39:30.556	-6	3	2	Second_test	Проект остановлен
2024-03-18 14:24:57.977	-6	0	2	Second_test	Проект запущен
2024-03-18 14:26:13.759	-5	1	2	127.0.0.1 (ACF4E94F611448C6B896855AE94FA23)	Клиент подключен

Figure B3. “Messages” table